Constraints on the Dynamical Environments of Supermassive Black-Hole Binaries Using Pulsar-Timing Arrays

Stephen R. Taylor, Joseph Simon, and Laura Sampson

Constraints On The Dynamical Environments Of Supermassive Black-hole Binaries Using Pulsar-timing Arrays

Stephen R. Taylor,[1, *] Joseph Simon,[2] and Laura Sampson[3]

[1]*Jet Propulsion Laboratory, California Institute of Technology,*
4800 Oak Grove Drive, Pasadena, CA 91106, USA
[2]*Center for Gravitation, Cosmology and Astrophysics,*
University of Wisconsin Milwaukee, PO Box 413, Milwaukee WI 53201, USA
[3]*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy,*
Northwestern University, 2131 Tech Drive, Evanston, IL 60208, USA

We introduce a technique for gravitational-wave analysis, where Gaussian process regression is used to emulate the strain spectrum of a stochastic background by training on population-synthesis simulations. This leads to direct Bayesian inference on astrophysical parameters. For PTAs specifically, we interpolate over the parameter space of supermassive black-hole binary environments, including 3-body stellar scattering, and evolving orbital eccentricity. We illustrate our approach on mock data, and assess the prospects for inference with data similar to the NANOGrav 9-yr data release.

INTRODUCTION

Expensive computer simulations are a common problem in astrophysics. Complicated studies, from the details of exoplanet formation to cosmological galaxy evolution, require numerical simulations that can take hours to months [1–3]. Each simulation initiates from a set of physical assumptions or input parameters. Bayesian estimation of these parameters using experimental data requires a full exploration of possible values, i.e. simulations at each point in parameter space, becoming prohibitively computationally expensive. In this Letter, we describe a method to avoid this problem using the technique of Gaussian process (GP) regression [4]. We demonstrate this method with astrophysical inference of supermassive black-hole binary (SMBHB) environments using the nanohertz gravitational waves (GWs) that will soon be detected by pulsar timing arrays (PTAs) [5].

The detection of the stochastic background of GWs from the mergers of SMBHBs throughout our universe will likely occur within the next ~ 10 years [6, 7]. As discussed in Arzoumanian *et al.* [8] (N16), and Sampson *et al.* [9], the details of the GW background are sensitive to the physics that drives mergers. This can include binary eccentricity, 3-body scattering encounters with stars in the galactic nucleus, accretion disk dynamics, and much more — all of which are theorized solutions to the *final parsec problem* [10], which arises because dynamical friction in post-merger galaxy remnants is insufficient to drive SMBHs to milliparsec orbital separations [11]. It is at these separations that GW emission dominates binary evolution, and ultimately drives the binaries to merger. Since we do not observe multiple SMBHs in the cores of galaxies [12, 13], we expect that most galactic mergers lead to BH mergers, and so some confluence of mechanisms must solve the final parsec problem [14].

In this Letter we show that it is possible to perform full Bayesian inference on the detailed physics of SMBHB mergers by training a GP on a small set of simulations, initialized on a grid in input parameter space. We then use the GP as a prior on the shape of the GW spectrum in PTA data. This prior is a function of astrophysical parameters, so our analysis directly samples the posterior distributions of these parameters.

Analytic spectral models exist to encapsulate the influences of stars/gas [9] or eccentricity *alone* [15, 16] on the strain spectrum — this type of analysis was performed in N16. Stellar scattering, though, impacts SMBHB eccentricity evolution [17], and so we must model their combined influence. GP regression gives us the first PTA technique capable of constraining the *combined* dynamical influences of a dense galactic-center stellar distribution and evolving SMBHB eccentricity.

METHODOLOGY

GP regression is a powerful interpolation scheme which treats (noisy) data as a random draw from a multivariate Gaussian process with a mean vector and covariance function. We use the initial training data to learn the GP's covariance structure, after which we make predictions about the outcome of hypothetical experiments between the training points (interpolation), and beyond them (extrapolation). Our assumption is that the data do not exhibit pathological discontinuities that would be poorly captured by a smooth covariance (kernel) function. GP regression allows us to interpolate and extrapolate beyond the initial training data, and additionally provides the uncertainty in the prediction which we can propagate forward to our final statistics. It has recently been investigated in a ground-based GW context to marginalize over the uncertainties between expensive numerical relativity waveforms and cheaper post-Newtonian waveforms [18–20], but has not been used in GW data-analysis to emulate the statistical properties of astrophysical populations.

