

Noise in Cyclic Spectroscopy (Revised: 13th May 2012)

Mark A. Walker, Manly Astrophysics

5/7/2012

This document was produced in response to questions about the signal-to-noise ratio attained in cyclic spectroscopy of radio pulsars. These questions relate to both the cyclic spectra themselves and, more importantly perhaps, the quantities derived from them, namely the impulse response functions (“filter functions”) and intrinsic pulse profiles. The variance of parameters describing the impulse response function is shown to be independent of either the lag under consideration or the period of the pulsar, but for lags which are small compared to the pulse width some strong parameter degeneracies are identified. These degeneracies are associated with low-frequency pure-phase perturbations to the filter.

1 Construction of the cyclic spectrum

We begin our development by quoting the relationship between a signal, $x(t)$, a function of time, with Fourier Transform $X(\nu)$, and the cyclic spectrum of that signal, S_x . At modulation frequency α we have (P. Demorest 2011, MNRAS, 416, 2821, and references therein)

$$S_x(\alpha, \nu) \equiv \langle X(\nu + \alpha/2) X^*(\nu - \alpha/2) \rangle. \quad (1)$$

Thus if we apply a filter, $H(\nu)$, such that the filtered signal is $Z(\nu) = H(\nu) X(\nu)$ then the cyclic spectrum of the filtered signal is

$$S_z(\alpha, \nu) = H(\nu + \alpha/2) H^*(\nu - \alpha/2) S_x(\alpha, \nu). \quad (2)$$

In the case of a radio pulsar the signals X, Z are just electric fields, and the frequency ν is the radio frequency. Filtering of the signal occurs as a result of propagation, notably dispersion and scattering in the ionised ISM, and in the process of reception, e.g. the bandpass filter.

Here we confine attention to the case of small fractional radio bandwidths, for which we expect the intrinsic cyclic spectrum to be approximately independent of ν :

$$S_x(\alpha, \nu) \rightarrow S_x(\alpha). \quad (3)$$

The quantity $S_x(\alpha)$ is already familiar to astronomers from conventional analysis of radio pulsar signals: it is just the Fourier Transform of the pulse profile. But we emphasise that it is the transform of the intrinsic pulse profile, rather than the transform of the measured pulse profile — the difference being that the latter includes the influence of scattering and other contributions to the filter H .

In general both S_z and S_x are complex quantities, but in the particular case $\alpha = 0$ we obtain the zero-modulation-frequency components of the filtered and unfiltered signals, respectively. These are non-negative real numbers: just the power-spectra of the signals.

We note that the cyclic spectrum is unchanged by a uniform phase rotation of the filter, so the overall phase of H is indeterminate.

2 Noise bias and variance

The computed cyclic spectrum includes measurement noise which we can characterise in the following way. Suppose that the recorded voltage is $Z(\nu) + N(\nu)$, then we expect the measured cyclic spectrum to be

$$\langle D(\alpha, \nu) \rangle = S_z(\alpha, \nu) + \langle |N(\nu)|^2 \rangle \delta(\alpha), \quad (4)$$

where the delta-function appears because the measurement noise is stationary. Thus our measured cyclic spectrum is free of noise bias except at $\alpha = 0$.

The actual data which we record, $D(\alpha, \nu)$, will differ from $\langle D(\alpha, \nu) \rangle$ because of measurement noise and because the signal we are interested in is stochastic in nature. If there is no averaging (see discussion following equation 8) the variance of the measured cyclic spectrum is given by (J. Antoni 2007 Mechanical Systems and Signal Processing 21, 597)

$$\text{var}\{D(\alpha, \nu)\} = \langle D(0, \nu - \alpha/2) \rangle \langle D(0, \nu + \alpha/2) \rangle. \quad (5)$$

We note that at zero modulation frequency, equation 5 recovers the familiar result for statistically stationary signals that the variance of a single sample of the power is just the square of the mean power.

For observations of radio pulsars with current instrumentation, receiver-plus-sky noise is usually expected to be the dominant contribution to $D(0, \nu)$ and in this case we have

$$\text{var}\{D(\alpha, \nu)\} \simeq \langle |N(\nu - \alpha/2)|^2 \rangle \langle |N(\nu + \alpha/2)|^2 \rangle. \quad (6)$$

If the measurement noise is white, as is often the case in practice, then equation 6 yields a uniform variance,

$$\text{var}\{D\} = \langle |N|^2 \rangle^2 = \sigma^2, \quad (7)$$

over the entire cyclic spectrum. Equation 7 is the model noise distribution which we employ here. It can readily be generalised, using the foregoing equations, to include the effect of filter roll-off at the edges of the band, cases where the system noise is not white, the influence of Radio-Frequency Interference and noise due to the pulsar itself.

It is straightforward to estimate σ , because at zero modulation frequency the cyclic spectrum is just a power spectrum, for which the noise level is just

$$\sigma = \frac{\mathcal{S}_{sys}}{\sqrt{t\Delta\nu}}, \quad (8)$$

where \mathcal{S}_{sys} is the system equivalent flux-density, t is the integration time, and $\Delta\nu$ the radio-frequency width of a single channel. Here we consider only a single polarisation state, but clearly the results can be generalised to combinations of different polarisations.

Equation 8 clarifies what is meant by the ‘‘no averaging’’ requirement immediately preceding equation 5. For cyclic spectroscopy of pulsars we typically have $\Delta\nu = \Delta\alpha$, and we always have $\Delta\alpha = 1/P$, where P is the pulse period. Thus for $t = P$ we have a time-bandwidth product of unity and $\sigma = \mathcal{S}_{sys}$. Equation 5 is then appropriate to a single pulse, and if the cyclic spectrum is averaged over N_p pulses then the variance on the average cyclic spectrum decreases as $1/N_p$.

3 Modelling the cyclic spectrum

In estimating H and S_x we want the models which best fit the data, so we have an optimisation problem. We introduce the residual between data and model:

$$R(\alpha, \nu) \equiv D(\alpha, \nu) - S_z(\alpha, \nu), \quad (9)$$

and we seek to minimise these residuals in some sense.

Suppose our data, D , have N_ν radio-frequency channels, and N_α modulation-frequency bins. In this case we are modelling a filter with N_ν complex unknowns, and an intrinsic cyclic spectrum with $N_\alpha/2$ complex unknowns. (The pulse profile is a real quantity, so the spectrum at negative modulation frequencies is simply the complex conjugate of that at positive frequencies.) Thus there are $N_\nu + N_\alpha/2$ complex unknowns and $\sim N_\nu \times N_\alpha/2$ complex constraints provided by the data, so for $N_\nu, N_\alpha \gg 1$ the model is over-determined. In this situation we cannot make the residuals zero everywhere and we simply aim to make them small.

