

Estimation of error on the cross-correlation, phase and time lag between evenly sampled light curves

Ranjeev Misra^a, Archana Bora^b, Gulab Dewangan^c

^a*IUCAA, Post Bag 4, Ganeshkhind, Pune-411007, India*

^b*Dept. of Applied Sciences, Gauhati University Institute of Science and Technology, Guwahati 781014, India*

^c*IUCAA, Post Bag 4, Ganeshkhind, Pune-411007, India*

Abstract

Temporal analysis of radiation from Astrophysical sources like Active Galactic Nuclei, X-ray Binaries and Gamma-ray bursts provide information on the geometry and sizes of the emitting regions. Establishing that two light-curves in different energy bands are correlated and measuring the phase and time-lag between them is an important and frequently used temporal diagnostic. In this work, we have presented expressions for estimate of the errors on the cross-correlation, phase and time-lag between two light-curves and the same have been tested using simulations. Earlier estimates depended upon numerically expensive simulations or on dividing the light-curves in large number of segments to find the variance. The estimates presented here allow for analysis of light-curves with relatively small (~ 1000) number of points, as well as to obtain information on the longest time-scales available. For testing the analytical expressions light-curves have been simulated from both white and $1/f$ stochastic processes with measurement errors. As a demonstration, we also apply this technique to the *XMM-Newton* light-curves of the Active Galactic Nucleus, Akn 564.

1. Introduction

Establishing that two light-curves, measured in different energy bands, are correlated with each other is an important temporal diagnostic for various kinds of Astrophysical sources, especially for Active Galactic Nuclei (AGN) and X-ray binaries. The detection and measurement of the level of correlation constrains the radiative processes active in the source and can be used to validate (or rule out) models based on spectral analysis. Phase and time-lags detected for correlated light-curves can provide further insight into the geometry and size of the emitting region. Often in these applications, the light curves available for analysis are of short duration and have measurement errors. The true temporal behaviour of a source can only be established if there are robust estimates of the errors on the cross-correlation, phase and time-lags.

It is important to emphasize that a cross-correlation analysis between two finite length light-curves will not provide an accurate measure of the correlation between them, even in the absence of measurement errors. Intrinsic stochastic fluctuations in the light curves will induce an error on the cross-correlation measured. An estimate of the significance and error of the cross-correlation detected, should take into account both, measurement errors as well as statistical fluctuations.

A standard method to estimate the error on the cross-correlation involves dividing the light curves into several equal segments and finding the cross-correlation for each. Then the net cross-correlation is given by the average of the different segments and

the variance is quoted as an error. For example, this technique is implemented by the function “crosscor” of the high energy astrophysics software *HEASOFT*¹. The method is reliable only if the light curves can be divided into a large number of segments ($\gg 10$) and each segment is sufficiently long and not dominated by measurement errors. The temporal behaviour of many astrophysical systems depends on the time-scales of the analysis and hence by using this method, one loses information on the behaviour of the system on time-scales comparable to the length of the original data. In AGN, the time scale involved is long comparable to the length of observation in many cases, hence it is not practical to divide the light curve in segments. Moreover, there does not seem to be any established way by which this method can be extended to get an estimate of the time-lag between the light curves and its error.

These deficiencies can be overcome by using a Monte Carlo technique where one simulates a large number of pairs of light curves having the same *assumed* temporal properties and with the same measurement errors as the original pair. The results of the original pair can be compared with the simulated ones to ascertain the confidence level of the cross-correlation and time-lag. The simulated light curves should take into account the stochastic fluctuations of the light curves and not just the measurements errors. Indeed, when the light curve is sampled unevenly and with measurement errors changing in time, the Monte Carlo technique may be the only way to obtain reliable estimates (for e.g. Peterson et al., 1998). Monte Carlo technique is numerically expensive and hence are not practical for analysis of a large sets of data. More importantly, the results depend on the subjectivity of the assumed temporal properties of

Email addresses: rmisra@iucaa.in (Ranjeev Misra),
 abora.80@gmail.com (Archana Bora), gulabd@iucaa.in (Gulab
 Dewangan)

¹<http://heasarc.gsfc.nasa.gov/docs/software/lheasoft/>

the system. For example, to ascertain the errors on an observed cross-correlation and time-lag value, the simulations are generally done with the assumption that these are the true intrinsic values. Similar assumptions have to be made on the shape of the power spectra of the light curves.

As pointed out and discussed extensively by Welsh (1999), an analytical estimate of the variance on cross-correlation is not straight forward. In the literature, there is an analytical estimate for the cross-correlation known as Bartlett's equation Bartlett (1955) which is not often used in Astronomical contexts. This method is available in the "crosscorrelation" function in the IMSL numerical libraries². The error is accurate only when the complete knowledge of the cross-correlation and auto-correlation functions are available. Its effectiveness for short duration light curves is uncertain. Moreover, this error estimate does not naturally translate into error estimates for the phase and time lag between the light curves.

Complete information regarding the temporal relation between two light curves can be obtained by computing the coherence and time-lag as a function of Fourier frequency. A detailed description of the technique as well as physical interpretation is given by Nowak et al. (1999). The two light curves are divided into many segments and for each segment a Fourier transform is undertaken and coherence and phase lag as a function of frequency is estimated. For the different segments, the coherence and phase lags are averaged and their errors can be estimated analytically. Such detailed information can only be obtained for long light curves which can be split into several segments. In the absence of such rich data, statistically significant results can be obtained by averaging over Fourier frequencies. Indeed, from this view point the cross-correlation, is in some sense, the average of the coherence over all frequencies. However, computing the error on the cross-correlation using the error estimates for the coherence is not straight forward. First, the averaging has to be appropriately weighted by the power in each frequency bin. Secondly, the error estimate for the coherence is reliable only if the error itself is small, which is the case when many segments are averaged and not necessarily true for the coherence at a single frequency bin obtained from a single segment.

In this work, we present an expression for the cross-correlation between two evenly sampled light curves. The error estimate is based on the Fourier transforms of the light curves then averaging over different frequency modes. §2 presents the expression and the same has been verified by simulations with and without measurement errors. §3 highlights the difficulties in estimating a time-lag and its error using the standard method of finding the peak of the cross-correlation function. The cross-Correlation phasor is introduced in §4 which leads to an estimate of the phase lag between the light curves. In the same section, a technique is introduced by which one can measure the time lag and its error. In this method the time-lag measured can be even smaller than the sampling time bin of the light curves. The complete fully self contained algorithm is presented in §5 for

easy reference. As an example, in §6, the technique is applied to the *XMM-Newton* light curves of the highly variable and well studied AGN, Akn 564. In §7, the summary and discussion includes a list of important assumptions on which the technique is based and provides examples when the assumptions may not be valid.

2. ANALYTICAL ERROR ESTIMATE OF CROSS-CORRELATION

2.1. Light curves without measurement errors

We first consider an idealised case of two light curves, X and Y , without measurement errors. The two light curves are of length N , which have recorded the count rates in $j = 0, 1, 2, \dots, N-1$ discrete equally spaced time intervals, Δt . The mean is subtracted from each of them. Further it is assumed that they are partially linearly dependent on each other by A , such that we have,

$$X = x_j \quad (1)$$

$$Y = z_j + Ax_j \quad (2)$$

where x_j and z_j are time-series produced by two independent stochastic processes. Each time series can be conveniently represented, in frequency domain k by its discrete Fourier transform, \tilde{X}_k , defined as

$$\tilde{X}_k = \sum_{j=0}^{N-1} X_j \exp(2\pi i jk/N) \quad (3)$$

and a power spectrum is estimated as $P_{Xk} \equiv (2/N)|\tilde{X}_k|^2$. Here the normalisation constant of the power spectrum is used as suggested by Leahy et al Leahy et al. (1983). For a stationary system, the ensemble average (i.e. average of an infinite number of realisations) of the power, $\langle P_{Xk} \rangle$, is a characteristic of the stochastic process. A power derived from a single time series, P_{Xk} is only an estimator of its value. In particular the real and imaginary parts of \tilde{X}_k varying independently can be derived from two independent Gaussian distributions Timmer and Koenig (1995). The standard deviation of P_{Xk} from $\langle P_{Xk} \rangle$ is roughly equal to $\langle P_{Xk} \rangle$ i.e the power estimate from a single light curve has nearly 100% sampling variation. The variance $\sigma_X^2 \equiv \sum P_{Xk}$ is again an estimate of the ensemble averaged variance $\langle \sigma_X^2 \rangle = \sum \langle P_{Xk} \rangle$ Van der Klis (1989) where $k = -N/2, \dots, N/2 - 1$ and $k \neq 0$.