Our training data are GW spectra built from synthesized populations of SMBHBs, evolved from different initial conditions. In this proof-of-concept study we restrict our attention to post–dynamical-friction binary eccentricity and stellar densities in the galactic core. Our technique can be trivially expanded to include a wider range of galactic-center environmental influences and assembly-history processes.

We build and train a GP capable of predicting the shape of the strain spectrum when binary populations have (i) non-zero eccentricity; (ii) significant orbital evolution driven by stellar hardening; (iii) all of the above. Our procedure is as follows:

1. **Simulate training data:** Build a bank of SMBHB populations by initializing simulations with different binary eccentricities and environments.
[Computationally expensive].

2. **Train the Gaussian-process model:** Model the strain distribution over population realizations as Gaussian with a mean and standard error. This noisy data is used to train a GP and optimize its kernel hyper-parameters.
[Computationally cheap].

3. **Analyze PTA data with Gaussian-process model:** The trained GP predicts the shape of the strain spectrum for our GW analysis.
[Computationally cheap].

We now briefly discuss steps 1 and 2. Step 3 is briefly described in the results section.

Population synthesis

The orbital eccentricity and environmental couplings of a population of SMBHBs do not directly impact the merger-rate density. They effect the evolution of the binaries and the frequency distribution of the characteristic strain emitted by each source, $h_c(f_{gw})$. Thus, all populations share a common binary merger-rate density prescription like that described in Simon and Burke-Spolaor [21], which utilizes observations of galaxy stellar mass functions (GSMFs), galaxy close-pair fractions ($f_{\rm pair}$), and black hole-host galaxy relations ($M-M_{\rm bulge}$) to infer the rate density. To ease computational burden, a single measurement of each observable was used: GSMF from Ilbert *et al.* [22], $f_{\rm pair}$ from Robotham *et al.* [23], and $M-M_{\rm bulge}$ from McConnell and Ma [24]. For each realization, the value of each observable was randomly drawn from a Guassian with widths equal to the cited one-sigma uncertainty regions, this method propagates observational uncertainties through model inference.

A SMBHB's orbital evolution from interactions with a fixed, isotropic, unbound, cuspy stellar background are described in Ravi *et al.* [25], which draws extensively on nu-
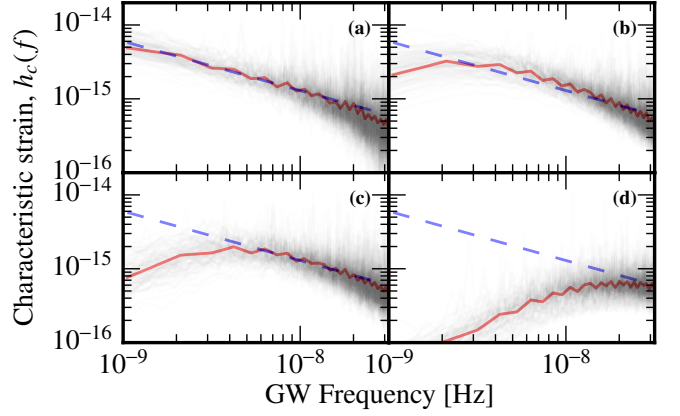

FIG. 1. Characteristic strain spectra for binary populations under various astrophysical conditions. Grey lines indicate single population realizations, while red is the mean over realizations. Dashed blue shows a $\propto f^{-2/3}$ strain spectrum for reference. (a) $\{e_0 = 0, \rho = 10\,M_\odot{\rm pc}^{-3}\}$; (b) $\{e_0 = 0.95, \rho = 10\,M_\odot{\rm pc}^{-3}\}$; (c) $\{e_0 = 0, \rho = 10^4\,M_\odot{\rm pc}^{-3}\}$; (d) $\{e_0 = 0.95, \rho = 10^4\,M_\odot{\rm pc}^{-3}\}$

merical simulations from Quinlan [26], Sesana *et al.* [27], and Sesana [28]. As in Ravi *et al.* [25], we anchor all binary evolution to a starting frequency of $f_{gw} = 10^{-12}$ Hz, which is sufficiently low that any GW emission is outside PTA sensitivity ranges. The initial eccentricity parameter, e_0, is set at this frequency, and is evolved across the PTA band in accordance with whichever mechanism (stars or GWs) dominates the orbital evolution.

