

MAXIMUM LIKELIHOOD FITTING OF X-RAY POWER DENSITY SPECTRA: APPLICATION TO HIGH-FREQUENCY QUASI-PERIODIC OSCILLATIONS FROM THE NEUTRON STAR X-RAY BINARY 4U1608-522

DIDIER BARRET¹ AND SIMON VAUGHAN²

¹ Institut de Recherche en Astrophysique et Planétologie & Université de Toulouse (UPS), 31028, Toulouse Cedex 4, France;
didier.barret@irap.omp.eu

² X-Ray and Observational Astronomy Group, University of Leicester, Leicester, LE1 7RH, UK
 Received 2011 September 12; accepted 2011 November 30; published 2012 February 1

ABSTRACT

High-frequency quasi-periodic oscillations (QPOs) from weakly magnetized neutron stars display rapid frequency variability (second timescales) and high coherence with quality factors up to at least 200 at frequencies about 800–850 Hz. Their parameters have been estimated so far from standard $\min(\chi^2)$ fitting techniques, after combining a large number of power density spectra (PDS), to have the powers normally distributed (the so-called Gaussian regime). Before combining PDS, different methods to minimize the effects of the frequency drift to the estimates of the QPO parameters have been proposed, but none of them relied on fitting the individual PDS. Accounting for the statistical properties of PDS, we apply a maximum likelihood method to derive the QPO parameters in the non-Gaussian regime. The method presented is general, easy to implement, and can be applied to fitting individual PDS, several PDS simultaneously, or their average, and is obviously not specific to the analysis of kHz QPO data. It applies to the analysis of any PDS optimized in frequency resolution and for low-frequency variability or PDS containing features whose parameters vary on short timescales, as is the case for kHz QPOs. It is equivalent to the standard χ^2 minimization fitting when the number of PDS fitted is large. The accuracy, reliability, and superiority of the method is demonstrated with simulations of synthetic PDS, containing Lorentzian QPOs of known parameters. Accounting for the broadening of the QPO profile, due to the leakage of power inherent to windowed Fourier transforms, the maximum likelihood estimates of the QPO parameters are asymptotically unbiased and have negligible bias when the QPO is reasonably well detected. By contrast, we show that the standard $\min(\chi^2)$ fitting method gives biased parameters with larger uncertainties. The maximum likelihood fitting method is applied to a subset of archival *Rossi X-ray Timing Explorer* data of the neutron star X-ray binary 4U1608-522, for which we show that the lower kHz QPO parameters can be measured on timescales as short as 8 s. To demonstrate the potential use of the results of the maximum likelihood method, we show that in the observation analyzed the time evolution of the frequency is consistent with a random walk. We then show that the broadening of the QPO due to the frequency drift scales as \sqrt{T} , as expected from a random walk (T is the integration time of the PDS). This enables us to estimate the intrinsic quality factor of the QPO to be ~ 260 , whereas previous analysis indicated a maximum value around 200.

Key words: accretion, accretion disks – methods: data analysis – methods: numerical – stars: neutron – X-rays: binaries

Online-only material: color figures

1. INTRODUCTION

The standard method of weighted least-squares (i.e., minimum χ^2 , hereafter $\min(\chi^2)$) model fitting is only equivalent to the maximum likelihood estimation (MLE) of the model parameters when the data to be fitted are normally (Gaussian) distributed. As is well known, the distribution of M averaged periodogram ordinates (from a stationary, linear stochastic process) follows a χ^2_{2M} distribution with $2M$ degrees of freedom; as M increases, the χ^2_{2M} tends toward a normal distribution (Groth 1975; Papadakis & Lawrence 1993). In X-ray timing the most commonly used method to reach the so-called Gaussian regime is Bartlett’s method (Bartlett 1948): segment the original time series into M non-overlapping segments, compute a periodogram for each segment, and average over all segments to produce a spectral estimate to be fitted (van der Klis 1989). Typically, one uses $M > 50$ to produce approximately Gaussian distributed averages (Papadakis & Lawrence 1993). The drawbacks of this method are a loss of frequency resolution by a factor M (reducing the sensitivity for narrow feature detection) and a suppression of the lowest frequencies. There is also a major drawback in the case of the analysis of high-frequency quasi-periodic

oscillations (QPOs) because their frequency varies rapidly with time, typically on timescales of seconds (Barret et al. 2005a). Since a kHz QPO can be a relatively narrow feature (FWHM ~ 2 –4 Hz), sub-Hz frequency resolution of the Power Density Spectra (PDS, equivalent to a segment duration of a few seconds) is required to have the QPO profile properly sampled. In order to use $\min(\chi^2)$ fitting one must average a large number of segments, M , which then requires a total integration time exceeding hundreds, even thousands, of seconds, depending on the strength of the QPO. On such a timescale, the drift of the QPO frequency will smear out the QPO profile, leading to biases in the fitted QPO parameters, in particular the width. Methods to minimize the contribution of the frequency drift to the fitted QPO width have been proposed, using for instance the centroid (without fitting) of the excess power, as a best guess for the QPO frequency (Barret et al. 2005a, 2006). Although optimized, those methods suffer from the fact that statistical fluctuations of the noise can be stronger than the QPO signal, hence biasing the QPO frequency determination, especially when M is small.

In this paper, we apply the maximum likelihood method to fit spectral models of high-frequency QPOs in the non-Gaussian regime. To our knowledge, this is the first time such

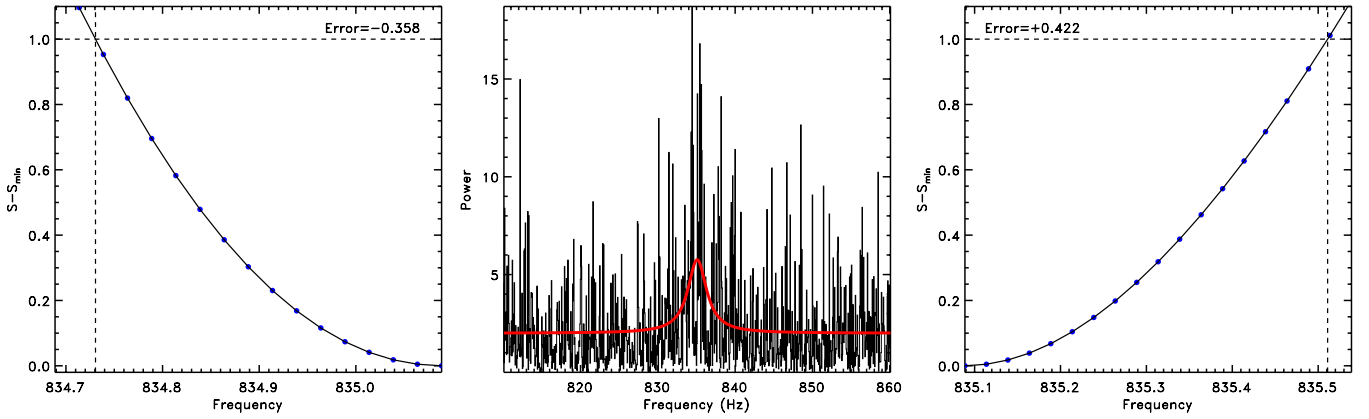


Figure 1. MLE fit of one single 16 s PDS (center panel), together with the ΔS curves for the estimate of the QPO frequency (left and right panels). Despite a very noisy PDS, the MLE picks up the QPO at the right frequency (835.08 Hz assumed in the simulation). The 1σ error on the QPO frequency is only about 0.4 Hz. (A color version of this figure is available in the online journal.)

a method is applied to the analysis of *Rossi X-ray Timing Explorer* (*RXTE*) Proportional Counter Array (PCA) data, although it was discussed early on by Stella et al. (1994) for the analysis of continuum PDS recorded with *EXOSAT*. Such a method has been used already in a variety of applications, such as the modeling of solar oscillation spectra (Anderson et al. 1990; Appourchaux 2011) or fitting broadband PDS from long *XMM-Newton* observations of variable Seyfert 1 galaxies (Vaughan 2010). In Section 2, we describe the maximum likelihood method. Then in Section 3, the reliability and robustness of the technique is demonstrated using simulations of synthetic PDS. Its superiority over $\min(\chi^2)$ fitting is illustrated in the case of $M = 4$, and a comparison between the MLE and the $\min(\chi^2)$ fitting is also presented for typical M values used in QPO analysis. In Section 4, the method is applied to a small subset of the archival *RXTE* data of the neutron star X-ray transient 4U1608-522, and the results of the MLE are used to infer the intrinsic width of the QPO.