One can define the cross-correlation estimate of the two time series as

$$C_{XY} = \frac{c_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} \quad (4)$$

where

$$c_{XY} = \frac{1}{N} \sum_{j=0}^{N-1} X_j Y_j = \frac{1}{N^2} \sum_{k=-N/2}^{N/2-1} \tilde{X}_k \tilde{Y}_k^* \quad (5)$$

²<http://www.vni.com>

Here \tilde{X}_k and \tilde{Y}_k are Discrete Fourier transforms of X_j and Y_j respectively and

$$\begin{aligned}\sigma_X^2 &= \frac{1}{N} \sum_{j=0}^{N-1} X_j^2 = \frac{1}{N^2} \sum_{k=-N/2}^{N/2-1} |\tilde{X}_k|^2 \\ \sigma_Y^2 &= \frac{1}{N} \sum_{j=0}^{N-1} Y_j^2 = \frac{1}{N^2} \sum_{k=-N/2}^{N/2-1} |\tilde{Y}_k|^2\end{aligned}\quad (6)$$

Their ensemble averages are $\langle c_{XY} \rangle = A < \sigma_x^2 \rangle$, $\langle \sigma_x^2 \rangle = \langle \sigma_x^2 \rangle$ and $\langle \sigma_y^2 \rangle = \langle \sigma_z^2 \rangle + A^2 \langle \sigma_x^2 \rangle$. It is to be noted that a normalisation factor of $\frac{1}{N}$, due to the DFT process, has been incorporated in frequency domain term of Eqn (6) C_{XY} has the useful property that its ensemble average

$$\langle C_{XY} \rangle = \frac{A \langle \sigma_x^2 \rangle}{\sqrt{\langle \sigma_x^2 \rangle (\langle \sigma_z^2 \rangle + A^2 \langle \sigma_x^2 \rangle)}} \quad (7)$$

is zero if the two light series are uncorrelated (i.e. $A = 0$) and ± 1 if they are completely correlated (i.e. when $\langle \sigma_z^2 \rangle = 0$). However, in absence of any a priori information about the stochastic process, quantities need to be estimated using the measured values only. Thus, C_{XY} is also an estimate of $\langle C_{XY} \rangle$ and its accuracy needs to be quantified.

To ascertain whether there is a detectable correlation between the two light-curves (i.e. $|C_{XY}| > 0$) it is first necessary to show that, at some confidence level, $|c_{XY}| > 0$. One can define a null hypothesis sigma level $\sigma_{NH} = |C_{XY}|/\Delta C'_{XY}$ and fix a criterion (a prudent one being $\sigma_{NH} > 3$) to ascertain whether any correlation has been detected. It is important to note that only if the criterion is satisfied should one proceed to estimate the degree of cross-correlation C_{XY} otherwise any such attempt will not only be incorrect but also meaningless.

If c_{XY} is uncorrelated with $\sqrt{\sigma_x^2 \sigma_y^2}$, then

$$\Delta C_{XY} = \frac{\Delta c_{XY}}{\sqrt{\sigma_x^2 \sigma_y^2}} \quad (8)$$

where the variation in $\sqrt{\sigma_x^2 \sigma_y^2}$ has been neglected. However, as discussed extensively by Welsh (1999), $\sqrt{\sigma_x^2 \sigma_y^2}$ is, in general, correlated with c_{XY} . In particular, Δc_{XY} depends on the variation of σ_x^2 through the term $A\sigma_x^2$.

A possible solution is to define a transformation, $P(C_{XY})$ whose terms are not correlated (or at least not so correlated). Then estimate the expected variation for that function, ΔP and use that to obtain an estimate for ΔC_{XY} . Below we describe such a transformation and subsequently test the results obtained from simulations. The transformation chosen for the analysis is

$$P = \frac{c_{XY}^2}{\sigma_x^2 \sigma_y^2 - c_{XY}^2} = \frac{C_{XY}^2}{1 - C_{XY}^2} \quad (9)$$

where the subtraction of c_{XY}^2 in the denominator may make it nearly independent of the numerator. The average deviation of P can be estimated to be,

$$\Delta P = \frac{2 \langle c_{XY} \rangle \Delta c_{XY}}{\langle \sigma_x^2 \sigma_y^2 - c_{XY}^2 \rangle} \quad (10)$$

where the variation of the denominator has been neglected. ΔP is related to ΔC_{XY} by

$$\Delta P \sim \left(\frac{dP}{dC_{XY}} \right) \Delta C_{XY} = \left(\frac{2C_{XY}}{(1 - C_{XY}^2)^2} \right) \Delta C_{XY} \quad (11)$$

giving us

$$\Delta C_{XY} = \frac{1 - \langle C_{XY} \rangle^2}{\sqrt{\langle \sigma_x^2 \rangle \langle \sigma_y^2 \rangle}} \Delta c_{XY} \quad (12)$$

Naturally, ΔC_{XY} depends on ensemble averaged quantities which characterise the stochastic processes that have produced the light curves. Typically, one does not have *a priori* information of the stochastic processes and the ensemble averaged quantities need to be estimated from the light curves. Thus the best estimate, $\Delta C'_{XY}$, of ΔC_{XY} can be obtained by replacing these ensemble averaged quantities with the measured ones. Also, we have $(\Delta c_{XY})^2 = \langle c_{XY}^2 \rangle - \langle c_{XY} \rangle^2 = \frac{1}{N^4} \sum_{k=-N/2}^{N/2-1} \langle |\tilde{X}_k|^2 \rangle \langle |\tilde{Y}_k|^2 \rangle$

Hence

$$\Delta C'_{XY} = \frac{1 - C_{XY}^2}{N^2 \sqrt{\sigma_x^2 \sigma_y^2}} \sqrt{\sum_{k=-N/2}^{N/2-1} |\tilde{X}_k|^2 |\tilde{Y}_k|^2} \quad (13)$$

For practical situations $\Delta C'_{XY}$ can be used as an estimate for the error on C_{XY} .

2.2. Comparison with results from simulations

We generated 200 independent light-curves using the method described by Timmer and Koenig (1995). Vaughan et al. (2003a) discuss the different methods to generate stochastic light curves and give arguments for favouring the one prescribed by Timmer and Koenig (1995). The intrinsic power spectrum of the stochastic process was assumed to be a power-law i.e. $P(f) \propto f^{-\alpha}$. The light-curves were generated of length $8N$ and rebinned to a length of N , to avoid aliasing effects. From these 200 light curves, 19900 pairs of the light curves were generated which obey,

$$X_j = x_j \quad (14)$$

$$Y_j = z_j + Ax_j \quad (15)$$

where x_j and z_j are two different simulated light-curves with $\langle \sigma_x^2 \rangle = \langle \sigma_z^2 \rangle = 1$. The cross-correlation, C_{XY} was computed for each pair. For $N = 1024$, $\alpha = 0$ and for three different values of $A = 0, 1$ and 5 , the histograms of the C_{XY} are plotted in the top panel of Figure 1. These histograms, H_j are normalised such that their summation $\sum H_j \delta = 1$ where δ is the bin size. They are compared with a normalised Gaussian distribution with a centroid value equal to the expected averaged cross-correlation of

$$\langle C_{XY} \rangle = \frac{A}{\sqrt{1 + A^2}} \quad (16)$$

and with width σ equal to ΔC_{XY} computed using Eqn (12). As can be seen, the normalised Gaussian distributions describe well the simulated results which validates the method and assumptions used to estimate ΔC_{XY} in the previous subsection.

