Notes on Kalman filter for PTA analysis

T. Kimpson

February 17, 2025

This note defines the Kalman "machinery" for the state-space formulation of PTA analysis in terms of timing residuals. It is heavily based on work by P. Meyers and the MINNOW package, with extensions to handle multiple pulsars and the influence of the stochastic GW background, with contributions from A. Vargas. This work builds on our previous papers, continuing from their state-space formulation in the frequency domain.

1. Requirements

For the purposes of running state-space algorithms, we define the following:

- $\boldsymbol{X}$: the state vector, dimension n_X
- $\boldsymbol{Y}$: the observation vector, dimension n_Y
- $\boldsymbol{F}$: the state transition matrix, dimension $n_X \times n_X$
- $\boldsymbol{H}$: the measurement matrix, dimension $n_Y \times n_X$
- $\boldsymbol{Q}$: the process noise covariance matrix, dimension $n_X \times n_X$
- $\boldsymbol{R}$: the measurement noise covariance matrix, dimension $n_Y \times n_Y$

The goal of this note is to define the above matrices.

2. Derivation of the Measurement Equation

2.1. Introduction

We seek to apply the Kalman filter method to real PTA data, rather than a measured frequency time series $f_{\mathrm{m}}^{(n)}(t)$, as done in previous work.

By "real data," we refer to a .TIM file (which contains the TOAs) and a .PAR file (which provides constrained *a priori* estimates of the pulsar parameters, such as sky position, spin frequency, etc.).

For the purposes of this note, we assume that the .TIM and .PAR files have been processed through a standard pulsar timing library (e.g., TEMPO or PINT) to produce timing residuals δt. These timing residuals define our measurement vector, $\boldsymbol{Y}$.

Pulsar timing libraries employ a deterministic, parameterized model that utilizes estimates of the timing ephemeris parameters, $\hat{\boldsymbol{\theta}}$, to predict the pulse TOA, $t_{\rm det}(\hat{\boldsymbol{\theta}})$. This deterministic model accounts for standard effects such as Shapiro delay, proper motions, and binary interactions.[1] The timing residuals are then defined as the difference between the actual TOAs and the predicted TOAs:

$$\delta t = t_{\rm TOA} - t_{\rm det}(\hat{\boldsymbol{\theta}}). \tag{1}$$

2.2. Definition of phases, $\phi(t)$

In previous works, we have a hidden state $\boldsymbol{X}$ which we identify with the intrinsic pulsar frequency $f_p(t)$ and a measurement $\boldsymbol{Y}$ which we identify with a measured frequency $f_m(t)$. The two are related via a measurement equation

$$f_{\rm m}(t) = f_{\rm p}(t)\left[1 - a(t)\right], \tag{2}$$

where $a(t)$ quantifies the influence of the GW (for simplicity, we omit here the superscript-(n) notation.)

We want to derive an equivalent measurement equation that relates the new measurement $\boldsymbol{Y} = \delta t$ with some hidden states $\boldsymbol{X}$.

To start, separate the intrinsic pulsar frequency into deterministic and stochastic parts $f_{\rm p}(t) = \bar{f}(t) + \delta f(t)$ and define the following phase variables:

$$\text{Measured phase:} \quad \phi_{\rm m}(t) = \int_0^t f_{\rm m}(t')dt', \tag{3}$$

$$\text{Intrinsic phase:} \quad \phi_{\rm p}(t;\boldsymbol{\theta}) = \int_0^t \bar{f}(t';\boldsymbol{\theta})dt', \tag{4}$$

$$\text{Model phase:} \quad \phi_{\rm p}(t;\hat{\boldsymbol{\theta}}) = \int_0^t \bar{f}(t';\hat{\boldsymbol{\theta}})dt'. \tag{5}$$

Here, $\boldsymbol{\theta}$ represents the true timing-ephemeris parameters, while $\hat{\boldsymbol{\theta}}$ are the best-fit model estimates (i.e., the true parameters satisfy $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} + \delta\boldsymbol{\theta}$). In a standard PTA analysis it is assumed that any difference between the true determinisic solution and the estimated

[1] One might reasonably question whether removing all known contributions first, before analyzing residuals, could inadvertently filter out part of the signal of interest. This potential issue is well known in the pulsar community and will be addressed later.

solution will be small (see e.g. Taylor 2021, Section 7.1) and we can therefore linearize as

$$\phi_{\mathrm{p}}(t;\boldsymbol{\theta}) = \phi_{\mathrm{p}}(t;\hat{\boldsymbol{\theta}}) + \mathbf{M}_{\phi}\delta\boldsymbol{\theta} \tag{6}$$

where $\mathbf{M}_{\phi}$ is the design matrix of partial derivatives (of the phases) with respect to the parameters i.e. $= \partial_{\theta}\phi(t)$.