Here we follow the usual practice of minimising the sum-of-squares of the residuals

$$M \equiv \sum_{\alpha \neq 0, \nu} R^* R, \quad (10)$$

with respect to all of the model parameters. The summation over α in equation 10 (and in all the subsequent development) excludes zero because of the noise bias (see eq. 4), which is not independently known. We then have

$$\frac{\partial M}{\partial q} = -2 \sum_{\alpha \neq 0, \nu} R \frac{\partial S_z^*}{\partial q}, \quad (11)$$

where q represents any of the model parameters which define H or S_x , and minimisation of M implies

$$\frac{\partial M}{\partial q} = 0 \quad (12)$$

for every parameter q .

We compute the derivative for each parameter in turn. Each coefficient of H and S_x is complex and thus involves two distinct real parameters; we take these to be the real and imaginary parts. Thus for $S_{rm} \equiv \text{Re}\{S_x(\alpha_m)\}$, $S_{im} \equiv \text{Im}\{S_x(\alpha_m)\}$, we have

$$\frac{\partial M}{\partial S_{rm}} + i \frac{\partial M}{\partial S_{im}} = \nabla_S M \Big|_{\alpha=\alpha_m}, \quad (13)$$

where

$$\nabla_S M := -4 \sum_{\nu} R(\alpha, \nu) H^*(\nu + \alpha/2) H(\nu - \alpha/2). \quad (14)$$

And for $H_{rk} \equiv \text{Re}\{H(\nu_k)\}$, $H_{ik} \equiv \text{Im}\{H(\nu_k)\}$, we have

$$\frac{\partial M}{\partial H_{rk}} + i \frac{\partial M}{\partial H_{ik}} = \nabla_H M \Big|_{\nu=\nu_k}, \quad (15)$$

where

$$\nabla_H M := -4 \sum_{\alpha \neq 0} R(\alpha, \nu - \alpha/2) H(\nu - \alpha) S^*(\alpha). \quad (16)$$

Alternatively we may represent the filter $H(\nu)$ by its Fourier transform, $h(\tau)$. Here we make use of the conventions

$$h_j \equiv h(\tau_j) = \sum_k H_k \exp[2\pi i \tau_j \nu_k], \quad (17)$$

and

$$H_k \equiv H(\nu_k) = \frac{1}{N_\nu} \sum_j h_j \exp[-2\pi i \tau_j \nu_k]. \quad (18)$$

Following the same pattern as before, we introduce $h_{rj} \equiv \text{Re}\{h_j\}$, $h_{ij} \equiv \text{Im}\{h_j\}$. And as h and H are alternative bases spanning the same space we can write

$$\frac{\partial M}{\partial h_{rj}} = \sum_k \left\{ \frac{\partial H_{rk}}{\partial h_{rj}} \frac{\partial M}{\partial H_{rk}} + \frac{\partial H_{ik}}{\partial h_{rj}} \frac{\partial M}{\partial H_{ik}} \right\}; \quad (19)$$

and similarly for the derivative with respect to h_{ij} . In this way we find

$$\frac{\partial M}{\partial h_{rj}} + i \frac{\partial M}{\partial h_{ij}} = \nabla_h M \Big|_{\tau=\tau_j}, \quad (20)$$

where

$$\nabla_h M := \frac{1}{N_\nu} \sum_\nu \exp[2\pi i \tau \nu] \nabla_H M, \quad (21)$$

is, up to a factor of N_ν , just the Fourier transform (equation 17) of the frequency-space gradient of M (equation 16).

4 Behaviour in the vicinity of the best fit

If we have found the best fit model, in the least-squares sense, then $\nabla_S M = \nabla_H M = \nabla_h M = 0$. The figure of demerit, M/σ^2 , is distributed like χ^2 , so at the best-fit point we expect $M = M_o \simeq N_{\text{dof}} \sigma^2$, where the number of degrees of freedom is $N_{\text{dof}} \simeq N_\nu N_\alpha - 2N_\nu - N_\alpha - 1$. We approximate the Hessian matrix $\partial^2 M / \partial q_i \partial q_j$ as diagonal, so in the immediate neighbourhood of the best fit point the form of M is quadratic in each of the various parameters, q_j , separately:

$$M - M_o \simeq \sum_j \frac{1}{2} \frac{\partial^2 M}{\partial q_j^2} (q_j - q_{jo})^2. \quad (22)$$

The fit becomes significantly worse if we move away from the optimum point to any other point such that $M - M_o = \sigma^2$ (Y. Avni 1976, ApJ 210, 642), so in isolation the individual parameters have a standard error δq_j given by

$$(\delta q_j)^2 = 2\sigma^2 \left(\frac{\partial^2 M}{\partial q_j^2} \right)^{-1}. \quad (23)$$

We now evaluate the curvature terms needed to estimate the parameter uncertainties.

It is straightforward to show, by differentiating equations 14 and 16, that

$$\frac{\partial^2 M}{\partial S_{rm}^2} = \frac{\partial^2 M}{\partial S_{im}^2} = 4 \sum_\nu |H(\nu + \alpha_m/2)|^2 |H(\nu - \alpha_m/2)|^2, \quad (24)$$

and

$$\frac{\partial^2 M}{\partial H_{rk}^2} = \frac{\partial^2 M}{\partial H_{ik}^2} = 4 \sum_{\alpha \neq 0} |H(\nu_k - \alpha)|^2 |S_x(\alpha)|^2. \quad (25)$$

For the lag-space representation of the filter we obtain

$$\frac{\partial^2 M}{\partial h_{rj}^2} = \frac{4}{N_\nu^2} \operatorname{Re} \sum_{\alpha \neq 0, \nu} |S_x(\alpha)|^2 H(\nu - \alpha) [H^*(\nu - \alpha) + H(\nu) \exp(2\pi i \tau_j (2\nu - \alpha))], \quad (26)$$

and

$$\frac{\partial^2 M}{\partial h_{ij}^2} = \frac{4}{N_\nu^2} \operatorname{Re} \sum_{\alpha \neq 0, \nu} |S_x(\alpha)|^2 H(\nu - \alpha) [H^*(\nu - \alpha) - H(\nu) \exp(2\pi i \tau_j (2\nu - \alpha))], \quad (27)$$

(having neglected the contribution from a sum over the residuals, whose expectation is zero).

5 Estimation of parameter uncertainties

It is clear that the uncertainties in our parameter estimates depend on the filter coefficients and intrinsic pulse profile. But for our purposes here it suffices to determine rough estimates of the parameter uncertainties. To proceed we therefore consider the particular case $H(\nu) = 1$. For this circumstance we obtain

$$\frac{\partial^2 M}{\partial S_{rm}^2} = \frac{\partial^2 M}{\partial S_{im}^2} = 4 N_\nu, \quad (28)$$

and

$$\frac{\partial^2 M}{\partial H_{rk}^2} = \frac{\partial^2 M}{\partial H_{ik}^2} = 4 F^2, \quad (29)$$

where F is the total pulsed flux, with

$$F^2 := \sum_{\alpha \neq 0} |S_x(\alpha)|^2. \quad (30)$$

For the lag representation of the filter we find

$$\frac{\partial^2 M}{\partial h_{rj}^2} = \frac{\partial^2 M}{\partial h_{ij}^2} = \frac{4}{N_\nu} F^2, \quad (\tau_j \neq 0), \quad (31)$$

and for $\tau_j = 0$ the curvature with respect to the real part of the coefficient h_j is twice this value, whereas there is no curvature with respect to the imaginary part. This last point, which implies a formally infinite uncertainty, should not cause concern because the overall phase of the filter is completely arbitrary.