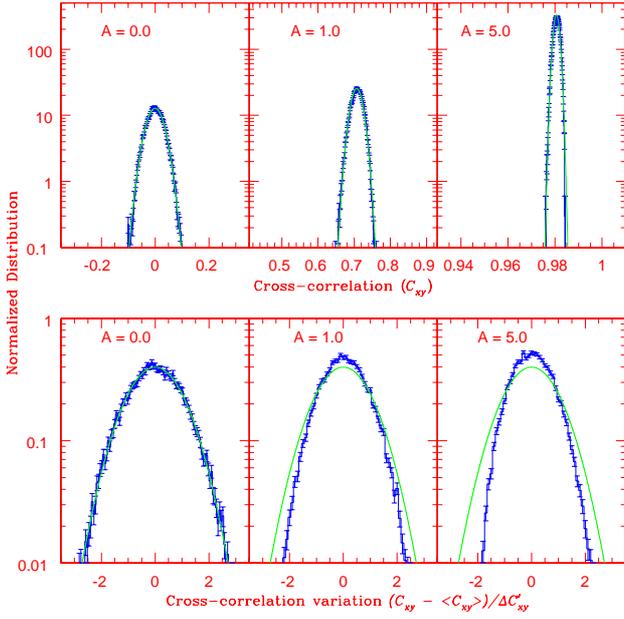


Figure 1: Comparison of simulation with analytical results. 19900 pairs of light-curves of length $N = 1024$ were created for $X_j = x_j$ and $Y_j = z_j + Ax_j$. x_j and z_j are independent time-series generated from a stochastic white noise process (i.e. power spectrum index $\alpha = 0$). Top Panel compares the normalised histogram of C_{XY} with a Gaussian with centroid at the expected $\langle C_{XY} \rangle$ and width given by $\sigma = \Delta C_{XY}$ (Eqn 12). Bottom panel shows the histogram of the cross-correlation variation $(C_{XY} - \langle C_{XY} \rangle)/\Delta C'_{XY}$. If $\Delta C'_{XY}$ (Eqn 13), which is estimated from the pair of light-curves, is a true measure of the variation, then the distribution should be a zero centred Gaussian with $\sigma = 1$ (solid line).

However, in practical situations one has to use the approximation $\Delta C'_{XY}$ to estimate the variance which in general will vary for each pair of light curves. In the bottom panel of Figure 1, we plot histograms of deviation of C_{XY} from the average $\langle C_{XY} \rangle$ normalised by the estimated deviation $\Delta C'_{XY}$ i.e. $(C_{XY} - \langle C_{XY} \rangle)/\Delta C'_{XY}$. If $\Delta C'_{XY}$ is an accurate measure of the variation of C_{XY} then the distribution of the normalised variation should be a zero centred Gaussian with $\sigma = 1$. The plot verifies this prediction by comparing the distribution with such a Gaussian shape. The distributions agree well with each other except for large A , where the Gaussian distribution is slightly broader than the simulation results. This implies that the $\Delta C'_{XY}$ is a slight overestimation of the true deviation for large values of A .

Figure 2 shows the same comparison as Figure 1, but for the case when the power-law index $\alpha = 1$. Qualitatively the comparison between the expected and obtained distribution are similar to the $\alpha = 0$ case, except for some quantitative differences. For the same length of the light curves, ΔC_{XY} is larger for $\alpha = 1$. Since C_{XY} is by definition constrained to be less than unity, the distribution differs from the symmetric Gaussian shape for large A . The bottom panel shows that $\Delta C'_{XY}$ is a better representation of the variation than it was for $\alpha = 0$.

For white noise (i.e. $\alpha = 0$), the dependence of ΔC_{XY} on the length of the light-curves is $\propto 1/\sqrt{N}$, while for $\alpha = 1$ the

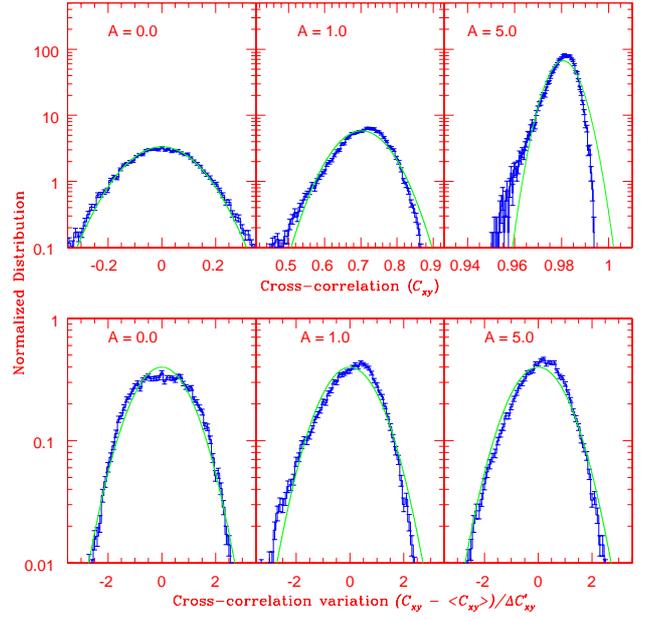


Figure 2: Same as Figure 1, except that x_j and z_j are independent time-series generated from a stochastic $1/f$ noise process (i.e. power spectrum index $\alpha = 1$). The width of the distributions are broader than for white noise.

dependence is weaker $\sim 1/\log N$ for large N Chatfield (2016); Keshner (1982). The original light-curves may be divided into M parts, and cross-correlations of each may be averaged. For $\alpha = 0$ this will not lead to any change in the accuracy with the final ΔC_{XY} being nearly the same. However, for $\alpha = 1$, $\Delta C_{XY} \propto 1/(\sqrt{M} \log(N/M))$ which would give a much better accuracy than finding the cross-correlation for the whole light-curve. However, such a cross-correlation will not have information about the behaviour of the system on timescales corresponding to duration of the original light curve.

2.3. Light curves with measurement errors

We next consider a more realistic case, where the light-curves have measurement errors. In particular,

$$\begin{aligned} X_j &= x_j + e_{Xj} \\ Y_j &= z_j + Ax_j + e_{Yj} \end{aligned} \quad (17)$$

where x_j and z_j are time-series produced by two independent stochastic processes as before and e_{Xj} and e_{Yj} are the known measurement errors for measuring X_j and Y_j respectively. The cross-correlation is now defined as

$$C_{XY} = \frac{c_{XY}}{\sqrt{(\sigma_X^2 - \sigma_{XE}^2)(\sigma_Y^2 - \sigma_{YE}^2)}} \quad (18)$$

where c_{XY} is the same as before (Eqn 5) and σ_{XE} and σ_{YE} are the rms variation of the measured errors i.e.

$$\sigma_{XE}^2 = \frac{1}{N} \sum_{j=0}^{N-1} e_{Xj}^2 \quad (19)$$

and similarly for σ_{YE}^2 .

The expressions for $\langle C_{XY} \rangle$ and ΔC_{XY} remain the same as for the measurement error free case discussed previously and following the same procedure as before, one can estimate

$$\Delta C_{XY} = \frac{(1 - \langle C_{XY} \rangle^2) \Delta C_{XY}}{\sqrt{(\langle \sigma_X^2 \rangle - \langle \sigma_{XE}^2 \rangle)(\langle \sigma_Y^2 \rangle - \langle \sigma_{YE}^2 \rangle)}} \quad (20)$$

analogous to Eqn (12). To this error estimate we have to add the fluctuations of σ_{XE} and σ_{YE} around their ensemble averaged values $\langle \sigma_{XE} \rangle$ and $\langle \sigma_{YE} \rangle$. Note that it is these ensemble averaged values $\langle \sigma_{XE} \rangle$ and $\langle \sigma_{YE} \rangle$ that are known *a priori* and not σ_{XE} and σ_{YE} . If the measurement errors are Gaussian white noise (as is generally the case) then, $\Delta \sigma_{X,YE}^2 = (1/\sqrt{N})\sigma_{X,YE}^2$. Moreover since the fluctuations are independent of the true signal, they can be added to ΔC_{XY} using standard error propagation techniques. Thus

$$\begin{aligned} \left(\frac{\Delta C_{XY}}{\langle C_{XY} \rangle}\right)^2 &= \left(\frac{\Delta C_{XY1}}{\langle C_{XY} \rangle}\right)^2 + \left(\frac{\Delta \sigma_{XE}^2}{\sqrt{2}(\langle \sigma_X^2 \rangle - \langle \sigma_{XE}^2 \rangle)}\right)^2 \\ &\quad + \left(\frac{\Delta \sigma_{YE}^2}{\sqrt{2}(\langle \sigma_Y^2 \rangle - \langle \sigma_{YE}^2 \rangle)}\right)^2 \\ &= \left(\frac{\Delta C_{XY1}}{\langle C_{XY} \rangle}\right)^2 + \left(\frac{\langle \sigma_{XE}^2 \rangle / \sqrt{2N}}{\langle \sigma_X^2 \rangle - \langle \sigma_{XE}^2 \rangle}\right)^2 \\ &\quad + \left(\frac{\langle \sigma_{YE}^2 \rangle / \sqrt{2N}}{\langle \sigma_Y^2 \rangle - \langle \sigma_{YE}^2 \rangle}\right)^2 \end{aligned}$$