To generate a population, we draw a finite number of sources whose binary parameters match the merger-rate density for each realization, with eccentricities that have been evolved according to the prescribed environmental conditions $\{e_0, \rho\}$, where ρ is the mass density of stars at the gravitational influence radius of the binary. We assume that all binaries have the same initial eccentricity and are embedded in a stellar distribution with the same density. In future work we will study how varying the distribution of environmental conditions across sources effects the performance of our method.

Almost all ($> 99\%$) of the GW strain in the PTA sensitivity band ($10^{-10} < f_{gw} < 10^{-7}$ Hz) comes from less than 2×10^5 sources. These sources are saved for each realization. The characteristic strain spectrum of the GW background, $h_c(f)$, is built up as the quadrature sum of the strain from each source. The eccentricity of each source determines how the strain is distributed across frequency, with quasi-circular sources emitting mostly in the $n = 2$ harmonic of the orbital frequency, while more eccentric sources emit across tens to hundreds of harmonics [29]. Figure 1 shows the characteristic strain spectra at the four corners of the combined parameter space $\{e_0, \rho\}$ explored in this work.

Training the Gaussian process

Our work uses the `George` GP regression library [30].[1] We train a separate GP at each sampling frequency of the GW spectrum, from a base frequency of $f = 1/T$ up in increments of $1/T$ for a pre-defined number of modes, where T is the span between the maximum and minimum TOAs in the PTA dataset.[2] While we assume that variation of the strain across astrophysical parameter space is smooth, we do not enforce that it is smooth across frequencies.

We construct the strain spectra for 100 population realizations at each of $14 \times 13 = 182$ combinations in $\{e_0, \rho\}$ parameter space, where $e_0 \in \mathcal{U}[0.0, 0.95]$ and $\log_{10}(\rho/M_\odot \mathrm{pc}^{-3}) \in \mathcal{U}[1, 4]$. We then form a data vector at each GW frequency, corresponding to the base-10 logarithm of the mean squared characteristic-strain, $\log_{10}\langle h_c^2 \rangle$, for every set of dynamical conditions. Each data point has an associated uncertainty corresponding to the log-space standard deviation of the squared strain over the 100 population realizations.

To characterize the GP, we choose a *Squared Exponential* (SE) kernel function with a flat, diagonal metric. This stationary, infinitely differentiable function has tunable length-scale parameters which set the correlation between data points across parameter space, and in our case has the form

$$k(\vec{x}, \vec{x'}) = \sigma_k^2 \exp\left(-(x_i - x_i')^2/2\sigma_i^2 - (x_j - x_j')^2/2\sigma_j^2\right), \quad (1)$$

where $\vec{x}$, $\vec{x'}$ are two coordinates in our 2-d $\{e_0, \log_{10}\rho\}$ parameter space, each dimension of the input coordinate can have a separate variance σ_i^2, σ_j^2, and the kernel has an overall variance scaling, σ_k^2.[3]

The data vector at each frequency is used to train the GP by mapping out the posterior distribution of the kernel function parameters. We set the kernel parameters to their *maximum-a-posteriori* values. An example is shown in Fig. 2 for a GW frequency of $1/(30\,\mathrm{yr})$, where the left panel shows the coordinates of the training data (red points) along with the prediction of the trained-GP. The right panel shows the uncertainty in this prediction.

Figure 3 shows tests of the interpolation and extrapolation fidelity of the GP at $f = 1/(30\,\mathrm{yr})$, achieved by removing some points from the training data. Interpolation performs very well, but extrapolation is poor, so we do not use our GP model outside of the boundaries of the original training data.[4] Figu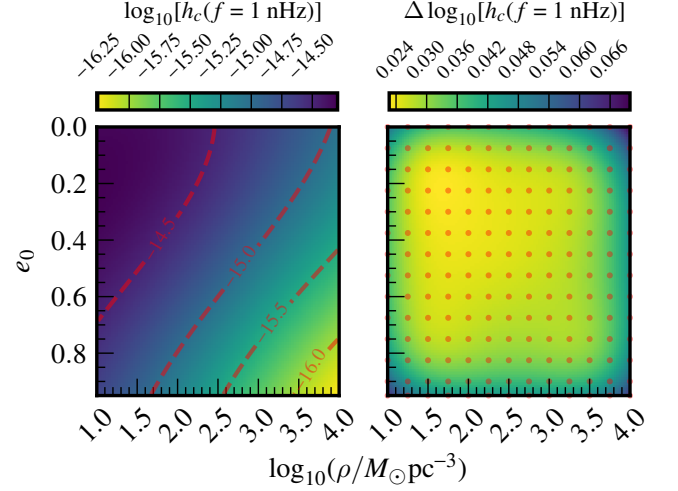re 3 also shows that there is no noticeable difference in performance when fixing the kernel parameters to their *maximum-a-posteriori* values compared to sampling from their posterior distribution. As we expand our GP model to more astrophysical influences in the future, we will monitor this to assess whether kernel parameter sampling is also needed as part of our hierarchical Bayesian pulsar-timing model.