2.3. Definition of the timing residual, δt

The timing residual is given by Equation (1). We can also provide an equivalent definition in terms of phases / frequencies as follows.

According to the timing-ephemeris model, the n-th pulse arrives at time t if

$$\phi_{\mathrm{p}}(t;\hat{\boldsymbol{\theta}}) = n. \tag{7}$$

for some integer n. The actual pulse arrives at time $t + \delta t$, satisfying:

$$\phi_{\mathrm{m}}(t + \delta t) = n. \tag{8}$$

A linear expansion gives

$$\begin{aligned}\phi_{\mathrm{m}}(t + \delta t) &\approx \phi_{\mathrm{m}}(t) + \frac{d}{dt}\phi_{\mathrm{m}}(t)\delta t \\ &= \phi_{\mathrm{m}}(t) + f_{\mathrm{m}}(t)\delta t.\end{aligned} \tag{9}$$

Rearranging, we express the timing residual in terms of phases:

$$\delta t(t) = \frac{\phi_{\mathrm{p}}(t;\hat{\boldsymbol{\theta}}) - \phi_{\mathrm{m}}(t)}{f_{\mathrm{m}}(t)}. \tag{10}$$

2.4. Expanding $\phi_m(t)$

From Equation (2) and Equation (3), we can express the measured phase as

$$\phi_m(t) = \int_0^t f_m(\tau)\, d\tau \tag{11}$$

$$= \int_0^t [1 - a(\tau)] \left[\bar{f}(\tau) + \delta f(\tau)\right] d\tau \tag{12}$$

$$= \underbrace{\int_0^t \bar{f}(\tau)\, d\tau}_{\text{timing solution}} + \underbrace{\int_0^t \delta f(\tau)\, d\tau}_{\text{red noise}} - \underbrace{\int_0^t a(\tau)\, \bar{f}(\tau)\, d\tau}_{\text{GW term}} - \underbrace{\int_0^t a(\tau)\, \delta f(\tau)\, d\tau}_{\text{small term}}. \tag{13}$$

We make the following observations

- The timing solution term is just the phase $\phi_{\mathrm{p}}(t;\boldsymbol{\theta})$, Equation (4)

- The red noise term integral is just a phase, a fluctuation from the determinisic timing model, which we will call $\delta\phi$.

- Because the spin-down derivative term is small, the determinstic part of the GW term can be approximated as constant, $\bar{f}(t) = f_0 + \dot{f}t \approx f_0$. We will define $r(t) = \int_0^t a(\tau)\, d\tau$, (see e.g. Equation 5 of Sesana & Vecchio 2010)

- The second order term is small and can be dropped.

We can therefore write Equation (13) more concisely as

$$\phi_m(t) = \phi_{\rm p}(t; \boldsymbol{\theta}) \ + \ \delta\phi(t) \ - \ f_0 r(t) \tag{14}$$

2.5. Putting it all together

Combining Equations (6), (10), and (14), and taking the denominator of Equation (10) to be constant ($= f_0$) we can write

$$\delta t = \frac{1}{f_0}\left[\phi_{\rm p}(t; \hat{\boldsymbol{\theta}}) - \phi_{\rm p}(t; \boldsymbol{\theta}) - \delta\phi(t) + f_0 r(t)\right] \tag{15}$$

$$= -\frac{1}{f_0} M_\phi \delta\boldsymbol{\theta} - \frac{\delta\phi}{f_0} + r(t) \tag{16}$$

In practice, TEMPO/PINT return a design matrix defined in terms of partial derivatives of the TOAs, rather than in terms of phases, so we can just write

$$\boxed{\delta t = \boldsymbol{M}_{\rm TOA} \delta\boldsymbol{\theta} - \frac{\delta\phi}{f_0} + r(t)} \tag{17}$$

Equation (17) is the measurement equation.

Summary of assumptions and approximations

- Linearisation to define the M-matrix in Equation (6)
- Linear expansion and neglecting higher-order terms in Equation (9).
- Approximating $\bar{f}(t) \approx f_0$ in the GW term of Equation (13).
- Approximating $f_{\rm m} \approx f_0$ in the denominator of Equation (10).
- Neglecting small second-order terms in Equation (13).

3. Single pulsar

To start, lets put the above into a state-space frame work with a single pulsar, $N = 1$.

3.1. State vector

The state vector is

$$\boldsymbol{X} = (\delta\phi, \delta f, r, a, \delta\epsilon_1, \delta\epsilon_2, \cdots, \delta\epsilon_{\mathrm{M}}) \tag{18}$$

The first two variables, $\delta\phi$ and δf, are deviations from the spin-down parameters in timing model fit. Since we know these deviations won't move us more than 1 turn away (because these are nice millisecond pulsars) these are reasonable state variables to use. So for example, δf is a fluctuation from the measured spin-down in the timing model.

