Stephen Taylor | Research Statement

Jet Propulsion Laboratory, 4800 Oak Grove Drive – Pasadena, CA 91109
+1 (626) 689-5832 • Stephen.R.Taylor@jpl.nasa.gov
stevertaylor.github.io • stevertaylor • stephen-taylor

My research interests centre around the gravitational-wave (GW) study of compact stellar systems, their progenitor populations, and their astrophysical environments. The LIGO/Virgo collaboration's ground-breaking detection of stellar-mass black hole (BHs) binary systems – a source class that could not have been observed through traditional electromagnetic means – recently inaugurated the field of observational GW astronomy. But much like the electromagnetic spectrum, the GW landscape is expansive and ripe for discovery. At frequencies 11 orders of magnitude lower than the LIGO sensitivity band is the pulsar-timing band, where the dominant sources are nanohertz-emitting supermassive ($M > 10^8 M_\odot$) black-hole binaries (SMBHBs) that are formed during the mergers of massive galaxies. My work has covered both extremes, but recently I have focused on bringing new inference strategies to pulsar-timing efforts as a means of accelerating the detection of nanohertz GWs, studying the ionized interstellar medium, and performing pulsar noise characterization.

Previous and current research accomplishments

I commenced my research career by investigating the prospects for catalogues of double neutron-star systems detected by ground-based laser interferometers to be used as "standard sirens". This exploited the potential to directly infer luminosity distances from their gravitational waveforms, rather than relying on the piecewise-constructed electromagnetic cosmic-distance ladder. Instead of searching for electromagnetic counterparts to obtain a redshift, I constructed a hierarchical Bayesian model that used the information in the redshifted chirp-mass to simultaneously infer the neutron-star mass distribution and the Hubble constant. I adapted this model to probe the dark-energy equation of state and the progenitor star-formation rate with the potential third-generation Einstein Telescope.

My interests then diversified to the low-frequency GW regime, where rigorous Bayesian analysis was relatively new to the field. A pulsar-timing detection of GWs will correspond to substantial evidence for a model where the pulse arrival-time data is correlated between pulsars by the distinctive Hellings and Downs overlap-reduction function. This is accurate only if the stochastic GW background (GWB) is isotropic, and while this assumption is sufficient for initial detection, detailed characterization of the GWB angular distribution will need more flexible models for the overlap-reduction function. I developed analysis tools to probe the spatial correlations between pulsars as a means of inferring the angular distribution of GWB power, and later authored a high-profile *Physical Review Letter* on behalf of the European Pulsar Timing Array (EPTA) collaboration that provided the first such angular power constraints. This was part of a series of publications by the EPTA, in which I formed part of the core analysis team to study GWB and individual binary constraints with the latest pulsar-timing datasets. I recently extended my research of GWB-power anisotropy to perform phase-coherent mapping of the nanohertz GW sky (not just stochastic signals), which in turn informed my further development of phase-coherent mapping methods for ground-based laser interferometers.

When galaxies merge, their centrally-located SMBHs sink via dynamical friction within the merger remnant to form a bound pair at parsec separations, whereupon stellar loss-cone scattering and viscous circumbinary-disk drag evolve the binary to milliparsec separations. At this point GW emission dominates the orbital evolution. There are three major research groups [EPTA (Europe), NANOGrav (North America), PPTA (Australia)] working to ensure that pulsar-timing arrays (PTAs)2 measure these nanohertz GWs via the precision monitoring of an ensemble of galactic millisecond pulsars. Deviations of pulse arrival times from an ephemeris model are searched for the influence of extragalactic GWs, which perturb the light

