Power-density spectrum of non-stationary short-lived light curves

Cristiano Guidorzi

Department of Physics, University of Ferrara, via Saragat 1, I-44122, Ferrara, Italy

Accepted 2011 April 27. Received 2011 April 27; in original form 2011 February 7

ABSTRACT

The power-density spectrum of a light curve is often calculated as the average of a number of spectra derived on individual time intervals the light curve is divided into. This procedure implicitly assumes that each time interval is a different sample function of the same stochastic ergodic process. While this assumption can be applied to many astrophysical sources, there remains a class of transient, highly non-stationary and short-lived events, such as gamma-ray bursts, for which this approach is often inadequate. The power spectrum statistics of a constant signal affected by statistical (Poisson) noise are known to be a $\chi^2_2$ in the Leahy normalization. However, this is no more the case when a non-stationary signal is also present. As a consequence, the uncertainties on the power spectrum cannot be calculated on the basis of the $\chi^2_2$ properties, as assumed by tools such as XRONOS POWSPEC. We generalize the result in the case of a non-stationary signal affected by uncorrelated white noise and show that the new distribution is a non-central $\chi^2_2(\lambda)$, whose non-central value $\lambda$ is the power spectrum of the deterministic function describing the non-stationary signal. Finally, we test these results in the case of synthetic curves of gamma-ray bursts. We end up with a new formula for calculating the power spectrum uncertainties. This is crucial in the case of non-stationary short-lived processes affected by uncorrelated statistical noise, for which ensemble averaging does not make any physical sense.

Key words: methods: statistical – gamma-ray burst: general.

1 INTRODUCTION

The study of the temporal variability of time series in various branches of science and engineering has propelled the development of several techniques, in both frequency and time domains. Variability studies in the case of astronomical sources are crucial to gain insight into the dynamical and microphysical time-scales, and therefore on the size of the emitting region as well as the nature of the emission process. This is of key importance in the X- and $\gamma$-ray domain, where remarkable flux variations are observed over time-scales from days to ms.

Fourier techniques are widely used in this field, as witnessed by the popular timing analysis package XRONOS\(^1\) (Stella & Angelini 1992) included in the NASA HEASOFT package.\(^2\) The Fourier spectral analysis is fundamental in the study of stationary processes, since it provides an immediate physical interpretation as a power–frequency distribution. The Fourier power-density spectrum (PDS) in particular decomposes the total variance of a given time series to the different frequencies thanks to Parseval’s theorem (e.g. van der Klis 1988, hereafter K88). The PDS analysis and related tools are suitable to both searching for possible periodic signals hidden in the data and characterizing the so-called ‘red noise’ connected with the presence of the aperiodic variability.

Practically in the most general case one divides the time series in multiple adjacent intervals, over which the corresponding PDS is calculated. Finally, for each frequency bin the resulting PDS value and uncertainty are the mean and standard deviation of the corresponding power distribution, as is routinely done by the dedicated XRONOS tool POWSPEC (Stella & Angelini 1992). However, the fundamental assumption behind this procedure is that each time interval represents a different sampling of the same stochastic stationary process. The operation of replacing ensemble averages with time averages of a single realization makes sense only if the process is ergodic (e.g. Priestley 1981).

Astronomical X-ray data rely on photon counting instruments. As such, measurement uncertainties are often dominated by the photon counting statistics, i.e. the Poisson distribution. This translates into white noise in the PDS; adopting the normalization introduced by Leahy et al. (1983) (hereafter L83), the power is known to follow a $\chi^2_2$ distribution with 2 degrees of freedom ($\chi^2_2$).

\(^1\)Available at http://heasarc.gsfc.nasa.gov/docs/xanadu/xronos/xronos.html.
\(^2\)Available at http://heasarc.gsfc.nasa.gov/docs/software/lheasoft/.

© 2011 The Author
Monthly Notices of the Royal Astronomical Society © 2011 RAS
When additional, genuine variability of the source is also present, the resulting power distribution changes correspondingly. For instance, there is a class of X-ray sources, such as Cyg X-1, whose time series are the result of a process characterized by the source-intrinsic correlated noise, for which time averages of PDSs from individual intervals are still meaningful. These time series belong to a class of random processes whose PDS is still compatible with a rescaled \( \chi^2 \) distribution (K88; Israel & Stella 1996).

In the case of highly non-stationary and short-lived events such as gamma-ray bursts (GRBs), the problem of a proper treatment of the Fourier PDS requires particular care. Dividing the light curve of a single GRB into several sub-intervals and deriving an average PDS do not make any physical sense, given that the phenomenon is everything but stationary. Therefore, only a single-sample PDS can be calculated over the entire observation duration. In these cases the statistical distribution of the PDS remains to be determined in the more general case of red deterministic variability due to the source. One has then to be careful with assuming the 100 per cent uncertainties on the PDS provided by standard tools such as POWSPEC: indeed, this relies on the \( \chi^2 \) distribution, which does not hold any more in the more general case of a variable source with red noise.

In the GRB literature, there have been different approaches. In one case, if one considers a set of different time series due to different GRBs as many realizations of the same stochastic process, then averaging the PDS of different GRBs still makes sense. However, it must be pointed out that a strong assumption lies behind this: i.e. there is a unique stochastic process giving rise to the variety of observed GRB time profiles. This way, one gains insight into the properties of this general process, whose PDS is found to be described with a power law, PDS \( \propto \omega^{-\gamma} \) (Beloborodov, Stern & Svensson 2000; Spada, Panaitescu & Mészáros 2000). In the other case, each GRB time profile is considered individually as the unique sample of a unique stochastic process, which is different from other GRBs. The statistics of its unique PDS are not known and are no more \( \chi^2 \). Monte Carlo simulations aimed at estimating the PDS uncertainties by generating other (synthetic) samples of the same process cannot make use of the observed sample curve (e.g. Ukwatta et al. 2011). Indeed, such a procedure increases the noise variance and changes the \( \chi^2 \) nature itself of the power distribution.