The subsequent two variables, r and a, quantify the effect of the gravitational wave. The variable $a(t)$ is the familiar redshift quantity from PTA P3, and r is the induced timing residual, obtained by integrating the redshift (see e.g. Equation 5 of Sesana & vecchio 2010).

The final M parameters, $\delta\epsilon_i$ are parameters of the design matrix. These parameters will be used to incorporate the effects of the timing model.

The observation vector is

$$\boldsymbol{Y} = (\delta t) \tag{19}$$

where δt is the timing residual.

3.2. Dynamical Equations (continuous time)

The above state variables evolve according to the following dynamical equations[2]

$$\frac{d}{dt}\delta\phi^{(n)} = \delta f^{(n)}, \tag{20}$$

$$\frac{d}{dt}\delta f^{(n)} = -\gamma_p^{(n)}\delta f^{(n)} + \chi_p^{(n)}(t;\sigma_p^{(n)}), \tag{21}$$

$$\frac{d}{dt}r^{(n)} = a^{(n)}, \tag{22}$$

$$\frac{d}{dt}a^{(n)} = -\gamma_a a^{(n)} + \chi_a^{(n)}(t;\sigma_a^{(n)}), \tag{23}$$

$$\frac{d}{dt}\delta\epsilon_m^{(n)} = \chi_\epsilon^{(n)}(t;\sigma_\epsilon) \quad \forall m \in [1, M^{(n)}]. \tag{24}$$

where χ_a is the usual white noise stochastic process.

The states are related to the observables via a measurement equation

$$\delta t = \frac{\delta\phi}{f_0} + \mathbf{M}\delta\epsilon - r \tag{25}$$

[2]TK: these might not be the equations to use. For instance, do we still need a γ term? I just use these as palce holders for now.

where $\mathbf{M}$ is the *design matrix* and f_0 is the pulsar rotation frequency [3].

Note that the dynamics of the $\delta\epsilon_m^{(n)}$ terms here are a bit different to how it is done in MINNOW. In MINNOW these terms have zero process noise, and they just get initialised with some values (and covariance) which can then be estimated by the inference procedure. In our case I think we can marginalise over these timing model parameters by including them in the state and setting the variance σ_ϵ to be very large. This is effecively what pulsar people already do when using Gaussian processes in ENTERPRISE; see e.g. Section IV of van Haasteran & Vallisneri, 2014. In this way, we can avoid having to estimate an extra $\sum_{i=1}^{N} M^{(i)}$ parameters. I include the process noise term here by default; if we want to remove it in the future and do the full parameter estimation, we can just set $\sigma_\epsilon = 0$ everywhere that follows.

3.3. Discretisation

3.3.1. F-matrix

For the $(\delta\phi, \delta f)$ block, the continuous system is

$$\frac{d}{dt}\begin{pmatrix}\delta\phi \\ \delta f\end{pmatrix} = \begin{pmatrix}0 & 1 \\ 0 & -\gamma_p\end{pmatrix}\begin{pmatrix}\delta\phi \\ \delta f\end{pmatrix} + \begin{pmatrix}0 \\ 1\end{pmatrix}\chi_p(t), \tag{26}$$

and the exact discretisation (matrix exponential) gives

$$F_p = \begin{pmatrix}1 & \dfrac{1 - e^{-\gamma_p \Delta t}}{\gamma_p} \\ 0 & e^{-\gamma_p \Delta t}\end{pmatrix}. \tag{27}$$

Similarly, for the (r, a) block we have

$$F_a = \begin{pmatrix}1 & \dfrac{1 - e^{-\gamma_a \Delta t}}{\gamma_a} \\ 0 & e^{-\gamma_a \Delta t}\end{pmatrix}. \tag{28}$$

For the timing model parameters $\delta\epsilon_m$, the dynamics are a random walk:

$$\delta\epsilon_{m,k+1} = \delta\epsilon_{m,k} + \eta_{\epsilon,m,k}, \tag{29}$$

so that the deterministic evolution is simply the identity.

Thus, the overall discrete state transition matrix is given by the block diagonal matrix

$$\boldsymbol{F} = \begin{pmatrix}F_p & 0 & 0 \\ 0 & F_a & 0 \\ 0 & 0 & I_M\end{pmatrix}, \tag{30}$$

[3] I need to check the sign of the residual term. Is it definitely a minus?