travel time between the Earth and the pulsar. The target signal is a stochastic GW background (SGWB) formed of the incoherent superposition of individual binary signals, although strategies are also in place to search for individual systems. The SGWB causes stochastic timing deviations that are correlated across different pulsars with a unique signature ? the ?Hellings and Downs curve?3. Pulsar-timing has recently constrained the GW strain amplitude at frequencies of 1/year to be below 10-15 with 95% credibility4,5,6, necessitating the galaxy evolution community to revise their models of the SGWB amplitude7. While this is a great success, near-future detection is impeded in two main ways, which influence GW searches at opposite ends of current modelling approaches. At the initial timing-analysis stage, forthcoming limits to pulsar timing precision will not necessarily be radiometer noise, but rather pulse ?jitter? arising from the difference between an ensemble averaged pulse profile-template and the observed average over a finite number of pulses8. Jitter and other ?profile-domain? processes are not currently modelled in SGWB searches, which deal with observation-averaged pulse time-of-arrival (TOA) data as the fundamental input. The large collecting areas of new telescopes (e.g. MeerKAT9) are such that jitter will likely set the future pulsar-timing noise floor, impeding these (and current) instruments from near-future detection. Secondly, search models of SGWB signals have heretofore followed very basic analytic prescriptions, usually with either a two-parameter power-law spectrum, or with an additional low-frequency turnover to model super-efficient binary hardening through couplings to the galactic environment10. Since PTAs are at the stage where stringent signal constraints are being placed, it is crucial to improve these by having physically-detailed and well-motivated spectral models. This is currently inhibited by the intractability of an analytic model that includes all the necessary effects.
Objectives: In light of these major impediments, this project aims to: RO1. Build a hierarchical Bayesian model that extends pulsar-timing inference backward from the current initial time-of-arrival data to the raw observed pulse profiles. RO2. Expand the RO1 model to include multiple pulsars such that SGWB searches can be performed simultaneously with pulse-shape inference and jitter mitigation. RO3. Exploit Bayesian model emulation to build GW-signal models that are informed by detailed numerical population synthesis calculations. RO4. Develop a web database to store simulated SMBHB populations for reproducibility of RO3 results, cross-validation of simulation techniques, and as a community resource.
Overview of the action: The successful outcome of GravPANTHER will result in a new Bayesian inference framework that accelerates the robust pulsar-timing detection of nanohertz gravitational waves. This framework will build on Dr Taylor?s already significant methodological advances as evidenced by his open-source suite of PTA analysis tools (NX01).
Abbott, B. P. et al., Physical Review Letters 116, 061102 (2016). 6 Shannon, R. M. et al., Science 349, 1522 (2015). 2 Foster, R. S. and Backer, D. C., ApJ 361, 300 (1990). 7 Sesana A. et al., (2016), arXiv:1603.09348. 3 Hellings, R. W. and Downs, G. S., ApJ 265, L39 (1983). 8 Shannon, R. W. et al., MNRAS 443, 1463 (2014). 4 Lentati, L., Taylor, S. R., et al., MNRAS 453, 2576 (2015). 9 Booth, R. S., and Jonas, J. L., African Skies 16, 101, (2012). 5 Arzoumanian, Z. et al., ApJ 821, 13 (2016). 10 Sampson, L., et al. Physical Review D 91, 084055 (2015). This action will directly result in several high-profile peer-reviewed publications, with Dr Taylor leading the International Pulsar Timing Array (IPTA)11 in implementing these methods to publish the tightest nanohertz SGWB constraints and the most detailed GW study of SMBHB dynamical environments. 1.1.2 Research methodology and approach This action is divided into the following work packages (WP), with associated tasks (T), and project coherence illustrated in Figure 1. WP1 (addressing RO1): An open-source Bayesian profile-domain analysis package Preliminary analysis has demonstrated the potential for ?generative pulsar timing analysis?12 to combat pulse jitter, wherein the pulse profile shape is inferred in tune with all ephemeris and noise parameters. This can additionally model band-dependent profile changes, epoch-to-epoch profile stochasticity, and secular profile evolution. Dr Taylor will collaborate on a high-profile IPTA project to produce a robust profile-domain analysis package. (external collaborators: Dr L. Lentati, Dr R. Shannon, Dr M. Vallisneri, Dr J. A. Ellis) T1.1: (Months: 0-2) Constructing the profile-domain likelihood