2. DESCRIPTION OF THE MAXIMUM LIKELIHOOD METHOD

For describing the method, we follow the formalism of Vaughan (2005, 2010). The periodogram of any linear stochastic process time series of length N , denoted $I_j = I(f_j)$ at Fourier frequency $f_j = j/N\Delta T$ (with $j = 1, \dots, N/2$), is exponentially distributed about the true spectral density $S_j = S(f_j)$ (S_j is also the expectation value at f_j),

$$p(I_j|S_j) = \frac{1}{S_j} \exp(-I_j/S_j) \quad (1)$$

(see, e.g., Groth 1975; Priestley 1981; Leahy et al. 1983; van der Klis 1989; Percival & Walden 1993; Bloomfield 2000). For simplicity, we assume even N . As pointed out by Vaughan (2005, 2010), this is valid only for Fourier frequencies $j = 1, 2, \dots, N/2 - 1$, as the power at the Nyquist frequency ($j = N/2$) follows a χ_1^2 distribution with one degree of freedom (unlike powers at the other frequencies which follow the χ_2^2 distribution).

Assuming a model $S(\theta)$, with parameters $\theta \equiv \{\theta_1, \theta_2, \dots, \theta_L\}$, the joint probability density of observing $N - 1$ periodogram points I_j , given the model values \hat{S}_j ($j = 1, N - 1$, ignoring the

Nyquist frequency), is

$$\mathcal{L} = \prod_{j=1}^{N-1} p(I_j|S_j) = \prod_{j=1}^{N-1} \frac{1}{S_j} \exp(-I_j/S_j), \quad (2)$$

where \mathcal{L} is the likelihood, a function of θ . This is sometimes known as the *Whittle* likelihood. The χ_2^2 distribution and hence the Whittle likelihood are the asymptotic expressions but are usually extremely good approximations for the range of N typically examined in these studies (e.g., $N = 8192$). Maximizing the likelihood \mathcal{L} is equivalent to minimizing $S \equiv -2 \ln \mathcal{L}$. S is then:

$$S = -2 \ln \mathcal{L} = 2 \sum_{j=1}^{N-1} \left\{ \frac{I_j}{S_j} + \ln S_j \right\}. \quad (3)$$

The above formula can in fact be generalized to the case of fitting the average of M PDS, as discussed in Appendix A.

Finding the model parameters that minimize S yields the MLEs of the model parameter, $\hat{\theta}$ (Anderson et al. 1990). The minimization can be achieved with a standard numerical optimization algorithm, such as *POWELL* or *AMOEB*A (e.g., Press et al. 1992, chapter 10).

It is easy to show that if the spectral model is a constant, $S(f) = a$, then its MLE, \hat{a} , can be found by minimizing

$$S = 2 \left\{ (N - 1) \ln a + 1/a \sum_{j=1}^{N-1} I_j \right\}. \quad (4)$$

The solution is found where the derivative of S against a is null:

$$\hat{a} = \frac{1}{N - 1} \sum_{j=1}^{N-1} I_j. \quad (5)$$

The MLE of the constant is simply the sample mean.

Confidence intervals can be computed from the so-called *Fisher matrix* (F), the expectation value of the Hessian (e.g., Andrae 2010; Heavens 2009, and references therein). The error on $\hat{\theta}_i$, σ_i is then computed as

$$\sigma_i^2 = (F)_{ii}^{-1}, F_{ij} = \left\langle -\frac{\delta^2 \ln \mathcal{L}}{\delta \theta_i \delta \theta_j} \right\rangle. \quad (6)$$

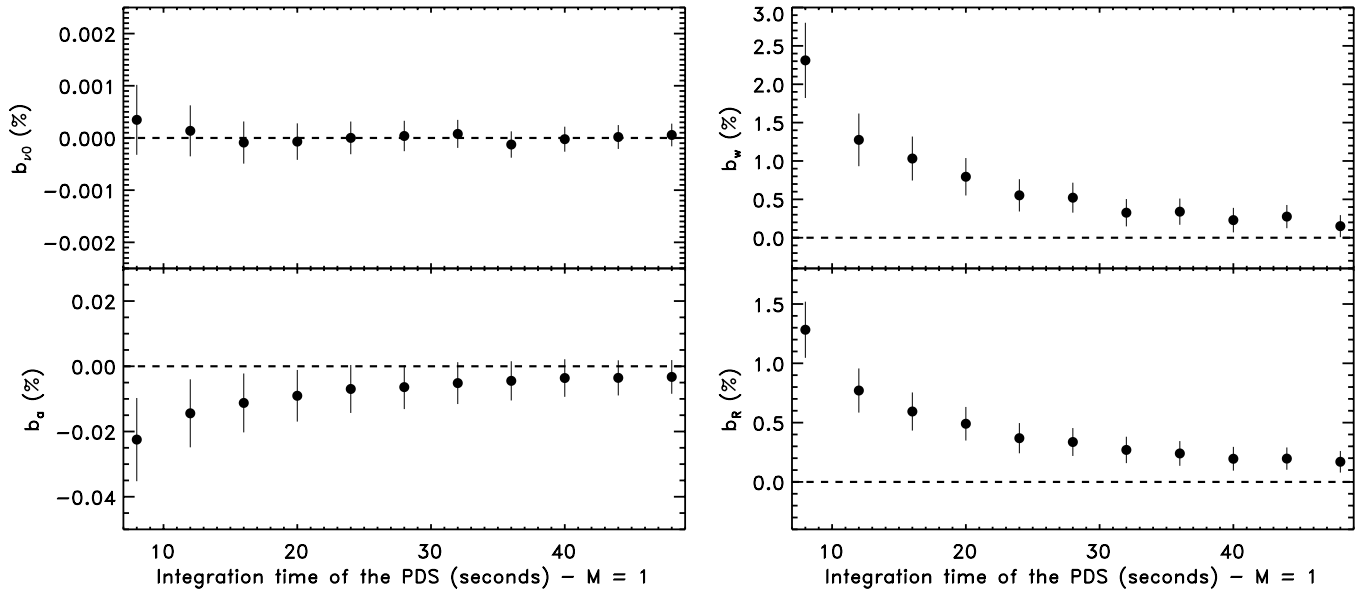


Figure 2. Measured biases on the four model parameters (top left, v_0 ; bottom left, a ; top right, w ; bottom right, R) as a function of the PDS integration time (T). Those biases are derived from the difference between the sample mean of the MLE estimates from single PDS ($M = 1$) and the best-fitted value of the average of the 16384 PDS generated for each integration time, and are expressed in percentage of the best-fitted averages. Error bars on the bias are derived from the standard error on the MLEs; they do not account for error on the fits to the averaged periodograms (which are generated from the same data and so are not independent of the averaged MLEs). As can be seen there is no bias in the MLEs of the QPO frequency and Poisson level, but there is a $\lesssim 2\%$ – 3% , hence negligible bias, in R and w at short PDS integration times (corresponding to a lower signal-to-noise ratio for the QPO). Note that 3% corresponds to a bias of ~ 0.1 Hz on w for the QPO considered here. As expected, these biases decrease when T increases.