ΔC_{XY} is in terms of ensemble averaged quantities which have to be estimated using the light curves. Hence

$$\begin{aligned} \Delta C'_{XY} &= \frac{(1 - C_{XY}^2)}{N^2 \sqrt{(\sigma_X^2 - \langle \sigma_{XE}^2 \rangle)(\sigma_Y^2 - \langle \sigma_{YE}^2 \rangle)}} \\ &\quad \times \sqrt{\sum_{k=-N/2}^{N/2-1} |\tilde{X}_k|^2 |\tilde{Y}_k|^2} \quad (22) \end{aligned}$$

and

$$\begin{aligned} \left(\frac{\Delta C'_{XY}}{C_{XY}}\right)^2 &= \left(\frac{\Delta C'_{XY1}}{C_{XY}}\right)^2 + \left(\frac{\langle \sigma_{XE}^2 \rangle / \sqrt{2N}}{\sigma_X^2 - \langle \sigma_{XE}^2 \rangle}\right)^2 \\ &\quad + \left(\frac{\langle \sigma_{YE}^2 \rangle / \sqrt{2N}}{\sigma_Y^2 - \langle \sigma_{YE}^2 \rangle}\right)^2 \quad (23) \end{aligned}$$

is the estimation of the variation in the presence of measurement errors.

To validate the above results we generated 19900 pairs of the light-curves which obeyed Eqn (17). The measurement errors were generated from a Gaussian distribution with

$$\langle \sigma_{XE}^2 \rangle / \langle \sigma_X^2 \rangle = \langle \sigma_{YE}^2 \rangle / \langle \sigma_Y^2 \rangle = 0.5$$

. Figures 3 and 4 show the comparison of the distribution with the expected Gaussian distribution for power spectral index $\alpha =$

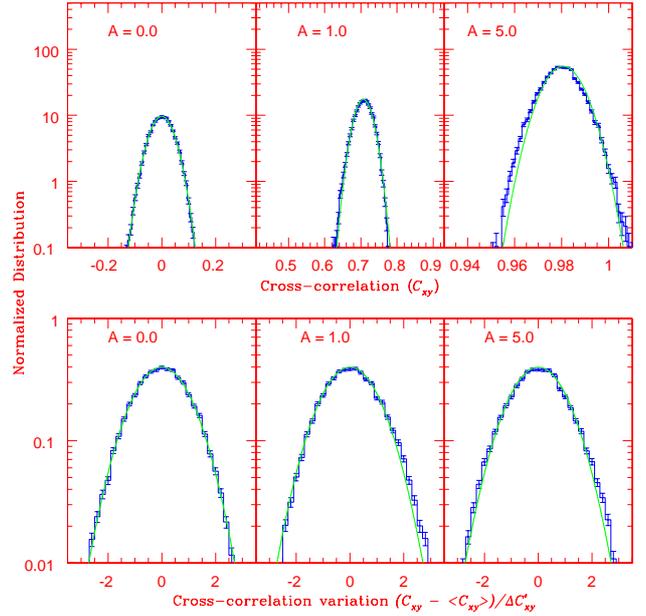


Figure 3: Comparison of simulation with analytical results in the presence of measurement errors. 19900 pairs of light-curves of length $N = 1024$ were created for $X_j = x_j + e_{Xj}$ and $Y_j = z_j + Ax_j + e_{Yj}$. x_j and z_j are independent time-series generated from a stochastic white noise process (i.e. power spectrum index $\alpha = 0$) The measurement errors were simulated from a Gaussian distribution such that $\langle \sigma_{XE}^2 \rangle / \langle \sigma_X^2 \rangle = \langle \sigma_{YE}^2 \rangle / \langle \sigma_Y^2 \rangle = 0.5$. Top Panel compares normalised histogram C_{XY} with a Gaussian with centroid at the expected $\langle C_{XY} \rangle$ and width given by $\sigma = \Delta C_{xy}$ (Eqn 21). Bottom panel shows the histogram of the cross-correlation variation $(C_{XY} - \langle C_{XY} \rangle) / \Delta C'_{XY}$. If $\Delta C'_{XY}$ (Eqn 22), which is estimated from the pair of light-curves, is a true measure of the variation, then the distribution should be a zero centred Gaussian with $\sigma = 1$ (solid line). Note that in the presence of measurement errors $\langle C_{XY} \rangle$ can be greater than one.

0 and 1 respectively. As expected the distribution of C_{XY} is broader in the presence of measurement errors. Note that in this case C_{XY} can be greater than unity. The figures illustrate that the variance estimations reasonably describe the simulated distributions.

The above results are for the case when the measurement errors are Gaussian distributions. We have verified that even when the mean counts per time bin is ~ 10 , similar results are obtained when the measurement errors are due to Poisson fluctuations. In order to correctly propagate the error and obtain Eqn (21), it is implicitly assumed that $\Delta \sigma_{X,YE}^2 = (2/\sqrt{N})\sigma_{X,YE}^2 \ll \sigma_{X,Y}^2 - \sigma_{X,YE}^2$. Note that these are also the criteria that any significant variability has been detected in each of the two light curves. In other words if the criterion is not satisfied for one of the light curves, this implies that there is no significantly excess variance than expected from the measurement errors and hence a cross-correlation analysis cannot be undertaken.

3. The cross-correlation function

In general two light-curves may be linearly related to each other with a time lag τ . To investigate such possibilities, one

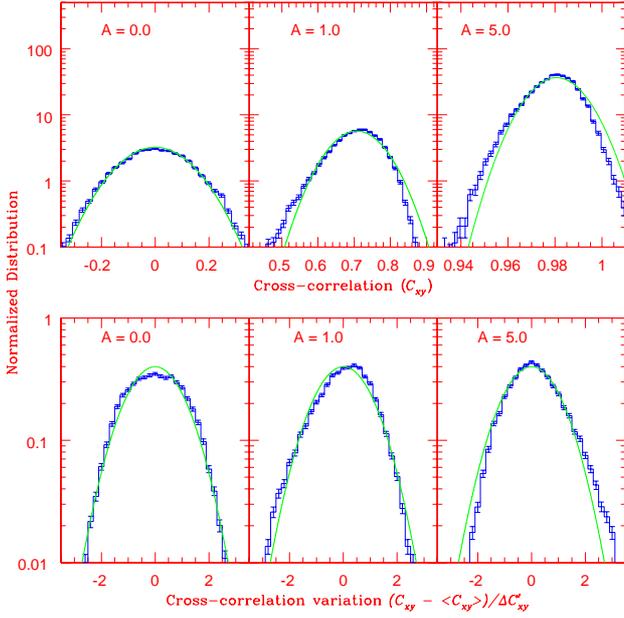


Figure 4: Same as Figure 3, except that x_j and z_j are independent time-series generated from a stochastic $1/f$ noise process (i.e. power spectrum index $\alpha = 1$).

