TOAs

stochastic

contains everything except timing model

$$t^{obs} = t^{det}(c_{true}) + n$$

Residuals

$$\delta t = t^{obs} - t^{det}(c_{est}) = t^{det}(c_{true}) - t^{det}(c_{est}) + n$$

let $c_{true} = c_{est} + \varepsilon$

$$\Rightarrow \delta t = t^{det}(c_{est} + \varepsilon) - t^{det}(c_{est}) + n$$

$$= t^{det}(c_{est}) + \left.\frac{\partial t^{det}(c_{est}+\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} \varepsilon - t^{det}(c_{est}) + n + \mathcal{O}(\varepsilon^2)$$

$$\approx \left.\frac{\partial t^{det}(c_{est}+\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} \varepsilon + n$$

$$= M\varepsilon + n$$

design matrix
only depends on derivatives of timing model

* in practice this procedure is iterated by using either a standard weighted least squares fit or a generalized least squares fit.

Likelihood

n follows Gaussian with unknown covariance matrix $C(\phi)$

$$p(n|\phi) = \frac{1}{\sqrt{\det(2\pi C)}} \exp\left(-\frac{1}{2} n^T C^{-1} n\right)$$

where $C(\phi) = \langle n n^T \rangle$

this is what is minimized in LS

$$\Rightarrow p(\delta t | \phi, \varepsilon) = \frac{1}{\sqrt{\det(2\pi C)}} \exp\left(-\frac{1}{2}(\delta t - M\varepsilon)^T C^{-1}(\delta t - M\varepsilon)\right)$$

let $g = M^T C^{-1} \delta t$ & $\Gamma = M^T C^{-1} M$

log-likelihood $\rightarrow$ $\log p(\delta t|\phi,\varepsilon) - p(\delta t|\phi, 0) = \varepsilon^T g - \frac{1}{2}\varepsilon^T \Gamma \varepsilon$
$= \log \Lambda$

maximize $\log \Lambda$ $\rightarrow$ $0 = \frac{\partial \log \Lambda}{\partial \varepsilon} = g - \Gamma \hat{\varepsilon}$

ML $\hat{\varepsilon} = \Gamma^{-1} g = (M^T C^{-1} M)^{-1} M^T C^{-1} \delta t$

$$\sigma_{\hat{\varepsilon}} = \sqrt{\mathrm{diag}(\Gamma^{-1})} = \sqrt{\mathrm{diag}\left((M^T C^{-1} M)^{-1}\right)}$$

so-called post-fit residuals

$$\delta t^{post} = \delta t - M\hat{\varepsilon} = \underbrace{\left(I - M(M^T C^{-1} M)^{-1} M^T C^{-1}\right)}_{R} \delta t = R\,\delta t$$

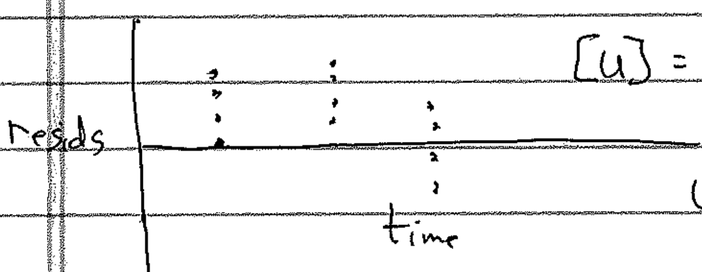
Rank-reduced noise model

residuals - $\delta t = \underbrace{M\varepsilon}_{\text{timing model}} + \underbrace{Fa}_{\text{red noise}} + \underbrace{Uj}_{\text{Jitter}} + \underbrace{n}_{\text{white noise}}$

$$Fa \rightarrow \sum_{n=1}^{N} \left[b_n \sin(2\pi n/T\,t) + c_n \cos(2\pi n/T\,t) \right]$$

$$F = \begin{bmatrix} \sin(2\pi t_0/T) & \cos(2\pi t_0/T) & \cdots & \sin(2\pi n t_0/T) & \cos(\ldots \\ \vdots & \vdots & & & \\ \sin(2\pi t_{ntoa}/T) & \cos(2\pi t_{ntoa}/T) & \cdots & & \end{bmatrix}$$

$$a = \begin{bmatrix} b_1 \\ c_1 \\ b_2 \\ c_2 \\ \vdots \\ \vdots \end{bmatrix}$$



