MNRAS 000, 1–8 (2023) Preprint 15 February 2024 Compiled using MNRAS LATEX style file v3.0

State-space analysis of a continuous gravitational wave source with apulsar timing array: inclusion of the pulsar terms

Kimpson[1], O'Neil[2], Melatos[2], O'Leary, Evans[3], Moran, others, etc. etc.⋆
[1]School of Physics, University of Melbourne, Parkville, VIC 3010, Australia
[2]OzGrav, University of Melbourne, Parkville, VIC 3010, Australia
[3]Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Victoria 3010, Australia

Last updated 15 February 2024

ABSTRACT

- PTAs can detect the nano-Hz GWs
- The detection procedure can be formulated as a state-space method
- Previous state-space methods use frequencies
- We lift this limitation

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

1 INTRODUCTION

State-space methods have been shown (Kimpson et al. 2024a,b) to be a promising and complementary alternative to standard data analysis techniques for the detection of nanohertz gravitational waves (GWs) by pulsar timing arrays (PTAs; Tiburzi 2018; Verbiest et al. 2021). State-space methods self-consistently track the intrinsic timing noise in PTA pulsars (e.g. Shannon & Cordes 2010a; Lasky et al. 2015; Caballero et al. 2016; Goncharov et al. 2021), enabling the specific time-ordered realisation of the timing noise to be disentangled statistically from the GW-induced modulations. Conversely, traditional PTA analysis techniques average over the ensemble of possible timing noise realizations to obtain the noise power spectral density (PSD), and discards time-ordered information when taking the modulus of the complex Fourier phase in each PSD frequency bin. State-space methods are then complementary to traditional PTA analysis techniques for two reasons: (i) they track the actual noise realization rather than an ensemble average; (ii) they preserve time ordering by implicitly preserving the Fourier phases, which the PSD discards. State-space methods follow the customary signal processing approach in many industrial and scientific electrical engineering applications. TK:CITATIONS HERE

For the purposes of simplicity, and in order to maintain consistency with prior works applying state-space methods within neutron star astrophysics (e.g. Meyers et al. 2021a,b), previous state-space PTA analysis techniques have accepted as an input a sequence of pulse frequencies. However, in order to analyse real data and make contact with traditional PTA analysis techniques it is necessary for state-space methods to accepted as an input traditional pulsar observables such as the pulsar phase or pulse time of arrival (ToA).

In this paper we take a step towards lifting the above limitation. We demonstrate how to extend previous PTA state-space methods to accept as an input the measured pulse phase. We show how the intrinsic pulse phase and the intrinsic frequency for each constituent pulsar in the PTA can be tracked using a Kalman filter (Kalman 1960). We combine the Kalman tracking of the intrinsic pulsar's states with a Bayesian nested samplers (Skilling 2006) to estimate the static parameters of a continuous, quasi-monochromatic GW source, and calculate the marginalized likelihood for model selection.

The paper is organised as follows. In Section 2 we... Throughout the paper we adopt natural units, with $c = G = \hbar = 1$

2 STATE-SPACE FORMULATION

In this section we formulate PTA detection of the stochastic GW background as a state-space problem. There are N pulsars in the array, labelled $1 \leqslant n \leqslant N$. The n-th pulsar's spin frequency $f_{\rm p}^{(n)}(t)$, as measured in the local, freely-falling rest frame of the pulsar's centre of mass, evolves according to a stochastic differential equation. We adopt the phenomenological model for $f_{\rm p}^{(n)}(t)$ as a mean-reverting Ornstein-Uhlenbeck process also used in previous state-space analyses (Vargas & Melatos 2023; Kimpson et al. 2024a,b). The n-th pulsar's phase $\phi_{\rm p}^{(n)}(t)$ is the time integral of $f_{\rm p}^{(n)}(t)$. Section 2.1 presents the dynamical equations for the evolution of the intrinsic states, $f_{\rm p}^{(n)}(t)$ and $\phi_{\rm p}^{(n)}(t)$. An observer at Earth measures the pulse phase $\phi_{\rm m}^{(n)}(t)$. The relation between the states $f_{\rm p}^{(n)}(t)$, $\phi_{\rm p}^{(n)}(t)$ and the measurement, $\phi_{\rm m}^{(n)}(t)$, is described by a measurement equation which quantifies the influence of a continuous GW source and PTA instrumental noise. The model of the measurement equation is presented in Section 2.2. In Section 2.3 we summarise the static parameters of the model.

⋆ Contact e-mail: tom.kimpson@unimelb.edu.au

2.1 Dynamical state evolution

The instantaneous phase of the n-th pulsar is the time integral of the rest frame spin frequency, which itself evolves according to a mean-reverting Ornstein-Uhlenbeck process, described by a Langevin equation with a time-dependent drift term (Vargas & Melatos 2023), viz.

$$\frac{d\phi_{\rm p}^{(n)}}{dt} = f_{\rm p}^{(n)} \tag{1}$$

$$\frac{df_{\rm p}^{(n)}}{dt} = -\gamma^{(n)}[f_{\rm p}^{(n)} - f_{\rm em}^{(n)}(t)] + \dot{f}_{\rm em}^{(n)}(t) + \xi^{(n)}(t)\,, \tag{2}$$

where $f_{\rm em}^{(n)}$ is the deterministic evolution, an overdot denotes a derivative with respect to t, $\gamma^{(n)}$ is a damping constant whose reciprocal specifies the mean-reversion timescale, and $\xi^{(n)}(t)$ is a white noise stochastic process which satisfies

$$\langle \xi^{(n)}(t) \rangle = 0\,, \tag{3}$$

$$\langle \xi^{(n)}(t)\xi^{(n')}(t') \rangle = [\sigma^{(n)}]^2 \delta_{n,n'}\delta(t-t')\,. \tag{4}$$