In this work we address the issue of a correct evaluation of the statistics of the PDS, and in particular of the power uncertainties, for a single sample of a non-stationary and short-lived signal, such as that of GRB time profiles. In this work time series are meant to be deterministic profiles affected by uncorrelated noise. We derived a formula for the variance of the PDS and show that it agrees with the known results of L83 in the pure white noise case. Finally, we test the validity of our results in the case of synthetic light curves of typical GRBs.

Hereafter, time series are assumed to be discrete, equispaced and with no data gaps. Our treatment assumes the noise to be purely statistical and uncorrelated, and we consider the two cases of Poisson or Gaussian statistics, suitable to photon counting detectors. PDSs are calculated assuming the Leahy normalization (L83). We therefore neglect dead time effects, which are known to affect the statistics and suppress the variance of the resulting time series (e.g. Müller 1973; K88).

## 2 DESCRIPTION

Let \( x_k \) \((k = 0, \ldots, N - 1)\) be a time series observed within a time window with a duration \( T \). The corresponding bin times are \( t_k = kT/N \). The observed series represents a sample function of the true time series we would have observed in the absence of statistical noise (e.g. with an infinite collecting area detector). Hereafter, random variables are written in bold. Each \( x_k \) is therefore a single sample of the random variable \( X_k \), whose expected value and variance are defined as \( E[X_k] = \eta_k \) and \( E[(X_k - \eta_k)^2] = \sigma_k^2 \). The random variables are assumed to be independent, \( E[X_kX_l] = \eta_k\eta_l \) \((k \neq l)\).

The discrete Fourier transform (DFT) amplitudes are defined by

\[
a_j = \sum_{k=0}^{N-1} x_k e^{2\pi j k/N} \quad j = -N/2 + 1, \ldots, N/2
\]

and

\[
x_k = \frac{1}{N} \sum_{j=-N/2}^{N/2-1} a_j e^{-2\pi j k/N} \quad k = 0, \ldots, N - 1,
\]

where the corresponding \( j \)th frequency is \( f_j = jT \). The highest frequency is the Nyquist frequency, \( f_{N/2} = N/2T \). In the Leahy normalization the power spectrum is defined by

\[
P_j = \frac{2}{N_{ph}} |a_j|^2 = \frac{2}{N_{ph}} \sum_{k,l} x_k x_l e^{2\pi i(k-l)j/N} \quad \left(j = 1, \ldots, N/2\right)
\]

or, equivalently,

\[
P_j = \frac{2}{N_{ph}} \sum_{k,l} x_k x_l \cos(2\pi(k-l)j/N).
\]

\(N_{ph} = \sum_{k=0}^{N-1} \eta_k^2\) is the expected total variance. As shown in Appendix B, this happens to coincide with the total number of counts in the specific case of Poisson statistics, which explains the reason for the choice of this name. From equation (3) the expected value of the random variable \( P_j \) is derived straightaway, and is

\[
E[P_j] = 2 + \frac{2}{N_{ph}} \sum_{k,l} \eta_k \eta_l e^{2\pi i(k-l)j/N},
\]

© 2011 The Author, MNRAS 415, 3561–3570
Monthly Notices of the Royal Astronomical Society © 2011 RAS
where we used \( E[x_i^2] = \sigma_i^2 + \eta_i^2 \). The first term in the right-hand side of equation (5) accounts for the statistical noise variance, also called ‘white’ because it does not depend on frequency. When the signal is constant, \( \eta_i = \eta (\forall k) \), \( P_j \) \((j = 1, \ldots, N/2 - 1)\) is known to be \( \chi_2^2 \) distributed, except for the Nyquist frequency, for which \( P_{N/2} \) is \( \chi_1^2 \) distributed and must be treated separately (L83). In the pure white noise case the second term vanishes, thus leaving \( E[P_j] = 2 \).

In general, the second term can also be seen as the DFT of the autocorrelation function (ACF) defined by

\[
R_j = \sum_{k=0}^{N-1} \eta_k \eta_{k+j}, \quad (j = 0, \ldots, N-1),
\]

assuming the periodic boundary condition \( \eta_{k+N} = \eta_k \). This is the discrete version of the Wiener–Kinchin theorem stating that the power spectrum of a time series is the Fourier transform of its ACF (e.g. Papoulis & Pillai 2002).

We define the PDS of the deterministic function \( P_j^{(0)} \) by

\[
P_j^{(0)} = \frac{2}{N \phi} \left| \sum_{k=0}^{N-1} \eta_k e^{2\pi ik j/N} \right|^2
\]

so that equation (5) can be written as

\[
E[P_j] = 2 + P_j^{(0)}.
\]

So far no assumption was made on the kind of distribution of the random variables \( x_k \). Hereafter, we distinguish between the Gaussian and the Poisson noise cases.