Using equation 23 we can immediately translate these curvatures into parameter uncertainties. The results are

$$\delta S_m = \frac{\sigma}{\sqrt{2N_\nu}}, \quad (32)$$

$$\delta H_k = \frac{\sigma}{F\sqrt{2}}, \quad (33)$$

and

$$\delta h_j = \frac{\sigma}{F} \sqrt{\frac{N_\nu}{2}}, \quad (\tau_j \neq 0). \quad (34)$$

In all these cases the coefficients are complex; the quoted uncertainty is the uncertainty in the real part of the coefficient, which is equal to the uncertainty in the imaginary part.

With the exception of one coefficient of h , we note that the noise (i.e. uncertainty in the parameter estimation) is uniform across each set of coefficients.

In practice the system noise, σ , is dependent on the total number of radio-frequency channels, N_ν , because we have a fixed total bandwidth, B , for the instrument. Thus $N_\nu \Delta\nu = B$, and equation 8 can be written

$$\sigma = \mathcal{S}_{sys} \sqrt{\frac{N_\nu}{tB}}. \quad (35)$$

The corresponding parameter uncertainties then become

$$\delta S_m = \frac{\mathcal{S}_{sys}}{\sqrt{2tB}}, \quad (36)$$

$$\delta H_k = \frac{\mathcal{S}_{sys}}{F} \sqrt{\frac{N_\nu}{2tB}}, \quad (37)$$

and

$$\delta h_j = \frac{\mathcal{S}_{sys}}{F} \frac{N_\nu}{\sqrt{2tB}}, \quad (\tau_j \neq 0). \quad (38)$$

In other words: these are the noise levels for measurements of S , H and h , respectively.

A further simplification is appropriate. For cyclic spectroscopy of a pulsar with period P , the pulsar's rotation frequency $\Omega = 1/P$ is necessarily equal to the spacing in modulation frequency, $\Delta\alpha$, and in turn this is the natural choice for channelisation, $\Delta\nu$. Thus the natural configuration is $PB = N_\nu$, and for this case equations 37 and 38 become

$$\delta H_k = \frac{\mathcal{S}_{sys}}{F} \sqrt{\frac{P}{2t}}, \quad (39)$$

and

$$\delta h_j = \frac{\mathcal{S}_{sys}P}{F} \sqrt{\frac{B}{2t}}, \quad (\tau_j \neq 0), \quad (40)$$

while equation 36 is independent of N_ν and thus unchanged.

6 Peak signal-to-noise ratios

To understand the meaning of these uncertainties, which gauge the noise level, it's helpful to compare them to the peak signal levels for the case under consideration. For the filter, which we have precisely specified, we have $H_k = 1$ and $h_j = N_\nu \delta_{j0}$. (We employ the Kronecker Delta: $\delta_{jk} = 1$ if $j = k$, and zero otherwise.) Remembering that the variance on the (real part of the) zero-lag coefficient is half that of the non-zero lags we see that

$$\text{Signal : Noise } \{H\} = \frac{F}{\mathcal{S}_{sys}} \sqrt{\frac{2t}{P}}, \quad (41)$$

and

$$\text{Signal : Noise } \{h\} = \frac{2F}{\mathcal{S}_{sys}} \sqrt{tB}. \quad (42)$$

We have not specified the harmonic strengths, S_m . But from equation 30 we see that we can take F as an order-of-magnitude estimate for the size of the strongest harmonic, so that

$$\text{Signal : Noise } \{S\} \sim \frac{F}{\mathcal{S}_{sys}} \sqrt{2tB}. \quad (43)$$

These results contain no surprises. The signal-to-noise on the harmonic strengths depends on the system equivalent-flux-density, and on the instrumental bandwidth and integration time in the usual fashion. There is no dependence on pulse-period. And the signal-to-noise on the lag-space representation of the filter depends on these quantities in exactly the same way. But the frequency-space representation of the filter splits up the signal into $N_\nu = PB$ channels, and the signal-to-noise on H is therefore lower by $\sqrt{N_\nu}$.

7 More general filters

The curvature of the demerit function with-respect-to the various model parameters depends on the structure in the filter functions, as manifest in equations 24–27, but we have so far considered only the simplest filter, $H(\nu) = 1$. We now consider how structure in the filter affects the noise level on various parameters.

It is, of course, possible to concoct bizarre examples of filters which imply correspondingly unusual noise properties. But we shall ignore such possibilities as our purpose here is to describe what one might normally expect to encounter in practice. To that end we will restrict our discussion to cases where $\langle |H(\nu)|^2 \rangle \sim \langle |H(\nu)|^4 \rangle \sim 1$, and we will characterise the impulse response function by a typical scattering time, τ_s , corresponding to a filter decorrelation bandwidth $\sim 1/\tau_s$.

Consider first the noise level for the pulse harmonic coefficients. For low harmonics the summation in equation 24 is approximately $N_\nu \langle |H(\nu)|^4 \rangle$. But at higher harmonics, where $|\alpha_m| \tau_s \sim 1$, there is some decorrelation between $|H(\nu - \alpha_m/2)|$ and $|H(\nu + \alpha_m/2)|$ and the sum declines. In the limit of complete decorrelation, $|\alpha_m| \tau_s \gg 1$, the summation yields $N_\nu \langle |H(\nu)|^2 \rangle^2$. Providing that both second- and fourth-order expectation values are of order unity, this is not a big effect. For example, in the random-phasor picture for the electric field the intensity statistics are exponential, so $\langle |H(\nu)|^2 \rangle = 1$ and $\langle |H(\nu)|^4 \rangle = 2$, yielding a noise level for high harmonics which is $\sqrt{2}$ larger than for low harmonics. In this picture, the noise level for high harmonics coincides with the value quoted in equation 36, for the case $H(\nu) = 1$.

Quite a different situation arises for the filter coefficients H_k . It is evident that the curvatures given in equation 25 may be much less than $4F^2$ in regions where the filter function is small, with correspondingly large errors on those coefficients. As with the noise on the pulse harmonics, there are two different limiting cases relating to the value of the typical scattering time. Most of the pulsed flux, F , is contributed by harmonics up to $|\alpha_m| \sim 1/w$, where w is the temporal width of the pulse. If $\tau_s \ll w$ then the filter function $H(\nu_k - \alpha)$ is almost constant over the range of α which contributes most to F , so the curvature in equation 25 becomes $4F^2 |H_k|^2$. Clearly this curvature could be very large (small) in comparison with the estimate given in equation 29, leading to correspondingly small (large) errors in the H_k estimates. In the opposite limit, where $\tau_s \gg w$, the filter coefficient $|H(\nu - \alpha)|$ changes rapidly with harmonic number and we obtain a curvature estimate $\sim 4F^2 \langle |H(\nu)|^2 \rangle \sim 4F^2$, comparable to that given in equation 29.