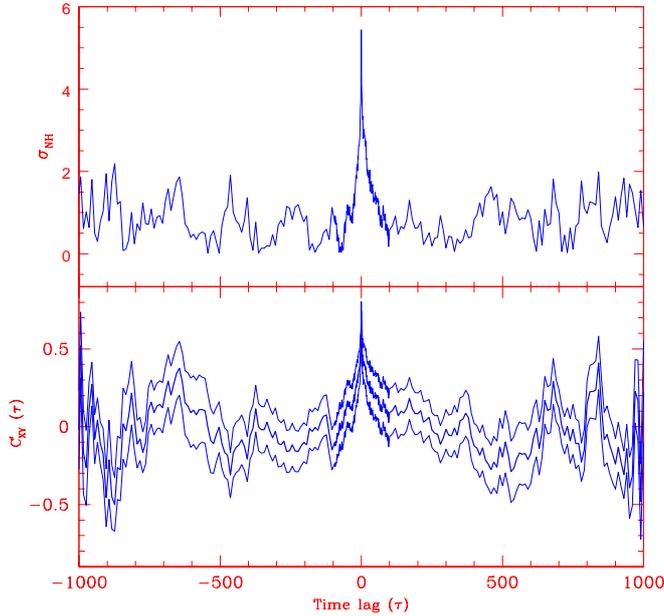


Figure 5: Significance and cross-correlation function for two simulated light-curves with measurement errors. Two pairs of light-curves of length $N = 1024$ were created for $X_j = x_j + e_{Xj}$ and $Y_j = z_j + Ax_j + e_{Yj}$. x_j and z_j are independent time-series generated from a stochastic $1/f$ noise process (i.e. power spectrum index $\alpha = 1$). The measurement errors were simulated from a Gaussian distribution such that $\langle \sigma_{XE}^2 \rangle / \langle \sigma_x^2 \rangle = \langle \sigma_{YE}^2 \rangle / \langle \sigma_z^2 \rangle = 0.5$. The top panel shows the significance $\sigma_{NH} = |C_{XY}| / \Delta C_{XY}$ as a function of τ . Note that $\sigma_{NH} < 3$ for all τ except when $|\tau| \sim 0$. The bottom panel shows the cross-correlation function, $C_{XY}(\tau)$ (Thick line), $C_{XY}(\tau) + \Delta C_{XY}$ and $C_{XY}(\tau) - \Delta C_{XY}$ (Thin lines).

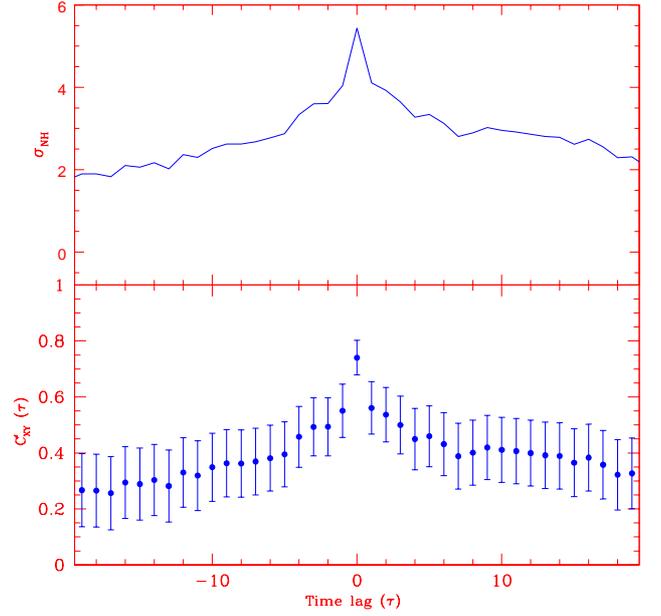


Figure 6: The blown up portion of Fig 5 near $\tau = 0$ shown for clarity.

can calculate the cross-correlation function, $C_{XY}(\tau)$ between a light curve X_j and the $Y_{j+\tau}$. For light-curves of length N , there will be only $N'(\tau) = N - \tau$ overlapping terms. $C_{XY}(\tau)$ can be computed based on these $N'(\tau)$ terms and then the definitions, error analysis of the previous sections follow through without modifications. Such a definition of $C_{XY}(\tau)$ has been called *locally defined cross-correlation function* (LDCCF) by Welsh (1999) and is different from the standard one. In the standard definition the length of the original light-curves N is preserved either by padding the unknown part of the light with zeros (for the time domain computation) or by repeating the series (in the Fourier domain computation). Here we consider only LDCCF for which the analysis mentioned in the earlier section holds.

For every time lag, τ , $\sigma_{NH} = |C_{XY}| / \Delta C_{XY}$ needs to be computed to ascertain whether there is any detectable correlation. In Fig 5, σ_{NH} is plotted against τ for two light curves generated using $1/f$ stochastic process, with measurement errors and with $A = 1$. In fact, the two light-curves are the first pair of light-curves used in the simulation described in Fig 4. Since C_{XY} is being computed for a large number of time lags τ , (although they are not independent, see below), it is prudent to keep a conservative criterion for correlation detection, $\sigma_{NH} > 3$. It can be seen from the figure that $\sigma_{NH} < 3$ for all values of τ except when $\tau \sim 0$. As seen in the bottom panel of Fig 5, there are couple of peaks in $C_{XY}(\tau)$ but none of them (except near $\tau = 0$) are significant.

The $C_{XY}(\tau)$ are not in general independent of each other. This is clearly seen as an example in Fig 6 where the region near $\tau = 0$ has been expanded for clarity. For $|\tau| < 3$, $\sigma_{NH} > 3$ and the dispersion of $C_{XY}(\tau)$ is much less than what is expected

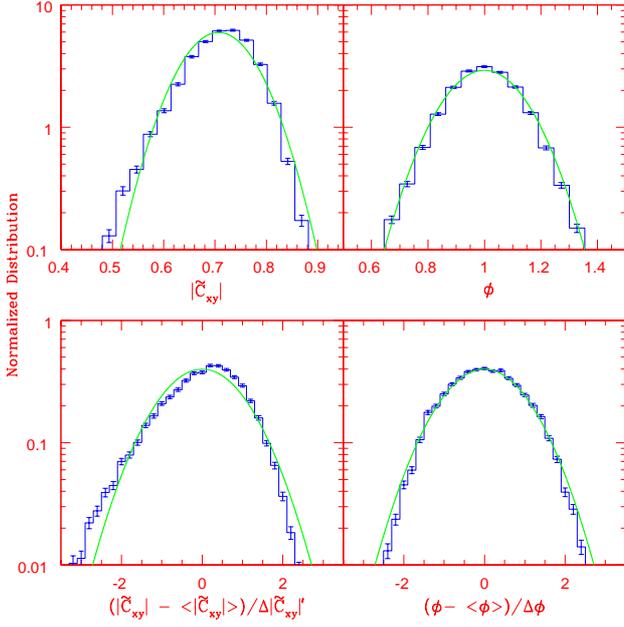


Figure 7: Comparison of simulation with analytical results. 19900 pairs of light-curves of length $N = 1024$ were created for $X_j = x_j$ and $Y_j = z_j + Ae^{i\langle\phi\rangle}x_j$. x_j and z_j are independent time-series generated from a stochastic $1/f$ noise process (i.e. power spectrum index $\alpha = 1$). The values of the parameters used for the simulations are $A = 1$ and $\langle\phi\rangle = 1$. Solid lines are the expected distribution for the error estimates $\Delta|C_{XY}|$ and $\Delta\phi$. $\Delta|C_{XY}|$ is given by Eqn (23) except that C_{XY} is replaced by $|C_{XY}|$ while $\Delta\phi$ is given by Eqn (28)

from the error bars $\Delta C'_{XY}$. This shows that it is unpreferable to rebin $C_{XY}(\tau)$ in τ space. Instead if there is a justification to rebin in time, then the original light-curves should be rebinned and not $C_{XY}(\tau)$. The data clearly shows a peak of $C_{XY}(\tau)$ at $\tau = 0$ indicating a time-lag consistent with zero. However, it is difficult to justify any error measurement on this time lag. In practice, one can represent the peak of the function as a Gaussian and take its width as an representative error for the lag. However, there are several un-attractive features of this technique. First, the width of $C_{XY}(\tau)$ represents the auto-correlation of the coherent signal rather than any error on the measured time-lag. Second, since the errors on $C_{XY}(\tau)$ are correlated a formal fit is not allowed. Finally, the Gaussian fit will in general depend on the number of points used to represent the ‘‘peak’’ of $C_{XY}(\tau)$. A justifiable technique to compute the error on the time-lag is required.