### 2.1 Gaussian noise

Each random variable \( x_k \) is distributed according to a normal \( N(\eta_k, \sigma_k) \). We can express equation (4) as the result of matrix products:

\[
P_j = X^T A X, \quad A_{kl} = \frac{2}{N \phi} \cos(2\pi(k - l)j/N)
\]

where \( X \) is the column vector whose \( k \)th row element is \( x_k \), and \( A_{kl} \) is the \((k, l)\) element of \( A \). The matrix \( A \) is a positive-definite quadratic form. As such, from the algebra of the quadratic forms (e.g. Ennis & Johnson 1993) the distribution of the random variable \( P_j \) is a non-central \( \chi_2^2(\lambda) \) if and only if the two following conditions are fulfilled:

\[
\begin{align*}
\text{rank}(A) &= r, \\
\lambda &= H^T A H = P_j^{(0)},
\end{align*}
\]

with \( H \) being the column vector whose \( k \)th element is \( \eta_k \). Equation (11) follows from the definition of \( P_j^{(0)} \) given in equation (7). \( \Sigma \) is the covariance matrix, so its \((k, l)\) element is given by

\[
\Sigma_{kl} = \text{Cov}(x_k, x_l) = E[x_k x_l] - E[x_k] E[x_l] = \delta_{kl} \sigma_k^2,
\]

where \( \delta_{kl} \) is Kronecker’s delta. To verify the first equation (10), we calculate the \((k, l)\) element of the matrix \( A \Sigma A \), which is found to be

\[
(A \Sigma A)_{kl} = A_{kl} + \frac{2}{N \phi} \sum_{m=0}^{N-1} \sigma_m^2 \cos(2\pi(k + l - 2m)j/N) \quad \left( j = 1, \ldots, N \right),
\]

The second term of the right-hand side is identically zero when all the variances \( \sigma_k^2 \) are equal. More generally, the second term of the right-hand side of equation (13) is \( O(1/N^2 \phi) \) times \( A_{kl} \). Therefore, equation (10) is approximately fulfilled in practical cases of interest.

The rank \( r \) of \( A \) does not depend on \( H \), so we can calculate it in the special case when \( \eta_k = \eta (\forall \eta) \), i.e. the case of a constant signal. We already know that \( P_j \) is \( \chi_2^2 \) distributed (L83), so it must be \( r = 2 \).

This proves that in the more general case of a varying signal described by a deterministic function whose discrete values are given by \( \eta_k (k = 0, \ldots, N - 1) \), and affected by pure uncorrelated Gaussian noise \( \sigma_k \), the Leahy-normalized power spectrum \( P_j \) distributes according to a non-central \( \chi_2^2(\lambda) \), where \( \lambda = P_j^{(0)} \) (equation 11). The expected value of \( P_j \) is already known from equation (5), while its variance is

\[
\text{Var}(P_j) = 2(2 + 2\lambda) = \left( 1 + \frac{2}{N \phi} \sum_{k,l} \eta_k \eta_l e^{2\pi ik(l-k)/N} \right) = \left( 1 + P_j^{(0)} \right) \quad \left( j = 1, \ldots, N \right).
\]

Appendix A reports the direct calculation of \( \text{Var}(P_j) \) and equation (A7) reports the exact formula for the variance. In the \( j = N/2 \) case \( P_{N/2} \) satisfies the conditions (10) with \( r = 1 \),

\[
\text{Var}(P_{N/2}) = 4 \text{Var}(P_{N/2}/2) = 8(1 + 2\lambda) = 8 \left( 1 + \frac{2}{N \phi} \sum_{k,l} \eta_k \eta_l e^{2\pi ik(l-k)} \right) = 8 \left( 1 + P_{N/2}^{(0)} \right) \quad .
\]

Appendix A reports the direct calculation of \( \text{Var}(P_j) \) and equation (A7) reports the exact formula for the variance. In the \( j = N/2 \) case \( P_{N/2} \) satisfies the conditions (10) with \( r = 1 \),

\[
\text{Var}(P_{N/2}) = 4 \text{Var}(P_{N/2}/2) = 8(1 + 2\lambda) = 8 \left( 1 + \frac{2}{N \phi} \sum_{k,l} \eta_k \eta_l e^{2\pi ik(l-k)} \right) = 8 \left( 1 + P_{N/2}^{(0)} \right) \quad .
\]

\(^3\) If \( y \) is \( \chi_2^2(\lambda) \) distributed, then \( E[y] = r + \lambda \), and \( \text{Var}(y) = 2(r + 2\lambda) \).
where we used $\lambda = P_{N/2}^{\text{MLE}}/2$. Equations (14) and (15) can also be written as a function of the expected power:

$$\text{Var}(P_j) = \begin{cases} 4 \left( E(P_j) - 1 \right) & (j = 1, \ldots, \frac{N}{2} - 1) \\ 8 \left( E(P_{N/2}) - 1 \right) & (j = \frac{N}{2}) \end{cases}. \quad (16)$$

Clearly, in the general case of uncorrelated noise and a significant power above the white noise level, i.e. when $E(P_j) > 2$, assuming a 100 per cent uncertainty on the resulting power spectrum $P_j$, as assumed by widespread tools in X-ray astronomy such as XRONOS POWSPEC, overestimates the uncertainty. While the 100 per cent uncertainty inherited from the $\chi^2_1$ distribution is correct under the assumption that all power is due to Poisson noise and intrinsic correlated noise, in the case of uncorrelated noise here considered from equation (16) the uncertainty is given by $2 \sqrt{E(P_j) - 1}$, and not $E(P_j)$. The two coincide only in the case of pure noise associated with a constant signal, so that $E(P_j) = 2$. The 100 per cent uncertainty assumed by XRONOS remains correct in the presence of the genuine stochastic variability due to correlated noise.