Finally we consider the effect of a structured filter on the errors associated with the lag-space filter coefficients, h_j . The curvatures of the merit function with respect to real and imaginary parts are (equations 26 and 27) made up of two terms. The first term is the same in both cases and we expect it to be $4F^2 \langle |H(\nu)|^2 \rangle / N_\nu \sim 4F^2 / N_\nu$. The second term differs in sign between the real and imaginary parts of the coefficients; it is the real part of a sum of complex numbers. In normal circumstances those complex numbers bear no particular phase relationship to each other, so the second term is typically small in comparison with the first. We therefore neglect it, and we conclude that in normal circumstances the curvatures given in equation 31 are appropriate to all lag-space filter coefficients.

8 Noise level in the delay-Doppler image

Components of the filter at non-zero delay, $\tau = \tau_n$, typically exhibit a non-zero Doppler-shift, ω_m . Therefore, with a sequence of N_t cyclic spectra we construct the delay-Doppler image, $d(\tau, \omega)$, which is the Fourier transform of the sequence of impulse response functions, $h(\tau, t_i)$. Following the same Fourier-transform conventions as in equations 17 and 18 yields the same relationship between the noise in the two domains exhibited by equations 33 and 34, but with the replacements $N_\nu \rightarrow N_t$ and $H \rightarrow d$. In other words the noise on d_{nm} must be smaller than that on h_n by a factor $\sqrt{N_t}$, so

$$\delta d = \frac{\mathcal{S}_{sys} P}{F} \sqrt{\frac{B}{2T}}, \quad (44)$$

where $T = t N_t$ is the total observing time. The Fourier transform over the sample time, t_i , sums the delayed component in phase at $\omega = \omega_m$, so that $d_{nm} = h_n$, and consequently the signal-to-noise ratio at d_{nm} is just $\sqrt{N_t}$ larger than the signal-to-noise on a single measurement of h_n .

9 A simple example

As a simple numerical example we can consider the case of the Green Bank Telescope, using the GUPPI backend with a bandwidth of 200 MHz, observing for 40 minutes at L-band. The system equivalent flux density for GBT at L-band is $\mathcal{S}_{sys} = 10$ Jy, and the noise floor on a delay-Doppler image is thus $10 \mu\text{Jy}$ after 40 minutes. If we are observing a source which has $F = 10$ mJy of pulsed flux, and the filter function contains very little structure ($H(\nu) \simeq 1$, so $h_j \simeq N_\nu \delta_{j0} = PB \delta_{j0}$), the dynamic range on the delay-Doppler image will amount to 60 dB of power.

In addition to the direct path to the source, we now imagine that there is a single image of magnification $\epsilon \sim 10^{-3}$ at delay τ_n . In this case there is an additional component $\sim \sqrt{|\epsilon|} N_\nu \delta_{jn}$ in the impulse response function. (The phase of this component has no influence on the signal-to-noise ratio of the measurement, so we ignore it here.) We thus have

$$d_{nm} = N_\nu \sqrt{|\epsilon|} = PB 10^{-3/2}, \quad (45)$$

and

$$\delta d = \frac{\mathcal{S}_{sys} P}{F} \sqrt{\frac{B}{2T}} = \frac{10 \mu\text{Jy}}{10 \text{ mJy}} PB = 10^{-3} PB, \quad (46)$$

and the signal-to-noise ratio on this faint image is $10^{3/2} \simeq 30$.

In the foregoing calculation we have assumed that the additional, faint image is perfectly localised in delay. In practice, interstellar scattering broadens all images and thus introduces a smearing of order τ_s in delay-space. Individual pixels in the delay-Doppler image have width $1/B$ in the delay dimension, implying a smearing of $\tau_s B$ pixels in that direction. To detect the image we need to combine the signal from all of these pixels, implying that in practice the signal-to-noise ratio may be smaller than the ideal case by a factor $\sqrt{\tau_s B}$.

10 The Hessian reloaded

In §4 we approximated the Hessian of M as diagonal and thus obtained a simple model for the variation of M around the minimum, M_o (equation 22). Here we examine deviations from that model, as revealed by the non-zero off-diagonal elements.

For those parameters which describe S_x it is easy to show that all off-diagonal elements are exactly zero, so our model involves no approximation at all in respect of those parameters.

The situation is not so simple for the filter-function coefficients, H_k . Starting from equations 15 and 16 we can take a further derivative with respect to H , leading to the diagonal elements given in equation 25. The mixed second-order derivative with respect to the parameters describing H_k is zero:

$$\frac{\partial^2 M}{\partial H_{ik} \partial H_{rk}} = 0. \quad (47)$$

But the mixed second-order derivatives with respect to the parameters describing H_j and H_k (i.e. with $j \neq k$) are all non-zero; they evaluate to

$$\frac{\partial^2 M}{\partial H_{rj} \partial H_{rk}} = 4 \operatorname{Re}\{|S_x(\nu_k - \nu_j)|^2 H(\nu_k) H(\nu_j) - S_x^*(\nu_k - \nu_j) R(\nu_k - \nu_j, \frac{\nu_k + \nu_j}{2})\}, \quad (48)$$

$$\frac{\partial^2 M}{\partial H_{rj} \partial H_{ik}} = 4 \operatorname{Im}\{|S_x(\nu_k - \nu_j)|^2 H(\nu_k) H(\nu_j) - S_x^*(\nu_k - \nu_j) R(\nu_k - \nu_j, \frac{\nu_k + \nu_j}{2})\}, \quad (49)$$

and

$$\frac{\partial^2 M}{\partial H_{ij} \partial H_{ik}} = -4 \operatorname{Re}\{|S_x(\nu_k - \nu_j)|^2 H(\nu_k) H(\nu_j) + S_x^*(\nu_k - \nu_j) R(\nu_k - \nu_j, \frac{\nu_k + \nu_j}{2})\}. \quad (50)$$

If the relevant harmonic (i.e. $\alpha = \nu_k - \nu_j$) is well above the noise, σ , then the first term in each of these expressions will typically dominate the second. We also note that the diagonal curvatures (equation 25) ought to be much larger than any of the off-diagonal terms, because the latter all involve a single harmonic whereas the former incorporates a sum over all the harmonics. Thus the Hessian is diagonal dominant in respect of the parameters H_k . Evidently the coupling between different coefficients of H declines significantly once the frequency separation between them becomes large compared to the inverse of the pulse width.

