REPORT ON PROGRESS

Pulsar timing arrays: the promise of gravitational wave detection

To cite this article: Andrea N Lommen 2015 Rep. Prog. Phys. 78 124901

View the article online for updates and enhancements.

Related content

- GRAVITATIONAL WAVES FROM INDIVIDUAL SUPERMASSIVE BLACK HOLE BINARIES IN CIRCULAR ORBITS: LIMITS FROM THE NORTH AMERICAN NANOHERTZ OBSERVATORY FOR GRAVITATIONAL WAVES
  Z. Arzoumanian, A. Brazier, S. Burke-Spolaor et al.

- THE NANOGRAV NINE-YEAR DATA SET: LIMITS ON THE ISOTROPIC STOCHASTIC GRAVITATIONAL WAVE BACKGROUND
  Z. Arzoumanian, A. Brazier, S. Burke-Spolaor et al.

- LIMITS ON THE STOCHASTIC GRAVITATIONAL WAVE BACKGROUND FROM THE NORTH AMERICAN NANOHERTZ OBSERVATORY FOR GRAVITATIONAL WAVES
  P. B. Demorest, R. D. Ferdman, M. E. Gonzalez et al.

Recent citations

- Spin memory effect for compact binaries in the post-Newtonian approximation
  David A. Nichols

- Continuous gravitational wave searches with pulsar timing arrays: Maximization versus marginalization over pulsar phase parameters
  Yan Wang et al

- Harmonic Space Analysis of Pulsar Timing Array Redshift Maps
  Elinore Roebber and Gilbert Holder
1. The promise of gravitational wave detection

Gravitational waves represent a new phenomenon we can use to understand the universe. The 20th century brought us radio (including microwaves and millimeter waves), x-rays, infrared, and gamma-rays, but all of these tools are electromagnetic phenomena. The 20th century also brought the first promise of non-electromagnetic tools for studying the universe: cosmic ray and neutrino experiments, and finally gravitational waves. Gravitational waves will yield a profound new tool for astronomers. With gravitational waves we expect to detect the acceleration of mass in the universe, rather than the electromagnetic signature of the mass. For example, we will discriminate between galaxy formation scenarios (Volonteri et al 2003), and we can measure the spins of binary black holes directly (Mingarelli et al 2012, Vitale et al 2014).

After putting both the fields of pulsars and gravitational waves in historical perspective (section 2), we introduce the concept of gravitational waves and describe briefly current gravitational wave experiments in section 3. Section 4 is the largest section of the manuscript and describes in detail how pulsars can function as gravitational wave detectors. It is divided into a number of subsections including the concept of a pulsar as a clock (section 4.1), the idea of an array of pulsars (section 4.2), current issues in pulsar timing (section 4.3), including the ways in which the fitting to the timing model absorbs gravitational waves (section 4.3.1), employing wide-bandwidth systems effectively (section 4.3.2), and characterizing and limiting noise contributions (section 4.3.3).

In section 5 we calculate the response of a pulsar arrival time to a gravitational wave. We discuss the sources of gravitational waves in section 6, and what sources we expect to detect first in section 7. Section 8 describes various detection schemes and the concerns around them, including the important question, ‘What constitutes a detection?’ (section 8.7). Section 9 examines the possibility of the expansion of pulsar timing arrays (PTAs) in the future and includes a discussion of the Astronomy 2010 Decadal Review’s praise for the experiment (section 9.1), the uncertain futures of both the Green Bank and Arecibo telescopes (section 9.2), the SKA (section 9.3.1), FAST (section 9.3.2), and CHIME, MeerKAT, Molonglo Radio Observatory, Deep Space Network and ARTS (section 9.3.3), and a telescope dedicated to Gravitational waves (section 9.3.4). Section 10
discusses the possibility of optimizing the experiment to maximize the signal. Finally, we conclude in section 11 with thoughts about the future of the enterprise of pulsar timing for gravitational wave detection.

2. Gravitational waves and pulsars: a lesson in parallel history

Gravitational wave detection using pulsars combines two fields that have rich histories, each with its own moment of discord. Both Anthony Hewish and Joseph Weber planned and built their experiments in the 1960s, with the most famous announcement from both fields coming at the end of that decade.

Hewish could not have called himself a pulsar astronomer, though that is how he is now known, since the field did not yet exist. Hewish’s ‘Interplanetary Scintillation Array’ was built with the help of Jocelyn Bell who saw the first pulsar on a chart recorder in 1967. Bell and Hewish announced the discovery in 1968 (Hewish et al 1968).

A similar pioneer, Joseph Weber, could not have called himself a gravitational wave physicist when he began working on gravitational waves, because the field did not yet exist. Weber constructed the first ‘Weber bar’ and announced evidence for the discovery of gravitational waves about the same time (Weber 1969).

Both discoveries caused discord for different reasons: Hewish’s because he was later (1974) awarded the Nobel prize for the discovery, and Jocelyn Bell was not; Weber’s because his detection could never be replicated.

These fields, however, have moved far beyond these initial detections, and both owe much to their pioneers. We now know of more than 2300 pulsars1. The first extrasolar planets were discovered around pulsars (Wolszczan and Frail 1992), and pulsars have been used to study galactic structure (e.g. Cordes and Lazio 2002), stellar evolution (e.g. Kiel et al. 2008), and the equation of state of very dense matter (e.g. Poddadiowski et al 2005, Miller 2013). The field of gravitational wave detection now includes multiple functioning ground based interferometers (e.g. LIGO Scientific Collaboration and Virgo Collaboration et al 2013, The LIGO Scientific Collaboration and Virgo Collaboration et al 2014), plans for space-based experiments (Amaro-Seoane et al 2012, the eLISA Consortium et al 2013), PTA experiments, and cosmic microwave background experiments (BICEP2 Collaboration et al 2014). See the GWIC roadmap at https://gwic.ligo.org/roadmap.

These two fields first began to share ground when Hulse and Taylor (1975) discovered the first binary pulsar and showed its orbital decay to be precisely consistent with the emission of gravitational waves. Now, in the first decades in the 21st century, PTAs will potentially be one of the first experiments to detect gravitational waves. In the remainder of this review we will explore how pulsars work as timekeepers, how PTAs aim to use them to detect gravitational waves, and how the international community is working together to make this happen in the next decade.

---

1 www.atnf.csiro.au/people/pulsar/psrcat

3. Gravitational waves and gravitational wave experiments

Gravitational waves are traveling perturbations in the fabric of space-time due to accelerating masses, much like electromagnetic waves are perturbations in electric and magnetic fields due to accelerating charges. Gravitational waves are a natural consequence of the Theory of General Relativity (GR). Einstein predicted the existence of them (Einstein 1916), but he also claimed that the perturbations would be so small that we should not bother looking for them. Now, 100 years later, several groups are on the brink of directly detecting the perturbations.

The only aspects of GR needed to predict the existence of gravitational-waves are (1) that information cannot travel faster than the speed of light and (2) that mass curves space-time. Consider two black holes, each perturbing space-time, orbiting around each other. The curvature at any particular point in space must change in response to the moving mass, but the space-time distortion cannot travel faster than the speed of light. The gravitational wave is the changing curvature propagating through space-time, a perturbation in the metric transmitting the news that the mass has moved.

Mathematically, for small perturbations, we can separate the metric describing the curvature of space $g_{\mu\nu}$ into two parts: $\eta_{\mu\nu}$, a flat spacetime metric, and $h_{\mu\nu}$, the gravitational wave perturbation.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}.$$  (1)

We write the spacetime metric as $g_{\mu\nu}$, $(\mu, \nu = 0, 1, 2, 3)$, where 0 is typically the temporal dimension, t, and 1, 2, 3 are the spatial dimensions x, y, z. We then say that for a small metric perturbation, $h_{\mu\nu}$, where $h_{\mu\nu} << g_{\mu\nu}$, we can approximate spacetime as being flat with a small perturbation. We write this as the sum of the Minkowski metric, $\eta_{\mu\nu}$, plus the perturbation, $h_{\mu\nu}$, such that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, called the ‘linear approximation’, which is correct to first order in $h_{\mu\nu}$. The perturbation (or gravitational wave) $h_{\mu\nu}$ affects the proper distance between points, and as such affects the light travel time between those points. In most cases, gravitational wave detectors look for changes in the light travel time between two points, known as the ‘arm’ of the detector. The changes in proper length that gravitational waves are expected to produce in the arms of the various experiments are between one part in $10^{25}$ to one part in $10^{15}$, depending on the source properties and frequency range being considered.

Many detectors utilize interferometric techniques to detect the difference in phase between two laser paths, aka ‘arms’. This includes ground-based detectors such as the Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO), aVIRGO, GEO, TAMA, (Ando and TAMA Collaboration 2002), Affeldt et al (2014), Acrernese et al (2015) and The LIGO Scientific Collaboration et al (2015)) the Kamioka Gravitational Wave Detector (KAGRA) (Somiya 2012), and planned space-based detectors such as the Evolved Laser Interferometer Space Antenna (eLISA) (Amaro-Seoane et al 2013), the Deci-Hertz Interferometer Gravitational wave...
Observatory (DECIGO) (Kawamura et al. 2011) and the Big Bang Observer (BBO) (Harry et al. 2006). Pulsar timing experiments such as those described just below in section 3.1 also look for phase changes as light travels along an ‘arm’ but the arm in this case is the path the light takes as it travels from a pulsar to the Earth. Every scheme creates a detector with long arms, such that the very small fractional change (1 part in $10^{22}$ for example) in the length of the arm is actually measurable. In LIGO for example, 1 part in $10^{22}$ of a 4 km arm is smaller than any atomic distance (it is 0.0004 femto-meters). It is not measured as a distance, but rather as a time ($10^{-27}$ s), an interval that is actually possible to measure via interferometric techniques. In pulsar timing the arms are longer ($10^{19}$ m) and the fractional change is larger (1 part in $10^{15}$) so the length change is larger ($10^4$ m). Owing to the short-wavelength regime of PTAs the typical time change we search for is only a fraction of the 30 μs it takes light to travel $10^4$ m. The time changes PTAs expect to detect are typically smaller than a microsecond. Measuring a delay of a microsecond represents a significant challenge for pulsar timing.

Figure 1 shows as a function of frequency the various sources of expected gravitational waves (in color) and the sensitivity of each of the gravitational wave experiments in black lines. The sensitivity curves for interferometers are appropriate for individual sources. For more information on the string models please see Sanidas et al (2013) and for more information on relic backgrounds see the summary in Lentati et al (2015a).

3.1. Existing PTAs

There are three principle PTA collaborations that have ongoing pulsar timing programs designed to detect gravitational waves. The European Pulsar Timing Array (EPTA) is comprised of researchers across Europe who use five 100 m class dishes: the Lovell Telescope, the Nancay Radio Telescope, The Sardinia Radio Telescope, the Westerbork Radio Synthesis Telescope, and the Effelsberg Radio Telescope (see for example Lentati et al (2015a)). Members of the EPTA run the telescopes separately but also in a phased array mode as part of the Large European Array of Pulsars (LEAP) program (Kramer and Champion 2013). When the five dishes are in tied-array mode, the sensitivity is better than the Arecibo Observatory 300 m telescope (see below).

The North American Nanohertz Observatory of Gravitational Waves (NANOGrav) is comprised of researchers at many institutions across the United States and Canada and uses the world’s two largest single-dish radio telescopes, the National Astronomy and Ionosphere Center Arecibo Observatory 300 m William E. Gordon Telescope (hereafter the Arecibo Telescope) and the National Radio Astronomy Observatory Robert C Byrd Green Bank Telescope with its 100 m dish (hereafter the Green Bank Telescope) (Jenet et al. 2009).
The Parkes Pulsar Timing Array (PPTA) is principally an Australian collaboration but collaborates internationally and uses the Parkes 64 m dish in Australia for its regular timing program (Hobbs 2013).