The *Fisher matrix* can be evaluated numerically. The confidence limits on the MLEs can also be calculated from $\Delta S = S(\theta) - S(\hat{\theta})$ in the same manner as the popular $\Delta\chi^2$ method (Bevington 1969; Cash 1979), e.g., $\Delta S = 1$ corresponds to the 68.3% confidence limits on one parameter (see discussion in Vaughan 2005). The Fisher matrix method is fast but gives reliable confidence intervals only when the likelihood surface is Gaussian; the ΔS method is slower but gives reasonable confidence intervals even for non-Gaussian likelihoods and can be used to search for local minima. We have verified through simulations and with real data that the errors computed by the two methods did not differ by more than a few percent.

One can also use standard tools of maximum likelihood analysis, such as the likelihood ratio test to test for additional free parameters in the model (see Vaughan 2010; Protasov et al. 2002 for the conditions of use). Unlike the $\min(\chi^2)$ method, the fit statistic used in the maximum likelihood (i.e., S) does not provide an automatic goodness-of-fit test. However, as discussed in Vaughan (2010), useful diagnostic statistics can be constructed and calibrated using Monte Carlo simulations that are sensitive to data-model mismatch.

The minimization process can be made more robust against local minima by (1) repeating the optimization from several (random) settings of the initial parameter estimates and (2) monitoring the ΔS values during the confidence interval calculations. Non-negative parameters such as w and R can also be logarithmically transformed³ to improve the reliability of the numerical methods (Anderson et al. 1990) (Equation (7) below must then be rewritten with $R \rightarrow \exp(\ln R)$ and $w \rightarrow \exp(\ln w)$). Immediately after fitting, R and w can then be transformed back in linear space for further processing. With IDL, we used the *POWELL* minimization routine, but checked that the downhill

simplex method implemented in the *AMOEB*A routine gave exactly similar results. A completely independent set of codes, from the generation of PDS to the MLE, was also developed in *R* (<http://www.r-project.org/>) on a different platform and yielded very similar and fully consistent results, with the IDL ones.

3. RESULTS

We now wish to test the reliability of the MLE with simulated data containing a QPO of known and realistic parameters. For this, we need to define a model for the PDS. This model is the sum of a constant a (to account for the Poisson noise, i.e., equal to 2 for Leahy normalized PDS; Leahy et al. 1983), plus a Lorentzian with three parameters, the normalization R , the width w (FWHM), and the centroid frequency v_0 :

$$S(v; \theta) = a + \frac{Rw}{2\pi((v - v_0)^2 + (w/2)^2)} \quad (7)$$

(where we have denoted the parameters $\theta = \{a, R, w, v_0\}$).

Random time series are generated from the PDS model by inverse Fourier transforming the randomized data (Timmer & König 1995). The time series are then processed to produce PDS of the required frequency resolution. As with the real data, the simulated time series should be much longer than the required PDS integration time. This method includes the well-known effect of the windowing-induced power leakage, e.g., narrow features are slightly broadened (see, e.g., Uttley et al. 2002; Vaughan 2010; see Appendix B). This effect is quite significant for narrow features and short integration times, as shown in Figure 8. The bias on the width of the QPO caused by leakage (hereafter b_w^{leakage}) goes approximately like T^{-1} . Besides this bias, we have checked that our simulations did not generate any biases on the other parameters, R , v_0 , and a . This was done by averaging the periodograms from a large number of simulations, as an approximation to the expectation of

³ As discussed by Appourchaux (2011 and references therein) these parameters tend to be lognormally distributed; hence the log-transformed parameters are approximately normally distributed.

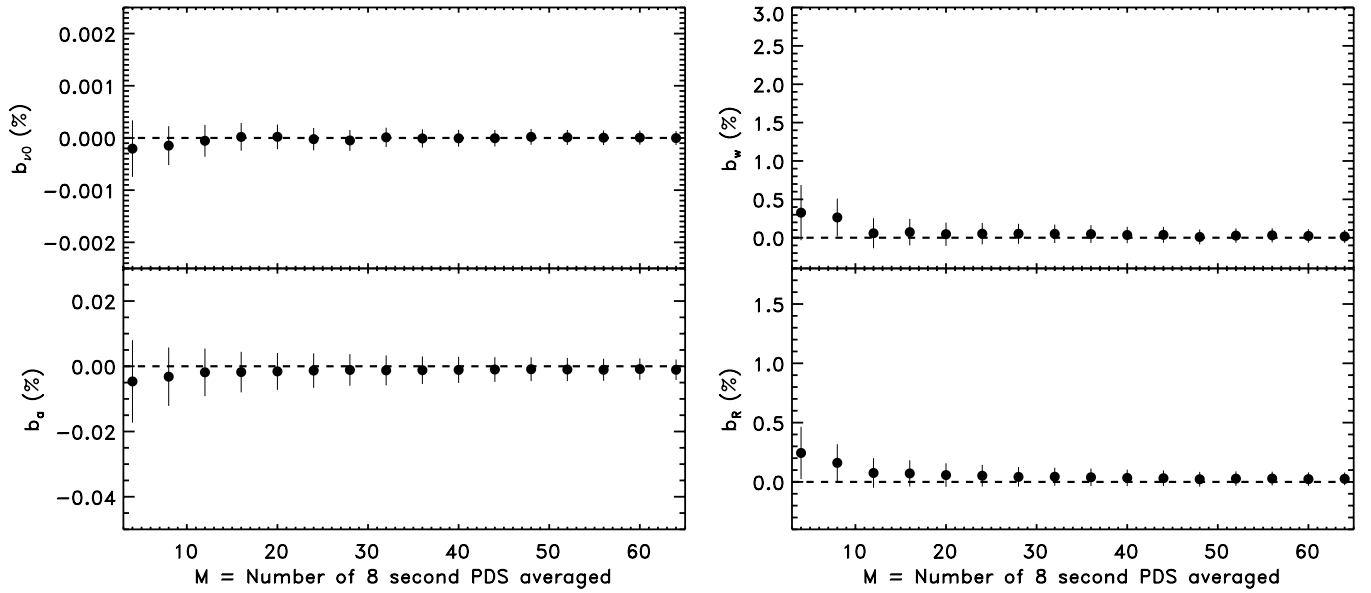


Figure 3. Measured biases on the four model parameters (top left, v_0 ; bottom left, a ; top right, w ; bottom right, R), as a function of the number of 8 s PDS averaged (same y-axis scale as Figure 2). Those biases are derived from the difference between the sample mean of 4096 MLEs (each one from fitting the average of M PDS), and an estimate of the average QPO parameters generated by the simulation over the entire set of $4096 \times M$ PDS simulated. The biases are expressed in percentage of these average values. Error bars on the bias are derived from the standard error on the MLEs; they do not account for error on the fits to the averaged periodograms (which are generated from the same data and so are not independent of the averaged MLEs). Clearly, all the biases converge to zero.

the periodogram (Appendix B), and fitted this (using $\min(\chi^2)$ or the maximum likelihood) to obtain the parameters of the expected periodogram. Within errors, the best-fit parameters, denoted R , \bar{v}_0 , \bar{a} , and \bar{w} , corrected for the leakage bias, were consistent with the input values used for the simulations.