Alternatively we can describe the filter in terms of the lag-space coefficients h_j . We have already determined the first derivatives of M with respect to these parameters (equations 19, 20, 21), and the diagonal components of the Hessian (equation 26 and 27). Analogous to equation 19, the curvatures may be evaluated by again noting the expansion

$$\frac{\partial}{\partial h_{rm}} = \sum_k \left\{ \frac{\partial H_{rk}}{\partial h_{rm}} \frac{\partial}{\partial H_{rk}} + \frac{\partial H_{ik}}{\partial h_{rm}} \frac{\partial}{\partial H_{ik}} \right\}, \quad (51)$$

and applying this to equation 20, using equations 21 and 16. In this way we obtain

$$\frac{\partial^2 M}{\partial h_{rm} \partial h_{rj}} + i \frac{\partial^2 M}{\partial h_{rm} \partial h_{ij}} = A_{mj} + C_{mj}, \quad (52)$$

where

$$A_{mj} \equiv \frac{4}{N_\nu^2} \sum_{\alpha \neq 0, \nu} |S_x(\alpha)|^2 |H(\nu - \alpha)|^2 \exp [2\pi i(\tau_j - \tau_m)\nu], \quad (53)$$

and

$$C_{mj} \equiv \frac{4}{N_\nu^2} \sum_{\alpha \neq 0, \nu} |S_x(\alpha)|^2 H(\nu - \alpha/2) H(\nu + \alpha/2) \exp [2\pi i(\tau_j + \tau_m)\nu + \pi i(\tau_j - \tau_m)\alpha]. \quad (54)$$

Similarly, for h_{im} we obtain

$$\frac{\partial^2 M}{\partial h_{im} \partial h_{ij}} - i \frac{\partial^2 M}{\partial h_{im} \partial h_{rj}} = A_{mj} - C_{mj}. \quad (55)$$

Clearly the various individual curvatures can thus be evaluated simply by taking real and imaginary parts of equations 52 and 55. From equations 53 and 54 we note the symmetries $A_{jm} = A_{mj}^*$ and $C_{jm} = C_{mj}$.

In the particular case $m = j$ we reproduce the diagonal curvature terms given in equations 26 and 27. Additionally, however, we now obtain the mixed second-order derivative between real and imaginary parts of h_j :

$$\frac{\partial^2 M}{\partial h_{rj} \partial h_{ij}} = \text{Im}\{C_{jj}\}, \quad (56)$$

which in general is not zero.

It is possible to proceed directly to evaluation of the curvatures from equations 52-55, once a specific filter is assumed. It is, however, helpful to rewrite equations 53 and 54 with the help of the expansion of $H(\nu)$ in terms of the coefficients of $h(\tau)$. For this we again use equation 18, leading to the results

$$A_{mj} = \frac{4}{N_\nu^3} \sum_{\alpha \neq 0} |S_x(\alpha)|^2 \cos [2\pi\alpha(\tau_j - \tau_m)] \sum_n h_n^* h_{n+j-m}, \quad (57)$$

and

$$C_{mj} = \frac{4}{N_\nu^3} \sum_{\alpha \neq 0, n} |S_x(\alpha)|^2 \cos [2\pi\alpha(\tau_j - \tau_n)] h_n h_{m+j-n}. \quad (58)$$

Both matrices still require the evaluation of a double-sum, but in the case of A_{mj} the expression factorises into two one-dimensional sums. In equations 57 and 58 the notation is such that h_{n+j-m} means $h(\tau_n + \tau_j - \tau_m)$, for example.

We now evaluate the curvatures for a simple model filter.

10.1 Example filter: a delta-function in lag-space

The simplest non-trivial model filter is $h_n := N_\nu \delta_{no}$. In this case there is only one non-zero element contributing to any sum over n . Using equations 57 and 58 we obtain

$$A_{mj} = \frac{4}{N_\nu} F^2 \delta_{mj}, \quad (59)$$

and

$$C_{mj} = \frac{4}{N_\nu} \sum_{\alpha \neq 0} |S_x(\alpha)|^2 \cos [2\pi\alpha(\tau_j - \tau_o)] \delta_{o,m+j-o}, \quad (60)$$

where, following equation 30, we have written F for the total pulsed flux. Evidently for this filter A_{mj} is just the identity matrix multiplied by $4F^2/N_\nu$.

On the other hand, C_{mj} is non-zero only where $m + j = 2o$ (i.e. $\tau_m + \tau_j = 2\tau_o$), corresponding to the reverse diagonal passing through $\tau_m = \tau_j = \tau_o$. At this point we also have $C_{oo} = 4F^2/N_\nu$.

These results clearly reduce to the diagonal curvatures given earlier (equation 31 and text immediately following) for the particular case $\tau_o = 0$ considered there.

Now, however, we have some additional curvatures in the non-zero off-diagonal elements of C_{mj} . Some properties of these elements can be seen immediately. First, both A_{mj} and C_{mj} are purely real for the filter we are considering, so that there are no mixed curvatures between real and imaginary components of the filter coefficients: the only non-zero off-diagonal curvatures are

$$\frac{\partial^2 M}{\partial h_{rm} \partial h_{rj}} = C_{mj} = -\frac{\partial^2 M}{\partial h_{im} \partial h_{ij}}, \quad (m \neq j), \quad (61)$$

and these are manifestly symmetric. Secondly, if the pulse width is w then the harmonic sum is dominated by frequencies such that $|\alpha| \lesssim 1/2\pi w$, hence for $|\tau_j - \tau_o| \ll w$ the harmonic sum evaluates to approximately F^2 , yielding values $\simeq 4F^2/N_\nu$ for the non-zero coefficients of C_{mj} in the immediate vicinity of C_{oo} . And as we move further away from the leading diagonal, the non-zero coefficients of C_{mj} decline to small values on lag separations $|\tau_j - \tau_o| \sim w$.

We therefore find that for each parameter describing the coefficients h_j ($j \neq o$) there is, in addition to the diagonal curvature, exactly one mixed curvature. The mixing occurs between coefficients which are symmetrically placed with respect to τ_o . And for lags $|\tau_j - \tau_o| \ll w$ the magnitude of the mixed curvatures is almost equal to that of the diagonal terms.

10.2 Interpretation of the off-diagonal elements

What do the off-diagonal elements imply for the uncertainties in the individual parameters? To explore that question we continue to use the lag-space delta-function as our model filter. For parameters which describe h far from the location of the delta-function, i.e. at $|\tau_j - \tau_o| \gtrsim w$, the hypersurface M can be approximated using the diagonal Hessian we introduced previously. But for $|\tau_j - \tau_o| \ll w$ the mixed curvature terms are comparable in size to the diagonal terms and there are major differences between the true structure of M and the diagonal-Hessian approximation (equation 22).