Code development for hierarchical profile-domain likelihood, incorporating all processes influencing the pulse shape, arrival-time, evolution, and band-dependence. T1.2: (Months: 1-2) Simulations on idealized datasets Concurrently with the later stages of Task 1.1, an internal mock data challenge will be performed, contrasting results with traditional TOA-domain analysis. Datasets will be idealized, ranging from single epoch to single radio-channel tests. Parallel-tempering MCMC techniques will sample from a form of the profile-domain likelihood that is marginalized over low-level linear parameters (such that the search dimensionality is manageable, e.g. O(10-102)). T1.3: (Months: 2-4) Operating on real datasets The full hierarchical (un-marginalized) profile-domain likelihood will be tested on real multi-channel datasets spanning decades of observations [e.g. Parkes J1909-3744 data]. The dimensionality of the likelihood (>O(103)) will require highly efficient sampling. We will use a recently available Hamiltonian Monte Carlo (HMC) No-U-Turn-Sampler (hereon referred to as NUTS) with custom coordinate transformations to avoid the ?Neal?s funnel? problem of hierarchical likelihoods13. As in Task 1.2, we will compare our results to traditional TOA-domain analysis. WP2 (addressing RO2): Profile-domain GW searches The methodology of WP1 will be extended to an array of pulsars for SGWB searches, where isolation of profile-domain processes will increase the analysis sensitivity (see Figure 2). (external collaborators: Dr M. Vallisneri, Dr J. A. Ellis) T2.1: (Months: 4-6) TOA-domain upper limits TOA-domain SGWB constraints are already possible with the marginalized TOA-domain likelihood, however an upper limit has never been derived with the full hierarchical form. We do so here with the latest IPTA data, again using the NUTS from Task 1.3. We will compare the resulting SGWB constraints to conventional marginalized likelihood results, which should be identical under optimized sampling. T2.2: (Months: 6-9) Profile-domain upper limits Following from Task 2.1, we will swap out the TOA-domain hierarchical likelihood for the profile-domain hierarchical likelihood of WP1. Since the Bayesian analysis is introduced at a stage closer to the raw data, and properly isolates profile-domain processes from signal and noise processes, these SGWB upper limits should be more sensitive than TOA-domain upper limits. We will once again analyse the latest IPTA data to obtain the best-possible upper limits

11 Verbiest, J. P. W., et al., MNRAS 458, 1267 (2016). 12 Lentati, L., and Shannon, R. M., MNRAS 454, 1058, (2015) 13 van Haasteren, R., ?Piccard?, (2016).

T2.3: (Months: 9-12) Profile-domain SGWB searches and detection significance Tasks 2.1 and 2.2 do not model the Hellings and Downs correlations between pulsars, since this is not necessary to obtain GW upper limits. The current HMC NUTS code shows inefficient (but not incorrect) parameter space exploration when these are modelled. We will explore alternative coordinate transformations for the low-level parameters, in addition to an adaptive mass matrix for the Hamiltonian trajectories. If these still show inefficiencies, we will explore methods such as Riemannian-manifold HMC or elliptical slice sampling. Regardless of inefficiencies, we will perform the first profile-domain GW search, with an associated posterior odds ratio for detection. The technique will be applied to the latest IPTA data for the best sensitivity and detection prospects. WP3 (addressing RO3): Bayesian model emulation for gravitational-wave searches Physically-detailed SGWB spectral models will be developed which can be rapidly computed within either the TOA-domain or profile-domain likelihood. These models will include all parameters that influence the dynamical evolution of SMBHBs through the final parsec of orbital separation, including the degree of stellar loss-cone refilling, binary eccentricity, and binary accretion rate from a circumbinary disk14. (external collaborators: Dr J. R. Gair, Dr L. Sampson) T3.1: (Months: 12-13) Developing the Bayesian model emulation formalism In Bayesian model emulation, we run a small number of expensive simulations with varying input parameter values (e.g. binary eccentricity at the time of hardening). The output are treated as draws from a Gaussian process (GP), which are used to train the GP kernel hyper-parameters, allowing predictions (with uncertainties) at arbitrary parameter locations15 (see Figure 3 for a demonstration). We will adapt this for SGWB searches, where the model will be a GP trained on synthesised black-hole populations, instead of an analytically-derived signal model. T3.2: (Months: 12-15) Detailed SMBHB population synthesis simulations During Task 3.1, Dr