3.1. Biases in the MLEs

In order to test the reliability of the method, we generated and fitted a large number of random data sets and computed the bias on the MLEs as the difference between the sample mean of the MLE estimates $\langle \theta_i^{\text{MLE}} \rangle$ and the best-fitted value of the average periodogram, such that for parameter i , we have $b_i = \langle \theta_i^{\text{MLE}} \rangle - \bar{\theta}_i$. Alternatively, the bias can be estimated from the true value put in the simulations (θ_i^0), so that we have $b_i = \langle \theta_i^{\text{MLE}} \rangle - \theta_i^0$. In this case, special care must be taken for computing the bias on w , as the measured value must be corrected for the leakage bias (b_w^{leakage}), such that $b_w = \langle w^{\text{MLE}} \rangle - b_w^{\text{leakage}} - w^0$. In this paper, we have used the first method to compute the bias. The MLEs are asymptotically unbiased, but when estimated from a finite data set, some may be biased. This is a reason why the analysis of real data (especially when attempting to fit low signal-to-noise ratio features) should be compared with simulated data sets, as closely representative of the data as possible. However, as we will show below, some MLEs are unbiased, and for the biased ones, the bias can be made negligible (i.e., much smaller than the scatter of the data), when the MLE is applied to QPOs that are reasonably well detected (see discussion below for detection criteria).

As we are primarily interested in QPOs of a few Hz breadth that can be detected and fitted on a few tens of seconds, we have generated periodograms with different integration times, corresponding to sub-Hz frequency resolutions (from $T = 8$ to $T = 48$ s, in step of 4 s). For each T , 16384 PDS (128 sets of 128 PDS) were generated from light curves $64 \times T$ long and fitted. A large number of simulations is required to give precise Monte Carlo estimates of the bias and variance of the MLEs. We

have explored a wide range of model parameters, but here we have assumed $R = 19.0$, $w = 3.23$ Hz, and $v_0 = 835.08$ Hz. $R = 19.0$ corresponds to a QPO rms amplitude of about $\sim 10\%$ for a source count rate of $2000 \text{ counts s}^{-1}$. Those parameters are consistent with the ones derived from the analysis of the 4U1608-522 data presented in Section 4. One example of an MLE for one single PDS ($T = 16$ s) is shown in Figure 1. To increase speed, yet to enable a good determination of the Poisson level, the fits were performed on a frequency range of 250 Hz on each side of the QPO peak. As can be seen, despite a very noisy periodogram, the maximum likelihood fitting picks up the QPO at the right frequency, with a 1σ uncertainty of only about 0.4 Hz.

The results on the biases are shown in Figure 2. As can be seen, the biases on the fitted Poisson level and on the QPO frequency are almost zero. There is a small remaining positive, but *negligible* bias on R and w that decreases with increasing T . The bias on R and w is much smaller than the typical 1σ of the estimates, e.g., by more than a factor of 10 for all T , and hence can only be measured using a very large number of PDS, as is the case in our simulations (16384). As expected, the biases on R and w are more severe when the signal-to-noise ratio of the QPO decreases (e.g., for PDS of a given T for a QPO of smaller R or for PDS of shorter T for a QPO at a given R ; see Figure 2). Bouvier et al. (2009) considered a significant QPO detection when the ratio between the fitted R and its 1σ error is larger than 3 or equivalently when the mean significance of the excess power fitted⁴ exceeds 6σ . In Figure 2, both criteria are met, even for $T = 8$ s, for which $\langle R/\sigma_R \rangle = 3.5$ and $\langle n_\sigma \rangle = 9.2$. This means that, provided that T is adjusted as to ensure that QPOs are reasonably well detected (i.e., one of the two criteria above is met), the MLE will not introduce any significant biases on

⁴ This is defined as $n_\sigma = \frac{\bar{P} - P_{\text{mean}}}{2/\sqrt{MW}}$, with \bar{P} the mean of MW powers, P_{mean} the Poisson level (close to 2), M the number of averaged PDS, and w the number of frequency bins averaged; see Boirin et al. (2000). A scanning algorithm maximizes n_σ with respect to w , allowing w to vary within a plausible range, say 2–10 Hz, for narrow QPOs.

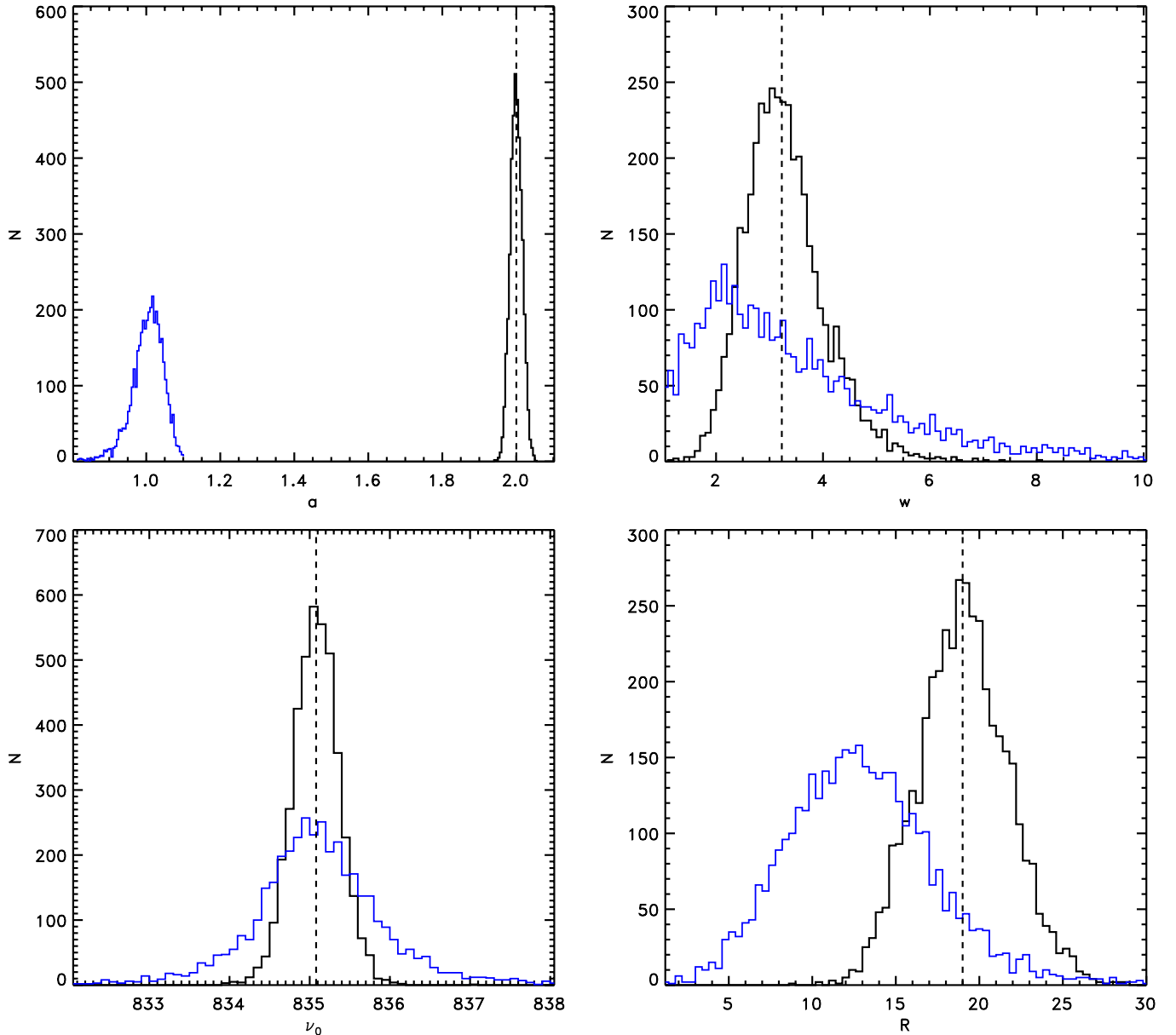


Figure 4. Results of fitting simulated data with a constant and a Lorentzian QPO model. Each of the four panels shows the histogram of the estimates for one of the four model parameters (top left, a ; bottom left, ν_0 ; top right, w ; bottom right, R). The parameters were estimated using the MLE (black line) and the $\min(\chi^2)$ method (blue line) by fitting the average of $M = 4$ periodograms (4096 samples). For $\min(\chi^2)$ fitting, outliers in the parameter distribution (e.g., fitted $R \geq 100$) were removed before building the histograms (about 2% of the total sample are outliers). The vertical dashed line indicates the true parameter values for the simulations. Clearly, the MLEs show no strong biases, but the $\min(\chi^2)$ results underestimate both the Poisson level and the QPO amplitude. The estimates have generally larger spreads from $\min(\chi^2)$ fitting than from the MLE.