At each one of the telescopes mentioned above the corresponding PTA collaboration has a regular timing program with a cadence between 1 week and 1 month. All have been timing pulsars for more than five years with the exception of the Sardinia Radio Telescope which was commissioned in 2012.

These three collaborations together comprise the International Pulsar Timing Array (IPTA) which is a consortium of consortia. The IPTA has multiple working groups that span the three collaborations (NANOGrav, the PPTA, and the EPTA). Under the IPTA umbrella the three collaborations share a common data set, with a constitution, data-sharing agreement, and authorship policy that governs the exchange of information and the subsequent publications that emerge (see www.ipta4gw.org). Not only does this maximize the scientific output to be gained from the three data sets, but it enhances global partnerships in science (see http://nanograv.org/pire/).

3.2. Other GW experiments

LIGO, its partner Virgo, and other planned gravitational wave experiments such as the New Gravitational-wave Observatory (aka eLISA) and KAGRA, are also actively pursuing the detection of gravitational waves in frequency bands complementary to each other and to pulsar timing (Abbott et al 2009, Kanda and the LCGT collaboration 2011, Amaro-Seoane et al 2012, the eLISA Consortium et al 2013). Advanced LIGO, scheduled to reach design sensitivity in 2018, may detect gravitational waves as early as the first year of operation (LIGO Scientific Collaboration et al 2013). LIGO is sensitive to frequencies from 10 Hz to 10kHz making it sensitive to black-hole binaries of up to 50 solar masses, black hole-neutron star binaries, supernovae, and spinning neutron stars, and has also been designed to be able to measure an astrophysical background at those frequencies (The LIGO Scientific Collaboration et al 2015). The frequency range to which PTAs are sensitive is 10 orders of magnitude away in the nanohertz range, as discussed in section 6.

4. Using pulsars to detect gravitational waves

4.1. Pulsars as clocks

We have only hinted so far at the concept of a pulsar as a clock. In this section we conduct a review of the concept of pulsars as cosmic timekeepers, and begin with a general introduction to pulsars.

A pulsar is a dead, dense, spinning remnant of what was formerly a massive star. The massive star exploded in a supernova and left behind about 1.4 solar masses of dense material consisting primarily of neutrons, thus a neutron star. A radio beacon is emitted out the star’s magnetic poles which are misaligned with the spin axis. Therefore, if the beacon is oriented a particular way, we see a pulse every time the beam sweeps by our line of sight. These pulses arrive with extraordinary regularity, and thus a pulsar can be thought of as a celestial clock.

Our ability to use a pulsar as a clock depends upon our ability to predict the arrival times of its pulses. In the best-timed millisecond pulsars we can predict the arrival time of the next pulse to within 100 ns over a five year period (e.g. Verbiest et al 2008). The remainder of this section is devoted to explaining what a challenge that is to achieve.

Starting from some initial time, \( t_0 \), e.g. the discovery of the pulsar, the objective is to predict the arrival time of future pulses. A simple model for the arrival time of a pulse is

\[
  t = t_0 + NP - \frac{1}{c} \hat{n} \cdot \mathbf{D} + R_{gw}(t),
\]

(3)

where \( P \) is its period, \( N \) is the number of turns the pulsar has made since \( t_0 \), \( \hat{n} \) is the unit vector pointing from the solar system barycenter (SSB) to the pulsar, \( \mathbf{D} \) is the vector pointing from the SSB to the telescope, and \( R_{gw}(t) \) is the hypothetical gravitational wave contribution (described in detail in section 5). In a real observation, the pulsar model is often much more complex, taking into account changes in the pulsar period as the pulsar loses energy, additional geometrical effects if the pulsar itself has an orbiting companion, etc. We ignore these terms to demonstrate the basic approach. For a complete treatment of equation (3) see Lorimer and Kramer (2012).

We expect \( R_{gw}(t) \) to be smaller than the other terms, so we need to carefully model the other terms in order to be sensitive to gravitational waves.

Earth’s motion around the sun enters into the equation in the dot product \( \hat{n} \cdot \mathbf{D} \). If the pulsar is near the north or south ecliptic pole, the dot product is close to zero, the light travel time from pulsar to Earth essentially does not change regardless of where the Earth is in its orbit. By contrast, if the pulsar is near the ecliptic plane, the light travel time changes by about two astronomical units over the course of a year. While this change is not large compared to the typical distances to pulsars (hundreds to thousands of light years), two astronomical units corresponds to 16 light-minutes, and 16 min is orders of magnitude larger than the typical pulsar period.

There is a fundamental chicken and egg problem in pulsar timing that goes as follows. The pulsed signal is buried in the noise for most pulsars, i.e. most single pulses are undetectable. Thus, in order to detect a pulse, one must add together many thousands of pulses over many thousands of turns of the pulsar. In order for this procedure to yield an accumulated pulse profile with a high signal to noise ratio, one must have a model that describes how to ‘fold’ the pulsar data onto itself. In order to find this model, one must be able to detect a pulse (Lorimer and Kramer 2012).

The chicken and egg problem is solved iteratively. When a pulsar is first discovered, the observer usually has a rough idea of its spin frequency by the location of the peak in the fast Fourier transform in which it was discovered. This approximate spin frequency is used to accumulate pulses in small batches and acquire three or four times of arrival (TOAs) for the pulsar. A model is fit to the TOAs and the new model is
used to fold the next set of data that comes in, and the process iterates. Depending on the pulsar “backend”\(^2\) one may be able to go back to the original data and re-fold the data using the improved model. Employing this iterative process results in a good timing model within a year, i.e. one that predicts the arrival time of the next pulse to within a few milli-seconds (0.001 times the pulse period) (e.g. Verbiest et al. 2009, Demorest et al. 2013).

The iterative process continues throughout the data-span of the pulsar, but the model is updated much less frequently than at first, perhaps every couple of years. For example, it is only after a couple of years of data-taking that one is able to fit for proper motion and parallax, so at that moment those parameters are added to the model and the model improves (Lommen et al. 2006).

The difference between the time the pulse is expected to arrive as determined from the model and the time it is measured to arrive is called the ‘residual.’ Thus, if a model is perfect and there is no noise and no gravitational wave signal in the data the residuals would all be zero. More realistically the residuals will have a mean of zero, and a Gaussian distribution that depends upon the amplitude of the measurement noise. (The observed noise is actually more complicated than this and is discussed further in section 4.3.3.) Some typical pulsar residuals are shown in figure 2. For much more detail on pulsar timing techniques see Lommen and Demorest (2013).

4.2. A pulsar timing array

A system of well-timed millisecond pulsars is a system of clocks, and can be used to measure space-time disturbances such as gravitational waves. Various things can perturb the pulse arrival times (the interstellar medium (ISM), calibration issues, the shape of the pulse profile changing, etc), but only gravitational waves will produce the specific arrangement of geometric timing offsets that we delineate in section 5. Such a system was termed a ‘pulsar timing array’ by Backer et al. (1983). According to current models PTAs will detect gravitational waves within the next decade (Demorest et al. 2009, Siemens et al. 2013). This prediction is based on current estimates of the amplitude of the stochastic background (see section 8) and on the ability of PTAs to increase their sensitivity over time (1) by increasing the baseline of the observations (by simply continuing to observe the pulsar regularly), (2) by tackling issues of noise reduction (see sections 4.3.3 and 10.1), and (3) by adding more pulsars to the array.

Pulsar timing is most sensitive to gravitational wave frequencies near 10 nHz ≈ f(Dataspan) (Demorest et al. 2013). The sources in this range are massive (>10^8 M☉, where M☉ is the mass of the sun) black hole binaries in stable orbits with periods of a few to tens of years (Rajagopal and Romani 1995, Jaffe and Backer 2003, Jenet et al. 2006). These would be continuous sources of gravitational waves, emitting over the span of many thousands of years, and may be detected either individually or in an ensemble as a stochastic background. Other sources of stochastic background gravitational radiation in the pulsar timing band include cosmic string networks and a primordial relic background (Grishchuk 2005, Sanidas et al. 2013). In addition PTAs may detect bursts of gravitational waves, where a ‘burst’ is anything shorter than the data span. These could come, for example, from highly eccentric black hole binaries near perigee (Finn and Lommen 2010) or from cosmic strings (Damour and Vilenkin 2001, Sanidas et al. 2013). We discuss sources in more detail in section 6. PTAs are shrewd gravitational wave detectors, capable of deciphering the direction and waveform of individual sources, and also capable of measuring the amplitude and spectrum of the gravitational wave background, thereby determining the rate at which galaxies merged as the universe evolved (e.g. Volonteri et al. 2003, Lommen 2012).

\(^2\) ‘Backend’ refers to the equipment at the telescope that is used to process the voltages that come in from the ‘front end’ or telescope. Descriptions of backends can be found in Cognard et al. (2009), DuPlain et al. (2008a), Karuppusamy (2011), Karuppusamy et al. (2008), Kramer et al. (1999), Sarkissian et al. (2011) and Volite et al. (2002).
4.3. Current issues in pulsar timing relevant for gravitational wave detection

There are several issues currently in pulsar timing that are also issues in using pulsars for gravitational wave detection: timing model fits, wide-bandwidth backends, and noise.

4.3.1. Fitting to the pulsar model absorbs gravitational waves. First, we cannot independently, i.e. apart from pulsar timing, measure the parameters of the pulsar system, so our fits to the pulsar model can absorb gravitational waves. For example, two of the most important fit parameters are the rotation period of the pulsar and the derivative of the rotation period (it turns out that pulsars slow down over time). A fit to period and period derivative is essentially a second order polynomial fit subtracted from the residuals. So in all cases we are subtracting a polynomial from the residuals, but we have no way of knowing whether the polynomial we are subtracting represents an increase in the precision of the pulsar timing model, or effectively a removal of the effects of a gravitational wave. While we do not expect any single source of gravitational waves to have a polynomial signature, the accumulation of many such sources in the gravitational wave background (see section 6.3) could indeed have a polynomial signature in the finite time-span of a pulsar’s residuals. The absorptive properties of fitting are discussed in the context of the stochastic background in section 8.2.

The principal solution to this problem is in the combination of multiple pulsars for any detection scheme. A signal in one pulsar can and will be inadvertently partially subtracted as described above, but the remaining signal present in multiple pulsars can be found by searching simultaneously in all the pulsars. This is the basis of several detection schemes described later in section 8. This scheme is complicated by the fact that a pulsar’s response to a gravitational wave has two terms, one that is associated with the gravitational wave at the Earth and is therefore in common among the pulsars and one that is associated with the gravitational wave at the pulsar and is therefore different for each pulsar. The latter, the so-called ‘pulsar term’ effectively adds noise to the scheme, because its phase depends upon the pulsar distance which we do not know well. These two terms are described mathematically in section 5.

There is another option for consideration, however, which is to determine the pulsar parameters in an independent way. Using VLBI Chatterjee et al. (2004) and Madison et al. (2013), for example, have demonstrated the ability to determine a pulsar’s position, i.e. RA and DEC, and its proper motion sufficiently well to constrain those parameters in the pulsar model (also see Brisken et al. 2002). These parameters can then be fixed in the pulsar timing model. However, there is no way (yet) to determine period and period derivative independently, and those will provide a larger effect in the pulsar residuals. The ability to do so would help the effort enormously. Simulations show that the gravitational wave background is likely to be adding tens of μs of signal to the residuals that is being ‘absorbed’ by the fit to period and period derivative. Please see the simulations in section 8.2 and figure 7 for an illustration of this.