In practice, when only a single-sample series $x_k$ is available of a short-lived process, such as the time profile of a GRB, the deterministic series $\eta_k$ is not known a priori. However, one could constrain the probability density function (PDF) of $P_j$, i.e. the unknown $\lambda$ (equation 11), with the only sampled value $P_j$. In principle, this makes no difference to taking the light curve with observed $x_k$ counts as $x_k \pm \sqrt{x_k}$ instead of the unknown $\eta_k \pm \sqrt{\eta_k}$ for a Poisson process (the Gaussian case is formally the same). We estimate $\lambda$, the best value for $\lambda$, adopting a Bayesian approach:

$$p(r, \lambda | P_j) = \frac{p(P_j | r, \lambda) p(r, \lambda)}{p(P_j)} \quad , \quad (17)$$

where $p(r, \lambda | P_j)$ is the posterior function of the parameters $(r, \lambda)$ ($r$ is fixed to either 1 or 2 depending on whether it is or not $j = N/2$) given the observed $P_j$. The likelihood function $p(P_j | r, \lambda)$ is merely the PDF of $P_j$, i.e. $\chi^2_1(\lambda, P_j)$. The prior $p(r, \lambda)$ may include the knowledge one might have on $\lambda$ prior to measuring $P_j$, e.g. when a specific shape of the deterministic PDS is expected. In the most general case, we assume a uniform prior. The term $p(P_j)$ normalizes the likelihood function. For a given $P_j$ $\lambda$ is chosen so as to maximize the posterior probability. In this case our approach is equivalent to a maximum likelihood estimation (MLE). In Appendix C, we show that it is $\lambda = 0$ ($P_j < 2$), and that $\lambda$ rapidly converges to $P_j$ ($j < N/2$) and to $P_{N/2}^{\text{MLE}}$ ($j = N/2$) for $P_j > 2$. We conservatively assumed $\lambda = P_j$ ($= P_{N/2}^{\text{MLE}}$ for $j = N/2$) for all values of $P_j$ to avoid the risk of underestimating the variance. Therefore, for a single time series $x_k$ consisting of an unknown deterministic function affected by uncorrelated white noise equations (14) and (15) are approximated to

$$\sigma(P_j) = \begin{cases} 2 \sqrt{P_j + 1} & (j < N/2) \\ 2 \sqrt{2} \sqrt{P_{N/2}^{\text{MLE}} + 1} & (j = N/2) \end{cases}. \quad (18)$$

Interestingly, the case of a constant variance (all $\sigma^2_j$ are equal), for which the first equation (10) is fulfilled exactly (see equation 13), was discussed by Groth (1975). Apart from a scalefactor of 2 in the definition of power, the PDF he derived is precisely that of a non-central chi square with $2n$ degrees of freedom, and a non-central parameter given by the deterministic (or ‘signal’, as he called it) power (see equations 12 and 14 therein). His treatment considered the case where the power is the sum of $n$ terms due to as many frequency bins, while here we consider the $n = 1$ case.

### 2.2 Poisson noise

Each random variable $x_k$ is distributed according to a Poisson distribution with expected value $\eta_k$. Since $x_k$ are no more normal, we cannot exploit the properties expressed by the conditions (10) for the Gaussian case. In Appendix B we calculate the corresponding variance of $P_j$.

Equation (B4) reports the exact formula for $\text{Var}(P_j)$.

The main results are the following.

(i) If the observed counts $x_k$ are so high as to ensure the Gaussian regime ($x_k \gg 1$) the results are the same as those discussed in Section 2.1, since the overall process essentially becomes Gaussian.

(ii) If $N_{\text{ph}} \gg 1$, even if the individual $x_k$ are in the low-count regime, equations (14)–(18) still hold, and in particular $P_j$ still distributes according to a non-central $\chi^2_1(\lambda)$, as in Section 2.1.

(iii) If $N_{\text{ph}} \sim \sim 1$, a few, equations (14)–(18) do not hold any more and the distribution of $P_j$ deviates from a $\chi^2_1(\lambda)$.

### 3 APPLICATIONS: GAMMA-RAY BURST LIGHT CURVES

We produced a number of GRB synthetic light curves to test the validity limits of the results we derived for nature of the power spectrum distribution and summarized by equations (14)–(18). GRBs are particularly suitable for this aim, given their nature of highly non-stationary and short-lived phenomena. We adopted the fast-rise exponential decay (FRED) profile as modelled by Norris et al. (1996). This function...
satisfactorily describes the temporal behaviour of the simplest example of the GRB light curve, which consists of a single pulse modelled as

\[
F(t) = \begin{cases} 
A \exp \left[-\left(\frac{t - t_{\text{max}}}{\tau_r}\right)^p\right], & t < t_{\text{max}} \\
A \exp \left[-\left(\frac{t - t_{\text{max}}}{\tau_d}\right)^p\right], & t > t_{\text{max}} 
\end{cases}
\]

(19)

where \(t_{\text{max}}\) is the peak time, \(\tau_r\) and \(\tau_d\) are the rise and decay times, respectively, \(A\) is the normalization and \(p\) is the peakedness (when \(v = 1\) the profile is a simple exponential, when \(v = 2\) it is a Gaussian). For a typical FRED it is \(\tau_r/\tau_d < 1\), with an average value of \(\sim 0.3–0.5\) (Norris et al. 1996). The continuous PDS of an FRED with \(p = 1\) (double exponential) can be calculated analytically, and apart from a normalization term is found to be

\[
P(v) = \left| \int_{-\infty}^{\infty} F_{p=1}(t) e^{2\pi i vt} dt \right|^2 = \frac{A^2 (\tau_r + \tau_d)^2}{\left[1 + (2\pi v \tau_r)^2\right] \left[1 + (2\pi v \tau_d)^2\right]}.
\]

(20)

Several examples of power spectra obtained for the more general case of an FRED with \(p \neq 1\) are discussed by Lazzati (2002).