(A color version of this figure is available in the online journal.)

the estimates of the QPO parameters. For lower signal-to-noise-ratio QPOs, as stressed above, it is recommended to quantify the residual biases through simulations representative of the data set analyzed.

The biases on the MLEs decrease with increasing T , as shown in Figure 2. We have extended the previous simulations to longer equivalent PDS integration times, adopting a different, though consistent, approach. We have generated 4096 M simulations of $T = 8$ s PDS (M varying from 4 to 64). The range of M considered corresponds to PDS integration times that are typical of the QPO analysis performed so far (Barret et al. 2005b). The 4096 averages of the M PDS were fitted by maximum likelihood (32 different sets of $128 \times M$ PDS), following Appendix A. The biases were again computed as the difference between

the sample mean of the 4096 MLEs and the sample mean of 32 MLEs derived from fitting the average $128 \times M$ PDS, as an estimate of the mean QPO parameters over the entire $4096 \times M$ PDS simulated. The residual biases are shown in Figure 3. First, to illustrate the equivalence of fitting the average of M 8 s PDS and fitting one single PDS of $M \times 8$ s integration time, it can be seen that the biases on w , R (about 0.2%–0.3%) measured for $M = 4$ ($T = 8$ s) are fully consistent with the biases measured for $M = 1$ and $T = 32$ s (see Figure 2). Second, as expected, all the biases converge smoothly to zero as M increases.

3.2. Comparison with $\min(\chi^2)$ Fitting

We have fitted simulated data with the MLE and $\min(\chi^2)$ for comparison, using the 8 s PDS generated for the bias

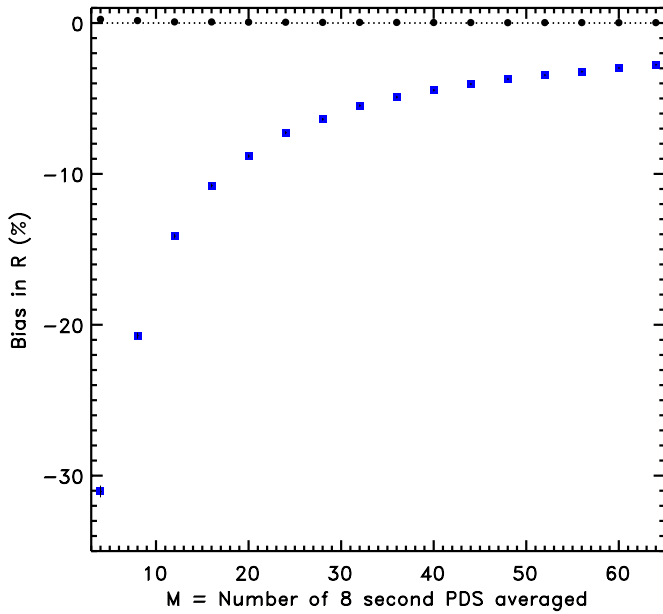


Figure 5. Biases of fitted QPO amplitude (R) with $\min(\chi^2)$ (blue filled squares) and MLE (black filled circles) fitting for different M , from 4 to 64, in steps of 4. Error bars on the MLEs are plotted but are in many cases smaller than the plot symbols. 4096 averages of M PDS were fitted by the two methods. The bias of the MLEs of R is lower than 0.3% ($M = 4$). At the opposite, biases of the R estimates from $\min(\chi^2)$ fitting can be as large as $\sim 30\%$ for $M = 4$ and remain as large as $\sim 2.5\%$ for $M = 64$. This figure alone demonstrates that MLE should always be preferred, when averaging a very large number of PDS is not possible.

(A color version of this figure is available in the online journal.)

analysis discussed above. For illustrative purpose, the averages of $M = 4$ PDS were fitted by both techniques (with the errors on each power I_j set to I_j/\sqrt{M} for $\min(\chi^2)$ fitting). We chose $M = 4$ because it corresponds to a timescale (32 s) on which the properties of kHz QPOs have been investigated previously (e.g., Barret et al. 2005a for 4U1608-522). To ensure a fair comparison of the MLE and $\min(\chi^2)$ methods the same starting values were used for minimization. The histograms of the estimates of each of the four model parameters, as derived from the ensemble of fits for the two methods, are shown in Figure 4. The strongest biases in the $\min(\chi^2)$ fitting clearly concern the Poisson level and the QPO amplitude. This is explained by the fact that the $\min(\chi^2)$ distribution is asymmetric, producing more data below than above the expected spectrum, and these low powers, which have smaller errors, drive the fitted parameters down. By contrast the QPO frequency estimates obtained by $\min(\chi^2)$ fitting are not significantly biased but do show a larger spread (i.e., larger uncertainty in the estimates) than the corresponding MLEs (the same applies to the QPO width). This shows the power of the MLE over $\min(\chi^2)$ when fitting a small number of PDS.

As can be seen from Figure 4, there is quite a large bias in R with $\min(\chi^2)$ fitting, and this bias is important as R gives the rms amplitude of the QPO, one of the QPO parameter that is often reported in the literature. It is interesting to see how it decreases when M increases. In the limit of a large M , the results from $\min(\chi^2)$ fitting should be unbiased and consistent with the ones derived from the MLE. Papadakis & Lawrence (1993) showed that for $M > 50$ the powers of the PDS are approximately normally distributed; hence $\min(\chi^2)$ fitting can be used. Using the same simulated data set as used to generate Figure 3, i.e., $4096 \times M$ PDS, we show how the bias on R varies with M with the two fitting techniques (see Figure 5). As expected the bias

decreases with increasing M with $\min(\chi^2)$ fitting. It is, however, interesting to note that for $M = 64$, the $\min(\chi^2)$ fitting leads to a bias of about 2.5%. This bias is larger than the one derived from the MLE when $M = 1$! (see Figure 2 for $T = 8$ s). With the MLE, the bias in the estimate of R decreases with T and is negligible for all M (it is always below $\sim 0.3\%$). This means that whenever averaging a very large number of PDS is not possible or desirable, MLE fitting should be preferred against $\min(\chi^2)$ fitting, especially if one is interested in measuring the QPO amplitude as accurately as possible.

4. APPLICATION TO RXTE/PCA DATA: THE CASE OF 4U1608-522

4U1608-522 is an X-ray transient and a prototypical high-frequency QPO source. The parameters of its high-frequency QPOs have already been reported (Berger et al. 1996; Mendez et al. 1998; Barret et al. 2005a, 2006). Here we are interested in a subset of the data recorded by the PCA on board RXTE on 1996 March 3 (ObsID 10072-05-01-00). This ObsID is split into three intervals. To produce the light curves, we have extracted all events in the detector channel range from 16 to 94. The smoothed dynamical PDS for the second interval of data is shown in Figure 6. As can be seen from that figure, both the frequency and amplitude of the QPO do not vary much during those observations. The MLE was applied by fitting one single periodogram of different integration times, down to 8 s. The results are shown in the right panel of Figure 6 for an integration time of 8 s. The mean significance of the excess power at the QPO frequency is about $\sim 8.7\sigma$, and the mean ratio of $R/\sigma_R \sim 3.4$, indicating that the MLEs on R and w are not biased by more than $\sim 2\%$ – 3% (see Figure 2 for $T = 8$ s). The QPO parameters are recovered for the first time on very short timescales, down to 8 s from direct fitting.