In Equation (4), $[\sigma^{(n)}]^2$ parametrizes the noise amplitude and produces characteristic root mean square fluctuations $\approx \sigma^{(n)}/[\gamma^{(n)}]^{1/2}$ in $f_{\rm p}^{(n)}(t)$ (Gardiner 2009). The deterministic evolution $f_{\rm em}^{(n)}$ is attributed to magnetic dipole braking for the sake of definiteness, with braking index $n_{\rm em} = 3$ (Goldreich & Julian 1969). PTAs are typically composed of millisecond pulsars (MSPs), for which the quadratic correction due to $n_{\rm em}$ in $f_{\rm p}^{(n)}(t)$ is negligible over the observation time $T_{\rm obs} \sim 10\,$yr. Consequently, $f_{\rm em}^{(n)}(t)$ can be approximated accurately by

$$f_{\rm em}^{(n)}(t) = f_{\rm em}^{(n)}(t_1) + \dot{f}_{\rm em}^{(n)}(t_1)t\,, \tag{5}$$

where t_1 labels the time of the first TOA.

2.2 Modulation of the pulsar states by a GW

In the presence of a GW, $\phi_{\rm p}^{(n)}(t)$ is related to $\phi_{\rm m}^{(n)}(t)$ and $f_{\rm m}^{(n)}(t)$ via a measurement equation,

$$\phi_{\rm m}^{(n)}(t) = \phi_{\rm m}^{(n)}\left[t - d^{(n)}\right] - f_{\rm p}^{(n)}\left[t - d^{(n)}\right] g^{(n)}(t) + \varepsilon^{(n)}(t)\,, \tag{6}$$

where $d^{(n)}$ labels the distance to the n-th pulsar, $\phi_{\rm m}^{(n)}$,$f_{\rm p}^{(n)}$ are evaluated at the retarded time $t - d^{(n)}$, and $\varepsilon^{(n)}(t)$ is a Gaussian measurement noise which satisfies

$$\langle \varepsilon^{(n)}(t) \rangle = 0\,, \tag{7}$$

$$\langle \varepsilon^{(n)}(t)\varepsilon^{(n')}(t') \rangle = \sigma_{\rm m}^2 \delta_{n,n'}\delta(t-t')\,. \tag{8}$$

In Equation (8), $\sigma_{\rm m}^2$ is the variance of the measurement noise at the telescope and is shared between all pulsars. In Equation (6) the measurement function $g^{(n)}(t)$ is,

$$g^{(n)}(t) = -\frac{H_{ij}[q^{(n)}]^i[q^{(n)}]^j}{2\Omega[1+\boldsymbol{n}\cdot\boldsymbol{q}^{(n)}]} \times \left[\sin\left(-\Omega t + \Phi_0\right) - \sin\left\{-\Omega t + \Phi_0 + \chi^{(n)}\right\}\right]\,, \tag{9}$$

where $[q^{(n)}]^i$ labels the i-th coordinate component of the n-th pulsar's position vector $\boldsymbol{q}^{(n)}$, Ω is the constant angular frequency of the GW, $\boldsymbol{n}$ is a unit vector specifying the direction of propagation of the GW, H_{ij} is the spatial part of the GW amplitude tensor, Φ_0 is the phase offset of the GW with respect to some reference time and $\chi^{(n)}$ is a pulsar-dependent phase correction (see Kimpson et al. 2024b, for additional discussion on $\chi^{(n)}$). The derivation of the measurement equations, Equation (6)–(9) is presented in Appendix A.

2.3 Static parameters

The state-space model described in Sections 2.1 and 2.2 comprises $5N$ static parameters, that are specific to the pulsars in the array, viz.

$$\boldsymbol{\theta}_{\rm psr} = \left\{\gamma^{(n)}, \sigma^{(n)}, f_{\rm em}^{(n)}(t_1), \dot{f}_{\rm em}^{(n)}(t_1), \chi^{(n)}\right\}_{1\leqslant n\leqslant N}\,. \tag{10}$$

It also comprises seven parameters, that are specific to the GW source, viz.

$$\boldsymbol{\theta}_{\rm gw} = \{h_0, \iota, \psi, \delta, \alpha, \Omega, \Phi_0\}\,, \tag{11}$$

where h_0 is the characteristic wave strain, ι is the orbital inclination, ψ is the polarisation angle, δ is the declination and α is the right ascension. These parameters enter the model through Equation (9), with $H_{ij} = H_{ij}(h_0, \iota, \psi, \delta, \alpha)$ and $\boldsymbol{n} = \boldsymbol{n}(\delta, \alpha)$. The complete set of $7 + 5N$ static parameters is denoted by $\boldsymbol{\theta} = \boldsymbol{\theta}_{\rm gw} \cup \boldsymbol{\theta}_{\rm psr}$. While we assume no prior information about $\boldsymbol{\theta}_{\rm gw}$, there are constraints on $\boldsymbol{\theta}_{\rm psr}$ from electromagnetic observations; for example estimates of $d^{(n)}$ are accurate to $\sim 10\%$ typically (Cordes & Lazio 2002; Verbiest et al. 2012; Desvignes et al. 2016; Yao et al. 2017).

2.4 Analysis scheme

State-space analysis methods for single-source continuous GWs observed by PTAs (e.g. Kimpson et al. 2024a,b) use a likelihood-based Bayesian framework to infer the static parameters of the model and select between models with and without a GW.

Specifically a Kalman filter is used....

which forms the basis for nested sampling.

We do not cover the details of the Kalman filter -nested sampling scheme used in this work and refer the reader to (e.g. Kimpson et al. 2024a,b) for details.