Taylor and Dr Sesana will run detailed population synthesis simulations across the full expanse of the physical parameter space influencing SMBHB evolution in galactic nuclei. Spectra will be constructed from these simulations to train a GP model, which is stored for subsequent inference. T3.3: (Months: 15-19) Constraining the final-parsec problem with a GP model The trained GP model will be used in an analysis of publicly available PTA data (delivering a high-profile short-author publication) and the latest IPTA data (delivering a high-profile collaboration result) to constrain the typical dynamical environment of SMBHBs. WP4 (addressing RO4): Web-portal and repository for SMBHB populations New PTA techniques require simulated binary populations to inject realistic GW signals into pulsar-timing data for validation. There are no standard repositories to find such populations, requiring researchers to obtain them via private communication with a handful of experts. This is an impediment to open science, and the proper cross-validation of the simulation routines used to obtain the populations. (external collaborators: Prof S. McWilliams, Mr J. Simon, Mr L. Kelley) T4.1: (Months: 15-18) Database design and prototype build Concurrently with Task 3.3, we will design a SQL database around the binary populations generated by Dr Sesana for WP3. The goals are to ensure that the results of Task 3.3 are reproducible, and more broadly that the PTA community has access to a standard repository of populations for robust signal recovery tests. The prototype database will permit browser queries and direct python queries. T4.2: (Months: 18-19) Hosting simulation output from other research groups Different groups employ different binary simulation techniques. We will host the simulation output from our external collaborators (listed above) to ensure a complete overview of methodologies are represented, and to allow straightforward comparison of simulation products between groups. T4.3: (Months: 19-24) Front-end development and paper release We will work closely with the UoB computer science department to deliver a user-friendly web front-end with a similar layout to the Illustris and EAGLE web-portals. Concurrently with this front-end development, we will draft a paper describing the hosted simulations and the database features.

14 Sesana, A., CQG, 30, 224014, (2013). 15 Rasmussen, C. E., and Williams, C. K. I., MIT Press, (2006).

1.1.3 Originality and innovative aspects of the research programme (including ER career expansion and new host collaborative opportunities) ? This action constitutes a huge advance in PTA Bayesian modelling techniques, creating a new framework for the forthcoming decade of nanohertz GW constraints and ultimately detection. ? It will provide the first open-source software for profile-domain timing analysis, replacing the TOA-domain status quo. This work is possible and timely, with the foundation already laid in the literature. ? We will exploit the timely development of an automatically-tuned HMC sampler that remains efficient in thousands of dimensions, having ramifications far beyond pulsar timing into other fields (e.g. CMB studies, and phase-space sampling of complex biological molecules). ? A key milestone will be performing the first profile-domain GW search, which will isolate profile-domain processes from the GW signal to increase analysis sensitivity. ? We will develop a Bayesian model-emulator of GW spectra to capture the complex interplay of SMBHB dynamical influences in galactic nuclei. Model emulation has far-reaching applications where analytic models are intractable, e.g. biological systems, weather prediction. ? We will publish the most sensitive GW constraints on final-parsec SMBHB dynamics, shedding much-needed light on the means by which SMBHs merge. ? We will create a community resource in the form of an online repository of simulated SMBHB populations. This will allow cross-validation of simulation techniques between research groups, and for realistic signals to be injected into new pipelines for robust efficacy studies. ? PTAs are producing longer and more precise datasets, but the time-to-detection with current analysis methods is still 7 years. These new techniques will increase analysis sensitivity, and accelerate detection significance by years, making pre-2020 detection a tantalizing prospect. ? The tools and experiences developed here will provide Dr Taylor with a foundation to forge an independent research path and adopt leadership roles within the European and broader international pulsar-timing community. ? The host will become a hub for EPTA and IPTA GW searches, fostering collaborative links with the North American, Chinese, South African, and Australian pulsar-timing communities.