FIG. 2. Mean (**left**) and standard deviation (**right**) of a GP trained on the strain spectrum at $f = 1/(30\,\mathrm{yr}) \sim 1\,\mathrm{nHz}$. The training spectra are constructed from 100 synthesized SMBHB populations at each of 182 different points in $\{e_0, \rho\}$ space (right, red points), with kernel length scales of 10 and 0.83, respectively. The uncertainty in the GP prediction is dominated by the intrinsic scatter in the simulations.

Finally, for this study we are interested in the influence of dynamical processes on the *shape* of the GW spectrum, and not on the overall amplitude implied by merger-rate density assumptions. Hence, before training we find the mean amplitude of the $\{e_0 = 0, \rho = 10\,M_\odot \mathrm{pc}^{-3}\}$ simulations at $f = 1/\mathrm{yr}$, then divide all spectra by this amplitude.

RESULTS

Our GP model fits into existing Bayesian analysis schemes as an extension of the hierarchical model. Current schemes expand the time-domain GWB signal onto a Fourier basis, then marginalize over the Fourier amplitudes with a Gaussian prior [32, 33]. This prior is parametrized to have a variance that obeys e.g. (i) a power-law spectrum with amplitude and spectral index parameters; (ii) a

---

[1] http://dan.iel.fm/george/current/
[2] For realistic datasets most of the spectral information is confined to the lowest few frequencies, but we model at least 20.
[3] We investigated the impact of kernel choices on our results, and found that both *Matérn* $3/2$ and *Matérn* $5/2$ kernels [31] give comparable posterior recoveries with respect to the SE kernel.
[4] Performance is graded on how far the removed points depart from the GP

prediction, in units of the GP uncertainty. For interpolation all removed points lie within the 1-σ region, while for extrapolation the removed points start at ~ 1-σ away, and depart farther from the GP prediction as we move farther outside the boundaries of the training set.
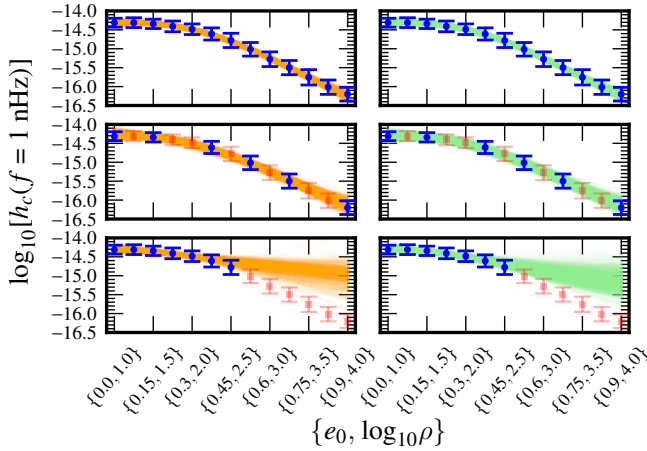
FIG. 3. Interpolation and extrapolation fidelity of the GP, trained on only the blue points. These points with uncertainties are combinations of $\{e_0, \log_{10}\rho\}$ along the downwards left-right diagonal of Fig. 2's left panel, where the strain is varying fastest. Faded red squares show data excluded from the training. Colored regions are random draws from the trained GP. The **left column** shows 10 random draws from the GP with each of 50 different kernel parameter values drawn from their posterior distribution. The **right column** is similar, except all 500 draws use the *maximum-a-posteriori* kernel parameters.

turnover model with three 'shape' parameters [9]; (iii) a free spectrum with a separate amplitude parameter per GW frequency.