The effect of the mixed term is most easily seen in the case where it is equal in magnitude to the diagonal terms (i.e. the limit $|\tau_j - \tau_o| \rightarrow 0$). In this case the variation in M around the minimum is given by

$$M(x, y) - M_o \simeq \frac{1}{2} \frac{\partial^2 M}{\partial x^2} (x^2 + y^2 \pm 2xy), \quad (62)$$

where we have focused on the variation of M with respect to a single pair of coupled parameters, x and y , and the sign of the final term is the sign of $\partial^2 M / \partial x \partial y$. It is trivial to reformulate the quadratic expression as a perfect square, $(x \pm y)^2$. That tells us that if we rotate from the basis (x, y) to the basis $(x + y, x - y) / \sqrt{2}$ then there is precisely zero curvature of M with respect to one of these new variables (whereas the curvature with respect to the other new variable is twice that in either x or y). In other words: in the fitting procedure there is a complete pairwise degeneracy in the parameter combination $(x + y) / \sqrt{2}$ or $(x - y) / \sqrt{2}$, depending on the sign of the mixed curvature.

Equation 61 and following text tells us that the sign of the mixed curvature is +ve for the real components and -ve for the imaginary components. Therefore it is the difference between h_{rj} and h_{rm} which is unconstrained, whereas it is the sum of h_{ij} and h_{im} which is unconstrained (for $|\tau_j - \tau_o| = |\tau_o - \tau_m| \ll w$). These degeneracies can be expressed more compactly in terms of the complex coefficients: $h_j + h_m^*$ is well constrained by the fitting, but $h_j - h_m^*$ is unconstrained. We note that this degeneracy is closely analogous to the confusion between an image and its conjugate in the case of in-line (Gabor) holography.

The presence of the degeneracies means that we can take our best-fit lag-space solution and add to it arbitrary real multiples of the functions $(\delta_{nj} - \delta_{nm})$ and $i(\delta_{nj} + \delta_{nm})$ without making the fit significantly worse. Together these two degeneracies give us the freedom to add to $H(\nu)$ a function

$$\exp[-2\pi i \tau_o \nu] i \{c_j \cos[2\pi(\tau_j - \tau_o)\nu] + s_j \sin[2\pi(\tau_j - \tau_o)\nu]\}, \quad (63)$$

where c_j and s_j are real numbers. These coefficients must satisfy $|c_j|, |s_j| \ll 1$ in order for our expansion of $M - M_o$ to be legitimately truncated at second-order. The freedom in $H(\nu)$ described by equation 63 holds for every strong pairwise degeneracy, so in general the poorly constrained addition to $H(\nu)$ is

$$\exp[-2\pi i \tau_o \nu] i \sum_n \{c_n \cos[2\pi \tau_n \nu] + s_n \sin[2\pi \tau_n \nu]\}, \quad (64)$$

where the sum is taken over the range $0 < \tau_n \lesssim w$. If we now recall that our model filter is $H(\nu) = \exp[-2\pi i \tau_o \nu]$ then we recognise that the parameter degeneracies identified above are equivalent to the application of a filter $\Psi(\nu)$, so that

$$H(\nu) \rightarrow H(\nu) \Psi(\nu), \quad (65)$$

where

$$\Psi(\nu) \equiv \exp \left[i \sum_n \{c_n \cos[2\pi \tau_n \nu] + s_n \sin[2\pi \tau_n \nu]\} \right], \quad (66)$$

with c_n and s_n being real numbers such that $|c_n|, |s_n| \ll 1$. In other words our model filter contains arbitrary low-frequency Fourier-mode phase structure, with one mode arising from each of the pairwise degeneracies in the parameter fitting.

As $\Psi(\nu)$ is effectively an additional filter acting on the signal, it follows that the degeneracies are equivalent to multiplication of the cyclic spectrum by

$$\Psi(\nu + \alpha/2) \Psi^*(\nu - \alpha/2) = \exp \left[i \sum_n 2 \sin(\pi \tau_n \alpha) \{s_n \cos[2\pi \tau_n \nu] - c_n \sin[2\pi \tau_n \nu]\} \right]. \quad (67)$$

We thus see that modifications to the cyclic spectrum associated with the filter Ψ are small because $|c_n|, |s_n|$ and $2\pi|\alpha|\tau_n$ are all small compared to unity. As formulated in equation 66 the fitting

degeneracies amount to pure-phase modifications of $H(\nu)$ which, in the small amplitude limit, approximate single Fourier modes.

As is evident from equations 65 and 66 $|H(\nu)|^2$ is unchanged by any of the degeneracies. Similarly, in the long pulse-period limit ($\alpha \rightarrow 0$) equation 67 approaches unity on the right-hand-side, so that the cyclic-spectrum becomes insensitive to the filter Ψ . In this circumstance, for the low-harmonics of the pulse (which contain most of the power), the cyclic spectrum is almost identical with the power spectrum. Thus our information on the phase of $H(\nu)$ gradually diminishes, leaving only information on the amplitudes $|H(\nu)|$. In this limit we are again faced with a phase-retrieval problem if we wish to determine $H(\nu)$.

The form of equation 65 shows that the degeneracies can also be thought of as convolving the impulse-response function, $h(\tau)$, by $\psi(\tau)$. Thus we expect that our best fit model h differs from the true impulse response not just because of additive noise, but also because of smearing by an unknown function.

11 Eliminating the degeneracies

Is it possible to break the degeneracies which are present in fitting the cyclic spectrum? There is at least one way in which it might be attempted: simply reduce the support of the filter-function by excluding negative lags from the solution. In other words demand that the filter be causal. That is one approach which has previously been proposed as a means of tackling the phase retrieval problem. In a general sense this approach is effective because it reduces the number of unknowns by a factor of 2, thus making the number of unknowns equal to the number of constraints (in the absence of any phase information whatsoever). In our particular context the effectiveness of the approach is clear because the degenerate parameters are all pairwise combinations $h(\tau) - h^*(-\tau)$. Thus if we force $h(\tau)$ to zero for all $\tau < 0$ then there can be no degeneracies.

In principle, then, demanding causality is a powerful way of suppressing the phase-errors (equations 65, 66) which would otherwise arise. However, there is a major problem in identifying which half of the lag-space should be removed from the support. After all, if we know the location of the lag corresponding to zero-delay then by definition we already know the pulse TOA. In particular there is a procedural issue here in that we need to know the TOA before we can start to solve for the filter. Iterative solution is one possible way around this difficulty, albeit unappealing.

Suppose that we make an attempt at restricting the support of h to one half of the lag range. To avoid missing genuine features of the filter at small lags we avoid setting the boundary of the support too close to where we think the zero delay actually is. In other words we err on the cautious side and set the boundary of the support at the point which we think corresponds to lag $-\Delta\tau$ (with $\Delta\tau > 0$). In this case we will only achieve suppression of the arbitrary Fourier-phase-modes at lags greater than $\Delta\tau$. It is therefore desirable to minimise $\Delta\tau$: set the support boundary as close as we can to zero delay.