In the left panel of Figure 7 we show the power spectrum of the frequency variations, using the estimates of the QPO frequency on a timescale of 16 s. A maximum likelihood fit of this power spectrum indicates that the frequency evolution is consistent with a simple random walk, as previously hypothesized (Belloni et al. 2005). This power spectrum is adequately fitted by a simple model comprising a power law of index 1.8 ± 0.2 down to ~ 1 mHz, plus a constant that is consistent with that expected from the uncertainties on the frequency estimates.

Under the random walk hypothesis, we expect the broadening of the QPO due to the variable frequency to increase with \sqrt{T} . We have fitted single PDS with increasing T and fitted the MLEs (corrected from the leakage bias) with a random walk model, assuming that the measured width for a given T (w_T) includes a contribution from the intrinsic QPO width w_{qpo} and a contribution from the drift (w_{drift}), such as $w_T^2 = w_{\text{qpo}}^2 + w_{\text{drift}}^2$, with $w_{\text{drift}} \propto \sqrt{T}$. As can be seen from Figure 7, a good fit is obtained when T varies from 8 to 60 s. This enables us to recover the intrinsic QPO width (in the limit of no frequency drift): $w_{\text{QPO}} = 3.20 \pm 0.03$ Hz. For a mean frequency of 829 Hz over the interval, this corresponds to a quality factor $Q = 259 \pm 5$. Correcting for a positive residual MLE (2%) bias (see Figure 2), this value goes up to $Q = 264 \pm 6$, not significantly different.

The mean of the MLEs of w ($\langle w^{\text{MLE}} \rangle = 3.29 \pm 0.08$ Hz) measured for $T = 8$ s (see Figure 7) is in fact already close to the intrinsic QPO width derived with the method above and could therefore be used to approximate a mean Q over a continuous observation (e.g., an ObsID in RXTE terminology). On the other hand, one should not take the weighted mean of

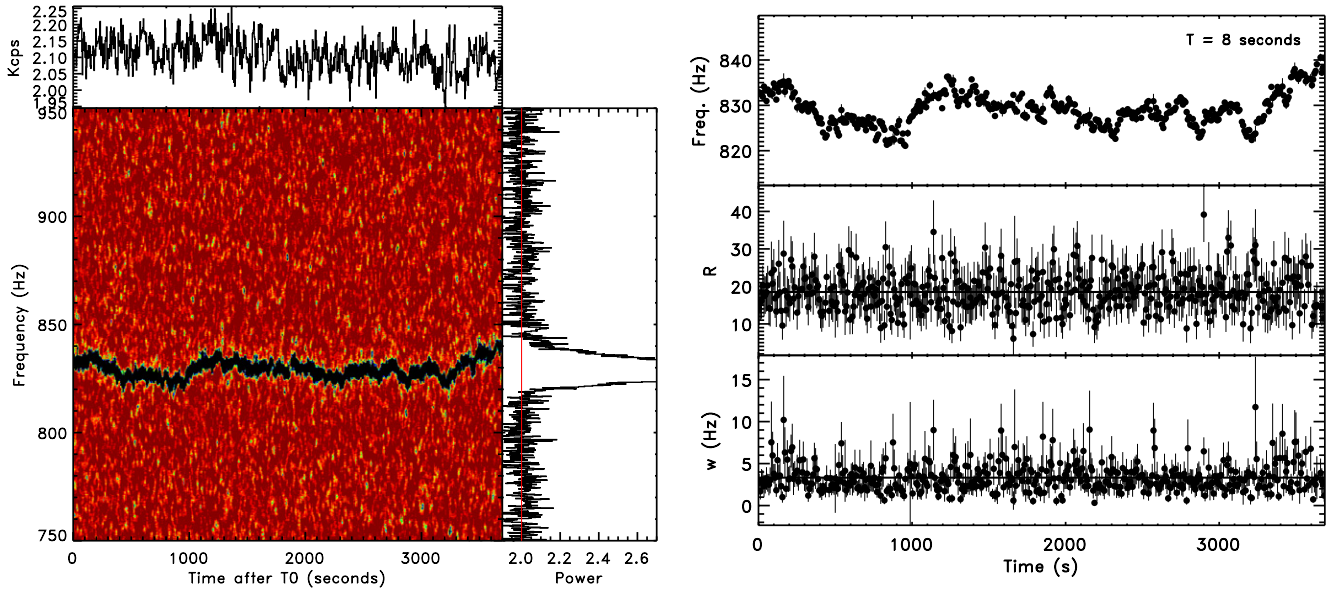


Figure 6. Left: dynamical PDS of 4U1608-522 as recorded during the second segment of the 1996 March 3 observations. The time resolution is 4 s. The X-ray light curve is shown at the top (in keps), and the average PDS over the whole segment is shown in the vertical right panel (cut to a maximum of 2.7). A strong QPO is visible with a mean frequency around 830 Hz. Right: the MLEs for this segment of data, as derived from fitting one single PDS with an integration time of 8 s (from the top to the bottom: the QPO frequency, R , and w).

(A color version of this figure is available in the online journal.)

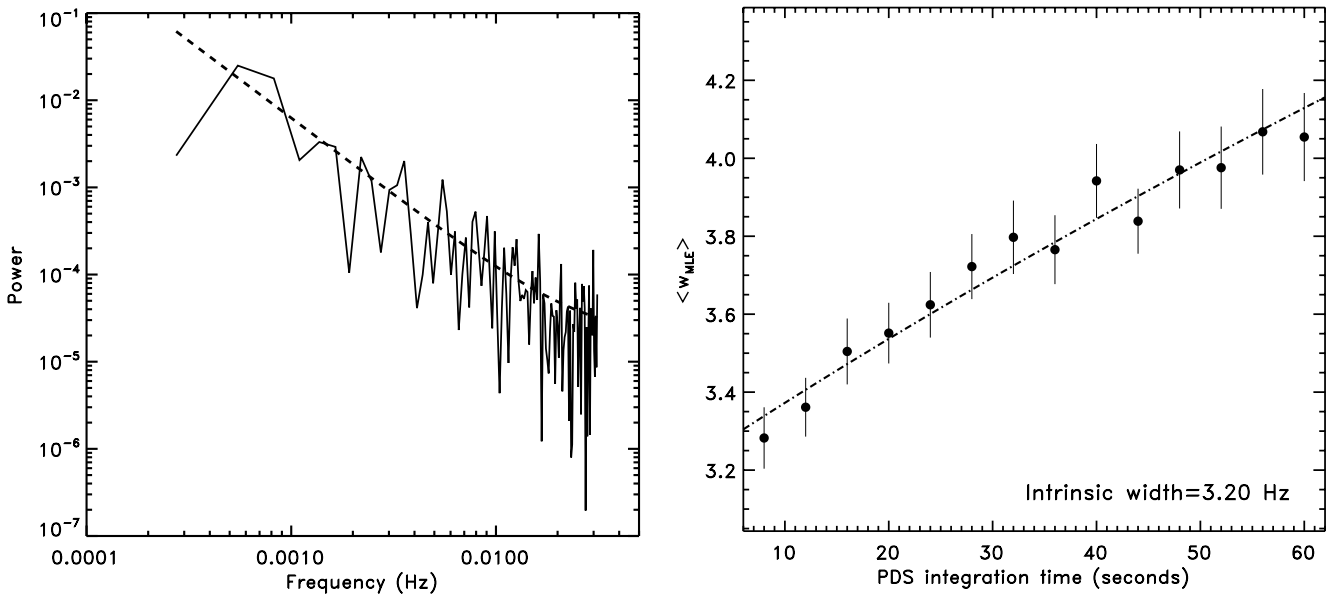


Figure 7. Left: periodogram of the frequency variations of 4U1608-522 as derived from the MLEs of the QPO frequency on a timescale of 16 s. The best power-law fit is shown with a dashed line (index 1.8 ± 0.2). Right: broadening of the QPO with increasing integration time of the fitted PDS. As expected from a random walk, the broadening goes like \sqrt{T} . The best fit is shown with a dot-dashed line. Such analysis enables us to derive the intrinsic QPO width to be 3.2 Hz, corresponding to a quality factor of 259 ± 5 at a frequency of 829 Hz.

w^{MLE} , as it underestimates the true width, due to the fact that smaller w^{MLE} have smaller error bars (see Figure 6). Similarly, as discussed in Appendix C, shifting and adding PDS to a reference frequency (e.g., the mean frequency over an ObsID), and fitting the resulting averaged and shifted PDS, leads also to a tiny negative bias on w (~ 0.1 – 0.2 Hz) and hence a slight overestimate of Q .