4.3.2. Wide-bandwidth backends. The second key issue in PTA’s derives from the enormous gains in receiver bandwidth that have recently been made at radio telescopes. For example, ten years ago the largest bandwidth detectable at the Green Bank Telescope was 64 MHz. Now the bandwidth is 800 MHz at most frequencies (DuPlain et al. 2008b). This is both a tremendous advantage and a tremendous challenge for pulsar timing (van Haasteren et al. 2011). The advantage comes from our ability to accumulate signal from a particular pulsar. The signal-to-noise ratio (ρ) in a particular pulsar follows a simple scaling relationship

\[ \rho \propto A \sqrt{\frac{\tau \Delta f}{T_{\text{sys}}}} \]  

(4)

where A is the collecting area of the telescope, \( \tau \) is the integration time, \( \Delta f \) is the bandwidth, and \( T_{\text{sys}} \) is the system temperature of the receiver (Lorimer and Kramer 2012). Therefore quadrupling the bandwidth is the same as doubling the collecting area of the telescope, or quadrupling the observing time, the latter two of which come at a very high cost. So the increase in the last decade produced the same effect as a 5-fold increase in our observing time, a gain that would have been impossible by requesting more time through the telescopes’ time allocation processes. In addition Shannon et al. (2014) demonstrate that wide-bandwidth timing will help reduce the effects of jitter noise, intrinsic to the pulse emission mechanism, as discussed in section 4.3.3.

The issue is that a pulsar’s shape changes with frequency, and over 800 MHz it can change substantially as shown in figure 3 (Pennucci et al. 2014). To understand why this is important we must describe the way an arrival time is determined. This is explained in detail in Lommen and Demorest (2013) and is sketched here. The pulse itself is not a delta-function but commonly occupies 3–10% of the pulse period and can occupy as much as 100% (see Henry and Paik (1969) and Maciesiak et al. (2011) and references therein). So does the arrival correspond to the arrival of the leading edge? the trailing edge? or something in between? In order to time the arrivals of these pulses to better than 0.1% of the pulsar period (which is in fact what we accomplish) we actually time the arrival of the entire shape of the pulse, and rely upon the shape being the same from one measurement to the next. The standard procedure had been to simply add up the pulse profile over all the bandwidth and to thus use an ‘average’ shape. As long as the average shape is constant, there is no issue. However, scintillation makes different parts of the bandwidth brighter at different times. In figure 3 imagine different ‘sub-bands’ being relatively brighter at different times. As a result the average shape changes over time. This can cause a change in arrival time of several microseconds, many times larger than the hundred-nanosecond gravitational wave experiment signal we look for.

3 Scintillation is a focusing and defocusing of radiation due to multi-path propagation through the ISM, and results from density inhomogeneities in the ISM (Stinebring et al. 1996). For a useful review of ISM effects on radio waves see Rickett (1990), and for a quantitative assessment of their effects on the timing stability of pulsars see Cordes (2013).
One solution is to break the large bandwidth into numerous sub-bands as shown in figure 3 (this is in practice often done by the observing backend) and time each sub-band separately. As long as the sub-band width is smaller than any scintillation feature the shape of the profile within each sub-band is stable and the TOA is not dependent on the scintillation. Instead of a single TOA the analysis yields one TOA for each sub-band. The problem of combining the numerous TOAs into a single TOA is substantial and has been addressed in numerous ways. Demorest et al. (2013) allowed the relative alignment of each of the sub-bands to be a fitted parameter in the global fit. More details can be found in Lommen and Demorest (2013).

Another recently explored idea is to create a two-dimensional template (intensity versus pulse phase versus frequency) rather than creating a one-dimensional template (intensity versus pulse phase) at a number of different frequencies. The phase of the template is adjustable smoothly as a function of frequency, but only adjustable perhaps according to the cold-plasma dispersion relation that governs the arrival time of pulses as a function of frequency. A model like this has been demonstrated to work on simulated data using Bayesian analysis by Messenger et al. (2011) and on real data by both Pennucci et al. (2014) and Liu et al. (2014) using $\chi^2$ minimization to optimize the solution. This scheme has the advantage of elegance, i.e. one only needs a single template for a single data scan no matter how wide the bandwidth of the scan.

4.3.3. Noise. The third important issue is that there exists noise intrinsic to the pulsars themselves, both white and red.

That there is white noise is not surprising, but that some of it results from the emission mechanism itself and is not all radiometer noise is a new finding by Shannon et al. (2014) and Osłowski et al. (2014). Both studies suggest that phase jitter resulting from bright single pulses plays an important role in timing models, and therefore in gravitational wave detections. Both show a significant improvement in goodness-of-fit $\chi^2$ when jitter noise is properly accounted for in fitting timing solutions.

The literature is not in agreement about the extent of the red noise. Handzo et al. (2015) found the pulsar residuals to be largely dominated by white noise, whereas (Keith et al. 2013) using PPTA data and a significantly different dispersion measurement technique found red noise to be present in nearly all pulsars.

In any case, all agree that red noise is present in some number of millisecond pulsars, and that the red noise decreases our ability to detect gravitational waves because its spectrum mimics that of the gravitational waves we aim to detect. For a nice summary of the effect of red noise on our ability to detect the stochastic background see figure 4 from Siemens et al. (2013), the most important conclusion of which is that the time required to detect gravitational waves is longer by several years.

However, red noise is not correlated among the pulsars in the array and the gravitational wave signal is, so our ability to distinguish the signal from the noise always involves multiple pulsars. For example for burst sources where the timescale of the burst is much shorter than the timescale of the red noise the red noise is negligible, we only need to combine the signal from a few pulsars to distinguish between the noise and the gravitational wave (Finn and Lommen 2010, Christy et al. 2014). The results of van Haasteren et al. (2009) and Siemens et al. (2013) demonstrate that for the stochastic background it is more important to time many pulsars.

5. Calculating the response of a pulsar to a gravitational waves

Here we describe mathematically the response of a pulsar to a gravitational wave, adapted from Jenet et al. (2004).

Assume a gravitational wave is propagating through space due to some distant source such as a supermassive black hole binary. The amplitude of the gravitational wave, $h$, is measured in units of dimensionless strain, where strain is the fractional change in length (or light travel time) that the gravitational wave causes. The gravitational wave strains we expect to detect in pulsar timing are of order $10^{-15}$. In other words, over a one-light-year path a gravitational wave with $h = 10^{-15}$ would change the light travel time by $\sim 30$ ns (assuming the wavelength of the gravitational wave were longer than a light year). The relationship between the amplitude $h$ and the

Figure 3. A simulated pulsar showing evolution of the pulse profile with frequency (in MHz), and two different possible alignments of the pulse profile over frequency. The choice of alignment of the sub-bands will render different average profiles. In addition, scintillation makes different sub-bands brighter in different observing epochs and will also change the average profile. Figure originally appeared in Lommen and Demorest (2013) and is used here in accordance with the Creative Commons Attribution 3.0 Unported License.
residual effect in a pulsar due to a gravitational wave $R_{gw}$ is as follows. The size of the change at time $t$ for a pulse from a pulsar (in other words, the residual due to the gravitational wave) is given by

$$R_{gw}(t) = \frac{1}{2}(1 + \cos(\mu))(r_+,(t) \cos(2\psi) + r_\times,(t) \sin(2\psi)),$$

where the gravitational wave source is located an angle $\mu$ away from the pulsar, and the source has polarization angle $\psi$. The subscripts $+$ and $\times$ refer to the two possible polarizations of the gravitational waves. The functions $r_+$ and $r_\times$, a geometric combination of which produces the residual effect, referred to collectively as $r_{+,\times}$, are related to the gravitational wave strain by

$$r_{+,\times}(t) = r_{+,\times}^e(t) - r_{+,\times}^p(t),$$

$$r_{+,\times}^e(t) = \int_0^t h_{+,\times}^e(\tau)d\tau,$$

$$r_{+,\times}^p(t) = \int_0^t h_{+,\times}^p(\tau - \frac{d}{c}(1 - \cos(\mu)))d\tau,$$

where $h_{+,\times}^\pm(t)$ is the gravitational wave strain at Earth (usually called the ‘Earth term’), $h_{+,\times}^p(t)$ is the gravitational wave strain at the pulsar (usually called the ‘Pulsar term’), $\tau$ is the time integration variable, $d$ is the distance between Earth and the pulsar, and $c$ is the speed of light.

Note that the pulsar term, $h_{+,\times}^p(t)$, is the same function as the Earth term evaluated at a geometrically delayed time. The delay is equal to the time between two events: (1) the gravitational wave arriving at the Earth and (2) the information that the gravitational wave has arrived at the pulsar arriving at the Earth. Hence, $r_{+,\times}^p$ is induced by waves emitted at a later epoch than those that induce $r_{+,\times}^e$.

Thus, $R_{gw}(t)$ is the perturbation in the residuals due to the gravitational wave, and in a perfect world $R_{gw}(t)$ would be the difference between the predicted arrival time and the measured arrival time, i.e. the residual, at time $t$. However, for reasons already elucidated such as the absorption of the gravitational waves into the timing model fits, red noise, and deviations caused by the issues of timing large bandwidths of signal, the residuals are not yet dominated by the gravitational wave contribution, $R_{gw}(t)$.

Except for a time delay between them, the Earth and pulsar terms are identical in the majority of the sources in the circular, slowly evolving, non-spinning approximation. The amount of the delay depends geometrically on the angle between the pulsar and the Earth and on the distance to the pulsar. So for a particular gravitational wave the Earth term is identical among all the pulsars in an array, while the pulsar terms are all delayed by different amounts. Therefore the pulsar term adds noise unless the delay can be determined and accounted for. This may be possible in a few cases (Corbin and Cornish 2010), the challenge being that pulsar distances are only known to about 20% (20–200 pc) even with SKA parallax measurements (Smits et al 2011), and for the distance to be useful we must know them to a fraction of a gravitational wavelength, roughly a fraction of a light-year. In some cases, though, we may be able to determine the pulsar distance using VLBI with sufficient accuracy (Chatterjee et al 2004, Madison et al 2013). Alternatively, a single very bright gravitational wave source would allow us to solve for the distances to all the pulsars.

It is important to get a sense of the shape of the response as delineated by equations (5) through (8), because the shape...
helps distinguish a gravitational wave from other possible perturbations of the residual. First, the left panel of figure 4 shows the absolute value of the Earth term of the pulsar response to a singular polarization of a gravitational wave traveling in the +z direction. Distance from the origin is proportional to the size of the response and the two different colors in the left panel have opposite signs. The response has an overall decrease in amplitude given by $(1 + \cos \mu)$ in going from a situation where $\mu = 0$ (the pulsar and the gravitational wave source are aligned), to a situation where $\mu = 180^\circ$ and the two are opposite each other in the sky. In the figure, pulsars nearly aligned with the source are down near the bottom of the shape where the response reaches a maximum value. The response decreases for pulsars toward the top of the figure, in other words, for pulsars away from the gravitational wave source. At $90^\circ$ away the amplitude is down by a factor of two and at $180^\circ$ away it goes to zero (when the pulsar and gravitational-wave source are anti-aligned). Therefore the pulsar being nearly aligned with the gravitational wave source is optimal for detection. However, when the two are precisely aligned $R_{GW}(t) = 0$ because the electromagnetic wave is essentially ‘surfing’ on the gravitational wave and no perturbation is produced. This effect is seen mathematically not in the $(1 + \cos \mu)$ term but in the difference between the Earth and pulsar terms (equation (6)) which is identically zero in the case of an exactly aligned source. For practical purposes this case is ignorable, because a small misalignment (an arc sec) renders this point moot. The misalignment only has to be large enough for the gravitational wave to pass some fraction of a wavelength away from the pulsar (i.e. about a light year) for $R$ to be substantially different from 0.

The effect of the cos $2\psi$ is seen in the four-lobed shape of figure 4 (again, only the + polarization is plotted so the reader can see the shape). There are nulls in the response at specific locations in the sky as a result of the polarization of the source. The implications of this shape and the way in which it gives rise to the so-called ‘Helling’s and Downs curve’ are described in section 8.1.