Although I have not analysed the timing error associated with the unconstrained phase-modes, I expect that the error will be non-zero and I guess it will be proportional to $\sum_n (c_n^2 + s_n^2) \tau_n$. From the timing perspective it would be useful to understand this aspect of the problem and how it relates to the total TOA error associated with our imperfect characterisation of the filter. Obviously there is an interaction here with the setting of the support boundary and the suppression of the arbitrary phase modes, as just discussed.

One possible path to enforcing causality is the following. Set all the notionally-negative lag coefficients to zero, while allowing an additional fit parameter – a temporal offset – which ensures that the filter is fully captured within the notionally-positive half-space of lags. (By “notionally” positive/negative I just mean the sign convention as per the usual FFT output ordering.) The appeal of this approach is that it would allow causality to be imposed without requiring any iteration. Obviously, though, the critical thing would be the criterion by which the temporal offset is determined — that needs some thought.

There may be other ways of suppressing the troublesome parameter combinations. For example, the true impulse-response function is narrower than any solution which is modified by convolution with $\psi(\tau)$. So we might think of minimising the width of the model $h(\tau)$, as well as requiring a match to the spectral measurements. But that is only a single additional constraint whereas I suspect we’d need to supply many additional constraints in order to overcome the problem of degeneracies. Also, I think we should avoid imposing a bunch of arbitrary conditions on our filters, because if we do that we’ll never really understand the properties of our solutions in any rigorous way.

12 Partial degeneracies at small lags

We have already noted (§10.2) that in the case of a delta-function filter there is complete degeneracy if the mixed second-order derivatives are equal in magnitude to the diagonal components. Equations 60 and 61 show that the off-diagonal terms are in fact always smaller in absolute value than the diagonal terms, but for lags much less than w the difference is not very great.

Expanding the cosine in equation 60 we find

$$C_{mj} \simeq A_{jj} \delta_{o,m+j-o} \left(1 - \frac{2\pi^2(\tau_j - \tau_o)^2}{F^2} \sum_{\alpha \neq 0} |\alpha S_x(\alpha)|^2 \right), \quad (68)$$

which tells us that the curvature of M with respect to the corresponding degenerate parameters is smaller than the diagonal value (equation 31) by a factor

$$2 \left(\frac{\pi(\tau_j - \tau_o)F_1}{F} \right)^2, \quad (69)$$

for small values of $|\tau_j - \tau_o|$, where we have introduced F_1 , with

$$F_1^2 := \sum_{\alpha \neq 0} |\alpha S_x(\alpha)|^2. \quad (70)$$

We thus expect that the uncertainty in the value of the degenerate parameter combinations is larger than the uncertainty in non-degenerate combinations by a factor of approximately

$$\frac{1}{\sqrt{2}} \frac{F}{\pi |\tau_j - \tau_o| F_1}, \quad (71)$$

for $|\tau_j - \tau_o| \ll w$. This factor can be very large.

An alternative way of estimating the effect of the parameter degeneracies is to consider the typical amplitude of the phase-modes (equation 66) associated with those degeneracies. The pure-phase-modes translate into a multiplication of the cyclic spectrum by the factor given in equation

67. With reference to the simplest model filter, $H(\nu) = 1$, that is equivalent to adding to the cyclic spectrum an amount

$$2i S_x(\alpha) \sum_n \sin(\pi\tau_n\alpha) \{s_n \cos[2\pi\tau_n\nu] - c_n \sin[2\pi\tau_n\nu]\}, \quad (72)$$

providing $|c_n|, |s_n| \ll 1$. It is convenient to take the Fourier Transform of this quantity over ν , which tells us that the pure-phase modes add to the cyclic correlation function an amount

$$\Delta C_z(\alpha, \tau) = i N_\nu S_x(\alpha) \sum_n \sin(\pi\tau_n\alpha) \{d_n \delta(\tau + \tau_n) + d_n^* \delta(\tau - \tau_n)\}, \quad (73)$$

where $d_n \equiv s_n + i c_n$. Thus each of the pure-phase Fourier modes affects all pulse-harmonics for two lags in the cyclic correlation.

Multiplying both sides of equation 73 by $S_x^*(\alpha) \sin(\pi\tau_n\alpha)$ and summing over harmonics, we see that

$$d_n N_\nu i \sum_\alpha |S_x(\alpha)|^2 \sin^2(\pi\tau_n\alpha) = \sum_\alpha \Delta C_z(\alpha, -\tau_n) S_x^*(\alpha) \sin(\pi\tau_n\alpha), \quad (74)$$

and similarly for d_n^* with ΔC_z evaluated at $\tau = \tau_n$. We can now determine the thermal noise amplitude in d_n and d_n^* , because we know that for white noise in S_z , with variance σ^2 , the variance in C_z is uniform and equal to $N_\nu \sigma^2$. Hence each of the harmonics on the right-hand-side of equation 74 has variance $\sigma^2 N_\nu |S_x(\alpha)|^2 \sin^2(\pi\tau_n\alpha)$ and we thus find

$$\text{var}\{d_n\} = \sigma^2 \left(N_\nu \sum_\alpha |S_x(\alpha)|^2 \sin^2(\pi\tau_n\alpha) \right)^{-1}. \quad (75)$$

We see from this result that for small lags, $\tau_n \ll w$, the uncertainty in the pure-phase Fourier mode amplitude is $\propto 1/\tau_n$, as we found in equation 71. The variances of the other coefficients are related to that of d_n via

$$\text{var}\{d_n\} = \text{var}\{d_n^*\} = 4 \text{var}\{c_n\} = 4 \text{var}\{s_n\}. \quad (76)$$

Analogous to the pure-phase modes of equation 66 we can consider also a set of pure amplitude Fourier modes with coefficients c'_n, s'_n , such that

$$H(\nu) \rightarrow H(\nu) \rho(\nu), \quad (77)$$

where

$$\rho(\nu) \equiv \exp \left[\sum_n \{c'_n \cos[2\pi\tau_n\nu] + s'_n \sin[2\pi\tau_n\nu]\} \right]. \quad (78)$$

In contrast to the pure-phase modes, we expect these modes to exhibit only small noise fluctuations. To verify this we proceed as per equations 72-76. The filter $\rho(\nu)$ modifies the cyclic spectrum by a factor

$$\rho(\nu + \alpha/2) \rho^*(\nu - \alpha/2) = \exp \left[\sum_n 2 \cos(\pi\tau_n\alpha) \{c'_n \cos[2\pi\tau_n\nu] + s'_n \sin[2\pi\tau_n\nu]\} \right], \quad (79)$$

and the associated variance on d'_n is found to be

$$\text{var}\{d'_n\} = \sigma^2 \left(N_\nu \sum_\alpha |S_x(\alpha)|^2 \cos^2(\pi\tau_n\alpha) \right)^{-1}. \quad (80)$$

This confirms that for small lags, $\tau_n \ll w$, the noise on pure-amplitude Fourier modes is much smaller than the noise on pure-phase Fourier modes. The two noise levels are plotted in figure 1 for the case of PSR B1937+21 observed at 430 MHz, with 6230 channels over 4 MHz.