3.3.2. Q-matrix

For the $(\delta\phi, \delta f)$ block with noise variance σ_p^2, the discretised noise covariance is

$$Q_p = \sigma_p^2 \begin{pmatrix} \dfrac{\Delta t}{\gamma_p^2} - \dfrac{2\left(1 - e^{-\gamma_p \Delta t}\right)}{\gamma_p^3} + \dfrac{1 - e^{-2\gamma_p \Delta t}}{2\gamma_p^3} & \dfrac{1 - e^{-\gamma_p \Delta t}}{\gamma_p^2} - \dfrac{1 - e^{-2\gamma_p \Delta t}}{2\gamma_p^2} \\ \dfrac{1 - e^{-\gamma_p \Delta t}}{\gamma_p^2} - \dfrac{1 - e^{-2\gamma_p \Delta t}}{2\gamma_p^2} & \dfrac{1 - e^{-2\gamma_p \Delta t}}{2\gamma_p} \end{pmatrix}. \tag{31}$$

For the (r, a) block with noise variance σ_a^2, we similarly have

$$Q_a = \sigma_a^2 \begin{pmatrix} \dfrac{\Delta t}{\gamma_a^2} - \dfrac{2\left(1 - e^{-\gamma_a \Delta t}\right)}{\gamma_a^3} + \dfrac{1 - e^{-2\gamma_a \Delta t}}{2\gamma_a^3} & \dfrac{1 - e^{-\gamma_a \Delta t}}{\gamma_a^2} - \dfrac{1 - e^{-2\gamma_a \Delta t}}{2\gamma_a^2} \\ \dfrac{1 - e^{-\gamma_a \Delta t}}{\gamma_a^2} - \dfrac{1 - e^{-2\gamma_a \Delta t}}{2\gamma_a^2} & \dfrac{1 - e^{-2\gamma_a \Delta t}}{2\gamma_a} \end{pmatrix}. \tag{32}$$

For each timing model parameter $\delta\epsilon_m$, modeled as a random walk,

$$\delta\epsilon_{m,k+1} = \delta\epsilon_{m,k} + \eta_{\epsilon,m,k}, \quad \eta_{\epsilon,m,k} \sim \mathcal{N}(0, \sigma_\epsilon^2\, \Delta t), \tag{33}$$

so that

$$Q_\epsilon = \sigma_\epsilon^2\, \Delta t\, I_M. \tag{34}$$

Thus, the overall process noise covariance is block diagonal:

$$\boldsymbol{Q} = \begin{pmatrix} Q_p & 0 & 0 \\ 0 & Q_a & 0 \\ 0 & 0 & \sigma_\epsilon^2\, \Delta t\, I_M \end{pmatrix}. \tag{35}$$

3.3.3. H-matrix

Recall that the measurement equation is

$$\delta t = \frac{\delta\phi}{f_0} + \boldsymbol{M}\, \boldsymbol{\delta\epsilon} - r. \tag{36}$$

Since the state vector is

$$\boldsymbol{X} = \begin{pmatrix} \delta\phi \\ \delta f \\ r \\ a \\ \delta\epsilon_1 \\ \vdots \\ \delta\epsilon_M \end{pmatrix}, \tag{37}$$

the measurement depends only on $\delta\phi$, r, and the $\delta\epsilon_m$. Therefore, the measurement matrix is

$$\boldsymbol{H} = \begin{pmatrix} \frac{1}{f_0} & 0 & -1 & 0 & M_1 & \cdots & M_M \end{pmatrix}, \tag{38}$$

so that

$$\delta t = \boldsymbol{H}\, \boldsymbol{X}. \tag{39}$$

3.3.4. R-matrix

Assuming the measurement noise is white with variance σ_t^2, and given that there is a single measurement per time step, we have

$$\boldsymbol{R} = \sigma_t^2. \tag{40}$$

4. Multiple pulsars

The formulation in Section 3 extends straightforwardly to multiple pulsars. There are two complications which must be handled with care.

1. The covariance between the $a^{(n)}$ terms for different pulsars. This is the Hellings-Downs effect.
2. The fact that in general all pulsars are observed at different times. This means that instead of having a nice observation vector $\mathbf{Y}$ of length N, in general we just have a single observation from a single pulsar at a given timestep.