We started with an FRED with the following parameters: \(\tau_r = 10\) s, \(\tau_d = 30\) s, \(p = 1.5\), \(A = 1000\) counts bin\(^{-1}\), \(t_{\text{max}} = 0\) s, superposed to a constant detector background with an average intensity of 1000 counts bin\(^{-1}\). We generated \(5 \times 10^3\) samples of this pulse with a bin time of 64 ms, assuming Poisson statistics. The total number of bins amounts to 4096 (\(=2^{12}\)). Given the large number of counts per bin, this is equivalent to the Gaussian case. Fig. 1 displays the synthetic curve of the deterministic model as well as one out of the simulated samples. The deterministic PDS is used for comparison to assess the goodness of the PDS of the sample curve, in particular of the uncertainties derived for the power adopting equation (18) as a function of frequency. This is shown by the ratio between the calculated \(\sigma\) and the scatter of the power of the synthetic PDSs observed for each corresponding frequency bin (bottom-right panel of Fig. 1). Apparently, the ratio ranges between 0.5 and 2. An overall comparison with the analogous ratio between the uncertainty provided by POWSPEC and the corresponding value determined from the MC simulations (mid-right panel of Fig. 1) shows that the improvement is noteworthy. Furthermore, the accuracy of equation (18) is evident when the PDS is dominated by the deterministic power of the signal (at low frequencies). At high frequencies, where the white statistical noise dominates, the power uncertainty is systematically overestimated up to a factor of 2. However, compared to the values provided by POWSPEC often underestimated by up to an order of magnitude, it still represents a significant improvement, particularly in the more conservative direction of overestimating rather than underestimating uncertainties.
Figure 2. Left: the Leahy power distribution at $\nu = 0.061$ Hz derived from 5000 synthetic sample curves of the FRED pulse of Fig. 1. The expected distribution is shown with a solid line, and corresponds to a non-central $\chi^2(\lambda)$, with $\lambda = 36.41$. Right: the same at $\nu = 0.076$ Hz, where $\lambda = 10.69$.

Figure 3. Same as in Fig. 1. Here the FRED has the same profile, but it is 200 times less intense. The background level is proportionally lower, 5 counts bin$^{-1}$. This example fits in the low-count rate Poisson regime ($N_{ph} = 2.3 \times 10^4$).

So far this proves that equation (18) overall provides a satisfactory means for estimating the uncertainty on the PDS of a sample curve. We go further and test whether the distribution of the individual $P_j$ is indeed a non-central chi-square distribution. To this aim, from the PDS of the sample curve considered above we chose two frequencies and derived the corresponding power distribution from the synthetic PDSs. The result is displayed in Fig. 2. The expected non-central parameter for the corresponding $\chi^2(\lambda)$ was calculated with equations (7, 14).

Fig. 3 shows the same FRED pulse 200 times fainter superposed to a correspondingly fainter constant background of 5 counts bin$^{-1}$. The single $x_t$ variables cannot be approximately assumed to be normally distributed, so that we may test the Poisson regime. Still, the total number of counts is still very large, $N_{ph} = 2.3 \times 10^4$. This means that the results obtained for the Gaussian case should still hold. Indeed, both the comparison between the calculated uncertainties and the scatter of the corresponding power from the simulated PDSs give similar results to the previous case (bottom-right panel of Fig. 3). Likewise, the power distributions of individual frequency bins are fully compatible with the corresponding expected non-central chi squares, as shown in Fig. 4.
4 SUMMARY AND CONCLUSIONS

We investigated the nature of the statistical distribution of the PDS of a single, non-stationary and short-lived time profile on a theoretical ground. This treatment assumes the time series to be deterministic profiles affected by uncorrelated noise. In other words, the time series here considered consist of a set of statistically independent, Poisson and normally distributed random variables, whose expected values represent the deterministic function of the varying signal to be studied.

We demonstrated that the PDF of the power is a non-central $\chi^2_\lambda$, whose non-central parameter $\lambda$ corresponds to the power of the deterministic function. This holds in the Gaussian case, as well as in the Poisson case, provided that the Gaussian limit of $N_{\text{ph}} \gg 1$ is fulfilled ($N_{\text{ph}}$ being the total number of counts). As a consequence, we provided a new formula for calculating the correct uncertainty of the power at each frequency as a function of the observed power itself. We finally showed the agreement with simulated light curves of typical GRB time profiles. These results provide a statistically solid basis to a proper treatment of power-density spectra in the case of non-stationary and short-lived time series affected by uncorrelated noise.

ACKNOWLEDGMENTS

The author is grateful to Mauro Orlandini and Raffaella Margutti for reading the manuscript and for their useful comments. The author also wishes to thank the referee, Michiel van der Klis, for providing very useful comments. The author acknowledges ASI for financial support (ASI-INAF contract I/088/06/0).