The right-hand panel of figure 4 shows the ‘antenna pattern,’ i.e. the response of a pulsar at the $-z$ location to an unpolarized gravitational wave coming from any direction in the sky. Distance from the origin indicates the magnitude of response to a wave from that direction.

6. The sources of gravitational waves in the PTA band

Gravitational waves are caused by the acceleration of mass as per the discussion in section 1 so there are many potential sources, but only a subset of them emit in the nanohertz regime where PTAs are most sensitive. We first describe the quadrupole approximation which will allow the reader to make estimates of the residual response to a variety of astrophysical sources. We go on to discuss what sets the frequency regime of PTAs, followed by a description of each of the types of sources in the PTA band: massive black hole binaries, cosmic strings, and a relic background.

6.1. The quadrupole approximation

Binary systems made of compact objects (neutron stars, black holes) are one of the most efficient ways to produce gravitational radiation, and we know that these sources exist at many mass ranges from stellar-mass neutron star binaries to billion-solar-mass black hole binaries (see for example Burgay et al (2003) and Merritt and Ferrarese (2001)). In the pulsar timing regime the main source is likely to be super-massive (billion-solar-mass) black hole binaries where the amplitudes in dimensionless strain are near $10^{-15}$ and the wave-periods are in years. To get a sense of the mechanics of creating such an amplitude, and a sense of the dependence of the amplitude on period and mass, we lead the reader through a simplified version of the derivation of the amplitude $h$ and then of the residual $R_{GW}$ for a simple binary system. The sources PTAs consider are far from coalescence and therefore in wide, non-relativistic orbits. In this limit, the energy density is dominated by the rest-mass density, and Hartle (2003) derives the gravitational wave metric perturbation far from a weak, non-relativistic source in the long wavelength approximation as

$$h = \frac{2G}{c^3d} I$$

(9)

where $h$ is the amplitude of a small perturbation of the metric of flat space-time, $d$ the distance to the source, $G$ the gravitational constant, and $c$ the speed of light. $I$ is the second mass moment and $\hat{I}$ is its second time derivative, also called the mass quadrupole moment. This is the so-called ‘quadrupole’ approximation.4

Most gravitational-wave astronomers would at this point move into geometrized units, where the base physical units are chosen such that both $G$ and $c$ can be set to unity, and mass and distance are in units of time. This is useful when working extensively with General and Special Theories of Relativity, but for the sake of clarity we will work exclusively in the International System of Units (SI).

A typical component of the second mass moment (a tensor) is

$$I \sim m_1 r_1^2 + m_2 r_2^2$$

(10)

where $m_1$ and $m_2$ are the two masses in orbit, and $r_1$ and $r_2$ their separations from the axis of rotation, presumed to be sinusoidal functions of time. To keep the bodies in motion, if $m_1 \leq m_2$, then $r_2 \leq r_1$. For a binary system the axis of rotation naturally coincides with the center of mass of the two bodies.

The second mass moment can also be written in terms of the total mass $M(= m_1 + m_2)$ of the system and the mass ratio $q(\equiv \frac{m_1}{m_2})$, as

$$I \sim qMr_1^2.$$

(11)

The two derivatives of $I$ give

$$\dot{I} \propto \frac{qMr_1^2}{P^2}$$

(12)

4 Note that both theoretical and numerical waveform calculation has progressed enough that nearly any waveform can be determined without the use of this approximation. See e.g. Chatziioannou et al (2013).
where \( P \) is the period of the sinusoid in \( r_1 \) and \( r_2 \). Since this system is non-relativistic we can use Newton’s form of Kepler’s Third Law \((a = \sqrt{\frac{GM}{4\pi^2}})\) to remove \( r_1 \) as this parameter can be very difficult to estimate for a distant SMBHB system. Keeping in mind that the radius of a circular orbit, is simply half the sum of \( r_1 \) and \( r_2 \), we can substitute \( r_1 = \frac{2a}{1 + q} \) into Kepler’s Law to get \( r_1 = \frac{\sqrt{GM}}{4\pi^2} r_2 \), which yields
\[
I \sim \frac{G^{5/3} M^{2/3}}{P^{2/3}} \frac{q}{(1 + q)^2}.
\] (13)

Rewriting \( h \) using this equation, we get
\[
h \sim \frac{G^{5/3} M^{2/3}}{P^{2/3} c^4} \frac{q}{(1 + q)^2}.
\] (14)

A more precise derivation yields:
\[
h \sim (2)^{5/3} 3^{2/3} \frac{G^{5/3} M^{2/3}}{P^{2/3} c^4} \frac{q}{(1 + q)^2}.
\] (15)

where we are using \( \sim \) here to indicate that this is just a typical component of the matrix \( h \).

Writing the above equation in a more useful form, and absorbing the constants \( G \) and \( c \), we get
\[
h \sim 7.6 \times 10^{-14} \left( \frac{M}{10^9 M_\odot} \right)^{5/3} \frac{1 \text{ yr}}{P} \frac{100 \text{ Mpc}}{d} \frac{q}{(1 + q)^2}.
\] (16)

The strain \( (h) \) produced by the periodic source of the gravitational-wave now induces a regular shift in the pulse arrival times of signals that travel through the gravitational radiation field. \( I \) and \( h \) are both \( 3 \times 3 \) matrices but in the precincts of this paper we are only dealing with the typical components of both. The residual or amplitude of perturbation in pulsar timing is the integral of the strain \( h \) with respect to time (for details on the full structure of the integral using the matrix \( h \) see Hartle (2003) and Jenet et al (2004)).

A face-on system will produce the maximum \( R \) and all other orientations will produce a smaller value, but for simplicity we choose the most auspicious orientation, and also a near alignment between the gravitational wave source and the pulsar so that the resulting expression represents the largest signal one can get from the black hole binary.

Finally, to obtain the residual amplitude we integrate \( h \) over time as shown in equation (5). We have assumed that the source is sinusoidal (true for circular binaries) and that integrating over \( h \) therefore brings out a factor of \( P/2\pi \). Our expression for the amplitude of perturbation is
\[
r \sim 190 \text{ ns} \left( \frac{M}{10^9 M_\odot} \right)^{5/3} \left( \frac{P}{1 \text{ yr}} \right)^{1/3} \left( \frac{100 \text{ Mpc}}{d} \right) \frac{q}{(1 + q)^2}.
\] (17)

This is the amplitude of the perturbation from a circular supermassive binary black hole system that would be detected in the timing residual of a known pulsar.

How is this \( r \) related to the \( R \) in equation (5)? The notation has been chosen suggestively to recall equation (6). This \( r \) is the approximate maximum amplitude of both \( r_{+}, r_{\times} \) and \( r_{+}, r_{\times} \). So by combining the geometrical information embedded in equation (5) that tells about the amplitude of the response in relation to the angle \( \mu \) between the pulsar and the gravitational-wave-source, with the expected maximum amplitude \( r \) above for a particular source oriented optimally and nearly in line with the pulsar, the reader can get a sense of how to estimate the effect for any source and any pulsar.

Any calculated amplitudes stronger than \( 10 \) ns should be cause for further investigation, since such a source could become detectable in the next decade.

6.2. What sets the frequency regime of PTAs?

The frequency regime of a PTA detector is set by two things: (a) the cadence of the observations and (b) the total timespan of the observations. The former, by the Nyquist theorem, sets the upper-limit on the frequency and the latter sets the lower-limit, the lowest frequency being 1 timespan. So if we could observe all the pulsars in the array once per day the high-frequency limit would be \( 10 \mu \text{Hz} \). This would require something like a dedicated pulsar telescope (discussed in section 9.3.4). There are smaller telescopes that are nearly dedicated to pulsars (the Jodrell Bank and Nancay Telescopes for example) but in order to get the required signal-to-noise ratio per point the sensitivity of the Green Bank Telescope is needed. There are some future planned experiments that could accomplish the required sensitivity on a few pulsars, which may be exactly what we need to look at higher-frequency single sources (see section 9). Burt et al (2011) showed that in order to maximize sensitivity to single sources we really should be concentrating on only a few pulsars. The more sensitivity PTAs have to higher frequencies the more sensitive they are to individual sources. Many efforts now concentrate on unresolved background detection because it will likely provide the first detection (Rosado et al 2015), but as the field becomes more mature, study of individual sources will provide critical astrophysical knowledge. The challenge is that a high-frequency source, even of the same mass, produces a smaller response in pulsar timing than does a low-frequency source. To see this, notice that the response (see equation (17)) depends upon the period, i.e. the longer the period the larger the response. In addition, high-mass sources do not last long at high-frequencies (the time to coalescence for a binary due to gravitational wave emission goes like \( M^{5/3} \) (Simon et al 2014) and thus most of the high-frequency (10’s of \( \mu \text{Hz} \)) sources will be at lower mass and therefore at lower gravitational wave amplitude.

6.3. Massive black hole binaries

Our current understanding of galaxy evolution suggests that there are many supermassive black hole binaries currently in existence in the cores of galaxies throughout the universe. It seems that every galaxy contains a supermassive black hole (see e.g. Magorrian et al (1998)). Further because these galaxies formed by merging with their constituent black holes
that also eventually merge, an estimated 10% of them are in binary systems currently with the fraction of binaries increasing with increasing redshift (Volonteri et al. 2003)\(^5\). We expect to eventually detect individual massive black hole binaries (section 8 discusses detection schemes) but the cumulative ‘crinkling’ of space-time from the combined signatures of many supermassive black hole binaries at different periods and from different directions is likely to be the first gravitational wave feature PTAs detect. The spectrum of these perturbations is most often assumed to be \( h_c \propto f^{-2/3} \) (Phinney 2001, Jaffe and Backer 2003), i.e. a very red spectrum, where \( h_c \) is the characteristic strain is the strain \( h \) averaged over many sources. A straight-forward calculation considering an ensemble of black hole binaries, each emitting gravitational waves of higher and higher frequency as it evolves yields the spectral index \((-2/3)\) (see for example Jenet et al. (2006) and Verbiest et al. (2009)). Several authors have performed more sophisticated versions of the same calculation including low frequency eccentricity (Enoki and Nagashima 2007, Sesana 2013, Ravi et al. 2015, Huerta et al. 2015) and environmental coupling (Kocsis and Sesana 2011, Ravi et al. 2014, McWilliams et al. 2014, Sampson et al. 2015) resulting in a flatter spectrum that in some cases changes slope.

The overall amplitude of the gravitational wave background is less certain, although (Phinney 2001) demonstrates that it is independent of the cosmology one chooses and only depends upon how the merger rate of massive black holes has evolved with redshift. Sesana et al. (2008) and Sesana (2013) specifically model the spectrum using a variety of models for the evolution of the merger rate and show that the amplitude and spectrum of the resulting background is relatively insensitive to such details (see figure 1).

In all cases, the spectrum of the gravitational waves is predicted to be red, and the noise (pulsar intrinsic spin noise, noise inflicted by the ISM, etc) displays a similarly red spectrum, so the signal is difficult to disentangle from the noise in the absence of other information. Thus, an observation of a \( h_c \propto f^{-2/3} \) spectrum in pulsar timing residuals would not be considered a detection of gravitational waves. We must look, therefore, for the particular spatial signature of gravitational waves, correlated between pulsars.

Sesana et al. (2009) show what the spectrum of sources is likely to look like using various models of galaxy evolution (see figure 5). At the highest frequencies (100s of nanoHz) the PTA band is likely to be dominated by a few sources, and as such may have significant anisotropies (Mingarelli et al. 2013). Simon et al. (2014) showed that the ‘hotspots’ of gravitational radiation are likely to have locations toward nearby galaxies and galaxy clusters. (See section 8.3 for discussion about whether the background is really a background). Nonetheless, the accumulation of the signal from multiple massive black holes remains the source most likely to be detected first.