4. The cross-Correlation Phasor

The cross-correlation phasor can be defined as

$$\tilde{C}_{XY} = \frac{\tilde{c}_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} \quad (24)$$

where

$$\tilde{c}_{XY} = \frac{2}{N^2} \sum_{k=1}^{N/2-1} \tilde{X}_k \tilde{Y}_k^* \quad (25)$$

Here the summation of phasor differs from that of the cross-correlation by the fact that for the phasor the summation is only over positive frequencies. They are related as $C_{XY} = \text{Re}(\tilde{C}_{XY})$.

For partially correlated light curves with no phase lag, the ensemble average $\langle \text{Im}(\tilde{C}_{XY}) \rangle = 0$ and the cross-correlation is given by $\langle \text{Re}(\tilde{C}_{XY}) \rangle = \langle C_{XY} \rangle$. By definition, the deviation of $\text{Re}(\tilde{c}_{XY})$ is the same as for c_{XY} whereas the deviation of $\text{Im}(\tilde{c}_{XY})$, is only due to the incoherent parts of the light curves and is given by the equation

$$\Delta \text{Im}(\tilde{c}_{XY}) = \Delta c_{XY} \sqrt{1 - |\tilde{C}_{XY}|^2} \quad (26)$$

If there is an intrinsic phase difference, $\langle\phi\rangle$ between the two light curves, then the ensemble average of \tilde{C}_{XY} will be a complex quantity given by $\langle |\tilde{C}_{XY}| \rangle e^{i\langle\phi\rangle}$. The standard deviation of $|\tilde{C}_{XY}|$ from $\langle |\tilde{C}_{XY}| \rangle$ can be estimated by Eqn (23) except that C_{XY} is to be replaced by $|\tilde{C}_{XY}|$. The phase difference between the two light curves can be estimated as

$$\sin\phi = \frac{\text{Im}(\tilde{c}_{XY})}{|\tilde{c}_{XY}|} \quad (27)$$

whose error, for small values of ϕ , can be estimated to be

$$\Delta\phi = \frac{\Delta c_{XY}}{|\tilde{c}_{XY}|} \sqrt{1 - |\tilde{C}_{XY}|^2} \quad (28)$$

To validate the above results we simulated the same set of light curves used for Figure (4) i.e. using 19900 pairs of light curves with measurement errors and generated from a stochastic $1/f$ noise process. We introduced a phase difference of $\phi = 1.0$ between the coherent parts of the light curves. The histograms of $|\tilde{C}_{XY}|$ and ϕ (and their deviations) are plotted against the expected estimates in Figure 7.

If the coherent parts of the light curves have a time-lag, τ , between them, then the cross-correlation phasor will have a non-zero phase. One can constrain the time-lag by shifting one of the light curves in time till the cross-correlation phase, $\phi = 0$. In other words, by defining a cross-correlation phasor function,

$$\tilde{C}_{XY}(\tau) = \frac{2}{N^2 \sqrt{\sigma_X^2 \sigma_Y^2}} \sum_{k=1}^{N/2-1} \tilde{X}_k \tilde{Y}_k^* e^{ik\tau/N} \quad (29)$$

one can obtain τ such the phase of $\tilde{C}_{XY}(\tau)$, $\phi(\tau) = 0$. The error on τ can be estimated by considering the range of τ for which $\phi(\tau) \pm \Delta\phi$ is consistent with zero. Note that τ need not be an integer and hence time-lags less than the time resolution of the light curves can be ascertained for good quality data.

The above analysis is valid only when there is a detected correlation between the two light curves. To ascertain whether there is a correlation (with phase lag) between the two, one needs to consider both the real and imaginary parts of \tilde{c}_{XY} and compare with Δc_{XY} . While one can compute the joint probabilities, a more prudent and simpler approach is to demand that a

correlation is detected only if $|\tilde{C}_{XY}|/\Delta C_{XY} > 3$. If the condition is not satisfied one can put an upper limit on the correlation as $3\Delta C_{XY}/\sqrt{\sigma_X^2\sigma_Y^2}$.

5. Algorithm to compute cross-correlation, phase and time lags

Based on the results of the earlier sections, we present here a complete self-contained description of the algorithm to compute and estimate errors for cross-correlation, phase and time lag between two light curves. It is assumed that there are two time series X_j and Y_j of length N and for each data there are associated known measurement errors e_{Xj} and e_{Yj} .

Step 1: Calculate variances for the light curves. Compute the Fourier transforms of each light curve using

$$\tilde{X}_k = \sum_{j=0}^{N-1} X_j \exp(2\pi i j k / N) \quad (30)$$

and similarly for \tilde{Y}_k . Compute intrinsic variances,

$$\sigma_{XI}^2 = \frac{2}{N^2} \sum_{k=1}^{N/2-1} |\tilde{X}_k|^2 - \sigma_{XE}^2 \quad (31)$$

where $\sigma_{XE}^2 = \frac{1}{N} \sum_{j=0}^{N-1} e_{Xj}^2$ and similarly for σ_{YI}^2 . If either $\sigma_{XI}^2 < 2\Delta\sigma_{XE}^2 = (4/\sqrt{N})\sigma_{XE}^2$ or $\sigma_{YI}^2 < (4/\sqrt{N})\sigma_{YE}^2$, then no significant variation has been detected in one of the light curves and further analysis is not possible. In such cases, an upper limit of $(4/\sqrt{N})\sigma_{XE}^2$ can be put for any intrinsic variation. If there is significant variation detected, then the total error (both stochastic and measurement) on the variance is

$$\Delta\sigma_{XI}^2 = \sqrt{\frac{2}{N^4} \sum_{k=1}^{N/2-1} (|\tilde{X}_k|^2)^2} \quad (32)$$

of which the measurement uncertainty contributes

$$\Delta\sigma_{XM}^2 = \sqrt{(\Delta\sigma_{XI}^2)^2 - \frac{2}{N^4} \sum_{k=1}^{N/2-1} (|\tilde{X}_k|^2 - |\tilde{X}_M|^2)^2} \quad (33)$$

where $|\tilde{X}_M|^2 = N\sigma_{XE}^2$ and is independent of k .

Step 2: Compute the cross-Correlation. The non-normalised cross-correlation phasor is

$$\tilde{c}_{XY} = \frac{2}{N^2} \sum_{k=1}^{N/2-1} \tilde{X}_k \tilde{Y}_k^* \quad (34)$$

and its error is given by

$$(\Delta c_{XY})^2 = \frac{1}{N^4} \sum_{k=-N/2}^{N/2-1} |\tilde{X}_k|^2 |\tilde{Y}_k|^2 \quad (35)$$

Check if $|\tilde{c}_{XY}|/\Delta c_{XY} > 3$. If not then no correlation is detected between the two light curves and the upper limit on the cross-correlation is $3\Delta c_{XY}/\sqrt{\sigma_{XI}^2\sigma_{YI}^2}$. If the condition is satisfied (i.e. the correlation is detected) then the cross-correlation is

$$|\tilde{C}_{XY}| = \frac{|\tilde{c}_{XY}|}{\sqrt{\sigma_{XI}^2\sigma_{YI}^2}} \quad (36)$$

with error

$$\left(\frac{\Delta|\tilde{C}_{XY}|}{|\tilde{C}_{XY}|}\right)^2 = \left(\frac{\Delta|\tilde{c}_{XY}|}{|\tilde{c}_{XY}|}\right)^2 + \left(\frac{\sigma_{XE}^2}{\sqrt{2N}\sigma_{XI}^2}\right)^2 + \left(\frac{\sigma_{YE}^2}{\sqrt{2N}\sigma_{YI}^2}\right)^2 \quad (37)$$

where

$$\Delta|\tilde{c}_{XY}| = \frac{(1 - C_{XY}^2)\Delta c_{XY}}{\sqrt{\sigma_{XI}^2\sigma_{YI}^2}} \quad (38)$$

Step 3: Compute the phase. The phase is given by

$$\sin\phi = \frac{\text{Im}(\tilde{c}_{XY})}{|\tilde{c}_{XY}|} \quad (39)$$

whose error can be estimated to be

$$\Delta\phi = \frac{\Delta c_{XY}}{|\tilde{c}_{XY}|} \sqrt{1 - |\tilde{C}_{XY}|^2} \quad (40)$$

Step 4: Compute the time delay between the light curves. Define

$$\tilde{C}_{XY}(\tau') = \frac{2}{N^2 \sqrt{\sigma_{XI}^2\sigma_{YI}^2}} \sum_{k=1}^{N/2-1} \tilde{X}_k \tilde{Y}_k^* e^{ik\tau'/N} \quad (41)$$

and solve for $\phi(\tau) = 0$ to get an estimate of the time delay τ . The error on τ , $\Delta\tau$ is to be estimated by considering the range of τ' for which $\phi(\tau') \pm \Delta\phi(\tau')$ is consistent with zero. Compute the significance of the cross-correlation $|\tilde{c}_{XY}|/\Delta c_{XY}$ at the two limits $\tau' = \tau \pm \Delta\tau$ and consider the limits to be bona-fide if the significance is > 3 , otherwise report that the particular limit on τ cannot be obtained.