We use and extend model (iii) by placing a Gaussian prior on the power spectrum at each GW frequency with a mean and standard deviation from the GP prediction. We numerically sample the power spectrum at each frequency (n parameters) and the hyper-parameters of the GP model ($\{A, e_0, \log_{10}\rho\}$), for a GWB model space with $n + 3$ dimensions. To ensure efficient numerical sampling, we perform a coordinate transformation such that the power spectrum at each frequency is sampled as a zero-mean unit-variance Gaussian [34]. The GP method is implemented within the NX01 Bayesian PTA package [35],[5] and we use a parallel-tempering MCMC sampler with several bespoke proposal schemes.[6]

We test the performance of our method in both the presence and absence of a detectable GW background. We use two types of data sets: (1) a simplified dataset consisting of an array of 18 pulsars whose positions and statistical noise properties match those used in N16. These pulsars are concurrently timed for 30 years with a total of 390 TOAs in each. (2) A realistic dataset consisting of an 18-pulsar array that emulates the N16 PTA in every way, with the same noise properties and an identical observation schedule over

a ~ 9 year baseline [36]. We calibrate this dataset to give the reported 95% upper limit of 1.5×10^{-15} on the amplitude of an $f^{-2/3}$ GWB power-spectrum at $f = 1/\mathrm{yr}$. This required us to inflate the TOA uncertainties by a factor of 1.4 across all pulsars.

Type-(1)— We create two type-(1) datasets with differing injected GWB spectra. First, we create 100 new population realizations with dynamical conditions $\{e_0 = 0.65, \log_{10}(\rho/M_\odot \mathrm{pc}^{-3}) = 3.35\}$, and inject the mean characteristic-strain spectrum. This combination of conditions was chosen to provide a strain-spectrum turnover within the frequency coverage of a 30 year data span, so that both dynamical properties could be constrained. We also deliberately excluded this simulation from the training data. Second, we inject a pure $f^{-2/3}$ power-law strain spectrum. In both cases, the overall amplitude scaling is set such that the Bayes factor for spatial correlations $\gtrsim 10^5$, because any inaccuracies in the GP model are best illuminated with parameter estimation in high signal-to-noise scenarios where the posterior uncertainty is small and modeling inaccuracies will show as systematic offsets.

The white noise was fixed at the level of the TOA uncertainties, while the GWB signal and per-pulsar red-noise parameter priors were uninformative. The resulting 2-d marginalized posterior distributions in $\{e_0, \rho\}$ space are shown in Fig. 4, where we see that the injected values lie within the 68% credible regions in both cases. The parameter degeneracy arises because a population of highly-eccentric binaries in a moderately-dense stellar environment can give a strain spectrum turnover at a similar frequency as a lower eccentricity population in a more dense stellar environment. Similar physical parameter degeneracies were observed in Chen *et al.* [37].

Type-(2)— No GWB signal is injected in this dataset, since N16 reported no evidence for GWs. Instead, we perform astrophysical inference on this fake dataset as was performed on the true dataset. The parameter priors were as in the Type-(1) datasets above. Instead of searching for a GWB amplitude with a prior that is uniform in log-space, we anchor the amplitude with priors that agree with two different astrophysical predictions at high GW frequencies. These predicted amplitude distributions derive from McWilliams *et al.* [38] (MOP14), and Sesana [39] (S13), where the assumptions that led to them are explained in the relevant papers, and contrasted in N16. They respectively constitute an optimistic ($\log_{10} A = -14.4 \pm 0.26$) and a median ($\log_{10} A = -15.0 \pm 0.2$) prediction of the GWB amplitude at high frequencies.