Current pulsar timing measurements already place meaningful limits on the dynamical history of the super-massive black hole binary population (Shannon et al. 2013, Arzoumanian et al. 2015b), i.e. a very red spectrum, where \( h_c \) is the characteristic strain is the strain \( h \) averaged over many sources. A straight-forward calculation considering an ensemble of black hole binaries, each emitting gravitational waves of higher and higher frequency as it evolves yields the spectral index \((-2/3)\) (see for example Jenet et al. (2006) and Verbiest et al. (2009)). Several authors have performed more sophisticated versions of the same calculation including low frequency eccentricity (Enoki and Nagashima 2007, Sesana 2013, Ravi et al. 2015, Huerta et al. 2015) and environmental coupling (Kocsis and Sesana 2011, Ravi et al. 2014, McWilliams et al. 2014, Sampson et al. 2015) resulting in a flatter spectrum that in some cases changes slope.

The overall amplitude of the gravitational wave background is less certain, although (Phinney 2001) demonstrates that it is independent of the cosmology one chooses and only depends upon how the merger rate of massive black holes has evolved with redshift. Sesana et al. (2008) and Sesana (2013) specifically model the spectrum using a variety of models for the evolution of the merger rate and show that the amplitude and spectrum of the resulting background is relatively insensitive to such details (see figure 1).

In all cases, the spectrum of the gravitational waves is predicted to be red, and the noise (pulsar intrinsic spin noise, noise inflicted by the ISM, etc) displays a similarly red spectrum, so the signal is difficult to disentangle from the noise in the absence of other information. Thus, an observation of a \( h_c \propto f^{-2/3} \) spectrum in pulsar timing residuals would not be considered a detection of gravitational waves. We must look, therefore, for the particular spatial signature of gravitational waves, correlated between pulsars.

Sesana et al. (2009) show what the spectrum of sources is likely to look like using various models of galaxy evolution (see figure 5). At the highest frequencies (100s of nanoHz) the PTA band is likely to be dominated by a few sources, and as such may have significant anisotropies (Mingarelli et al. 2013). Simon et al. (2014) showed that the ‘hotspots’ of gravitational radiation are likely to have locations toward nearby galaxies and galaxy clusters. (See section 8.3 for discussion about whether the background is really a background). Nonetheless, the accumulation of the signal from multiple massive black holes remains the source most likely to be detected first.

Current pulsar timing measurements already place meaningful limits on the dynamical history of the super-massive black hole binary population (Shannon et al. 2013, Arzoumanian et al. 2015b), i.e. a very red spectrum, where \( h_c \) is the characteristic strain is the strain \( h \) averaged over many sources. A straight-forward calculation considering an ensemble of black hole binaries, each emitting gravitational waves of higher and higher frequency as it evolves yields the spectral index \((-2/3)\) (see for example Jenet et al. (2006) and Verbiest et al. (2009)). Several authors have performed more sophisticated versions of the same calculation including low frequency eccentricity (Enoki and Nagashima 2007, Sesana 2013, Ravi et al. 2015, Huerta et al. 2015) and environmental coupling (Kocsis and Sesana 2011, Ravi et al. 2014, McWilliams et al. 2014, Sampson et al. 2015) resulting in a flatter spectrum that in some cases changes slope.

6.4. Cosmic strings

The most controversial sources in the PTA band are the cosmic strings. A loop with a subsequent cusp of a string could cause a burst of gravitational waves [e.g. Damour and Vilenkin 2005] that could be detected by any of the gravitational wave experiments depending on the string tension and the dominant frequency of the resulting gravitational radiation. Pulsar timing results already place a stricter limit on string tension than does the Planck Collaboration on its own (Planck Collaboration et al. 2014, Arzoumanian et al. 2015b, Lentati et al. 2015).

6.5. Relic background

A relic background of gravitational radiation was likely left by a phase transition early in the universe (much like the phase transition that caused the cosmic microwave background radiation) (Jenet et al. 2006, Sanidas et al. 2012, 2013). Though widely believed to exist, a relic background will probably not be detected. Unless our estimates of the background from massive black hole binaries turn out to be wildly optimistic, the relic background is many orders of magnitude below the background from supermassive black hole binaries (SMBHB).
7. What will we detect first?

As I write, the field of gravitational waves is without a detection, but has many probable sources, as described above. Which of these will be detected first?

The expected amplitude of the massive black hole binary background is many orders of magnitude above estimates for the relic background. And the binary background, because it is the accumulation of many individual sources, is expected to be detected first before any single individual source (Rosado et al 2015). However, Rosado et al (2015) acknowledge that the detection probability of single sources is not negligible. It would only take one bright source to make the prediction wrong (see Rosado and Sesana (2014) for a quantitative estimate of the likelihood of detecting a single source). With a single source, one has the possibility of actually disentangling the Earth term from the pulsar term in each of the pulsars and definitely determining the direction of the source (see the next section). In a background detection, however, many pulsar terms from many sources are added to many Earth terms, and disentanglement is impossible. The stochastic background analysis uses this fact to an advantage; as we will discuss in section 8.1 it is only the summation of many signals that will bring forth the Hellings and Downs curve. See section 8.7 for this discussion.

There is one prediction of gravitational wave burst rates in the pulsar timing band by Ravi et al (2015) in which they calculate the rates from eccentric SMBHB systems at 0.06 bursts every ten years above a 40 ns threshold for highly eccentric systems and 0.12 burst for every ten years for lower eccentricity sources with the same threshold. The rates get higher as the PTA sensitivity gets better.

The literature so far is in agreement that a stochastic background is likely to be detected before any individual sources.

8. Detection schemes

Roughly speaking any detection scheme will look for disturbances in arrival times that are correlated among the pulsars in the array, the idea being that many things can disturb pulsar arrival times (the ISM, calibration issues, the shape of the pulse profile changing, etc) but these other phenomena will not cause correlated residuals among the pulsars, and no other disturbances will show the specific arrangement of geometric timing offsets that is mathematically delineated in section 5. There are multiple methods of searching for each gravitational wave source type. The biggest problem that every analysis method must solve is that the pulsar term (see section 5) essentially provides an amplitude of noise that is equal to the amplitude of the correlated part of the signal, the Earth term. It is instructive to pay attention to how each detection method deals with the pulsar term. Sometimes differences are determined by the difference in the signal considered, and sometimes by a the philosophy of the researchers. For example, Ellis et al (2013) search for the stochastic background and regard the pulsar term as noise, while (Corbin and Cornish 2010) put forth the conditions under which the pulsar distance could be determined using the pulsar term. Lee et al (2011), Cordes and Janet (2012) and Mingarelli et al (2012) use the pulsar term as an asset to increase the signal strength and the number of signals, respectively.

8.1. Detection by looking for correlations

The expected response, equation (5), shows the pulsar term and the Earth term, and, as previously mentioned, shows that the Earth term is correlated among all the pulsars in the array, while the pulsar term is not. The two terms are of roughly equal amplitude as long as the evolution of the binary is slow. We could imagine an abstract situation in which we can separate the Earth and pulsar terms in each pulsar. For any particular pulsar pair, the expectation value of the correlation between the Earth-term residuals of the two pulsars has the value shown in the Hellings and Downs curve (figure 6), which is given by Hellings and Downs (1983b)

\[ C = \frac{3}{2} \left( \frac{1 - \cos \gamma}{2} \ln \frac{1 - \cos \gamma}{2} - \frac{1}{6} \frac{1 - \cos \gamma}{2} + \frac{1}{3} \right) \]  

(18)

where \( C \) is the correlation between pairs of pulsars’ residuals and \( \gamma \) is the angle between the pulsars. This value is based solely on the angular separation of the two pulsars. Notice that when the pulsars have zero angular separation their ‘Earth-term’ residuals are maximally correlated. This correlation, \( C \), passes through \( C = 0 \) at 49° and reaches a minimum at 82° and comes back up to exactly half its initial zero-separation value at 180° (Mingarelli and Sidery 2014). This particular signature results from the prefactor \( 1 + \cos \mu \) in equation (5) and the polarization term \( \cos 2\psi \) or \( \sin 2\psi \). When the pulsar term is included, the correlation for a pair of pulsars is 1.0 at zero separation (that provides the normalization for the function) and is well approximated by a delta function in the short wavelength approximation, falling off quickly with increasing angular separation to the value shown in equation (18). For the full mathematical treatment see Mingarelli and Sidery (2014). The following paragraphs describe the relationship between equations (5) and (18) for three different values of \( \gamma \): 0°, 90°, and 180°.

![Figure 6](https://example.com/figure6.png)
For $\gamma = 0$, i.e. for a given pair of pulsars at the same sky location, their two prefactors $1 + \cos \mu$ will be the same and thus pulsars separated by 0 degrees will exhibit the maximum correlation.

For pulsars separated by 90° the situation is not so simple because the two different prefactors depend on the location of the gravitational wave source and the polarization angle $\psi$ becomes important. The anti-correlation that is present in the Hellings and Downs curve near 90° comes from the $\cos 2\psi$ (or $\sin 2\psi$) terms in equation (5). Because of this term a 90° rotation of the sky about the gravitational wave path yields the opposite of the original response, and therefore on average yields a negative correlation between two pulsars separated by 90°. The fact that the minimum in the Hellings and Downs curve is at 82° rather than 90° comes from the average over all the possible gravitational wave source locations, i.e. when the two pulsars are separated by 90° in the sky, they are not always 90° away with respect to a rotation about the gravitational wave propagation direction.

A similar argument can be made for pulsars separated by 180°. The $\cos 2\psi$ term is likely to render the two pulsars roughly correlated (if one pulsar is on a positive ‘lobe’ of figure 4 the other pulsar will be on the opposite lobe with a response of the same sign). However, the correlation will be diminished owing to the mitigating property of the $1 + \cos \mu$ factor from equation (5) (for pulsars opposite each other in the sky the angle $\mu$ between the gravitational wave source and the pulsar will be large for at least one of the sources). See Mingarelli and Sidery (2014) for a mathematical treatment.

The Hellings and Downs curve is a tell-tale signature of gravitational waves and can be used to distinguish gravitational waves from other sources of deviation in the pulsar timing residuals.

The Hellings and Downs curve is a curve of expectation values, and thus is achieved experimentally only by the averaging of many signals. In this case each ‘signal’ is the imprint of a gravitational wave on a pulsar pair, so we can acquire many signals by considering many gravitational waves, in other words, a stochastic background of gravitational waves, or we can acquire many signals by considering many pulsar pairs (Cornish and Sesana 2013). Stochastic background detection methods aim to do both.

However, thus far in this section we have assumed that the pulsar term does not exist, and now we need to come back to a realistic situation. To the correlated residuals described in the previous paragraph we add a large number of pulsar terms, i.e. a large number of sinusoids, whose amplitudes are given by the same equation that yields the Hellings and Downs curve shown above, but whose phases are now essentially completely random. (Random because the geometric delay in the pulsar term can take on any value between 0 and $2\pi$ and all values are equally likely. Even if we include what we know about the pulsar distances, the uncertainties in those distances are typically 20%, which is many wavelengths of the gravitational wave, so again all values between 0 and $2\pi$ are equally likely. For more information on pulsar distance uncertainties see Verbiest et al (2012).)