Step 5: For multiple light curves or for a lightcurve divided in to segments find the weighted average of σ_{XI}^2 , σ_{YI}^2 and \tilde{c}_{XY} , using their error estimates as weights. Then if the cross-correlation is significant, find the phase and time lags as in stjpg 3 and 4 above.

6. Application to AGN light curves

To test and validate the effectiveness of the scheme, we analyse the lightcurve of a well studied Active Galactic Nucleus, Akn 564. The temporal and spectral properties of the source was studied using an *XMM-Newton* observation of the source by Dewangan et al. (2007). They computed the cross-correlation function for different energy bands and estimated a possible time-lag between the hard and soft bands to be ~ 1768 secs using the peak of the function as a measure. Using Monte Carlo simulations they estimated an error on the time lag to be

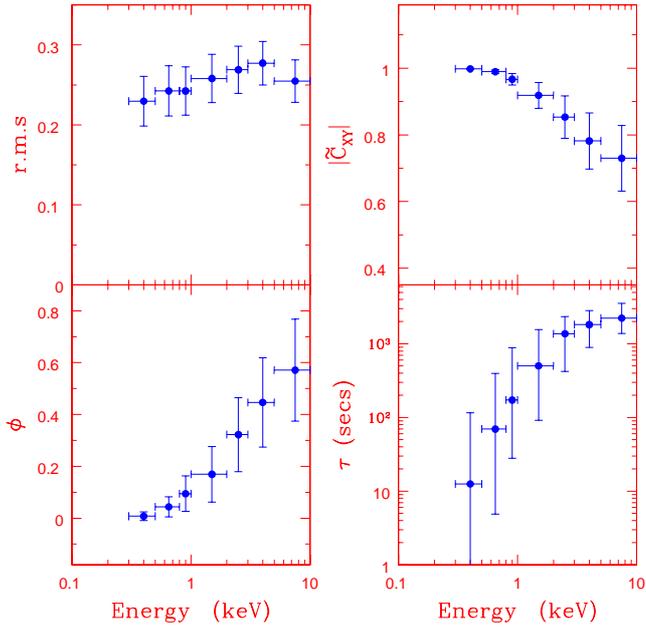


Figure 8: The variability property of Akn 564. Lightcurves of the source in different energy bands were used for the analysis. The time-bin for the light curves is 64 seconds and the number of data points is 1426. The rms, cross-correlation ($|\tilde{C}_{XY}|$), the phase difference (ϕ) and the time lag (τ) are plotted with energy. The reference energy band is 0.2-0.3 keV.

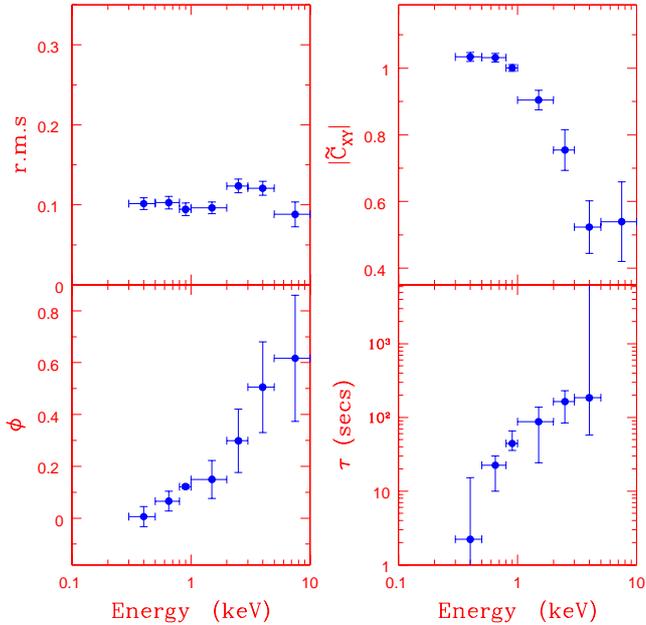


Figure 9: Same as Figure 8, except that the light curves were divided into ten segments and the results averaged.

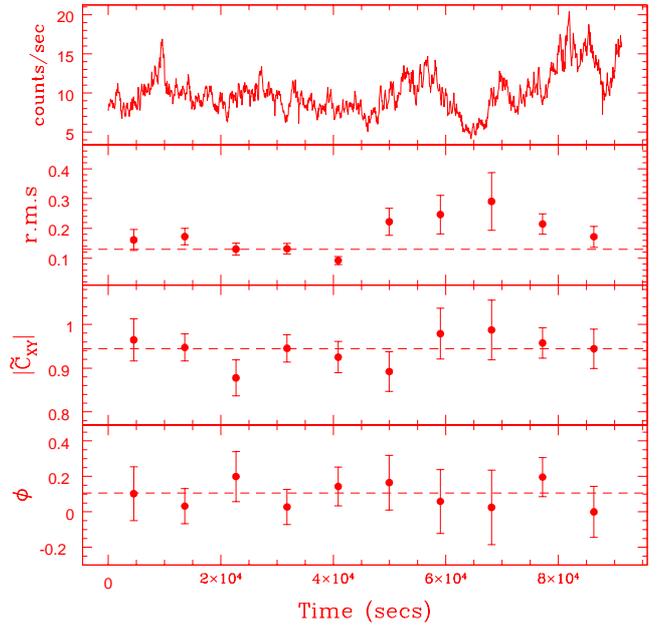


Figure 10: Checking the stationarity of the X-ray lightcurve of Akn 564. The complete lightcurve in the 1-2 keV band (shown in the top panel) has been divide into ten segments. For each segment the r.m.s is shown with errors in the second panel. The dashed line represents the average value. The cross-correlations, $|\tilde{C}_{XY}|$, and phase lags, ϕ between 0.2 – 1. and 1-2 keV bands for each segment are shown in the bottom two panels. The dashed lines represent the average values. It can be seen that the cross-correlation and the phase lag are consistent with being a constant showing that the system is stationary in these time-scales.

~ 100 secs due to measurement error. Arévalo et al. (2006) and McHardy et al. (2007) computed time-lags as function of frequency for *ASCA* and *XMM-Newton* observations of the source and found that there is a sharp drop in time lag for frequencies greater than 10^{-4} Hz.

We extracted light curves of the source using the *XMM-Newton* observation, in different energy bands. Details of the extraction process are given in Dewangan et al. (2007). The usable continuous time duration for the observation is for $\sim 10^5$ secs. Our motivation here is not to analyse in detail the temporal properties of the source and their physical interpretation, but instead to show as an example and validate the method described in this work. Thus, while finer time binning of the data is possible, we restrict our analysis to 64 sec bins, which resulted in light curves with length $N = 1426$. Figure 8 shows the results of our analysis. Note that the cross-correlation, phase and time-lag are well constrained as a function of energy. Figure 9 shows the results of the analysis when the light curves were divided into ten segments and the results averaged as described in the last section. Note that again the physical quantities are well constrained and while the phase difference is relatively unchanged between the two analysis, the time-lag decreases by nearly an order of magnitude. This is consistent with earlier results that the time-lag decreases with increasing Fourier frequency Papadakis et al. (2001); Vaughan et al. (2003b).