The resulting 2-d marginalized posterior distributions in $\{e_0, \rho\}$ space are shown in Fig. 5 for the different prior assumptions. Under the assumptions of the MOP14 and S13 amplitude priors, the emulated NANOGrav 9-yr data favors eccentric SMBHBs that are coupled to dense stellar distributions. This is because an $f^{-2/3}$ power-law spectrum with MOP14 or S13 amplitudes would produce a
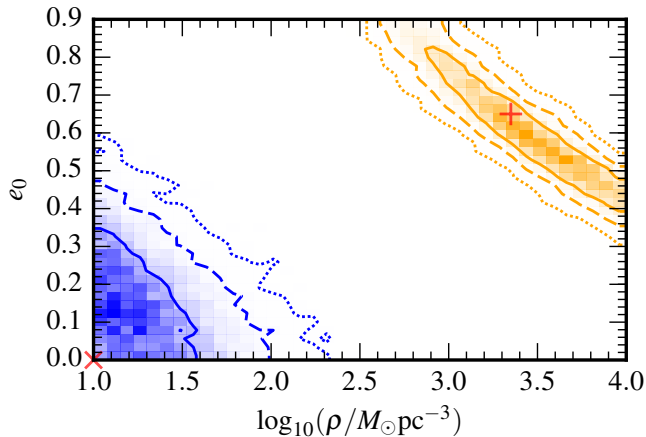
FIG. 4. Type-(1) datasets (see text) are created with injected GWB signals whose spectra are influenced by the dynamical environments indicated by $+$ and $\times$. The recovered posterior distributions from employing our GP search model are shown as orange and blue density regions respectively, with associated 68% (solid), 95% (dashed), and 99.7% (dotted) credible intervals.
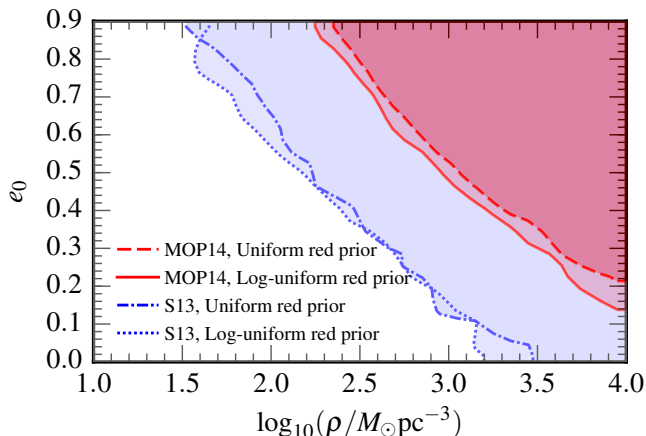

FIG. 5. Astrophysical inference on an emulated NANOGrav 9-yr dataset. The GWB amplitude at high frequencies is anchored to the MOP14 and the S13 priors. The lines indicate the boundaries of the 68% credible regions under different prior assumptions. See the text for details and discussion.

GWB signal in excess of current constraints at low frequencies. To be consistent with the absence of a GW signal, dynamical influences must attenuate the low-frequency GW spectrum. Because it is larger, the MOP14 prior necessitates stronger spectral attenuation and thus more extreme dynamical conditions. With the more recent Sesana *et al.* [40] amplitude prior (centered on $\log_{10} A = -15.4$) the posterior distribution in e_0 and $\log_{10} \rho$ would be flat and uninformative. The choice of red-noise prior has a minimal influence.

DISCUSSION

We have introduced a powerful new method for GW data analysis, where a model for the stochastic GW strain spectrum is built entirely through Gaussian-process emulation of population synthesis simulations. This is the first PTA method capable of searching for the *combined* SMBHB dynamical influences of stellar scattering encounters from the loss-cone and evolving binary eccentricity. These effects are important for understanding SMBHB dynamics [41], but are poorly constrained by current observations. PTAs offer a more direct measurement of sub-parsec binary evolution than is currently available. Further, our technique can be trivially expanded to incorporate additional effects, such as accretion from a circumbinary disk, different $M-\sigma$ relations, or dynamical-friction timescales.

Our results for a dataset that precisely emulates the sensitivity of the NANOGrav 9-year dataset indicate that the MOP14 and S13 GWB amplitude predictions mildly favor a population of eccentric SMBHBs evolving within dense stellar distributions. The true dataset gives quantitatively similar results, and will be explored further in a forthcoming analysis [42].

Beyond pulsar-timing analysis, this method can be adapted for LIGO or LISA population inference. Current schemes perform demographic analysis to recover the distributions of compact-system properties with either a parametric function [43, 44], or using a histogram with bin heights constrained by a GP prior [40, 44]. Linking these distributions back to progenitor properties or evolutionary channels has so far only been performed for discrete population synthesis simulations [45]. With our technique, one could use the histogram model with a GP prior trained on a set of population synthesis simulations. This would allow sampling of the continuous posterior distribution of progenitor properties, such as metallicity or the efficiency of common-envelope hardening. We will investigate this in future work.

A notebook that allows the user to read in simulations and construct their own GP model is available at https://github.com/stevertaylor/gw_pta_emulator.