REFERENCES

Beloborodov A. M., Stern B. E., Svensson R., 2000, ApJ, 535, 158
Ennis D. M., Johnson N. L., 1993, Comm. Statist., 22, 897
Groth E. J., 1975, ApJS, 29, 285
Israel G. L., Stella L., 1996, ApJ, 468, 369
Lazzati D., 2002, MNRAS, 337, 1426
Leahy D. A., Darbro W., Elsner R. F., Weisskopf M. C., Sutherland P. G., Kahn S., Grindlay J. E., 1983, ApJ, 266, 160 (L83)
Müller J. W., 1973, Nuc. Instr. Meth., 112, 47
Norris J. P., Nemiroff R. J., Botticelli B. T., Scargle J. D., Kouveliotou C., Paciesas W. S., Fishman G. J., 1996, ApJ, 459, 393
Papoulis A. S. U. Pillai, 2002, Probability, Random Variables and Stochastic Processes, 4th edn. McGraw Hill, New York
Priestley M. B., 1981, Spectral Analysis and Time Series. Academic Press, London, UK
Spada M., Panaitescu A., Mészáros P., 2000, ApJ, 537, 824
Stella L., Angelini A., 1992, in Di Gesù V., Scarsi L., Bucccheri R., Crane P., Maccarone M. C., Zimmerman H. V., eds, Data Analysis in Astronomy IV. Plenum, New York, p. 59
Ukwatta T. et al., 2011, MNRAS, 412, 875
van der Klis M., 1988, in Ögelman H., van den Heuvel E. P. J., eds, Timing Neutron Stars. NATO ASI, Series C, Kluwer, Dordrecht, p. 27 (K88)

APPENDIX A: VARIANCE OF THE POWER (GAUSSIAN CASE)

The central moments of a random variable $x_i$ normally distributed as $N(\eta_i, \sigma_i)$ are

$$E\{ (x_i - \eta_i)^{2k+1} \} = 0 \quad (\forall i), \quad E\{ (x_i - \eta_i)^2 \} = \sigma_i^2, \quad E\{ (x_i - \eta_i)^4 \} = 3\sigma_i^4 \quad (A1)$$

© 2011 The Author, MNRAS 415, 3561–3570
Monthly Notices of the Royal Astronomical Society © 2011 RAS
where the different variables at different values of $k$ are meant to be independent, so $E[x_i x_j] = \eta_l \eta_j$, ($k \neq l$). They can be used to calculate the corresponding non-central moments through the following equation:

$$E\{x_i^n\} = \sum_{i=0}^{n} \binom{n}{i} E\{x_i - \eta_l\}^i \eta_l^{n-i}. \quad (A2)$$

The non-central moments are given by

$$E\{x_i\} = \eta_l, \quad E\{x_i^2\} = \eta_l^2 + \sigma_l^2, \quad E\{x_i^4\} = \eta_l^4 + 3\eta_l^2\sigma_l^2 + 3\sigma_l^4. \quad (A3)$$

The variance of $P_j$ is calculated directly from equation (3):

$$\left(\frac{N_{ph}}{2}\right)^2 \operatorname{Var}(P_j) = \left(\frac{N_{ph}}{2}\right)^2 \left( E\{P_j^2\} - E^2\{P_j\} \right)$$

$$= \sum_{k,l,m,n} E[x_k x_l x_m x_n] e^{2\pi i(k-l-m-n)/N} - \left( \sum_k \sigma_l^2 + \sum_{k,l} \eta_l \eta_k e^{2\pi i(k-l)/N} \right)^2,$$

where we used the definition of $N_{ph}$. The terms of equation (A4) must conveniently be separated based on the different kind of moments. To simplify the notation, we define $\omega = 2\pi j/N$.

$$\left(\frac{N_{ph}}{2}\right)^2 \operatorname{Var}(P_j) = \sum_k E\{x_k^n\} + 2 \sum_{k \neq l} E\{x_k\} E\{x_l\} (e^{2\pi i(k-l)} + e^{-2\pi i(k-l)}) + \sum_{k,l} E\{x_k^2\} E\{x_l^2\} (2 + e^{2\pi i(k-l)})$$

$$+ \sum_{k,l,m \neq n} E\{x_k\} E\{x_l\} E\{x_m\} (4e^{2\pi i(l-m)} + e^{2\pi i(2k-l-m)} + e^{-2\pi i(2k-l-m)})$$

$$+ \sum_{k,l,m \neq n} E\{x_k\} E\{x_l\} E\{x_m\} e^{2\pi i(k-l-m-n)} - \sum_{k,l} \sigma_l^2 \sigma_k^2 - \sum_{k,l,m,n} \eta_l \eta_k \eta_m \eta_n e^{2\pi i(k-l-m-n)}$$

$$- 2 \sum_{k,l,m} \sigma_k^2 \eta_l \eta_m e^{2\pi i(k-l)}.$$

We adopted the following notation: $\sum_{k,l,m,\neq n}$ is a sum over $k$, $l$ and $m$, and is meant to exclude all the equality cases ($k \neq l$, $k \neq m$, $l \neq m$). Replacing the values of the corresponding moments,

$$= \sum_k \left[ 6\eta_l^4 \sigma_l^2 + 3\sigma_l^4 \right] + 6 \sum_{k \neq l} \eta_l \eta_k \left( e^{2\pi i(k-l)} + e^{-2\pi i(k-l)} \right) + \sum_{k,l} \left( \sigma_l^2 \sigma_k^2 + \eta_l^2 \sigma_k^2 + \eta_k^2 \sigma_l^2 \right) (2 + e^{2\pi i(k-l)})$$

$$+ \sum_{k,l,m \neq n} \sigma_k^2 \eta_m \left( 4e^{2\pi i(l-m)} + e^{2\pi i(2k-l-m)} + e^{-2\pi i(2k-l-m)} \right) - \sum_{k,l,m} \sigma_k^2 \sigma_l^2 - 2 \sum_{k,l,m} \sigma_k^2 \eta_l \eta_m e^{2\pi i(k-l)}.$$

After a few passages, by adding and subtracting the terms excluded in the sums in equation (A6), one ends up with the following equation:

$$\operatorname{Var}(P_j) = 4 \left( 1 + \frac{2}{N_{ph}} \sum_{k,l} \eta_l \eta_k e^{2\pi i(k-l)/N} \right)$$

$$+ 4 \frac{N_{ph}}{N} \sum_{k,l} \sigma_k^2 \eta_l \eta_m \left( e^{2\pi i(2k-l-m)/N} + e^{-2\pi i(2k-l-m)/N} \right). \quad (A7)$$