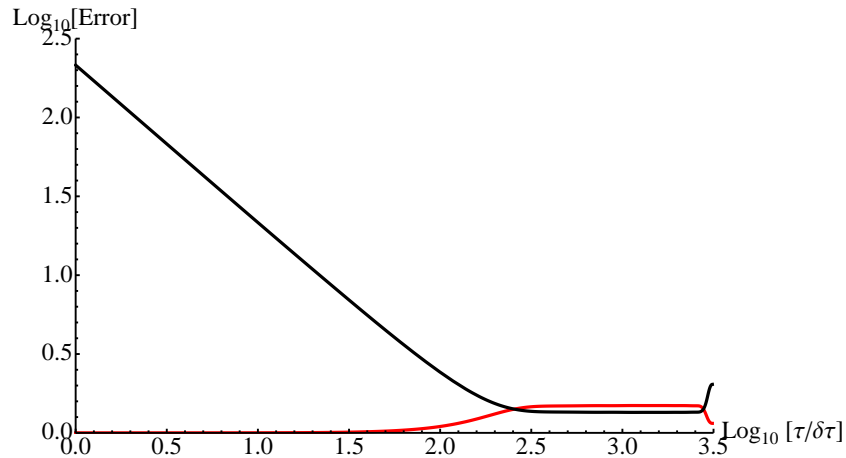


Figure 1: The standard deviations of pure-phase Fourier-mode coefficients (black) and pure-amplitude Fourier-mode coefficients (red), for PSR B1937+21 observed at 430 MHz, with 6230 channels over 4 MHz. These standard deviations, which arise entirely from white noise in the cyclic spectrum, are normalised to the pure-amplitude noise at zero-lag. The lag is plotted in units of the FFT lag increment. For the largest lags shown here we have $\tau \simeq P/2$, so that all the even-numbered pulse harmonics (which contain most of the pulsed power) satisfy $\sin(\pi\tau\alpha) \simeq 0$, $\cos(\pi\tau\alpha) \simeq 1$. Hence the noise on the pure-phase (pure-amplitude) modes increases (decreases) markedly at the largest lags.

13 Results for more general filters

Thus far we have used the simplest possible filter to analyse the parameter degeneracies associated with fitting the cyclic spectrum. However, their interpretation in terms of pure-phase modifications to the filter argues strongly that parameter degeneracies are not restricted to the particular model filter that we used.

Unfortunately it's difficult to work quantitatively with more complicated filter functions. For example, if the filter contains two lag-space delta-functions rather than one, then each lag-space coefficient h_j is coupled to five other coefficients. More generally we can imagine an impulse response which is significant only over a range $\sim \tau_s$, the typical scattering time, around lag τ_o . For simplicity we imagine that the non-zero impulse response coefficients all have a similar magnitude; it is then straightforward to show that the magnitude is $|h_n| \sim N_\nu/\sqrt{s}$, where $s := \tau_s/\delta\tau$ with $\delta\tau$ being the lag resolution. The diagonal coefficients of A do not vary much with the assumed properties of the filter and our estimate in §7 stands, but in addition we now have non-zero off-diagonal elements of A . For small $|\tau_j - \tau_m|$ the element A_{mj} is of magnitude $1/\sqrt{s}$ relative to the diagonal, and so A is diagonal dominant if the scattering time is sufficiently long. Moving away from the leading diagonal the coefficients decline further on the scale $|\tau_j - \tau_m| \sim \tau_s$ or w , whichever is smaller. As before, the matrix B is characterised by a reverse-diagonal passing through $m = j = o$, with a length of order the pulse width, w . But now the width of the non-zero region is of order τ_s or w , whichever is smaller.

Following the pattern of equations 72-75 it is possible to work out the noise level appropriate to any given pure-phase Fourier-mode, for any given filter, but it is unclear whether that figure is representative of the amplitudes of the degenerate modes. The problem here is that for non-delta-function filters we have no guarantee that the individual pure-phase Fourier-modes are the noisiest of all the possible filter perturbation modes. Indeed it seems very likely that they are not the noisiest.

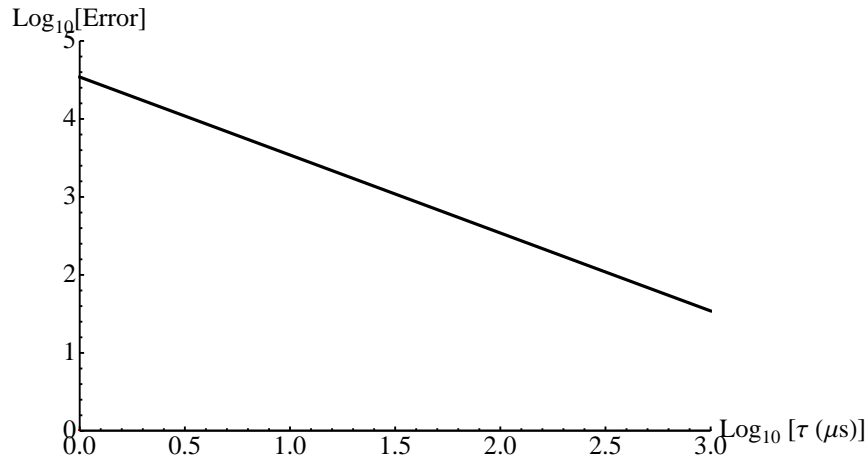


Figure 2: Estimate of the noise-level on pure-phase Fourier-modes for PSR B0834+06, for lags up to one millisecond. The noise is expressed in units of the noise on pure-amplitude Fourier-modes at zero lag.

Although it appears that the noisiest modes will always be pure-phase modes, there are an infinite number of possible basis functions with which we could expand the phase structure, and presumably the single-Fourier-mode emerges as the noisiest mode only in the case of a single-Fourier-mode filter.

For these reasons I expect that the results shown in figure 1 are broadly representative, even though they are computed for one particular filter, and that low-frequency phase noise of similar amplitude will typically be present for B1937+21 in the lag range shown. One consequence of this is that we should expect our current best-fit impulse-response functions for B1937+21 to be smeared by an unknown function of width $\sim w$. We should therefore consider imposing a causality requirement on all our solutions.

14 The case of B0834+06 versus B1937+21

The pulsar B0834+06 has a pulse period that is longer than B1937+21 by a factor of roughly 800. Although the harmonic structure of B0834+06's pulse is plainly different from that of B1937+21, the fractional FWHM of the two profiles is probably similar. Therefore, as a crude model for B0834+06 we can adopt the harmonic structure of B1937+21, but with all the harmonic frequencies scaled down by a factor of 800. The pure-phase Fourier-mode and pure-amplitude Fourier-mode noise-levels for B0834+06 in this model are shown in figure 2. The amplitude noise in this plot is essentially unity over the whole lag range shown. The phase-noise, however, is so large that over the entire range of this plot the cyclic spectrum does not really yield any significant phase information on $H(\nu)$.