---

* Stephen.R.Taylor@jpl.nasa.gov

[1] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. F. Snyder, S. Bird, D. Nelson, and L. Hernquist, Nature **509**, 177 (2014), arXiv:1405.1418 [astro-ph.CO].

[2] A. Fitts *et al.*, (2016), arXiv:1611.02281 [astro-ph.GA].

[3] T. K. Suzuki, M. Ogihara, A. Morbidelli, A. Crida, and T. Guillot, ArXiv e-prints (2016), arXiv:1609.00437 [astro-ph.EP].

[4] C. K. Williams and C. E. Rasmussen, the MIT Press **2**, 4 (2006).

[5] R. S. Foster and D. C. Backer, ApJ **361**, 300 (1990).

[6] S. R. Taylor, M. Vallisneri, J. A. Ellis, C. M. F. Mingarelli, T. J. W. Lazio, and R. van Haasteren, ApJ **819**, L6 (2016), arXiv:1511.05564 [astro-ph.IM].

[7] X. Siemens, J. Ellis, F. Jenet, and J. D. Romano, Classical and Quantum Gravity **30**, 224015 (2013), arXiv:1305.3196 [astro-ph.IM].

[8] Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. J. Chamberlin, S. Chatterjee, B. Christy, J. M. Cordes, N. J. Cornish, K. Crowter, P. B. Demorest, X. Deng, T. Dolch, J. A. Ellis, R. D. Ferdman, E. Fonseca, N. Garver-Daniels, M. E. Gonzalez, F. Jenet, G. Jones, M. L. Jones, V. M. Kaspi, M. Koop, M. T. Lam, T. J. W. Lazio, L. Levin, A. N. Lommen, D. R. Lorimer, J. Luo, R. S. Lynch, D. R. Madison, M. A. McLaughlin, S. T. McWilliams, C. M. F. Mingarelli, D. J. Nice, N. Palliyaguru, T. T. Pennucci, S. M. Ransom, L. Sampson, S. A. Sanidas, A. Sesana, X. Siemens, J. Simon, I. H. Stairs, D. R. Stinebring, K. Stovall, J. Swiggum, S. R. Taylor, M. Vallisneri, R. van Haasteren, Y. Wang, W. W. Zhu, and NANOGrav Collaboration, ApJ **821**, 13 (2016), arXiv:1508.03024.

[9] L. Sampson, N. J. Cornish, and S. T. McWilliams, Phys. Rev. D **91**, 084055 (2015), arXiv:1503.02662 [gr-qc].

[10] M. Milosavljević and D. Merritt, in *The Astrophysics of Gravitational Wave Sources*, American Institute of Physics Conference Series, Vol. 686, edited by J. M. Centrella (2003) pp. 201–210, astro-ph/0212270.

[11] M. C. Begelman, R. D. Blandford, and M. J. Rees, Nature **287**, 307 (1980).

[12] S. Burke-Spolaor, MNRAS **410**, 2113 (2011), arXiv:1008.4382.

[13] L. Wang, J. E. Greene, W. Ju, R. R. Rafikov, J. J. Ruan, and D. P. Schneider, ApJ **834**, 129 (2017), arXiv:1611.00039.

[14] F. M. Khan, D. Fiacconi, L. Mayer, P. Berczik, and A. Just, ApJ **828**, 73 (2016), arXiv:1604.00015.

[15] E. A. Huerta, S. T. McWilliams, J. R. Gair, and S. R. Taylor, Phys. Rev. D **92**, 063010 (2015), arXiv:1504.00928 [gr-qc].

[16] S. Chen, A. Sesana, and W. Del Pozzo, ArXiv e-prints (2016), arXiv:1612.00455.

[17] C. Roedig and A. Sesana, Journal of Physics Conference Series **363**, 012035 (2012), arXiv:1111.3742.

[18] C. J. Moore and J. R. Gair, Physical Review Letters **113**, 251101 (2014), arXiv:1412.3657 [gr-qc].

[19] J. R. Gair and C. J. Moore, Phys. Rev. D **91**, 124062 (2015), arXiv:1504.02767 [gr-qc].

[20] C. J. Moore, C. P. L. Berry, A. J. K. Chua, and J. R. Gair, Phys. Rev. D **93**, 064001 (2016), arXiv:1509.04066 [gr-qc].

[21] J. Simon and S. Burke-Spolaor, ApJ **826**, 11 (2016), arXiv:1603.06577.