To conclude, using a simple random walk model and the MLE for different PDS integration times, we have derived a Q factor of about 260, significantly larger than the value of 200 reported previously (Barret et al. 2005a). Unless the QPO shows rapid frequency variability below 8 s, the above value is

likely to be close to the intrinsic QPO width. The higher quality factor derived, implying a longer the lifetime for the underlying oscillator, puts even more stringent constraints on QPO models, as discussed in Barret et al. (2005a). It also shows that the MLE by enabling the unbiased estimation of QPO parameters on the shortest timescales permitted by the statistical quality of the data provides a clear improvement over previous methods. More data will be analyzed to test the random walk hypothesis, to measure the maximum Q factor of kHz QPOs in neutron star systems and to determine more accurately the dependency on the quality factor with frequency, in particular around the frequency, where Q drops very rapidly: a feature that has been interpreted as

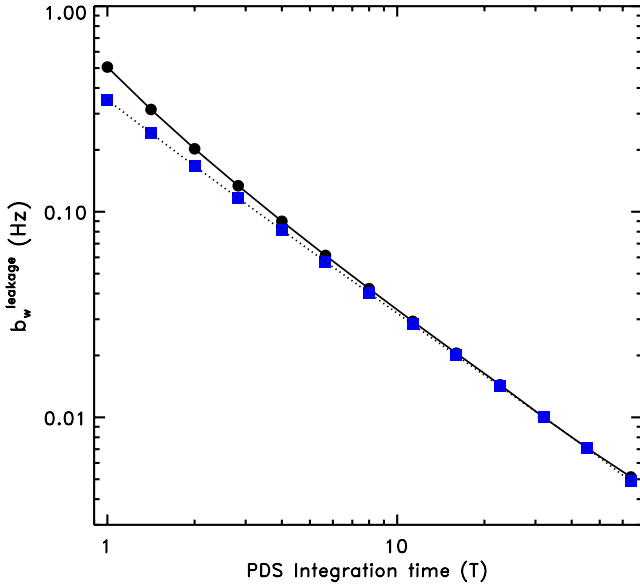


Figure 8. Magnitude of the bias on w (b_w^{leakage}) as a function of time series duration T evaluated for two different values of w (1 and 5 Hz; lower and upper curves, respectively). The leakage bias was computed numerically by comparing of the Lorentzian profile before and after convolution (computed on a fine grid of frequencies, $\delta f = 2^{-13}$). Roughly, the bias decays as $b_w^{\text{leakage}} \propto T^{-1}$. (A color version of this figure is available in the online journal.)

a possible signature of the approach to the innermost stable circular orbit (Barret et al. 2006, 2007).

5. DISCUSSION AND CONCLUSIONS

We have demonstrated the application of maximum likelihood fitting to periodogram data showing high-frequency QPOs. The method is simple to implement and, importantly, remains valid in the non-Gaussian regime, e.g., when applied to raw periodogram data. This means it can be used in situations where averaging many periodograms is not desirable, e.g., when short time series are to be analyzed at the highest frequency resolution. The simulation tests discussed in Section 3 show that the method gives accurate and unbiased estimates of the QPO parameters, and can be used to probe QPO properties on timescales as short as 8 s (for strong kHz QPOs). By contrast the standard $\min(\chi^2)$ fitting method gives biased parameters with larger uncertainties when applied to such short sections of data because the spectral estimates are not Gaussian distributed.

We have applied the MLE to an *RXTE* observation of the ~ 830 Hz QPO of the neutron star low-mass X-ray binary 4U1608-522. For one section of the observation we are able to recover the QPO frequency, width, and amplitude on timescales down to 8 s. We have shown that the frequency variations are consistent with a random walk in the observation analyzed, and we have inferred the intrinsic width of the QPO to correspond to a quality factor of ~ 260 at 830 Hz: a value significantly higher than previous estimates (Barret et al. 2005a). Of course, MLE fitting of time-resolved periodograms is not the only approach available for investigating the variability of the kHz QPOs. Another is to fit time-dependent models directly to the time–frequency data using hierarchical models (Gelman & Hill 2007). This is beyond the scope of this paper whose aim is to highlight the improvement (in, e.g., bias and variance of the fitted QPO parameters, and in some cases time resolution) that can be gained by using the Whittle MLE rather than the $\min(\chi^2)$ method that is currently the standard procedure.

The method described here was applied for the first time to *RXTE* data. By enabling to measure the QPO parameters on very short timescales with negligible biases, this technique opens the way for new types of analysis to be performed on the *RXTE* archive, as we have shown in the case of 4U1608-522 for a limited subset of data. In most analysis so far, the QPO properties have been derived from averaging data sets, hence leading to a potential loss of information. The MLE does not require averaging. Hopefully, this will lead to a better understanding of the high-frequency variability observed in accreting compact objects, both neutron stars and black holes. This is required to fully exploit the potential of this variability as a probe of strong gravity and dense matter (van der Klis 2000).

We thank Thierry Appourchaux, Martin Boutelier, Dacheng Lin, Mariano Mendez, and Cole Miller for useful comments on an earlier version of the manuscript. We are also grateful to an anonymous referee for a helpful report on the statistical aspects of the analysis presented in this paper. This research has made use of data obtained through the High Energy Astrophysics Science Archive Research Center Online Service, provided by the NASA/Goddard Space Flight Center.

APPENDIX A

MAXIMUM LIKELIHOOD FITTING OF AVERAGED PERIODOGRAMS

The periodogram is scattered around its expectation following a $\chi^2/2$ distribution. By averaging M periodograms we find that the average follows a $\chi^2_{2M}/2M$ distribution. From the properties of the χ^2 distribution it is straightforward to derive the log likelihood formula in this more general case (see also Appourchaux 2003):

$$S = -2 \ln \mathcal{L} = \nu \sum_{j=1}^{N-1} \left\{ \frac{I_j}{S_j} + \ln S_j + \left(\frac{2}{\nu} - 1 \right) \ln I_j + c(\nu) \right\}, \quad (\text{A1})$$

where $\nu = 2M$ is the degrees of freedom of the relevant χ^2 distribution and $c(\nu)$ is a constant for fixed ν (not a function of I_j or S_j). In the case of no averaging, i.e., $M = 1$ ($\nu = 2$), the above formula reduces to the usual log likelihood for the periodogram (since $c(\nu = 2) = 0$). The above formula may be used for maximum likelihood fitting of averaged periodograms with any M ; in the limit of very large M this is equivalent to the usual $\min(\chi^2)$ method.

We have verified that fitting the average of M PDS gives very similar results compared to fitting simultaneously M PDS (minimizing the sum of S over M) (see Figures 2 and 3). It is obviously as easy to implement, but runs much faster by a factor of $\sim M$. The errors on the MLEs can be computed the exact same way as in the case of $M = 1$, e.g., directly from the Fisher matrix (see Section 2).