Equation (A7) is exact. For the Nyquist frequency, equation (A7) becomes

$$\operatorname{Var}(P_{N/2}) = 8 \left( 1 + \frac{2}{N_{ph}} \sum_{k,l} \eta_l \eta_k e^{\pi i(k-l)} \right), \quad (A8)$$

in agreement with equation (15). For $j \neq N/2$ the second term in the right-hand side of equation (A7) is identically zero when all $\sigma_k = \sigma$, and it can be neglected when the $\sigma_k$'s are comparable with each other, being $O(1/N_{ph})$ times the first term. Equation (A7) can be approximated by

$$\operatorname{Var}(P_j) \approx 4 \left( 1 + \frac{2}{N_{ph}} \sum_{k,l} \eta_l \eta_k e^{2\pi i(k-l)/N} \right) \quad \left( j = 1, \ldots, \frac{N}{2} - 1 \right)$$

$$= 8 \left( 1 + \frac{2}{N_{ph}} \sum_{k,l} \eta_l \eta_k e^{\pi i(k-l)} \right) \quad \left( j = \frac{N}{2} \right), \quad (A9)$$

in agreement with equations (14) and (15).

**APPENDIX B: VARIANCE OF THE POWER (POISSON CASE)**

The case of a process $x_k$ affected by the Poisson noise is a more general case than that of a Gaussian noise, since the latter corresponds to the former in the high count rate regime, i.e. when it is $E[x_k] = \eta_l \gg 1$. For a Poisson process, the central moments are

$$E[(x_k - \eta_l^2)^2] = \eta_l, \quad E[(x_k - \eta_l)^3] = \eta_l, \quad E[(x_k - \eta_l)^4] = \eta_l + 3\eta_l^2. \quad (B1)$$

© 2011 The Author, MNRAS 415, 3561–3570

Monthly Notices of the Royal Astronomical Society © 2011 RAS
From equation (A2) the corresponding non-central moments are
\[ E[x_i^4] = \eta_x, \quad E[x_i^2] = \eta_x^2 + \eta, \quad E[x_i^3] = \eta_x^3 + 3\eta_x \eta + \eta, \quad E[x_i^4] = \eta_x^4 + 6\eta_x^3 + 7\eta_x^2 + 3\eta. \] (B2)

In the Poisson case the normalization constant \( N_{ph} = \sum \eta_k \) is the expected total counts. The definition of the power spectrum \( P_j \) is the same as that of equation (3). Using the first two moments of equation (B2), calculating the expected value of \( P_j \) is straightforward, and is found to be the same as equation (5).

Calculating the variance of \( P_j \) in the Poisson case is formally the same as the Gaussian case up to equation (A5), i.e. prior to substituting the specific values of the moments. At this point the two cases must be treated separately. Replacing the moments of equation (B2) in (A5), and defining \( \omega = 2\pi j / N \) as the limit. In the special case of a constant signal \( \eta \), equation (B4) becomes

\[
\left( \frac{N_{ph}}{2} \right)^2 \text{Var}(P_j) = \sum_k \left( 6\eta_k^2 + 7\eta_k^2 + \eta \right) + \sum_{k,l,m} \eta_k \eta_l \eta_m \left[ \eta_k \eta_l \eta_m \left( 4e^{i\omega(k-l-m)} + e^{i\omega(2k-l-m)} + e^{-i\omega(2k-l-m)} \right) \right] - 2 \sum_k \eta_k \eta_l \eta_m e^{i\omega(j-l-m)}. \] (B3)

Similarly to what was done in Appendix A, adding and subtracting the excluded terms in the sums, after a few passages one ends up with

\[
\text{Var}(P_j) = 4 \left( 1 + \frac{1}{N_{ph}} \right) + \frac{8}{N_{ph}} \sum_{k,l} \eta_k \eta_l e^{2\pi i(k-l)/N} + \frac{4}{N_{ph}^2} \left[ \sum_{k,l} \eta_k \eta_l e^{4\pi i(k-l)/N} + \sum_{k,l,m} \eta_k \eta_l \eta_m \left( e^{i2\pi(2k-l-m)/N} + e^{-i2\pi(2k-l-m)/N} \right) \right]. \] (B4)

As done for the Gaussian case, we have to treat the Nyquist frequency separately. When \( j = N/2 \), equation (B4) becomes

\[
\text{Var}(P_{N/2}) = 4 \left( 2 + \frac{1}{N_{ph}} \right) + \frac{8}{N_{ph}} \sum_{k,l} \eta_k \eta_l e^{2\pi i(k-l)/N} \] (B5)

which is equivalent to equation (A8) in the \( N_{ph} \gg 1 \) limit. In the special case of a constant signal \( \eta_k = \eta (\forall k) \), equation (B4) reduces to

\[
\text{Var}(P_j) = \begin{cases} 
4 \left( 1 + \frac{1}{N_{ph}} \right) & (j = 1, \ldots, \frac{N}{2} - 1) \\
4 \left( 2 + \frac{1}{N_{ph}} \right) & (j = \frac{N}{2}) 
\end{cases} \] (B6)

in agreement with the results of L83. In the more general case of a non-stationary signal \( \eta_k \), in the limit \( N_{ph} \gg 1 \), equation (B4) can be approximated by

\[
\text{Var}(P_j) \approx 4 \left( 1 + \frac{2}{N_{ph}} \sum_{k,l} \eta_k \eta_l e^{2\pi i(k-l)/N} \right) + \frac{4}{N_{ph}^2} \left[ \sum_{k,l} \eta_k \eta_l e^{4\pi i(k-l)/N} + \sum_{k,l,m} \eta_k \eta_l \eta_m \left( e^{i2\pi(2k-l-m)/N} + e^{-i2\pi(2k-l-m)/N} \right) \right]. \] (B7)