$[U] = N_{toa} \times N_{epoch}$

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$ → maps TOAs to epochs

can describe jitter

White noise covariance

$N = \langle n n^T \rangle \rightarrow$ pdf of noise

$$p(n|\phi_{white}) = \frac{1}{\sqrt{\det(2\pi N)}} \exp\left(\frac{-1}{2} n^T N^{-1} n\right)$$

components of $N \rightarrow N_{ij,k} = \underbrace{E_k^2}_{\text{EFAC}} \underbrace{W_{ij}}_{\text{diag}(\sigma_i^2)} + \underbrace{Q_k^2}_{\text{EQUAD}} \delta_{ij}$

ij - label TOA number, k labels backend

full Likelihood function

$$p(\delta t|\varepsilon, a, j, \phi_{white}) = \frac{1}{\sqrt{\det(2\pi N)}} \exp\left(\frac{-1}{2}(\delta t - F_a - M_\varepsilon - U_j)^T N^{-1}(\delta t - F_a - M_\varepsilon - U_j)\right)$$

Priors

$$p(j|J_k) = \frac{\exp\left(\frac{-1}{2} j^T J^{-1} j\right)}{\sqrt{\det(2\pi J)}}, \quad p(a|\varphi_n) = \frac{\exp\left(\frac{-1}{2} a^T \varphi^{-1} a\right)}{\sqrt{\det(2\pi\varphi)}}$$

$$p(\varepsilon|X) = \frac{\exp\left(\frac{-1}{2}\varepsilon^T X^{-1}\varepsilon\right)}{\sqrt{\det(2\pi X)}} \qquad X \sim \begin{bmatrix} \infty & & & \\ & \infty & & \\ & & \infty & \\ & & & \ddots \end{bmatrix}$$

$$J_{ij,k} = J_k \delta_{ij}, \quad \varphi_{ij,n} = \begin{bmatrix} 10^{\rho_1} & & & & & & \\ & 10^{\rho_1} & & & & & \\ & & 10^{\rho_2} & & & & \\ & & & 10^{\rho_2} & & & \\ & & & & \ddots & & \\ & & & & & 10^{\rho_n} & \\ & & & & & & 10^{\rho_n} \end{bmatrix}$$

posterior

$$p(\varepsilon, a, j, \phi_{white}, J_K, \rho_n) \propto p(\delta t \mid \varepsilon, a, j, \phi_{white})\, p(j \mid J_K)\, p(a \mid \rho_n)$$

ε, j, a are nuisance parameters

let $T = \begin{bmatrix} M & F & U \end{bmatrix}$ $\quad b = \begin{bmatrix} \varepsilon \\ a \\ j \end{bmatrix}$

prior is now

$$p(b \mid \phi) = \frac{\exp\left(-\frac{1}{2} b^{T} B^{-1} b\right)}{\sqrt{\det(2\pi B)}}$$

$$B = \begin{bmatrix} \infty & & \\ & \varphi & \\ & & J \end{bmatrix}$$

likelihood $\quad p(\delta t \mid b, \phi) = \dfrac{\exp\left(-\frac{1}{2} (\delta t - Tb)^{T} N^{-1} (\delta t - Tb)\right)}{\sqrt{\det(2\pi N)}}$

~~posterior: $p(b, \phi \mid \delta t) \propto$~~

$$\times \frac{\exp\left(-\frac{1}{2} b^{T} B^{-1} b\right)}{\sqrt{\det(2\pi B)}}$$

just like before, $\quad d = T^{T} N^{-1} \delta t$

$$\Sigma = \left(T^{T} N^{-1} T + B^{-1}\right)$$

$$\hat{b} = \Sigma^{-1} d$$

$$\sigma_{\hat{b}} = \sqrt{\operatorname{diag}(\Sigma^{-1})}$$

can write likelihood as

$$p(\delta t | b, \phi) = \frac{\exp\left(-\frac{1}{2}\left[\delta t^T N^{-1} \delta t - d^T \Sigma^{-1} d\right]\right)}{\sqrt{\det(2\pi N)}\sqrt{\det(2\pi B)}}$$

$$\times \exp\left(-\frac{1}{2}(b-\hat{b})^T \Sigma^{-1} (b-\hat{b})\right)$$

Marginalize

$$p(\delta t | \phi) = \int db \; p(\delta t | b, \phi) = \sqrt{\frac{(2\pi)^{\dim b}}{\det \Sigma}}$$

$\therefore$ Marginalized likelihood

$$p(\delta t | \phi) = \frac{\exp\left(\frac{-1}{2}\left[\delta t^T N^{-1} \delta t - d^T \Sigma^{-1} d\right]\right)}{\sqrt{(2\pi)^{\dim N} \det(N) \det(B) \det(\Sigma)}}$$