If this ‘large’ number of sinusoids is in fact large enough, the combined effect of all the added sine-waves will contribute very little to the overall signal. Provided we also have enough Earth terms to yield something close to the expectation value, we will see Hellings and Downs curve. The larger the number of randomly-phased sine waves in the sum, the greater the suppression of the pulsar-term noise. We simulated 10 randomly phased equal-amplitude pulsar term sine waves. Across 1000 realizations the average of the maximum amplitude in the accumulated pulsar term signal was 29% of the maximum amplitude in the accumulated Earth term signal. With 100 randomly phased sine waves the maximum accumulated pulsar term signal was 9% of the accumulated Earth term signal. The number of sources contributing is likely to be large, but the problem is that a few sine waves are likely to dominate, those from the largest and closest sources. (See for example recent work by Sesana et al (2009) and Ravi et al (2012) that shows that the spectrum is likely to be dominated by a few bright sources.) So instead, we aim to average the contributions from multiple pulsar pairs. For example, if we have three pairs that all have separations near 30° we can average the correlations from those three pairs and get a value that is closer to the Hellings and Downs curve value. The more pairs we can add into that one angular ‘bin’ the closer we are likely to get to the Hellings and Downs predicted value. It is for this reason that including many pulsars in the array is so important. (See section 10.1 for a quantitative discussion on the relationship between sensitivity and the number of pulsars. Also see Siemens et al (2013).)
yet achieved the small separations (1 parsec or less) required in order for a PTA to observe it. The most popular solution is to invoke gas dynamics in the galaxy, i.e. dynamical friction, but that requires some ‘tuning’ of the universe in the following way. If we want there to be substantial signal in the PTA regime, we need the dynamical friction to be effective in bringing the black holes closer together (to about a pc across) but not so effective that it makes the black holes essentially skip the phase in which their gravitational wave emission is at frequencies of 1/yr.

The lowest limit on the stochastic background is currently set by the PPTA at \( A = 2.4 \times 10^{-15} \) 1/yr where \( A \) (given by equation (2)) is the amplitude of the background at a period of a year (Shannon et al. 2013). This can be readily translated into a deviation of timing residual by multiplying by 1 year, which gives \( 6 \times 10^{-8} \text{s} \) or 60 ns. We expect the spectrum to follow \( h_c = f^{-2/3} \) as described in section 6.3 so the strain amplitude at 10 years should be nearly \( 10^{2/3} \) \( 5 \) times what it was at one year, and the residual is the integral of the strain as described in equations (7)–(8) which adds another factor of the wave-period, so we expect the residual response at wave-periods of 10 years to be \( 10^{22/3} \times 10 \times 60 \text{ns} \approx 3 \mu \text{s} \). This amplitude is large compared to the rms values of our pulsars, so why is this the best limit we can obtain? It has to do with the fitting of the pulsar model described in section 4.3.1. The fitting procedure unfortunately absorbs gravitational-wave signal. Absent the need to fit, i.e. if we knew the timing parameters a priori, we would have detected gravitational waves several years ago. In other words, the gravitational wave background may very well be contributing to pulsar timing residuals at amplitudes near 3 \( \mu \text{s} \), but once those signatures are fit to the pulsar timing models, their amplitudes are reduced, sometimes by an order of magnitude. Shown in figure 7 are three realizations of the signature of a gravitational-wave background computed from the combined effect of 10 000 black hole binaries with the frequency spectrum shown in equation (2) using a background amplitude of \( A = 2.4 \times 10^{-15} \text{1/yr} \) and \( \alpha = -2/3 \), and then fit to a period and period derivative. The reader should note that due to the pulsar-model fitting, the signal amplitude is far below the 3 \( \mu \text{s} \) signal we expect to be contributing at periods of 10 years.

8.3. Issue #2: Is the background really stochastic?

Is the background of gravitational waves really stochastic? Results by Sesana et al. (2009), Mingarelli et al. (2013) and Cornish and Sesana (2013) suggest that the spectrum, especially at the high frequency end of the PTA bandwidth (near 100 nHz) is dominated by a few bright sources (see figure 5 for one realization of a simulated background according to the model of Sesana et al. (2009)). These sources are nearby large SMBHBs. It turns out, however, that the search methods described above which rely on the assumption that there are many gravitational-waves coming from many different directions are still useful (but sub-optimal, see Cornish and Sesana (2013)) for finding an accumulation of gravitational wave signal from the combined effects of many SMBHBs, largely because the anisotropy is very small throughout all but the higher-frequency part of the detectable spectrum (Mingarelli et al. 2013). In all cases (stochastic background or not) the scatter around the Hellings and Downs curve is caused by a small number of sources contributing to background signal ensure that a sufficient number of pulsar pairs contribute to each bin in the Hellings and Downs curve as discussed in section 8.1. (See the second term in equation (6).)

Both Mingarelli et al. (2013) and Gair et al. (2014) demonstrate possible new ways of conducting the analysis in conjunction with considering the anisotropy of the background.

8.4. Detecting individual ‘continuous wave’ sources

Various groups have worked on the detection of these single sources rather than on the detection of the ensemble of them. Both Yardley et al. (2010) and Arzoumanian et al. (2014) have used pulsar timing data to limit the number of coalescing binary systems of a given chirp mass as a function of redshift. These two groups have placed limits on single sources using real data. Other groups have published algorithms and techniques for detection, as described below.

Sesana and Vecchio (2010) looked at continuous gravitational-wave sources such as binary black holes, where one must include the pulsar term, and they are able to localize the gravitational-wave source to within 40 square degrees for a 100-pulsar array and a signal to noise ratio of 10. The error box is big because they assumed they could not know the distances to the pulsars, so the 100% variation of the pulsar term makes it very hard to pin down the direction as described at the beginning of this section. Ellis et al. (2012) and Wang et al. (2014) have also developed techniques for single source detection.

Corbin and Cornish (2010) have demonstrated the possibility of searching over and recovering the pulsar distances from the chirp signal. They assumed white noise for the timing residuals which we may in fact have (see section on detector characterization below), but their technique has not been studied in the presence of red noise. They are able to localize the source to less than three square degrees for strong sources. Lee et al. (2011) point out that if the timing parallax (a distance...
measurement independent of the chirped signal) is estimated at the same time as the gravitational-wave parameters are estimated then the pulsar term can be used as a great asset in increasing signal strength in single source cases. They carefully predict the statistical uncertainty that PTAs can expect to achieve in determining characteristics of gravitational-wave sources such as orbital inclination angle, source position, frequency, and amplitude. They predict that PTA source localization ability will range from a radian down to several microradians depending on the strength of the source.

8.5. Detecting bursts

Finn and Lommen (2010) and Pitkin (2012) looked at bursts of gravitational waves, sources whose duration is shorter than the data span. Pulsar timing data spans are tens of years, so a burst could be a month-long source. Bursts would be caused by any quick, close encounter of two massive objects, such as the perigee in an eccentric orbit of a black hole binary or a ‘fly-by’ in a hyperbolic orbit. Bursts would also be called by cusps of cosmic strings.

Finn and Lommen (2010) put forth that when looking at bursts, one can ignore the pulsar term because it is unlikely to enter the dataset in a human lifetime. Note that because the distance between the Earth and the pulsar is hundreds or thousands of light years, the delay between the Earth term and the pulsar term is hundreds or thousands of years. Thus, any source that produces a burst of coherent response in all the pulsars (i.e. the Earth term) will produce the second (incoherent) part of its response, the pulsar term, many hundreds of years later. Finn and Lommen (2010) were able to localize a strong source to less than 1 square degree. For a moderate source it was hundreds of square degrees. For a weak source it was thousands of square degrees.

Pitkin (2012) investigated the possibility of extended an ‘Earth-term’ search as described above, by looking for coincidences in pulsar terms. Pitkin concluded that the Earth term search is much more sensitive, owing to the correlation of many pulsars rather than just two, but that a pulsar-term search could be used to extend the effective time-span of a search and to enhance the search toward a particular location.

There is another way to think about the Earth and pulsar terms. The response that we calculated in section 5 is actually the integral over the entire interaction between the electromagnetic wave and the gravitational wave, i.e. the gravitational wave influences the electromagnetic wave along the entire path in between the pulsar and the Earth. Thus, one can think of the Earth term as the gravitational wave beginning to influence the path of the electromagnetic wave, and the pulsar term as the gravitational wave ending its influence over the path of the electromagnetic wave. Both events can be detected, the former being the coherent burst discussed above, and the latter being the incoherent echo or ‘memory’ of the former. Pshirkov et al (2010), van Haasteren and Levin (2010), Cordes and Jenet (2012) and Madison et al (2014) call the latter ‘bursts with memory’ and have demonstrated that searching for such signals increases the number of possible detectable sources. We will miss the Earth term of some bursts, but still be able to observe the burst with memory when the pulsar term of the response occurs.

8.6. Issue #3: Bayesian versus Frequentist detection

This is a favorite topic of debate, but Bayesian techniques and Frequentist techniques are different, and the field benefits from the ability to compare between them. We can recategorize the techniques described above into the two very broad categories of Bayesian detection methods (Finn and Lommen 2010, van Haasteren et al 2011) and frequentist detection methods (Jenet et al 2006, Yardley et al 2010, Corbin and Cornish 2010, Lee et al 2011, Cordes and Jenet 2012, Ellis et al 2012, Arzoumanian et al 2014). Both methods require the mathematical articulation of all the known elements of the data set, the ‘priors’, such as the response of the pulsars to gravitational waves, any known preferential directions for the gravitational waves to be traveling, the noise levels in the pulsars, etc. In Bayesian methods (e.g. van Haasteren et al (2011)) these priors are used together with the data to determine the probability that a particular signal is present in the data. In frequentist methods (e.g. Jenet et al (2006)) these priors are used to simulate many realizations of the data to which the actual data are compared in order to determine the level at which the signal is present in the data. The fact that both techniques are used on the same data will eventually strengthen a claim of detection.

8.7. Issue #4: What constitutes a Detection?

There are several things that would clearly constitute a detection such as a plot of the correlation between pairs of pulsars as a function of each pair’s angular separation that shows a clear Hellings and Downs curve. Alternatively if a sinusoidal source is observed in multiple pulsars with the same frequency, and a coincident electromagnetic counterpart is found, that would also be a clear detection. But what if the data show less than that? First, for example, imagine a stochastic background analysis in which no Hellings and Downs curve is clear by eye, but the Bayesian analysis reveals that the odds of a gravitational wave background existing in the data are 100:1. Second, imagine a single source detection where instead of a sinusoidal source being observed in multiple pulsars, suppose that an analysis such like that in Finn and Lommen (2010) yields no waveform, and no source position is recovered, but the data suggest that the odds of a source existing are 9:1. Are either of these examples publishable detections?

The question of detection may not be a useful one as it represents a false line, i.e. in science we never actually prove anything exists, but rather continue to rule out possible alternatives until the detection becomes commonly accepted (Newton and Frost 1863). Some people think gravitational waves were already detected in 1981 when the orbital decay rate of the binary pulsar was shown to be exactly consistent with the emission of gravitational waves (Weisberg et al 1981). Others think gravitational waves were detected
again when the BICEP2 team announced the detection of the imprint of gravitational waves on the polarization of the cosmic microwave background (BICEP2 Collaboration et al 2014)\(^6\). Some people allow for these as ‘indirect’ detections, a designation I also do not find useful because even the interferometric and PTA results will be measuring the effect of the gravitational wave on something else, namely an electromagnetic wave.

So rather than ask ‘was it detected or not’ we could ask ‘with what probability was it detected?’ The question of whether something will be published at various points is an issue most collaborations must confront. LIGO has a detection protocol, i.e., a series of things that transpire in order for the collaboration to put its name on a paper that suggests detection. (The interested reader can investigate a LIGO detection exercise ‘the big dog’ in which a signal was injected and then allowed to proceed through the detection protocol\(^7\)). The International Pulsar Timing Array began to consider a detection protocol at the IPTA meeting in Banff 2014. Other less probable detections can and should be published, with complete candor about the circumstances of their detection. The articles should explain what data were used, how they were reduced, and what results were obtained. If the analysis represents a borderline detection the article should say so. This is the best procedure for science and for the field in general. More information is always better for the whole field.