Regarding (1), the ensemble statistics of $\chi_{\mathrm{a}}^{(n)}$ are

$$\langle \chi_{\mathrm{a}}^{(n)}(t) \rangle = 0 \ , \tag{41}$$

$$\langle \chi_{\mathrm{a}}^{(n)}(t)\, \chi_{\mathrm{a}}^{(n')}(t') \rangle = \left[\sigma_{\mathrm{a}}^{(n,n')} \right]^2 \delta(t-t') \ . \tag{42}$$

with

$$\left[\sigma_{\mathrm{a}}^{(n,n')} \right]^2 = \frac{\langle h^2 \rangle}{6} \gamma_{\mathrm{a}}\, \Gamma \left[\theta^{(n,n')} \right] \ , \tag{43}$$

where $\langle h^2 \rangle$ is the mean square GW strain from the M summed sources at the Earth's position, $\theta^{(n,n')}$ is the angle between the n-th and n'-th pulsars, and one has

$$\Gamma \left[\theta^{(n,n')} \right] = \frac{3}{2} x_{nn'} \ln x_{nn'} - \frac{x_{nn'}}{4} + \frac{1}{2} + \frac{1}{2} \delta_{nn'} \ , \tag{44}$$

with $x_{nn'} = \left[1 - \cos\theta^{(n,n')} \right] /2$.

4.1. Discretisation for multiple pulsars

For N pulsars, we define the stacked state vector

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}^{(1)} \\ \boldsymbol{X}^{(2)} \\ \vdots \\ \boldsymbol{X}^{(N)} \end{pmatrix}, \qquad \text{with} \qquad \boldsymbol{X}^{(n)} = \begin{pmatrix} \delta\phi^{(n)} \\ \delta f^{(n)} \\ r^{(n)} \\ a^{(n)} \\ \delta\epsilon_1^{(n)} \\ \vdots \\ \delta\epsilon_{M^{(n)}}^{(n)} \end{pmatrix} . \tag{45}$$

The state evolution is as in the single–pulsar case, except that the gravitational-wave noise processes $\chi_a^{(n)}(t)$ are correlated among pulsars. Their ensemble statistics are

$$\langle \chi_a^{(n)}(t) \rangle = 0, \tag{46}$$

$$\langle \chi_a^{(n)}(t)\, \chi_a^{(n')}(t') \rangle = \left[\sigma_a^{(n,n')}\right]^2 \delta(t-t'), \tag{47}$$

with

$$\left[\sigma_a^{(n,n')}\right]^2 = \frac{\langle h^2 \rangle}{6}\, \gamma_a\, \Gamma\big[\theta^{(n,n')}\big], \tag{48}$$

and

$$\Gamma\big[\theta^{(n,n')}\big] = \frac{3}{2} x_{nn'} \ln x_{nn'} - \frac{x_{nn'}}{4} + \frac{1}{2} + \frac{1}{2}\delta_{nn'}, \qquad x_{nn'} = \frac{1-\cos\theta^{(n,n')}}{2}. \tag{49}$$

1. Discretised State Transition Matrix ($\boldsymbol{F}$)

For pulsar n, the $(\delta\phi, \delta f)$ block discretises as

$$F_p^{(n)} = \begin{pmatrix} 1 & \dfrac{1-e^{-\gamma_p^{(n)}\Delta t}}{\gamma_p^{(n)}} \\ 0 & e^{-\gamma_p^{(n)}\Delta t} \end{pmatrix}, \tag{50}$$

and the (r, a) block is

$$F_a = \begin{pmatrix} 1 & \dfrac{1-e^{-\gamma_a\Delta t}}{\gamma_a} \\ 0 & e^{-\gamma_a\Delta t} \end{pmatrix}, \tag{51}$$

(with the same gravitational–wave damping rate γ_a assumed for all pulsars). The timing model parameters evolve by the identity. Hence, the state transition matrix for pulsar n is

$$F^{(n)} = \begin{pmatrix} F_p^{(n)} & 0 & 0 \\ 0 & F_a & 0 \\ 0 & 0 & I_{M^{(n)}} \end{pmatrix}. \tag{52}$$

Stacking the pulsar states, the overall state transition matrix is block diagonal:

$$\boldsymbol{F} = \mathrm{diag}\Big\{F^{(1)}, F^{(2)}, \ldots, F^{(N)}\Big\}. \tag{53}$$

2. Discretised Process Noise Covariance ($\boldsymbol{Q}$)

For pulsar n, the $(\delta\phi, \delta f)$ block has the discretised covariance

$$Q_p^{(n)} = \sigma_p^{(n)2} \begin{pmatrix} \dfrac{\Delta t}{\gamma_p^{(n)2}} - \dfrac{2\left(1-e^{-\gamma_p^{(n)}\Delta t}\right)}{\gamma_p^{(n)3}} + \dfrac{1-e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)3}} & \dfrac{1-e^{-\gamma_p^{(n)}\Delta t}}{\gamma_p^{(n)2}} - \dfrac{1-e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)2}} \\ \dfrac{1-e^{-\gamma_p^{(n)}\Delta t}}{\gamma_p^{(n)2}} - \dfrac{1-e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)2}} & \dfrac{1-e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)}} \end{pmatrix}. \tag{54}$$