Whilst previous state-space analyses have used a linear Kalman filter, the non-linear relation between $\boldsymbol{Y}(t)$ and $\boldsymbol{X}(t)$ necessitates the use of an extended (non-linear) Kalman filter (Zarchan & Musoff 2000) in this paper. Appendix B demonstrates how to discretise the continuous dynamical equations, Equations (2)–(5) and Equations (**??**)–(**??**), and measurement equations, Equation (**??**)–(**??**), and how to solve them recursively using an extended Kalman filter.

3 VALIDATION TESTS WITH SYNTHETIC DATA

3.1 Construction of synthetic data

3.2 Parameter estimation

In this section we outline how synthetic data $\boldsymbol{Y}(t)$ are generated in order to validate the analysis scheme in Section **??**. Application of the analysis scheme to synthetic data enables validation to occur systematically and under controlled conditions. In this section, by way of preparation, we explain how to generate the synthetic data employed in the tests, namely noisy frequency time series $\boldsymbol{Y}(t) = f_{\rm m}^{(n)}(t)$ for $1 \leqslant n \leqslant N$.

In order to synthesize $\boldsymbol{Y} = \{f_{\rm m}^{(1)}(t), \ldots, f_{\rm m}^{(N)}(t)\}$, we integrate Equations (2)–(5) numerically using a Runge-Kutta Itô integrator

Set	Parameter	Injected value	Units	Prior
$\boldsymbol{\theta}_{\rm psr}$	$f_{\rm em}^{(n)}(t_1)$	$f_{\rm ATNF}^{(n)}$	Hz	Uniform$\left[f_{\rm ATNF}^{(n)} - 10^3\eta_f^{(n)}, f_{\rm ATNF}^{(n)} + 10^3\eta_f^{(n)}\right]$
	$\dot{f}_{\rm em}^{(n)}(t_1)$	$\dot{f}_{\rm ATNF}^{(n)}$	$\rm s^{-2}$	Uniform$\left[\dot{f}_{\rm ATNF}^{(n)} - 10^3\eta_{\dot{f}}^{(n)}, \dot{f}_{\rm ATNF}^{(n)} + 10^3\eta_{\dot{f}}^{(n)}\right]$
	$\sigma^{(n)}$	$\sigma_{\rm SC}^{(n)}$	$s^{-3/2}$	LogUniform$\left[10^{-2}\sigma_{\rm SC}^{(n)}, 10^{2}\sigma_{\rm SC}^{(n)}\right]$
	$\gamma^{(n)}$	10^{-13}	$\rm s^{-1}$	—
$\boldsymbol{\theta}_{\rm gw}$	γ_a	10^{-2}	$\rm years^{-1}$	LogUniform$\left[10^{-4}, 10^{2}\right]$
	h	10^{-12}	–	LogUniform$\left[10^{-14}, 10^{-9}\right]$

Table 1. Injected static parameters used to generate synthetic data to validate the analysis scheme in the representative analysis of Section 4. The prior used for Bayesian inference is also displayed (rightmost column). The top and bottom sections of the table contain $\boldsymbol{\theta}_{\rm psr}$ and $\boldsymbol{\theta}_{\rm gw}$ respectively. The subscript "ATNF" denotes values obtained from the ATNF pulsar catalogue as described in Section 3. The subscript "SC" on $\sigma^{(n)}$ indicates that the injected value is calculated from Equation (12) and the empirical timing noise model for MSPs in Shannon & Cordes (2010a). The quantities $\eta_f^{(n)}$ and $\eta_{\dot{f}}^{(n)}$ are the uncertainties in $f_{\rm em}^{(n)}(t_1)$ and $\dot{f}_{\rm em}^{(n)}(t_1)$ respectively, as quoted in the ATNF catalogue. We do not infer $\gamma^{(n)} \sim 10^{-5}T_{\rm obs}$ for simplicity, so no prior is set. The priors on $\boldsymbol{\theta}_{\rm psr}$ are justified in detail in Kimpson et al. (2024a,b). TK:text on GW parameters...

implemented in the `sdeint` python package [1]. This produces random realizations of $f_{\rm p}^{(n)}(t)$ for $1 \leqslant n \leqslant N$, which we convert to $f_{\rm m}^{(n)}(t)$ via Equations (6)–(9). The numerical solutions depend on how we choose $\boldsymbol{\theta}_{\rm psr}$, $\boldsymbol{q}^{(n)}$ and $\sigma_{\rm m}$ or, equivalently, how we specify the configuration of a synthetic PTA. In Section **??** we describe how we choose the remaining elements of $\boldsymbol{\theta}$, namely $\boldsymbol{\theta}_{\rm gw}$. This latter step is equivalent to specifying the synthetic SMBHB source and differs from one test to the next according to the goal of the test.