9. Possibility for expansion of PTAs in the future

9.1. PTAs ranking in Decadal review

The National Academy’s Decadal Review recently (2010) ranked pulsar timing as the top priority in Particle Astrophysics and Gravitation in the ‘medium size’ category. (See table B.1 of the report at www.nap.edu/catalog/12951.html.) In the same report NANOGrav itself is listed as one of the eight ‘projects thought compelling for the mid-scale innovations program.’ (See table 7.1 in www.nap.edu/catalog/12951.html.) NANOGrav has recently been awarded an NSF Physics Frontiers Center; jointly funded through the PHYS and AST divisions.

9.2. Challenges for Arecibo and green bank telescopes

The two major telescopes used by NANOGrav, the Arecibo Telescope and the Green Bank Telescope, have been recommended for divestiture by the 2006 report of the Senior Review\(^8\) and the 2012 Portfolio Review\(^9\) respectively, both conducted by the National Science Foundation. Those are the two largest single-dish telescopes in the world, both excellent for pulsar timing, both with advocacy groups that are attempting to come up with new funding models. The Arecibo Telescope has now been run for four years by a consortium including SRI International, Universities Space Research Association (USRA), and the Universidad Metropolitana. But the transitional nature of these two telescopes is a great challenge to Pulsar Timing Array work, and in particular to the work of NANOGrav for whom these two telescopes are primary instruments.

9.3. New instruments

There are many new telescopes coming on-line elsewhere. The Square Kilometer Array (SKA) is currently under construction in South Africa and Australia, where SKA pre-cursor arrays MeerKAT in South Africa and The Australia Square Kilometer Array Project (ASKAP) are under construction, and where their precursors KAT-7 and ASKAP BETA are already in use. The Five-Hundred-Meter Aperture Spherical Telescope (FAST) in China, and Canadian Hydrogen Intensity Mapping Experiment (CHIME) in Canada, the Molonglo Radio Observatory in Australia, the Deep Space Network (one each in the US, Australia and Spain), and the Apertif Radio Transient System (ARTS) at the Westerbork Synthesis Radio Telescope in the Netherlands will all have substantial impacts on the field. We describe these and what they will be able to do for pulsar timing in more detail below. The following sections describe a number of telescopes being built that have collecting area very similar to that of the Green Bank Telescope’s 100 m dish (MeerKAT, CHIME). The Arecibo Observatory’s 300 m antenna has larger collecting area, but the telescope has limited declination range so it is only useful for part of the sky. Smaller telescopes (e.g. the Molonglo Radio Observatory, the Deep Space Network, and ARTS) will play supporting roles to these larger telescopes, leveraging and optimizing the available time on the larger telescopes by taking care of parts of the observations that require less sensitivity.

9.3.1. SKA

The SKA collecting area will be roughly 100 times that of Green Bank (and thus, could in principle obtain the same timing point in 1/100th the time that Green Bank does). However, the SKA will not be complete until 2025 and until then the SKAs precursor in South Africa, MeerKAT is in fact very similar to Green Bank. MeerKAT in particular is well-optimized for observing pulsars and a number of IPTA members have pulsar timing projects underway there.

The potential benefit of the SKA to the gravitational wave detection effort is enormous, but also carries with it significant challenges. The possibility is that in 2025 when the SKA is commissioned, we will quickly find about 6000 MSPs and find about 250 of them suitable for PTA inclusion (Smits et al 2009). Smits et al (2009) estimates that we will need five minutes for each pulsar. After including slew time and calibration Smits estimates we need 24 h to get a timing point on all 250 pulsars. The collecting area is so enormous (roughly a square kilometer) that if the noise scales like one over the

---

\(^6\) This detection was called into question by more recent foreground dust estimates (Cheng et al 2014, Planck Collaboration et al 2014), but my point about ‘indirect’ detection remains regardless.

\(^7\) www.ligo.org/science/GW100916/


\(^9\) Report from the National Science Foundation (NSF) Division of Astronomical Sciences Portfolio Review Committee entitled: Advancing Astronomy in the Coming Decade; Opportunities and Challenges: www.nsf.gov/mps/ast/portfolioreview/reports/ast_portfolio_review_report.pdf
square root of the collecting area we will have rms timing accuracy of about 10 ns on each of the pulsars. We obviously cannot acquire 100% of the telescope time, but one day of observing each week is realistic, and therefore the upper frequency limit of the gravitational radiation we could detect would be 0.5/week or about a μHz. A realistic estimate of the total time required to get a timing point on all 250 pulsars is about 24 h. The SKA era of PTA science is especially important for single source detection (Ravi et al 2015). Ravi et al (2015) makes estimates with a significantly more pessimistic next-generation telescope that has 100 pulsars with 100 ns RMS. (One possible reason for pessimism is discussed in the next paragraph). Ravi et al (2015) estimates that there is perhaps a 50% chance of seeing a single binary SMBH with such an array. Using scaling relations from Arzoumanian et al (2014) we can gain a sense of how the sensitivity would change with 250 pulsars and 10 ns RMS. Arzoumanian et al (2014) show that in the white noise regime the S/N scales as

\[ \rho \propto \sqrt{\frac{\sum T_\alpha a_\alpha^2}{\sigma_\alpha^2}} \]  

where \( \sigma_\alpha \) is the rms of pulsar \( \alpha \), \( T_\alpha \) is the timespan of the data, \( a_\alpha \) is a geometrical factor that depends upon the relationship of the pulsar to the gravitational waves, and \( c_\alpha \) is the cadence of observations. So the S/N of what was a ‘barely detectable’ source in the Ravi et al (2015) estimation now would have a signal to noise ratio of \( 10 \times \sqrt{2.5} \approx 16 \) times what it had before. In other words, many more sources would be observed.

However, whether or not we are sensitivity limited (e.g. whether or not the aforementioned 10 ns is achievable with the SKA) is the subject of much debate. Red noise has been observed in some pulsars, and if we cannot correct for it, increasing the collecting area of the telescope will not help us (see section 4.3.3). The good news, however, is that the scaling relations in Siemens et al (2013) show that in the intermediate-signal limit (where the gravitational-wave signal in the lowest-frequency bin is larger than the noise) the signal to noise of the background scales only very weakly with the rms of the pulsars (see section 10.1). Its dominant dependence is linear with the number of pulsars. We must note, however, that these scaling relations were derived assuming the noise in each pulsar is white.

9.3.2. FAST. China is building the Five-hundred-meter Aperture Spherical Telescope (FAST) which will essentially be the Arecibo Telescope but 60% larger in diameter. A principle science driver for the project is gravitational wave detection using pulsars. Commissioning is currently scheduled for 2016. What can a 500 m telescope do that a 300 m telescope cannot? It can get high signal-to-noise on any source in a shorter time, and therefore make all observations faster. A telescope’s gain, or power to accumulate signal from astrophysical sources, can be measured in m^2/K or in other words the effective area of the telescope divided by the system temperature of the receiver. This number, which we call ‘modified gain’, is related to the more formal gain (Jy/K) but more readily comparable for various dish/receiver combinations. FAST has 500 m diameter, 36% aperture efficiency, and a 20 K system temperature and so has a modified gain value of 884 m^2 K^-1 and so is three times more sensitive than GBT or MeerKAT.

9.3.3. Smaller telescopes that leverage the larger ones: CHIME, MeerKAT, molonglo radio observatory, deep space network, and ARTS. All of these telescopes, CHIME, MeerKAT, the Molonglo Radio Observatory, the Deep Space Network, and ARTS will provide low-frequency timing points on pulsars, in some cases (CHIME and MeerKAT) every day. These data can be used to refine the higher-frequency observations at the other telescopes and potentially improve timing by about a factor of two (private communication, Ingrid Stairs, Paul Demorest, Scott Ransom). CHIME and MeerKAT are largest of this collection of telescopes. In CHIME the 10 000 m^2 collecting area is broken up in five parabolic cylinders each 100 m × 20 m, but the total collecting area is very similar to GBT and MeerKAT. MeerKAT, the first phase of which is scheduled for completion in 2015, the full array being completed in 2017, will be very similar to the Green Bank Telescope for the purposes of observing pulsars.

MeerKAT is 105 m in diameter, has a 65% aperture efficiency, and a system temperature of 18 K so its modified gain is 312 m^2 K^-1.

For comparison the Green Bank Telescope has a diameter of 100 m, a system temperature of 18 K, and an aperture efficiency of 72% so its modified gain is 314 m^2 K^-1.

Thus, the two telescopes are remarkably similar in capability with one key difference. Because MeerKAT is an array of 64 dishes, it can synthesize about 100 beams (100 different directions in the sky) at once as long as the beams are within 1 square degree, the field of view of MeerKAT. This represents both an advantage over Green Bank and a computing problem that Green Bank does not have.

9.3.4. A telescope dedicated to pulsar timing for gravitational waves? The above paragraphs describe opportunistic uses of existing projects that were not necessarily designed with optimization of the pulsar timing experiment in mind. They are wise uses of these projects, but what sort of telescope would the PTA experimenters design if they were to design a dedicated PTA telescope. Where would it be? Would this be a valuable use of resources?

Because in pulsar timing one needs collecting area and not angular resolution, a single dish would be the best. Because the telescope would need to cover a multitude of different angular separations of pulsars (note that this is the important parameter for background detection, not necessarily sky coverage) the dish would need to be steerable. Green Bank is close to what would be designed, but a telescope designed from scratch would be closer to the equator.

Should funding be dedicated to this enterprise? Answering that question requires a statement of priorities or goals, and a statement of constraints on resources, both of which are beyond the purview of this review. However, here are some things to consider.
If the aim is to detect a gravitational-wave signal in the nanohertz regime then a dedicated telescope is unneeded. Existing resources will eventually fulfill that goal. Most recent estimates indicate the detection of the stochastic background will be first and that this will happen between 2017 and 2025. (See Siemens et al (2013).) (The large range in date comes from (a) our uncertainty about the level at which the background will occur, (b) the limits of our understanding of the amount of red noise in the timing residuals). This means that in 2017–2025 we will have confirmation that gravitational-waves exist in the nanohertz regime, and the first number we will get will be the amplitude of the background. After a couple more years of operation with existing instruments we may obtain the slope of the spectrum. If we are lucky we may detect a nearby single source (Rosado and Sesana 2014).

However, to move into a regime where we are using PTAs as a gravitational wave telescope to determine properties of individual black hole binary systems such as mass, spin, etc then we require a substantial increase in cadence of the experiment, namely, we need a TOA every day on each of the pulsars in the array. A daily cadence increases the range of sensitivity of the experiment into the frequency regime of 1/day or \(\sim 10 \mu\)Hz which is a wavelength regime not covered by any other experiment including eLISA/NGO. This frequency regime allows the possibility of detecting the inspiral-merger-ringdown of massive black hole binaries (the last days of their merging), not just their relatively stable 1 year orbits. Turning toward more controversial sources, it would also allow the detection of waveforms from cosmic string cusps, the detection of which would certainly be exciting. The daily cadence also increases the overall sensitivity of the experiment in the lower frequency (nHz) regime just by the increase in the amount of time spent on each pulsar.

As previously described, an unreasonably large fraction of SKA time would be required to provide daily cadence on all the millisecond pulsars, but the IPTA, employing the combined resources of existing and new telescopes, could help take up the slack. In an attempt to think practically about a future IPTA let us consider that we have six large telescopes in addition to the SKA, that have come together in the IPTA, and that we need those six telescopes collectively to observe 30 pulsars per day, leaving 70 for the SKA, for a total of 100 pulsars observed every day. In this speculative realm let us not fixate on exactly which telescopes make up the six, but roughly speaking it could be the Arecibo Telescope, the Green Bank Telescope, FAST, LEAP, MeerKAT, and a combination of the telescopes designated ‘supporting’ above. For the purpose of our approximation let each of the six telescopes be able to acquire a timing point for a single pulsar in one half hour. So to observe 30 pulsars every day we would need each of the telescopes for 18 h per week. (Note that no telescope can see the whole sky, so we would have to distribute the observing time cleverly.) This is a larger time allocation than pulsars currently have on any of the large telescopes, but only by a factor of two or three. Note that both the Lovell Telescope and the Nancay Telescopes have larger pulsar allocation times than this.