For the gravitational–wave (or a) block, note that the continuous noise processes are correlated among different pulsars. Thus, the cross–covariance between pulsars n and n' is discretised as

$$Q_a^{(n,n')} = \left[\sigma_a^{(n,n')}\right]^2 \begin{pmatrix} \dfrac{\Delta t}{\gamma_a^2} - \dfrac{2\left(1-e^{-\gamma_a \Delta t}\right)}{\gamma_a^3} + \dfrac{1-e^{-2\gamma_a \Delta t}}{2\gamma_a^3} & \dfrac{1-e^{-\gamma_a \Delta t}}{\gamma_a^2} - \dfrac{1-e^{-2\gamma_a \Delta t}}{2\gamma_a^2} \\ \dfrac{1-e^{-\gamma_a \Delta t}}{\gamma_a^2} - \dfrac{1-e^{-2\gamma_a \Delta t}}{2\gamma_a^2} & \dfrac{1-e^{-2\gamma_a \Delta t}}{2\gamma_a} \end{pmatrix} . \tag{55}$$

For the timing model parameters of pulsar n, modeled as a random walk,

$$Q_\epsilon^{(n)} = \sigma_\epsilon^2 \, \Delta t \, I_{M^{(n)}} . \tag{56}$$

The block coupling pulsars n and n' is then assembled as

$$\boldsymbol{Q}^{(n,n')} = \begin{pmatrix} \delta_{nn'} \, Q_p^{(n)} & 0 & 0 \\ 0 & Q_a^{(n,n')} & 0 \\ 0 & 0 & \delta_{nn'} \, Q_\epsilon^{(n)} \end{pmatrix} , \tag{57}$$

where $\delta_{nn'}$ is the Kronecker delta (i.e. the spin and timing noise are uncorrelated between different pulsars). Finally, the full process noise covariance is the block matrix

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{Q}^{(1,1)} & \boldsymbol{Q}^{(1,2)} & \cdots & \boldsymbol{Q}^{(1,N)} \\ \boldsymbol{Q}^{(2,1)} & \boldsymbol{Q}^{(2,2)} & \cdots & \boldsymbol{Q}^{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{Q}^{(N,1)} & \boldsymbol{Q}^{(N,2)} & \cdots & \boldsymbol{Q}^{(N,N)} \end{pmatrix} . \tag{58}$$

3. Measurement Matrix ($\boldsymbol{H}$)

For pulsar n, the measurement equation is

$$\delta t^{(n)} = \frac{\delta\phi^{(n)}}{f_0^{(n)}} + \boldsymbol{M}^{(n)} \, \boldsymbol{\delta\epsilon}^{(n)} - r^{(n)} , \tag{59}$$

so that the measurement matrix is

$$\boldsymbol{H}^{(n)} = \begin{pmatrix} \frac{1}{f_0^{(n)}} & 0 & -1 & 0 & M_1^{(n)} & \cdots & M_{M^{(n)}}^{(n)} \end{pmatrix} . \tag{60}$$

In a multi–pulsar analysis the overall measurement vector $\boldsymbol{Y}$ is built by stacking the individual measurements. In practice, since pulsars are generally observed at different times, only a subset of the pulsars contribute at any given time; the full measurement matrix is then formed by selecting the corresponding rows from the block–diagonal matrix

$$\boldsymbol{H}_{\text{full}} = \text{diag}\left\{ \boldsymbol{H}^{(1)}, \boldsymbol{H}^{(2)}, \ldots, \boldsymbol{H}^{(N)} \right\} . \tag{61}$$

4. Measurement Noise Covariance ($\boldsymbol{R}$)

Assuming that the measurement noise for pulsar n is white with variance $\sigma_t^{(n)2}$ and that different pulsars have independent measurement noise, then if at a given time step the set of observed pulsars is $\mathcal{O} \subset \{1, \ldots, N\}$, the measurement noise covariance is

$$\boldsymbol{R} = \mathrm{diag}\left\{\sigma_t^{(n)2}\right\}_{n \in \mathcal{O}} . \tag{62}$$

A. A worked example for the Q-matrix

It can help intution to "see" what the Q-matrix looks like explicitly. Consider the case where $N = 2$.

Recall that for each pulsar the state vector is partitioned into three sectors:

1. **Spin noise sector** (for $\delta\phi$ and δf): a 2×2 block.
2. **Gravitational–wave (redshift/residual) sector** (for r and a): a 2×2 block.
3. **Timing model (design matrix) sector** (for $\delta\epsilon$): a block of size $M^{(n)} \times M^{(n)}$.