In this paper we adopt for consistency the same $\boldsymbol{\theta}_{\rm psr}$ values as in K24, i.e. the $N = 47$ MSPs in the 12.5-year NANOGrav dataset (Arzoumanian et al. 2020). We assume all pulsars are observed with cadence $T_{\rm cad} = 1$ week over a 10 year period. Fiducial values for $\boldsymbol{q}^{(n)}$, $d^{(n)}$, $f_{\rm em}^{(n)}(t_1)$, and $\dot{f}_{\rm em}^{(n)}(t_1)$ are read from the Australia Telescope National Facility (ATNF) pulsar catalogue (Manchester et al. 2005) using the `psrqpy` package (Pitkin 2018). No direct measurements exist for $\gamma^{(n)}$. The mean reversion timescale typically satisfies $[\gamma^{(n)}]^{-1} \gg T_{\rm obs}$ (**?**Meyers et al. 2021a,b; Vargas & Melatos 2023); in this paper, for the sake of simplicity, we fix $\gamma^{(n)} = 10^{-13}$ $\rm s^{-1}$ for all n. No direct measurements exist for $\sigma^{(n)}$ either. We relate $\sigma^{(n)}$ to the root mean square TOA noise $\sigma_{\rm TOA}^{(n)}$ accumulated over an interval of length $T_{\rm cad}$ by

$$\sigma^{(n)} \approx \sigma_{\rm TOA}^{(n)} f_{\rm p}^{(n)}(t_1) {T_{\rm cad}}^{-3/2} \,. \tag{12}$$

As in K24, the empirical timing noise model for MSPs from Shannon & Cordes (2010b), applied to the 12.5-year NANOGrav dataset, implies median$[\sigma^{(n)}] = 5.51 \times 10^{-24}$ $\rm s^{-3/2}$, min$[\sigma^{(n)}] = 1.67 \times 10^{-26} \rm s^{-3/2}$ for PSR J0645+5158 and max$[\sigma^{(n)}] = 2.56 \times 10^{-19}$ $\rm s^{-3/2}$ for PSR J1939+2134.

In a similar vein, $\sigma_{\rm m}^{(n)}$ can be related to, $\sigma_{\rm TOA}^{(n)}$, by

$$\sigma_{\rm m}^{(n)} \approx f_{\rm p}^{(n)}(t_1)\sigma_{\rm TOA}^{(n)} {T_{\rm cad}}^{-1} \,. \tag{13}$$

For an MSP with $f_{\rm p}^{(n)} \sim 0.1$ kHz, $T_{\rm cad} = 1$ week, and $\sigma_{\rm TOA}^{(n)} \sim 1\mu$s, Equation (13) implies $\sigma_{\rm m}^{(n)} \sim 10^{-10}$ Hz. The most accurately timed pulsars have $\sigma_{\rm TOA}^{(n)} \sim 10$ ns and $\sigma_{\rm m}^{(n)} \sim 10^{-12}$ Hz. In this paper, for simplicity and the sake of definiteness, we fix $\sigma_{\rm m}^{(n)} = 10^{-11}$ Hz for all n, and take it as known *a priori* rather than a parameter to be inferred. When analysing real data this assumption is easily relaxed.

[1] https://github.com/mattja/sdeint

Although $\sigma_{\rm m}^{(n)}$ is assumed to be the same for every pulsar, $f_{\rm m}^{(n)}$ is constructed from a different random realisation of $\varepsilon^{(n)}(t)$ for each pulsar.

4 REPRESENTATIVE ANALYSIS

In this section we apply the Kalman filter in conjunction with nested sampling in order to calculate the model evidence (marginalized likelihood) and the joint posterior distribution for the static parameters. In Section 4.1 we consider the problem of detecting a GW in noisy PTA data in terms of a Bayesian model selection procedure, following the lead of other PTA analyses (e.g. Agazie et al. 2023; Antoniadis et al. 2023; Reardon et al. 2023; Xu et al. 2023). In Section 4.2 we infer the joint posterior distribution for the static parameters. TK: More sentences and detail here once we have decided on the definitive results to show

4.1 Detection

In this section we frame the problem of claiming a GW detection as a model selection procedure. We define $\mathcal{M}_{\rm GW}$ as the model that assumes a stochastic GW background signal exists in the data. This is equivalent to using the model outlined in Section **??**. We define $\mathcal{M}_{\rm null}$ as the model that assumes no GW background signal exists in the data. This is equivalent to setting $g^{(n)}(t) = 1$ in Equation (6). $\mathcal{M}_{\rm GW}$ is parameterised by $\boldsymbol{\theta}$. $\mathcal{M}_{\rm null}$ is parameterised by $\boldsymbol{\theta}_{\rm psr}$. The evidence integral $\mathcal{Z}$ returned by nested sampling, Equation (**??**), is the probability of the data $\boldsymbol{Y}$ given a model $\mathcal{M}_{\rm M}$. The support in the data for the presence of a GW signal over the absence of a GW signal is quantified via the Bayes factor

$$\beta = \frac{\mathcal{Z}(\boldsymbol{Y}|\mathcal{M}_{\rm GW})}{\mathcal{Z}(\boldsymbol{Y}|\mathcal{M}_{\rm null})} \,. \tag{14}$$

The Bayes factor β is plotted as functions of h_0 in Figure 1 for the representative source in Table 1. We vary the source amplitude from $h_0 = 10^{-15}$ (undetectable) to $h_0 = 10^{-12}$ (easily detectable).

4.2 Parameter estimation

In this section we...
Figure 2shows...