Splitting the lightcurve into segments and taking the average assumes that the during the time-scale, the source was stationary. This can be now explicitly tested using the analytical error estimates for each segment. This is demonstrated in Fig 10, where for each ten segments, the r.m.s, cross-correlations and phase-lags (between the energy bands 0.2-1 and 1-2 keV) are shown. The cross-correlations and phase lags are consistent with being a constant equal to the averaged value (shown as a dashed line). Formally the χ^2/dof for the data points to be constant are 5.3/9 and 3.3/9 for the cross-correlation and phase lag respectively. The slightly lower value of χ^2 than expected indicates only a slight overestimation of the error bars, especially for the phase lags. This is probably because for each segment, the error on the phase lag $\Delta\phi$ is large and hence the error distribution maybe slightly different than a Gaussian. Nevertheless, the figure clearly shows that not only can stationarity be tested but also reconfirms that the error estimates are reliable. Perhaps, not surprisingly, given the shape of the total light curve, the r.m.s of the 1-2 keV energy band is formally not consistent with being a constant, with a $\chi^2/dof = 28/10$. This is primarily due to the fifth segment where the r.m.s and its error is small. Note that the error is computed by assuming that the power spectrum of the segment is representative of the ensemble average. Perhaps a more prudent approach would be to estimate the error on the r.m.s using the averaged power spectrum rather than for each individual segment. However, whether such deviations are a significant indication of departure from stationarity is arguable and subjective. Hence, we recommend that only large deviations (for e.g. $\chi^2/dof > 5$) should be taken as serious evidence for non-stationarity.

7. Summary and discussion

In this work an estimate for the significance and error on the cross-correlation, phase and time lag between two light curves is presented. The error estimates take into account the stochastic fluctuations of the lightcurve as well as any known measurement errors. The technique has been verified using simulations of light curves generated from both white and $1/f$ stochastic processes with and without intrinsic correlation between them. The entire analysis consists of five algorithmic steps which are described in §6. The technique is ideally suited for short light curves of length $N \sim 1000$ and is an improvement over earlier methods which were based on numerically expensive simulations or by dividing the data into number of segments to find the variance.

The estimate presented in this work is based on several assumptions and hence is reliable only when they are valid. We emphasize this point by enumerating some of the main assumptions.

- Both the light curves have been generated from stochastic processes. Technically, this means that the phase of the different Fourier components are unrelated to each other i.e. $\langle \tilde{X}_k \tilde{X}_l \rangle = \delta_{lk}$. This assumption will be violated if the generation mechanism is a non-linear one. In general, it is difficult to ascertain the degree of non-linearity in a short lightcurve and it requires sensitive analysis like Bi-coherence measure and/or non-linear time series analysis. Thus, in most cases, the stochastic nature of the light curves have to be assumed. It is prudent to be aware that this assumption has been made and its validity is unknown, like for example, for the prompt emission of Gamma-ray bursts. A simple case where this assumption will be violated is if the power spectra have dominant harmonic features, where the power in the harmonics is comparable to that of the primary.
- The measurement errors are uncorrelated and have Gaussian distributions. The essential assumption is that the power spectrum of the measurement errors is independent of frequency (i.e. a white noise) and their phases are independent of each other. If the power spectrum has a different shape, then the appropriate changes have to be made and the basic results of this work need to be re-derived. For most practical purposes if the measurement errors are known, they usually are Gaussian distributions and hence this assumption is valid. If there are unknown systematic errors in the light curves then of course the analysis will not be applicable. Poisson distributions have the white noise property but in general the phases of the different Fourier components may be related. We have verified that for counts per time bin ~ 10 , the results of this analysis is valid. For counts less than that, caution is advised. However, for such low counts, meaningful results can only be obtained for long time series and it may better to obtain frequency dependent coherence and lag measurements.
- The light curves are evenly sampled without gaps. For unevenly sampled data the cross-correlation can be estimated Edelson and Krolik (1988), but there does not seem to be an analytical way to estimate the significance and error. One needs to use either Monte Carlo simulations or more practically estimate the error by dividing the light curves into several segments and finding their

variance.

- The light curves are stationary. As shown in the example of the light curves of Akn 564, this assumption can be tested by dividing the light curves into segments and checking whether the r.m.s, cross-correlation and phase lags are consistent to be a constant for different segments.

While this technique is useful for short duration light curves, coherence and frequency dependent time lags provide naturally more information and should be preferentially computed for long data streams. This technique may not be unique or optimal and hence there is a possibility and need for development of better methods provided they give robust and physically interpretable results. Finally, while cross-correlation, phase and time lags provide a quantitative measure of the system, their physical interpretation has to be done in terms of the physical geometrical and radiative model assumed for the system.

8. Acknowledgement

This work has made use of observational data obtained with XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and the USA (NASA). AB would like to thank the Inter-University Centre for Astronomy and Astrophysics, Pune for associateship programme.

References

- Arévalo, P., Papadakis, I., Uttley, P., McHardy, I., Brinkmann, W., 2006. Spectral-timing evidence for a very high state in the narrow-line seyfert 1 ark 564. *Monthly Notices of the Royal Astronomical Society* 372, 401–409.
- Bartlett, M.S., 1955. *An Introduction to stochastic processes: with special reference to methods and applications*. The University Press.
- Chatfield, C., 2016. *The analysis of time series: an introduction*. CRC press.
- Dewangan, G., Griffiths, R., Dasgupta, S., Rao, A., 2007. An investigation of the origin of soft x-ray excess emission from ark 564 and mrk 1044. *The Astrophysical Journal* 671, 1284.
- Edelson, R., Krolik, J., 1988. The discrete correlation function—a new method for analyzing unevenly sampled variability data. *The Astrophysical Journal* 333, 646–659.
- Keshner, M.S., 1982. $1/f$ noise. *Proceedings of the IEEE* 70, 212–218.
- Van der Klis, M., 1989. Fourier techniques in x-ray timing, in: *Timing Neutron Stars*. Springer, pp. 27–69.
- Leahy, D., Darbro, W., Elsner, R., Weisskopf, M., Kahn, S., Sutherland, P., Grindlay, J., 1983. On searches for pulsed emission with application to four globular cluster x-ray sources—ngc 1851, 6441, 6624, and 6712. *The Astrophysical Journal* 266, 160–170.
- McHardy, I., Arévalo, P., Uttley, P., Papadakis, I., Summons, D., Brinkmann, W., Page, M., 2007. Discovery of multiple lorentzian components in the x-ray timing properties of the narrow line seyfert 1 ark 564. *Monthly Notices of the Royal Astronomical Society* 382, 985–994.
- Nowak, M.A., Vaughan, B.A., Wilms, J., Dove, J.B., Begelman, M.C., 1999. Rossi x-ray timing explorer observation of cygnus x-1. ii. timing analysis. *The Astrophysical Journal* 510, 874.
- Papadakis, I., Nandra, K., Kazanas, D., 2001. Frequency-dependent time lags in the x-ray emission of the seyfert galaxy ngc 7469. *The Astrophysical Journal Letters* 554, L133.
- Peterson, B.M., Wanders, I., Horne, K., Collier, S., Alexander, T., Kaspi, S., Maoz, D., 1998. On uncertainties in cross-correlation lags and the reality of wavelength-dependent continuum lags in active galactic nuclei. *Publications of the Astronomical Society of the Pacific* 110, 660.
- Timmer, J., Koenig, M., 1995. On generating power law noise. *Astronomy and Astrophysics* 300, 707.

Vaughan, S., Edelson, R., Warwick, R., Uttley, P., 2003a. On characterizing the variability properties of x-ray light curves from active galaxies. *Monthly Notices of the Royal Astronomical Society* 345, 1271–1284.

Vaughan, S., Fabian, A., Nandra, K., 2003b. X-ray continuum variability of mcg-6-30-15. *Monthly Notices of the Royal Astronomical Society* 339, 1237–1255.

Welsh, W., 1999. On the reliability of cross-correlation function lag determinations in active galactic nuclei. *Publications of the Astronomical Society of the Pacific* 111, 1347.