[22] O. Ilbert, H. J. McCracken, O. Le Fèvre, P. Capak, J. Dunlop, A. Karim, M. A. Renzini, K. Caputi, S. Boissier, and et al., A&A **556**, A55 (2013), arXiv:1301.3157.

[23] A. S. G. Robotham, S. P. Driver, L. J. M. Davies, A. M. Hopkins, I. K. Baldry, N. K. Agius, A. E. Bauer, J. Bland-Hawthorn, and et al., MNRAS **444**, 3986 (2014), arXiv:1408.1476.

[24] N. J. McConnell and C.-P. Ma, ApJ **764**, 184 (2013), arXiv:1211.2816.

[25] V. Ravi, J. S. B. Wyithe, R. M. Shannon, G. Hobbs, and R. N. Manchester, MNRAS **442**, 56 (2014), arXiv:1404.5183.

[26] G. D. Quinlan, New A **1**, 35 (1996), astro-ph/9601092.

[27] A. Sesana, F. Haardt, and P. Madau, ApJ **651**, 392 (2006), astro-ph/0604299.

[28] A. Sesana, ApJ **719**, 851 (2010), arXiv:1006.0730 [astro-ph.CO].

[29] P. C. Peters and J. Mathews, Phys. Rev. **131**, 435 (1963).

[30] S. Ambikasaran, D. Foreman-Mackey, L. Greengard, D. W. Hogg, and M. O'Neil, ArXiv e-prints (2014), arXiv:1403.6015 [math.NA].

[31] M. L. Stein, *Interpolation of spatial data: some theory for kriging* (Springer Science & Business Media, 2012).

[32] L. Lentati, P. Alexander, M. P. Hobson, S. Taylor, J. Gair, S. T. Balan, and R. van Haasteren, Phys. Rev. D **87**, 104021 (2013), arXiv:1210.3578 [astro-ph.IM].

[33] R. van Haasteren and M. Vallisneri, Phys. Rev. D **90**, 104012 (2014), arXiv:1407.1838 [gr-qc].

[34] R. M. Neal, Annals of statistics (2003).

[35] S. R. Taylor, "NX01: Updated release," (2016).

[36] Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. Chamberlin, S. Chatterjee, B. Christy, J. M. Cordes, N. Cornish, K. Crowter, P. B. Demorest, T. Dolch, J. A. Ellis, R. D. Ferdman, E. Fonseca, N. Garver-Daniels, M. E. Gonzalez, F. A. Jenet, G. Jones, M. L. Jones, V. M. Kaspi, M. Koop, M. T. Lam, T. J. W. Lazio, L. Levin, A. N. Lommen, D. R. Lorimer, J. Luo, R. S. Lynch, D. Madison, M. A. McLaughlin, S. T. McWilliams, D. J. Nice, N. Palliyaguru, T. T. Pennucci, S. M. Ransom, X. Siemens, I. H. Stairs, D. R. Stinebring, K. Stovall, J. K. Swiggum, M. Vallisneri, R. van Haasteren, Y. Wang, and W. Zhu, ApJ **813**, 65 (2015), arXiv:1505.07540 [astro-ph.IM].

[37] S. Chen, H. Middleton, A. Sesana, W. Del Pozzo, and A. Vecchio, ArXiv e-prints (2016), arXiv:1612.02826 [astro-ph.HE].

[38] S. T. McWilliams, J. P. Ostriker, and F. Pretorius, ApJ **789**, 156 (2014).

[39] A. Sesana, MNRAS **433**, L1 (2013), arXiv:1211.5375.

[40] A. Sesana, F. Shankar, M. Bernardi, and R. K. Sheth, MNRAS **463**, L6 (2016), arXiv:1603.09348.

[41] A. Sesana, Classical and Quantum Gravity **30**, 224014 (2013), arXiv:1307.2600.

[42] Z. Arzoumanian *et al.*, (forthcoming).

[43] M. R. Adams, N. J. Cornish, and T. B. Littenberg, Phys. Rev. D **86**, 124032 (2012), arXiv:1209.6286 [gr-qc].

[44] W. M. Farr, N. Sravan, A. Cantrell, L. Kreidberg, C. D. Bailyn, I. Mandel, and V. Kalogera, ApJ **741**, 103 (2011), arXiv:1011.1459.

[45] S. Stevenson, F. Ohme, and S. Fairhurst, ApJ **810**, 58 (2015), arXiv:1504.07802 [astro-ph.HE].