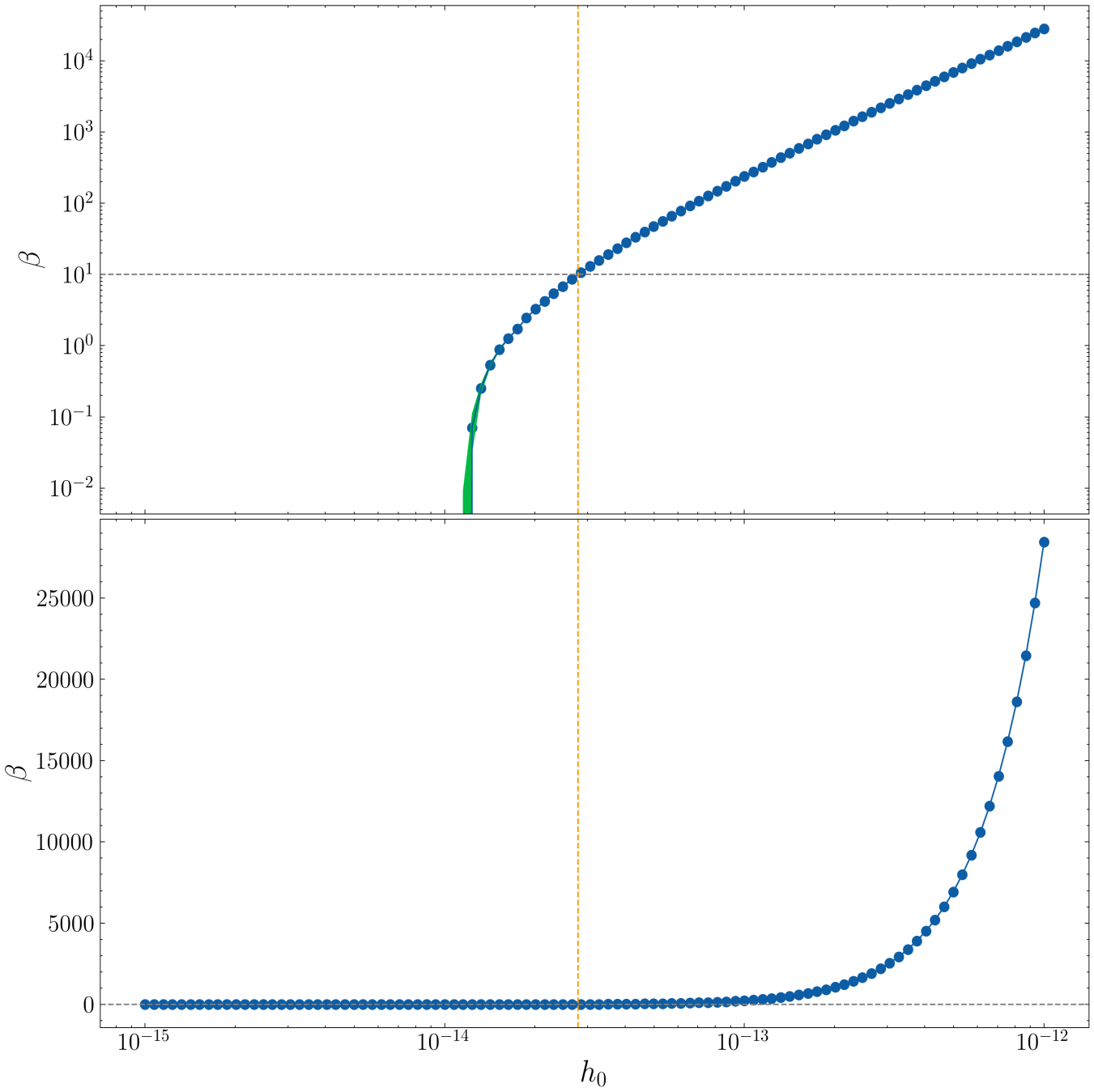

Figure 1. Bayes factor (odds ratio) β_{M} between the competing models $\mathcal{M}_{\mathrm{GW}}$ a(GW present in data) and $\mathcal{M}_{\mathrm{null}}$ (GW not present in data) as a function of the signal amplitude, h_0, for the representative example in Table 1. The horizontal grey dashed line labels an arbitrary detection threshold, $\beta_{\mathrm{M}} = 10$. The minimum detectable strain for which $\beta < 10$ equals X. The axes are plotted on logarithmic scales. TK:Caption TBD. Top panel is the same as bottom, just with a log y-axis. Green region is the uncertainty.