For pulsar n ($n = 1, 2$), the individual blocks are defined as follows:

Spin Noise Sector

For pulsar n, the discretised spin noise covariance is

$$Q_p^{(n)} = \sigma_p^{(n)2} \begin{pmatrix} \dfrac{\Delta t}{\gamma_p^{(n)2}} - \dfrac{2\left(1 - e^{-\gamma_p^{(n)}\Delta t}\right)}{\gamma_p^{(n)3}} + \dfrac{1 - e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)3}} & \dfrac{1 - e^{-\gamma_p^{(n)}\Delta t}}{\gamma_p^{(n)2}} - \dfrac{1 - e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)2}} \\ \dfrac{1 - e^{-\gamma_p^{(n)}\Delta t}}{\gamma_p^{(n)2}} - \dfrac{1 - e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)2}} & \dfrac{1 - e^{-2\gamma_p^{(n)}\Delta t}}{2\gamma_p^{(n)}} \end{pmatrix}. \tag{63}$$

Gravitational–Wave Sector

For the gravitational–wave noise, the continuous noise processes $\chi_a^{(n)}(t)$ are correlated among pulsars. Thus, for pulsars n and n' the discretised covariance is given by:

$$Q_a^{(n,n')} = \left[\sigma_a^{(n,n')}\right]^2 \begin{pmatrix} \dfrac{\Delta t}{\gamma_a^2} - \dfrac{2\left(1 - e^{-\gamma_a\Delta t}\right)}{\gamma_a^3} + \dfrac{1 - e^{-2\gamma_a\Delta t}}{2\gamma_a^3} & \dfrac{1 - e^{-\gamma_a\Delta t}}{\gamma_a^2} - \dfrac{1 - e^{-2\gamma_a\Delta t}}{2\gamma_a^2} \\ \dfrac{1 - e^{-\gamma_a\Delta t}}{\gamma_a^2} - \dfrac{1 - e^{-2\gamma_a\Delta t}}{2\gamma_a^2} & \dfrac{1 - e^{-2\gamma_a\Delta t}}{2\gamma_a} \end{pmatrix}. \tag{64}$$

For the **autocovariance** (i.e. $n = n'$) we denote this as $Q_a^{(n,n)}$, and for the **cross-covariance** (i.e. $n \neq n'$) we have

$$Q_a^{(1,2)} = Q_a^{(2,1)}.$$

Timing Model Sector

For pulsar n, the timing model parameters are assumed to follow a random walk. Their covariance is given by

$$Q_\epsilon^{(n)} = \sigma_\epsilon^2 \, \Delta t \, I_{M^{(n)}}, \tag{65}$$

where $I_{M^{(n)}}$ is the identity matrix of dimension $M^{(n)}$.

Constructing the Full Q Matrix for Two Pulsars

For two pulsars ($N = 2$), we construct the overall process noise covariance matrix $\boldsymbol{Q}$ as a 2×2 block matrix:

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{Q}^{(1,1)} & \boldsymbol{Q}^{(1,2)} \\ \boldsymbol{Q}^{(2,1)} & \boldsymbol{Q}^{(2,2)} \end{pmatrix} . \tag{66}$$

Each block $\boldsymbol{Q}^{(n,n')}$ is itself block–structured into three sectors (spin noise, gravitational–wave noise, and timing model noise).

Diagonal Block for Pulsar 1 ($\boldsymbol{Q}^{(1,1)}$)

$$\boldsymbol{Q}^{(1,1)} = \begin{pmatrix} Q_p^{(1)} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(1)}} \\ \mathbf{0}_{2\times 2} & Q_a^{(1,1)} & \mathbf{0}_{2\times M^{(1)}} \\ \mathbf{0}_{M^{(1)}\times 2} & \mathbf{0}_{M^{(1)}\times 2} & Q_\epsilon^{(1)} \end{pmatrix} . \tag{67}$$

Here:

- The upper-left 2×2 block is $Q_p^{(1)}$ (spin noise for pulsar 1).
- The middle 2×2 block is $Q_a^{(1,1)}$ (gravitational–wave autocovariance for pulsar 1).
- The lower-right $M^{(1)} \times M^{(1)}$ block is $Q_\epsilon^{(1)}$ (timing model noise for pulsar 1).

Diagonal Block for Pulsar 2 ($\boldsymbol{Q}^{(2,2)}$)

$$\boldsymbol{Q}^{(2,2)} = \begin{pmatrix} Q_p^{(2)} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{2\times 2} & Q_a^{(2,2)} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{M^{(2)}\times 2} & \mathbf{0}_{M^{(2)}\times 2} & Q_\epsilon^{(2)} \end{pmatrix} . \tag{68}$$

Here:

- The upper-left 2×2 block is $Q_p^{(2)}$ (spin noise for pulsar 2).
- The middle 2×2 block is $Q_a^{(2,2)}$ (gravitational–wave autocovariance for pulsar 2).
- The lower-right $M^{(2)} \times M^{(2)}$ block is $Q_\epsilon^{(2)}$ (timing model noise for pulsar 2).