On the other hand, on a dedicated telescope similar in size to the Green Bank Telescope we could obtain one timing point in 1/2 h, and thus could observe about 40 pulsars a day (once calibration and maintenance time are allotted). Such a telescope could observe 100 pulsars every three days (or so), or perhaps be used to observe the 40 best pulsars daily, with other instruments and the SKA observing the other 60.

The point is that even considering all existing and planned telescopes, the needs of the post-detection era still stretch the resources. Though no gravitational wave has yet been detected, this is the time for the community to think beyond detection and to plan for an era of characterization of gravitational wave sources. A dedicated instrument should be considered carefully as should the possible ramifications of not having a dedicated facility, i.e. the burden placed on existing resources.

10. Optimization of the experiment

If the PTA were a table-top experiment rather than a galactic-scale observatory, we would adjust the relative positions, amplitudes, powers, etc of the elements until we got the maximum throughput of the experiment. Though adjusting the individual detector elements (i.e. the pulsars) is impossible, it is possible to alter the amount of telescope time devoted to each pulsar. Whenever we optimize an experiment we optimize for a particular result, so in our case optimizing for detection of the stochastic background is different from optimizing for detection of single sources. And in fact, optimizing for characterization of single sources (i.e. their polarization properties, or waveforms) is different from optimizing for detection, i.e. for the likelihood of a source existing in the data (Finn and Lommen 2010).

10.1. Optimizing for a stochastic background

The punchline of this section is that for times in the future, when the PTA moves from the weak signal to the strong signal regime, it matters less and less what the noise from each of the pulsars is, and the signal-to-noise ratio becomes highly dependent on how many pulsars the array contains. This feature has to do with the random nature of the signal being produced by the background. Roughly, each gravitational wave source in the background adds two sine waves to the pulsar timing residuals, one from the Earth term and one from the pulsar term. The value of the Hellings and Downs curve at any particular angular separation is the expectation value of the correlation at that separation. In other words, it is the value of the correlation would have if an infinite number of sources contributed to the background. For long integration (dwell) times the observer effectively beats down the intrinsic noise in the pulsars but cannot beat down the scatter in the Hellings and Downs curve except by increasing the timing baseline or by averaging over many pulsar pairs. Recall from section 8.1 that averaging over many pulsar pairs is the physically realizable method.
Here is how long-term timing plays out specifically for the different regimes.

In the low signal to noise regime, which may or may not be where we are now (we may actually be in a transition region between low and intermediate signal), Siemens et al (2013) calculates the signal to noise ratio to be:
\[
\langle \rho \rangle \propto \frac{M c^2 T^3}{\sigma^2},
\]
where \(A\) is the amplitude of the stochastic background, \(c\) is the cadence (observations per time), \(M\) is the number of pulsars, \(\sigma\) is the white-noise level of the pulsars, and \(\beta\), the spectral index of the background, is 13/3 for a SMBBH background.

When the signal in the lowest frequency bin exceeds the white noise (and we enter the intermediate regime) the signal to noise ratio is
\[
\langle \rho \rangle \propto M \left( \frac{A}{\sigma c} \right)^{1/3} T^{1/2}.
\]

These equations assume that the noise is all white. Notice that in the intermediate regime the dependence on \(\sigma\) is very weak, i.e. the signal-to-noise ratio is inversely proportional to \(\sigma^{3/13}\). For example, if the rms of the pulsars decreases by a factor of two, the signal-to-noise ratio only increases by a factor of 1.2, whereas if the number of pulsars in the array doubles the signal-to-noise increases by a factor of two.

What is the optimization strategy for detecting the stochastic background? One obvious option is to add pulsars to the array. In the intermediate regime, where we will find ourselves eventually (if indeed we are not there now), this is the best strategy for detecting the background.

The pulsar timing for gravitational wave experiment is unique in that the overall timing baseline is important, e.g. the stochastic background scaling relations for the low-signal and intermediate-signal regimes (equations (21) and (22)) show that a ten-year data set with a 200 ns rms residual will detect the background with more significance than a five-year data set with 50 ns rms residual in either regime all other things being equal. So if possible, we need to plan our observing now around the detection strategy we will have in ten or twenty years.

The next question is how to improve the \(\sigma\) of each pulsar. If the pulsar noise is white then integrating or ‘dwell’ on each pulsar for longer times will decrease \(\sigma\) by the square root of the amount of dwell time (Handzo et al 2015). For the same reason, work by Lee et al (2012) shows that optimization, i.e. redistributing the time amongst the pulsars, does not improve stochastic background detection much when red noise is present, because adding more integration time per pulsar does not change the rms \(\sigma\) of each pulsar. Whether or not pulsars display significant red noise is the subject of much debate (see section 4.3.3).

10.2. Optimization for individual sources

For individual sources the optimal strategy involves accumulating the shape of the gravitational waveform rather than maximizing the number of pulsar pairs as in the stochastic background detection (Burt et al 2011, Deng and Finn 2014). In other words, the Hellings and Downs curve (figure 6) is much less important and the individual gravitational wave signal to noise ratio in individual pulsars becomes paramount. (Note that (Cornish and Sesana 2013) showed that it is possible, but not optimal, to use the Hellings and Downs curve to detect a single source). So for single source detection, optimization amounts to spending most of the allocated telescope time on the best pulsars. The answers are slightly different depending on whether the goal is waveform characterization, polarization characterization, source location, or simply source detection (Koop and Finn 2014).

10.3. Issue #5 What do we do about the fact that optimization for single sources is different from optimization for the stochastic background?

This is tricky, because given that the two optimization strategies have distinct differences, we have to be careful that we do not decide to take some middle ground between the two situations, essentially un-optimizing for both signals, i.e. ensuring that we are not doing the right thing for any of our potential signals while trying to be equitable to all. The answer comes by having to consider what is the more likely source now and the ways in which the experiment will evolve. Ravi et al (2015) and Rosado et al (2015) suggest that the stochastic background is more likely to be detected first and that pre-SKA optimization efforts should be geared toward a stochastic background detection.

However, if we want to detect single sources at some time in the future, then we should start planning for it now. The current NANOGrav observing scheme takes this into account. NANOGrav is spending extra time each observing session on two of its best pulsars (PSRs J1713+0747, 1909–3744), for the sake of single source detection. It may be that a few slightly smaller telescopes could be dedicated to a few pulsars in order to increase our sensitivity to single sources. For example, the original attempt to form a PTA utilized one of the 70 m antennas of NASA’s Deep Space Network (DSN, Hellings and Downs 1983). Two of these antennas are now in the process of being fitted with higher sensitivity receiving systems and appropriate digital backends to conduct high precision pulsar timing.

11. Conclusion

Pulsar timing will detect gravitational waves, the question is when, and what will it be able to do both in the short term and in the long term, and what do we need to be doing now in order to ensure that we can take full advantage of the array in the future. Because of the long timescale of the experiment, decisions we make now will have an impact on the science we can do 20 years in the future. Some things are clear, and some
things are still open questions. What is clear is that efforts to find more pulsars to put in the array should continue. The sensitivity of the array in the future will depend upon, in the case of stochastic background detection, the shear number of pulsars in the array, and in the case of individual sources, on the quality and time-ability of the pulsars we have. Finding even just one more pulsar like the reliable and precise PSR J0437–4715 would help single source detection immensely.

There are still questions about the amount of red noise in the pulsars. How much more time can we spend on each pulsar and still expect an increase in sensitivity of the array? The answer to this question is critical for making decisions about how to employ our telescopes. For example if we knew that the sensitivity of the array would increase by a factor of five if we spent 25 times as much time on each pulsar, then I think we would be building a dedicated pulsar timing telescope precisely for this purpose. But it is not actually clear that the sensitivity would increase that much. Recent work by Handzo et al (2015) shows that we can at least expect to increase the sensitivity by roughly $\sqrt{8}$ by increasing the dwell time by a factor 8.

The other question is how much to rely on the SKA and wait for this telescope, which will bring us more multiple pulsars, and more observing time.

Returning to what is clear, sometime in the next decade pulsar timing will detect gravitational waves in the nanohertz regime. Despite uncertainty about when, it will happen unless the background is at a much lower level than all models predict. Detecting and studying gravitational waves in the nanohertz regime will be a critical piece of the unveiling of the gravitational wave universe.

Acknowledgments

The author is grateful for discussions with B Christy, J Lazio, F Jenet, C Mingarelli, J Romano, C Rose and A Sesana. She gratefully acknowledges the support of National Science Foundation grants PIRE OIP 09-68296 and AST CAREER 07-48580.

References

Acernese F et al 2015 Class. Quantum Grav. 32 024001
Affeldt C et al 2014 Class. Quantum Grav. 31 224002
Amaro-Seoane P et al 2012 Class. Quantum Grav. 29 124016
Amaro-Seoane P et al 2013 GW Notes 6 4–110
Ando M and TAMA Collaboration 2002 Class. Quantum Grav. 19 1409
Burgay M et al 2003 Nature 426 531
Corbin V and Cornish N J 2010 (arXiv: 1008.1782)
Cordes J M 2013 Class. Quantum Grav. 30 224002
Cordes J M and Jenet F A 2012 Astrophys. J. 752 54
Cornish N J and Sesana A 2013 Class. Quantum Grav. 30 224005
Damour T and Vilenkin A 2001 Phys. Rev. D 64 064008
Damour T and Vilenkin A 2005 Phys. Rev. D 71 063510
Demorest P B et al 2013 Astrophys. J. 762 94
Finn L S and Lommen A N 2010 Astrophys. J. 718 1400
Grishchuk L P 2005 Phys.-USP. 58 1
Handzo E, Christy B, Lommen A N and Perrodon D 2015 (arXiv: 1510.09084)
Harry G M, Fritschel P, Shaddock D A, Folkner W andphinney E S 2006 Class. Quantum Grav. 23 4887
Henry G R and Paik H-J 1969 Nature 224 1188
Hobbs G 2015 Class. Quantum Grav. 30 224007
Kanda N and the LCGT collaboration 2011 (arXiv: 1112.3092K)
Kawamura S et al 2011 Class. Quantum Grav. 28 094011
Koop M and Finn L S 2014 In preparation
Kramer M and Champion D J 2013 Class. Quantum Grav. 30 224009
LIGO Scientific Collaboration et al 2015 Class. Quantum Grav. 32 074001
Lommen A N and Demorest P 2013 Class. Quantum Grav. 30 224001
Lommen A N et al 2006 Astrophys. J. 642 1012
Magorrian J et al 1998 Astronom. J. 115 2285
McConnell N J and Ma C-P 2013 Astrophys. J. 764 184
Messinger C, Lommen A, Demorest P and Ransom S 2011 Class. Quantum Grav. 28 055001
Mingarelli C M F and Sidery T 2014 Phys. Rev. D 90 062011
Pennucci T T, Demorest P B and Ransom S M 2014 Astrophys. J. 790 93
Phinney E S 2001 (arXiv: 0108028v1)
Planck Collaboration et al 2014 (arXiv: 1411.2271)
Rosado P, Sesana A and Gair J 2015 in preparation
Sesana A 2013 Mon. Not. R. Astron. Soc. 433 L1
Sesana A and Vecchio A 2010 Phys. Rev. D 81 104008
Shannon R M et al 2013 Science 342 334
Siemens X, Ellis J, Jenet F and Romano J D 2013 Class. Quantum Grav. 30 224015
Somija K 2012 Class. Quantum Grav. 29 124007
Volonteri M, Haardt F and Madau P 2003 Astrophys. J. 582 559