Not surprisingly, equation (B7) is the same as (A7) upon replacing \( \sigma^2 \) with \( \eta_k \), as expected for a Poisson variable in the Gaussian limit. As discussed in Appendix A, the result in equation (B7) can be approximated by equation (A9), provided that \( N_{ph} \gg 1 \). When the Gaussian limit is not satisfied, i.e. when \( N_{ph} \) is just a few, we note that the relation between expected value and variance for a non-central chi-square distributed random variable with \( r = 2 \) degrees of freedom \( (j \neq N/2) \) is not fulfilled:

\[
E(P_j) = 2 + \frac{2}{N_{ph}} \sum_{k,l} \eta_k \eta_l e^{2\pi i(k-l)/N} = r + \lambda \] (B8)

\[
\text{Var}(P_j) = 2 \left( 2 + \frac{4}{N_{ph}} \sum_{k,l} \eta_k \eta_l e^{2\pi i(k-l)/N} \right) + \frac{4}{N_{ph}} \left( 1 + \frac{4}{N_{ph}} \sum_{k,l} \eta_k \eta_l e^{2\pi i(k-l)/N} \right) \]

\[
= \left( r + 2\lambda \right) + \frac{2}{N_{ph}} \left( r + 4\lambda \right) + \frac{4}{N_{ph}^2} \] \( \ldots \) \( \neq 2 \left( r + 2\lambda \right). \] (B9)

We conclude that for \( N_{ph} \sim a \) few, the distribution of the power spectrum at a given frequency is not a non-central \( \chi^2(\lambda) \) as found in the Gaussian limit, and the expression for the variance to be used is given by equation (B9). In the same regime of \( N_{ph} \sim a \) few, \( P_{N/2} \) also deviates from a non-central \( \chi^2(\lambda) \) distribution, given that equation (B5) does not fulfil any more the corresponding relation between expected value and variance. The exact formula for the variance of the Nyquist frequency power therefore remains equation (B5).
**APPENDIX C: MAXIMUM LIKELIHOOD ESTIMATION**

The $j < N/2$ and $j = N/2$ cases must be treated separately, given that the PDFs of the corresponding random variables $P_j$ are non-central chi squares with $r = 2$ and $r = 1$ degrees of freedom, respectively. The purpose is to find the value for the non-central parameter $\lambda$ which maximizes the likelihood function $\chi^2_r(\lambda, P_j)$ for a given measured value $P_j$.

First, let us consider the $j < N/2$ case, for which it is $r = 2$. It is

$$\chi^2_2(\lambda, P_j) = \frac{1}{2\pi} e^{-(P_j+\lambda)/2} \int_0^{\pi} e^{\sqrt{\lambda P_j} \cos \theta} d\theta.$$  \hfill (C1)

When $P_j = 0$ equation (C1) reduces to a simple exponential, so that $\lambda = 0$. For $P_j \leq 2$, it is $\lambda = 0$. For $P_j > 2$, $\lambda$ is found by requiring

$$\frac{\partial \chi^2_2(\lambda, P_j)}{\partial \lambda} \bigg|_{\lambda=0} = 0$$ \hfill (C2)

equivalent to

$$\sqrt{\lambda} P_j \frac{I_1(\sqrt{\lambda} P_j)}{I_0(\sqrt{\lambda} P_j)}$$ \hfill (C3)

where $I_n$ is the modified Bessel function of the first kind and index $n$. At $P_j \gg 1$ the solution is $\lambda = P_j$. Numerical solutions to equation (C3) are displayed in Fig. C1 (solid line), which shows how rapidly $\lambda$ converges to $P_j$ as a function of $P_j$.

In the $j = N/2$ case it is $r = 1$ and the interested random variable is $P_{N/2}/2$. To simplify the notation, let us define $x = P_{N/2}/2$, so it is

$$\chi^2_1(\lambda, x) = \frac{1}{\sqrt{2\pi x}} e^{-(x+\lambda)/2} \coth(\sqrt{x}\lambda).$$ \hfill (C4)

When $0 \leq x \leq 1$ equation (C4) monotonically decreases for $\lambda > 0$, so it is $\lambda = 0$. When $x > 1$, the maximum is found analogously to equation (C2), thus

$$\frac{\partial \chi^2_1(\lambda, x)}{\partial \lambda} \bigg|_{\lambda=x} = 0$$ \hfill (C5)

equivalent to

$$e^{x\sqrt{x}} = \frac{1 + \sqrt{x}}{1 - \sqrt{x}}.$$ \hfill (C6)

Analogously to the $j < N/2$ case, the solution is $\lambda \lesssim x$ and rapidly converges to $\lambda = x$, as shown in Fig. C1 (dashed line).

Summing up, in both cases it is $\lambda = 0$ for $P_j \leq 2$, and for $P_j > 2$ it asymptotically tends to $P_j (P_{N/2}/2)$ for $j < N/2$ ($j = N/2$).

In the process of estimating the variance of $P_j$, we conservatively assume $\lambda = P_j (j < N/2)$, and $\lambda = P_{N/2}/2 (j = N/2)$, so

$$\lambda = \begin{cases} P_j & (j < N/2) \\ P_{N/2}/2 & (j = N/2) \end{cases}.$$ \hfill (C7)

By replacing equation (C7) into equations (14) and (15) equation (18) is obtained.