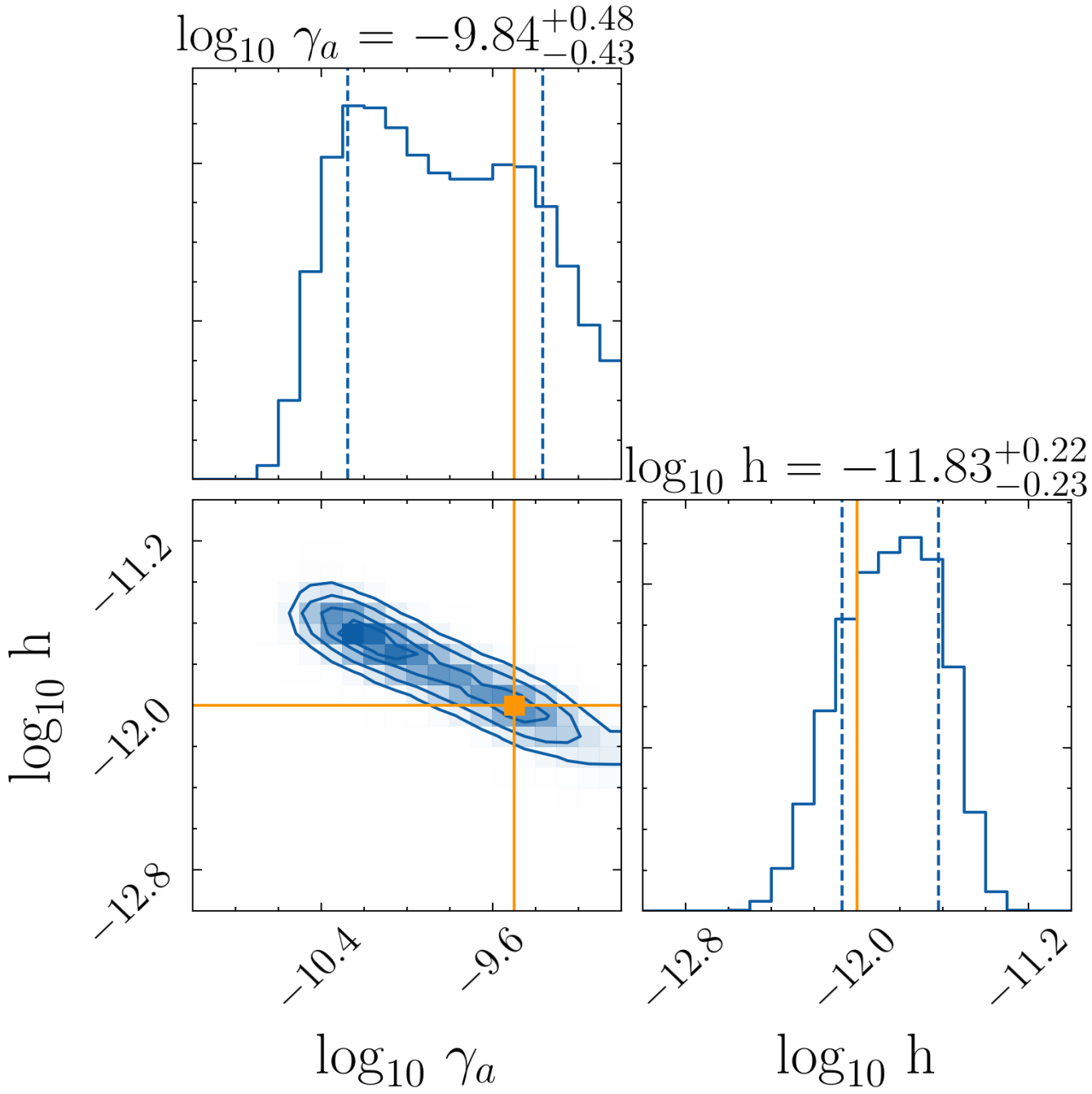

Figure 2. TK:Caption TBD. Just GW parameters shown, but nested sampling inference is run over all $\boldsymbol{\theta}$

REFERENCES

Agazie G., et al., 2023, ApJ, 951, L8

Antoniadis J., et al., 2023, arXiv e-prints, p. arXiv:2306.16214

Arzoumanian Z., et al., 2020, ApJ, 905, L34

Caballero R. N., et al., 2016, MNRAS, 457, 4421

Cordes J. M., Lazio T. J. W., 2002, arXiv e-prints, pp astro–ph/0207156

Desvignes G., et al., 2016, Monthly Notices of the Royal Astronomical Society, 458, 3341

Gardiner C., 2009, Stochastic Methods: A Handbook for the Natural and Social Sciences. Springer Series in Synergetics, Springer Berlin Heidelberg, https://books.google.com.au/books?id=otg3PQAACAAJ

Goldreich P., Julian W. H., 1969, ApJ, 157, 869

Goncharov B., et al., 2021, MNRAS, 502, 478

Kalman R. E., 1960, Journal of Basic Engineering, 82, 35

Kimpson T., et al., 2024a, Submitted to MNRAS

Kimpson T., et al., 2024b, Submitted to MNRAS

Lasky P. D., Melatos A., Ravi V., Hobbs G., 2015, MNRAS, 449, 3293

Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, AJ, 129, 1993

Meyers P. M., Melatos A., O'Neill N. J., 2021a, MNRAS, 502, 3113

Meyers P. M., O'Neill N. J., Melatos A., Evans R. J., 2021b, MNRAS, 506, 3349

Pitkin M., 2018, Journal of Open Source Software, 3, 538

Reardon D. J., et al., 2023, ApJ, 951, L6

Shannon R. M., Cordes J. M., 2010a, ApJ, 725, 1607

Shannon R. M., Cordes J. M., 2010b, ApJ, 725, 1607

Skilling J., 2006, Bayesian Analysis, 1, 833

Tiburzi C., 2018, Publ. Astron. Soc. Australia, 35, e013

Vargas A., Melatos A., 2023, TBD, 1, 1

Verbiest J. P. W., Weisberg J. M., Chael A. A., Lee K. J., Lorimer D. R., 2012, ApJ, 755, 39

Verbiest J. P. W., Osłowski S., Burke-Spolaor S., 2021, in , Handbook of Gravitational Wave Astronomy. p. 4, doi:10.1007/978-981-15-4702-7_4-1

Xu H., et al., 2023, Research in Astronomy and Astrophysics, 23, 075024

Yao J. M., Manchester R. N., Wang N., 2017, ApJ, 835, 29

Zarchan P., Musoff H., 2000, Fundamentals of Kalman Filtering: A Practical Approach. Progress in astronautics and aeronautics, American Institute of Aeronautics and Astronautics, https://books.google.com.au/books?id=AQxRAAAAMAAJ

APPENDIX A: DERIVATION OF THE PHASE MEASUREMENT EQUATION

TBD

APPENDIX B: EXTENDED KALMAN FILTER

TK: not an EKF!!

In this appendix, we describe the extended (non-linear) Kalman filter algorithm used in this paper. The Kalman filter accepts a temporal sequence of noisy measurements $\boldsymbol{Y}(t)$ as an input and recovers a temporal sequence of stochastically evolving system state variables, $\boldsymbol{X}(t)$, which are not observed directly. In this paper $\boldsymbol{Y}(t) = \{f_{\rm m}^{(n)}(t_1), \ldots, f_{\rm m}^{(n)}(T_{\rm obs})\}_{1\leqslant n\leqslant N}$ and $\boldsymbol{X}(t) = \{f_{\rm p}^{(n)}(t_1), a^n(t_1), \ldots, f_{\rm p}^{(n)}(T_{\rm obs}), a_{\rm p}^{(n)}(T_{\rm obs})\}_{1\leqslant n\leqslant N}$. Whilst previous state-space analyses of single continuous GW sources have used a linear Kalman filter (Kimpson et al. 2024a,b), the non-linear relation between $\boldsymbol{Y}(t)$ and $\boldsymbol{X}(t)$ (c.f. Equation (6)) necessitates the use of an extended (non-linear) Kalman filter (Zarchan & Musoff 2000) in this paper. General recursion relations for the discrete-time extended Kalman filter are written down for an arbitrary dynamical system in Section B1. The mapping onto the specific continuous-time state-space model in Section **??** is written down in Section B2.