Off-Diagonal Block Between Pulsar 1 and Pulsar 2 ($\boldsymbol{Q}^{(1,2)}$)

$$\boldsymbol{Q}^{(1,2)} = \begin{pmatrix} \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{2\times 2} & Q_a^{(1,2)} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{M^{(1)}\times 2} & \mathbf{0}_{M^{(1)}\times 2} & \mathbf{0}_{M^{(1)}\times M^{(2)}} \end{pmatrix} . \tag{69}$$

In this off-diagonal block:

- The spin noise sectors (upper-left 2×2) are zero because spin noise is uncorrelated between pulsars.
- The timing model sectors are zero.
- Only the gravitational–wave sector (the middle 2×2 block) is nonzero and given by $Q_a^{(1,2)}$.

Off-Diagonal Block Between Pulsar 2 and Pulsar 1 ($\boldsymbol{Q}^{(2,1)}$)

By symmetry (assuming the cross–covariance is symmetric), we have

$$\boldsymbol{Q}^{(2,1)} = \begin{pmatrix} \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(1)}} \\ \mathbf{0}_{2\times 2} & Q_a^{(2,1)} & \mathbf{0}_{2\times M^{(1)}} \\ \mathbf{0}_{M^{(2)}\times 2} & \mathbf{0}_{M^{(2)}\times 2} & \mathbf{0}_{M^{(2)}\times M^{(1)}} \end{pmatrix}, \tag{70}$$

with $Q_a^{(2,1)} = Q_a^{(1,2)}$.

Full $\boldsymbol{Q}$ Matrix for $N = 2$

Assembling the blocks, the full $\boldsymbol{Q}$ matrix is given by

$$\boldsymbol{Q} = \left(\begin{array}{ccc|ccc} Q_p^{(1)} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(1)}} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{2\times 2} & Q_a^{(1,1)} & \mathbf{0}_{2\times M^{(1)}} & \mathbf{0}_{2\times 2} & Q_a^{(1,2)} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{M^{(1)}\times 2} & \mathbf{0}_{M^{(1)}\times 2} & Q_\epsilon^{(1)} & \mathbf{0}_{M^{(1)}\times 2} & \mathbf{0}_{M^{(1)}\times 2} & \mathbf{0}_{M^{(1)}\times M^{(2)}} \\ \hline \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(1)}} & Q_p^{(2)} & \mathbf{0}_{2\times 2} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{2\times 2} & Q_a^{(2,1)} & \mathbf{0}_{2\times M^{(1)}} & \mathbf{0}_{2\times 2} & Q_a^{(2,2)} & \mathbf{0}_{2\times M^{(2)}} \\ \mathbf{0}_{M^{(2)}\times 2} & \mathbf{0}_{M^{(2)}\times 2} & \mathbf{0}_{M^{(2)}\times M^{(1)}} & \mathbf{0}_{M^{(2)}\times 2} & \mathbf{0}_{M^{(2)}\times 2} & Q_\epsilon^{(2)} \end{array}\right). \tag{71}$$

Component Summary:

- $\boldsymbol{Q}^{(1,1)}$ (Pulsar 1):
  - **Spin noise:** $Q_p^{(1)}$ (upper-left 2×2)
  - **Gravitational–wave noise:** $Q_a^{(1,1)}$ (middle 2×2)
  - **Timing model noise:** $Q_\epsilon^{(1)}$ (lower-right $M^{(1)} \times M^{(1)}$)
- $\boldsymbol{Q}^{(2,2)}$ (Pulsar 2):
  - **Spin noise:** $Q_p^{(2)}$
  - **Gravitational–wave noise:** $Q_a^{(2,2)}$
  - **Timing model noise:** $Q_\epsilon^{(2)}$

- $\boldsymbol{Q}^{(1,2)}$ (Between Pulsar 1 and Pulsar 2):
  - **Spin noise:** Zero (2×2 zero matrix)
  - **Gravitational–wave noise:** $Q_a^{(1,2)}$ (middle 2×2)
  - **Timing model noise:** Zero
- $\boldsymbol{Q}^{(2,1)}$ (Between Pulsar 2 and Pulsar 1):
  - **Spin noise:** Zero
  - **Gravitational–wave noise:** $Q_a^{(2,1)}$ (which equals $Q_a^{(1,2)}$)
  - **Timing model noise:** Zero

This explicit layout shows how the overall $\boldsymbol{Q}$ matrix incorporates:

- The individual (uncorrelated) spin noise and timing model noise (appearing in the diagonal blocks).
- The cross–correlations due to the gravitational–wave background (appearing in the off-diagonal blocks).