APPENDIX B

SPECTRAL LEAKAGE

The expectation value of the periodogram at Fourier frequency f_j is

$$E[I_j] = \int_{-f_N}^{+f_N} F(f_j - f') S(f') df' = S(f) - b_l(f), \quad (\text{B1})$$

where $S(f)$ is the true spectral density function, $F(f)$ is the Fejer kernel, and f_N is the Nyquist frequency. The convolution by the Fejer kernel distorts the periodogram away from the true spectrum and results from the finite sampling of the time series (see, e.g., Brillinger 1975, chapter 5; Priestley 1981, chapter 6). The expected difference between the true spectrum and the periodogram is the bias on the periodogram: $b_I(f_j) = E[S(f_j) - I_j]$.

The Fejer kernel depends on the time series duration, T , and has a main lobe of width $\Delta f = 1/T$ plus oscillatory side lobes that decay as $\sim 1/f^2$ either side. It is these side lobes that cause spectral leakage—the transfer of power between distant Fourier frequencies. Assuming a Lorentzian spectrum $S(f)$, typical of kHz QPOs, we can compute the effect of leakage by performing the above convolution numerically (on a fine grid of frequencies, $\delta f = 2^{-13}$). For well-resolved QPOs (i.e., $w > 1/T$) the convolved spectrum, i.e., the expectation of the periodogram, is very likely a Lorentzian but with slightly larger width w_C . The bias on w due to leakage, $b_w^{\text{leakage}} = w - w_C$, is itself a function of T . The convolution conserves the total power (integral under the curve) and so leakage does not directly bias R . However, a very small bias on R , almost always negligible, may be introduced by fitting a simple Lorentzian model to the distorted QPO profile, which can deviate slightly from a pure Lorentzian.

Figure 8 shows this broadening effect, b_w^{leakage} as a function of time series duration T , based on numerical calculation of the Fejer convolution. For longer time series the Fejer kernel becomes more concentrated and distorts the spectrum less, leading to smaller b_w^{leakage} . For $T = 4$ s the bias is ≈ 0.08 Hz, within the precision of some estimates of QPO widths, whereas for T longer than ~ 32 s the bias is so small as to be practically insignificant.

APPENDIX C

A SMALL BIAS WITH THE SHIFT-AND-ADD TECHNIQUE

The so-called shift-and-add technique, as introduced by Mendez et al. (1998), has been used extensively to study the power spectral properties of kHz QPOs. The method involves computing PDS estimates on short timescales (e.g., $T = 32$ s), fitting for the QPO frequency, and using the best-fitting frequency to shift the frequency scale of each spectrum to one in which the QPO frequency is constant. This method can improve sensitivity to QPOs with rapidly variable frequencies that would get washed out in a straight average of the short timescale PDS estimates.

In the ideal case of the QPO frequency being known with zero error, the averaged spectrum will be broadened slightly by leakage due to the finite length of the time series intervals used to compute the short timescale PDS estimates (as discussed above and in Appendix B). But the method of shifting using fitted frequencies introduces another tiny bias to the estimation of QPO width. The fitted frequencies (that determine the shifts to be applied) for each interval are scattered around the true frequency without bias, but are closer to the center of mass of each randomly sampled QPO spectrum than is the true frequency. On some occasions there will be more power above or below the true QPO frequency and these will contribute toward the profile of the QPO in the averaged spectrum. But if the data are shifted to reduce the spread of power around the fitted

QPO location, this must result in less power in the tails and more power in the center of the averaged QPO profile—i.e., a bias toward narrower QPOs in the average made using the shift-and-add technique. To estimate the magnitude of the bias introduced by the shift-and-add, we used the fitted frequencies of the simulated data set generated above (i.e., taking 128 sets of 128 PDS with an integration time T of 16 s). For each set, we have fitted the average PDS and the shifted-and-added PDS using $\min(\chi^2)$ fitting. The estimates of R , ν_0 , and a are consistent between the two sets of fits (straight average and shift-and-average), but the QPO width derived from the shifted-and-added PDS is significantly smaller than the one derived from the non-shifted averaged PDS by about 0.15 Hz. This is the reason why we should use the sample mean of the estimates of w^{MLE} to estimate the mean quality factor (Section 4). The shift-and-add bias decreases with T because the error on the fitted frequency decreases to zero: the width bias is less than 0.1 Hz for $T = 24$ s.

REFERENCES

- Anderson, E. R., Duvall, T. L., Jr., & Jefferies, S. M. 1990, *ApJ*, **364**, 699
 Andrae, R. 2010, arXiv:1009.2755
 Appourchaux, T. 2003, *A&A*, **412**, 903
 Appourchaux, T. 2011, Lecture Notes of the 22nd Canary Island Winter School (arXiv:1103.5352)
 Barret, D., Kluźniak, W., Olive, J. F., Paltani, S., & Skinner, G. K. 2005a, *MNRAS*, **357**, 1288
 Barret, D., Olive, J.-F., & Miller, M. C. 2005b, *MNRAS*, **361**, 855
 Barret, D., Olive, J.-F., & Miller, M. C. 2006, *MNRAS*, **370**, 1140
 Barret, D., Olive, J.-F., & Miller, M. C. 2007, *MNRAS*, **376**, 1139
 Bartlett, M. S. 1948, *Nature*, **161**, 686
 Belloni, T., Méndez, M., & Homan, J. 2005, *A&A*, **437**, 209
 Berger, M., van der Klis, M., van Paradijs, J., et al. 1996, *ApJ*, **469**, L13
 Bevington, P. R. 1969, *Data Reduction and Error Analysis for the Physical Sciences* (New York: McGraw-Hill)
 Bloomfield, P. 2000, *Fourier Analysis of Time Series: An Introduction* (2nd ed.; New York: Wiley)
 Boirin, L., Barret, D., Olive, J. F., Bloser, P. F., & Grindlay, J. E. 2000, *A&A*, **361**, 121
 Boutelier, M., Barret, D., & Miller, M. C. 2009, *MNRAS*, **399**, 1901
 Brillinger, D. R. 1975, *Data Analysis and Theory* (New York: Holt, Rinehart & Winston)
 Cash, W. 1979, *ApJ*, **228**, 939
 Gelman, A., & Hill, J. 2007, *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Analytical Methods for Social Research (Cambridge: Cambridge Univ. Press)
 Groth, E. J. 1975, *ApJS*, **29**, 285
 Heavens, A. 2009, arXiv:0906.0664
 Leahy, D. A., Darbro, W., Elsner, R. F., et al. 1983, *ApJ*, **266**, 160
 Mendez, M., van der Klis, M., van Paradijs, J., et al. 1998, *ApJ*, **494**, L65
 Papadakis, I. E., & Lawrence, A. 1993, *MNRAS*, **261**, 612
 Percival, D. B., & Walden, A. T. 1993, *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge: Cambridge Univ. Press)
 Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in FORTRAN. The Art of Scientific Computing* (2nd ed.; Cambridge: Cambridge Univ. Press)
 Priestley, M. B. 1981, *Spectral Analysis and Time Series* (London: Academic Press)
 Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2002, *ApJ*, **571**, 545
 Stella, L., Arlandi, E., Tagliaferri, G., & Israel, G. L. 1994, arXiv:astro-ph/9411050
 Timmer, J., & König, M. 1995, *A&A*, **300**, 707
 Uttley, P., McHardy, I. M., & Papadakis, I. E. 2002, *MNRAS*, **332**, 231
 van der Klis, M. 1989, in *Timing Neutron Stars: Proc. NATO Advanced Study Institute on Timing Neutron Stars held 1988 April 4–15*, in Cesme, Izmir, Turkey, ed. H. Ogelman & E. P. J. van den Heuvel (New York: Kluwer), **27**
 van der Klis, M. 2000, *ARA&A*, **38**, 717
 Vaughan, S. 2005, *A&A*, **431**, 391
 Vaughan, S. 2010, *MNRAS*, **402**, 307