B1 Recursion equations

The extended Kalman filter operates on temporally discrete, noisy measurements $\boldsymbol{Y}_k$, which are related to a set of unobservable discrete system states $\boldsymbol{X}_k$, via a general differentiable function

$$\boldsymbol{Y}_k = h\left(\boldsymbol{X}_k\right) + \boldsymbol{v}_k \,, \tag{B1}$$

where $h(x)$ is the measurement function or observation model, $\boldsymbol{v}_k$ is a zero-mean Gaussian measurement noise, $\mathcal{N} \sim (0, \boldsymbol{R}_k)$ with covariance $\boldsymbol{R}_k$, and the subscript k labels the time-step. Similarly, the states evolve as

$$\boldsymbol{X}_k = f\left(\boldsymbol{X}_{k-1}\right) + \boldsymbol{w}_{k-1} \,, \tag{B2}$$

where $f(x)$ is the state transition function, and $\boldsymbol{w}_k$ is a zero-mean Gaussian measurement noise, $\mathcal{N} \sim (0, \boldsymbol{Q}_k)$ with covariance $\boldsymbol{Q}_k$.

The extended Kalman filter is a recursive estimator with two distinct stages: a "predict" stage and an "update" stage. The predict stage predicts $\hat{\boldsymbol{X}}_{k|k-1}$, the estimate of the state at discrete step k, given the state estimates from step $k-1$. Specifically, the predict step proceeds as

$$\hat{\boldsymbol{X}}_{k|k-1} = f\left(\hat{\boldsymbol{X}}_{k-1|k-1}\right) \tag{B3}$$

$$\hat{\boldsymbol{P}}_{k|k-1} = \boldsymbol{F}_k \hat{\boldsymbol{P}}_{k-1|k-1} \boldsymbol{F}_k^{\intercal} + \boldsymbol{Q}_k \,, \tag{B4}$$

where $\hat{\boldsymbol{P}}_{k|k-1}$ is the covariance of the prediction and $\boldsymbol{F}_k$ is the Jacobian

$$\boldsymbol{F}_k = \left.\frac{\partial f}{\partial \boldsymbol{X}}\right|_{\hat{\boldsymbol{X}}_{k|k-1}} \,. \tag{B5}$$

The predict stage is independent of the measurements; the measurement $\boldsymbol{Y}_k$ is included to update the prediction during the update stage as follows:

$$\boldsymbol{\epsilon}_k = \boldsymbol{Y}_k - h\left(\hat{\boldsymbol{X}}_{k|k-1}\right) \,, \tag{B6}$$

$$\boldsymbol{S}_k = \boldsymbol{H}_k \hat{\boldsymbol{P}}_{k|k-1} \boldsymbol{H}_k^{\intercal} + \boldsymbol{R}_k \,, \tag{B7}$$

$$\boldsymbol{K}_k = \hat{\boldsymbol{P}}_{k|k-1} \boldsymbol{H}_k^{\intercal} \boldsymbol{S}_k^{-1} \,, \tag{B8}$$

$$\hat{\boldsymbol{X}}_{k|k} = \hat{\boldsymbol{X}}_{k|k-1} + \boldsymbol{K}_k \boldsymbol{\epsilon}_k \,, \tag{B9}$$

$$\hat{\boldsymbol{P}}_{k|k} = \left(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k\right) \hat{\boldsymbol{P}}_{k|k-1} \,. \tag{B10}$$

where $\boldsymbol{H}_k$ is the Jacobian

$$\boldsymbol{H}_k = \left.\frac{\partial h}{\partial \boldsymbol{X}}\right|_{\hat{\boldsymbol{X}}_{k|k-1}} \,. \tag{B11}$$

For a full review of the extended Kalman filter, including its derivation, we refer the reader to Zarchan & Musoff (2000).

To apply the Kalman filter in practice means specifying the two functions $h(x)$ and $f(x)$, their associated Jacobians $\boldsymbol{H}_k$ and $\boldsymbol{F}_k$, the covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, and the vectors $\boldsymbol{X}_k$ and $\boldsymbol{Y}_k$. In Section B2 we describe how these components are determined for the state-space model in Section 2.

For the form of equations, whilst the measurement are non linear the states are linear. In this case the Jacobian is X and it is sufficient to update X according to the usual linear equation

$$a = b \tag{B12}$$

B2 State space representation of a PTA analysis

We apply the Kalman recursion relations in Section B1 to the state-space model of a PTA with N pulsars described in Section 2 as follows.

In practice, to avoid errors that arise from finite precision arithmetic, rather than operating directly on $f_{\rm i}^{(n)}$ and $\phi_{\rm i}^{(n)}$ for $i \in \{\rm p, m\}$, we undertake a change of variables,

$$f_{\rm i}^{*(n)}(t) = f_{\rm i}^{(n)} - f_{\rm em}^{(n)} \tag{B13}$$

$$\phi_{\rm i}^{*(n)}(t) = \phi_{\rm i}^{(n)} - \phi_{\rm em}^{(n)} \tag{B14}$$

where

$$\phi_{\rm em}^{(n)} = \frac{df_{\rm em}^{(n)}}{dt} = \dot{f}_{\rm em}^{(n)}(t_1) \tag{B15}$$

In this case Equation (1) and Equation (2) become,

$$\frac{d\phi_{\rm p}^{*(n)}}{dt} = f_{\rm p}^{*(n)} \tag{B16}$$

$$\frac{df_{\rm p}^{*(n)}}{dt} = -\gamma^{(n)} f_{\rm p}^{(n)} + \xi^{(n)}(t) \,. \tag{B17}$$

Similarly, Equation (6) becomes

$$\phi_{\rm m}^{*(n)} = \phi_{\rm m}^{*(n)} - \left[f_{\rm p}^{*(n)} + f_{\rm em}^{(n)}\right] g^{(n)}(t) + \varepsilon^{(n)}(t) \,. \tag{B18}$$

We identify $\boldsymbol{X}(t)$ with a vector of length N composed of the intrinsic scaled pulsar frequency and phase, i.e.

$$\boldsymbol{X}(t) = \left(f_{\rm p}^{*(1)}(t), \phi_{\rm p}^{*(1)}(t), ..., f_{\rm p}^{*(N)}(t), \phi_{\rm p}^{*(N)}(t)\right) \,. \tag{B19}$$

Analogously, we package the measured, scaled pulsar phases as

$$\boldsymbol{Y}(t) = \left(\phi_{\rm m}^{*(1)}(t), ..., \phi_{\rm m}^{*(N)}(t)\right) \,. \tag{B20}$$

Going forward we will derive the relevant Kalman matrices for $N = 1$. Accordingly we drop the (n)-superscript notation. The extension to a general number of N pulsars then follows trivially through the construction of block diagonal matrices.

The states evolve according to the continuous stochastic differential equation (c.f. Equations (B16) and (B17))

$$d\boldsymbol{X} = \boldsymbol{A}\boldsymbol{X}\,dt + \boldsymbol{\Sigma} d\boldsymbol{B}(t)\,, \tag{B21}$$

where $\boldsymbol{A}$ is a 2×2 matrix,

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & -\gamma^{(1)} \end{pmatrix} \tag{B22}$$

and the 2×2 matrix $\boldsymbol{\Sigma}$ governs the magnitude of the increments of Brownian motion (Wiener process) $d\boldsymbol{B}(t)$, with

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^{(1)} \end{pmatrix}\,. \tag{B23}$$

The general solution to Equation (B21) is (Gardiner 2009)

$$\boldsymbol{X}(t) = e^{\boldsymbol{A}t}\boldsymbol{X}(0) + \int_0^t e^{\boldsymbol{A}(t-t')}\boldsymbol{\Sigma} d\boldsymbol{B}(t')\,. \tag{B24}$$

From Equation (B24) we construct the discrete, recursive solution for $\boldsymbol{X}(t_k) = \boldsymbol{X}_k$ in the form of Equation (B12), with

$$\boldsymbol{F}_k = \begin{pmatrix} 1 & 1 - e^{-\gamma^{(1)}\Delta t}/\gamma^{(1)} \\ 0 & e^{-\gamma^{(1)}\Delta t} \end{pmatrix}\,, \tag{B25}$$

$$\boldsymbol{w}_k = \int_{t_k}^{t_{k+1}} e^{\boldsymbol{A}(t_{k+1}-t')}\boldsymbol{\Sigma} d\boldsymbol{B}(t')\,, \tag{B26}$$

and $\Delta t = t_{k+1} - t_k$. From Equation (B26) the process noise covariance matrix is

$$\boldsymbol{Q}_k \boldsymbol{\delta}_{kj} = \langle \boldsymbol{w}_k \boldsymbol{w}_j^\intercal \rangle = \begin{pmatrix} Q_\mathrm{a} & Q_\mathrm{b} \\ Q_\mathrm{b} & Q_\mathrm{c} \end{pmatrix}\,, \tag{B27}$$

with

$$Q_\mathrm{a} = \frac{\sigma^2}{\gamma^3}\left(\gamma\Delta t - \kappa - \frac{\kappa^2}{2}\right) \tag{B28}$$

$$Q_\mathrm{b} = \frac{\sigma^2}{2\gamma^2}\kappa^2 \tag{B29}$$

$$Q_\mathrm{c} = \frac{\sigma^2}{2\gamma^2}\left(1 - e^{-2\gamma\Delta t}\right) \tag{B30}$$

and,

$$\kappa = 1 - e^{-\gamma\Delta t} \tag{B31}$$

Note lack of k dependence here on f and q. The measurement matrix $\boldsymbol{H}_k$ can be read straightforwardly from Equation (B18) as,

$$\boldsymbol{H}_k = \begin{pmatrix} 1 & -g(t_k) \end{pmatrix} \tag{B32}$$

with control term

$$\boldsymbol{B} = -f_\mathrm{em}(t_k)g(t_k) \tag{B33}$$

The measurement covariance satisfies $\boldsymbol{R}_k = E\left[\boldsymbol{v}\boldsymbol{v}^\intercal\right] = \sigma_\mathrm{m}^2$ for all k.

B3 Optimised representation

The explicit form of the discretised Kalman matrices given in Appendix B2 is straightforward to implement computationally. However, as N grows large there is additional computational burden when performing matrix algebra (especially matrix inversion) outlined in Equations (B6)–(B10). For the purposes of nested sampling, we require our likelihood evaluations to be as fast as possible. Moreover, because our matrices are block diagonal matrices constructed from the $N = 1$ components their is a high degree of sparsity.

To circumvent this computational demand for large N, we take advantage of the recognition that XYZ.

This paper has been typeset from a TEX/LATEX file prepared by the author.