The rapid X-ray variability of NGC 4051

S. Vaughan, P. Uttley, K. A. Pounds, K. Nandra and T. E. Strohmayer

1X-ray & Observational Astronomy Group, Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH
2School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ
3Max-Planck-Institut für extraterrestrische Physik (MPE), Giessenbachstr. 1, 85748 Garching, Germany
4Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ
5Astrophysics Science Division, NASA’s Goddard Space Flight Center, Greenbelt, MD 20771, USA

Accepted 2011 January 10. Received 2011 January 7; in original form 2010 November 22

ABSTRACT

We present an analysis of the high-frequency X-ray variability of NGC 4051 (MBH ∼ 1.7 × 10⁶ M☉) based on a series of XMM–Newton observations taken in 2009 with a total exposure of ∼570 ks (EPIC pn). These data reveal the form of the power spectrum over frequencies from 10⁻⁴ Hz, below the previously detected power spectral break, to ≥10⁻² Hz, above the frequency of the innermost stable circular orbit (ISCO) around the black hole (νISCO ∼ 10⁻³–10⁻² Hz, depending on the black hole spin parameter j). This is equivalent to probing frequencies of ≥1 kHz in a stellar mass (MBH ∼ 10 M☉) black hole binary system. The power spectrum is a featureless power law over the region of the expected ISCO frequency, suggesting no strong enhancement or change in the variability at the fastest orbital period in the system. Despite the huge amplitude of the flux variations between the observations (peak-to-peak factor of ≥50), the power spectrum appears to be stationary in shape and varies in amplitude at all observed frequencies following the previously established linear rms–flux relation. The rms–flux relation is offset in flux by a small and energy-dependent amount. The simplest interpretation of the offset is in terms of a very soft spectral component that is practically constant (compared to the primary source of variability). One possible origin for this emission is a circumnuclear shock energized by a radiatively driven outflow from the central regions and emitting via inverse-Compton scattering of the central engine’s optical–UV continuum (as inferred from a separate analysis of the energy spectrum). A comparison with the power spectrum of a long XMM–Newton observation taken in 2001 gives only weak evidence for non-stationarity in power spectral shape or amplitude. Despite being among the most precisely estimated power spectra for any active galaxy, we find no strong evidence for quasi-periodic oscillations (QPOs) and determine an upper limit on the strength of a plausible QPO of ≲2 per cent rms in the 3 × 10⁻³–0.1 Hz range and ∼5–10 per cent in the 10⁻⁴–3 × 10⁻³ Hz range. We compare these results to the known properties of accreting stellar mass black holes in X-ray binaries, with the further aim of developing a ‘black hole unification’ scheme.

Key words: galaxies: active – galaxies: individual: NGC 4051 – galaxies: Seyfert – X-rays: galaxies.

1 INTRODUCTION

If strong gravity dominates the dynamics of the inner accretion flows around black holes, then an elementary consequence is scale invariance: many important aspects of accretion on to supermassive black holes (MBH ∼ 10⁹ M☉) in active galactic nuclei (AGN) should be fundamentally the same as for stellar mass black holes (MBH ∼ 10 M☉) in Galactic black hole systems (GBHs). (See e.g. Shakura & Sunyaev 1976; Mushotzky, Done & Pounds 1993.) Over the past few years, several similarities have been observed in the X-ray variability of nearby AGN and GBHs, supporting the idea of black hole unification (Fender et al. 2006). In particular, the power spectrum (often power spectral density) and inter-band X-ray time lags appear to have similar frequency dependence and relative amplitudes in AGN and GBHs but with characteristic frequencies scaled (to first order) inversely with mass, ν ∝ 1/MBH (see e.g. Lawrence et al. 1987; Hayashida et al. 1998; Edelson & Nandra 1999; Uttley, McHardy & Papadakis 2002; Markowitz et al. 2003; Vaughan, Fabian & Nandra 2003a; McHardy et al. 2004, 2006.)
2 OBSERVATIONS AND DATA ANALYSIS

NGC 4051 was observed by *XMM–Newton* 15 times over 45 d during 2009 May–June and previously during 2001 and 2002 (see Table 1). The total duration of the useful data from the 2009 campaign is 572 ks, giving ~6 × 10^6 EPIC pn source counts (using PATTERN 0–4) or ~9 × 10^6 EPIC source counts (using PATTERN 0–12 pn and MOS combined).

The raw data were processed from observation data files following standard procedures using the *XMM–Newton* Science Analysis System (SAS v10.0.2). The EPIC data were processed using the standard SAS processing chains to produce calibrated event lists. For each observation, source events were extracted from these using a 40-arcsec circular region centred on the target, and background events were extracted from a rectangular region on the same chip but not overlapping with the source region. Examination of the background time series showed the background to be relatively low and stable throughout most of the observations, with the exception of the final few kiloseconds of each observation where the background rate increased as the spacecraft approached the radiation belts at perigee (each of the observations occurred towards the end of a spacecraft revolution). The EPIC data were taken using the small window mode of each camera and the medium blocking filter to reduce pile-up and optical loading effects, respectively. Nevertheless, during bright intervals NGC 4051 is sufficiently bright to cause moderate pile-up, especially in the MOS, which can affect power spectrum estimation (Tomsick, Kalemci & Kaaret 2004). In order to mitigate any distorting effects all results are based on pn data PATTERN 0–4 only, but where possible they were checked for consistency with the MOS data. Response matrices were generated using *RMFGEN* v1.55.2 and ancillary responses were generated with *ARFGEN* v1.77.4.

Time series were extracted from the filtered events for the source and background regions using a range of time bin sizes and energy ranges. These were all corrected for exposure losses, as catalogued in the Good Time Information (GTI) header contained in each event file, except where less than 30 per cent of a bin was exposed, in which case the data were calculated by linear interpolation from the nearest ‘good’ data either side and appropriate Poisson noise was added (this was needed for only 0.3 per cent of the total exposure).

### Table 1. Observation log

<table>
<thead>
<tr>
<th>Obs ID</th>
<th>rev.</th>
<th>Obs. date</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0109141401</td>
<td>0263</td>
<td>2001-05-16</td>
<td>121958</td>
</tr>
<tr>
<td>0157560101</td>
<td>0541</td>
<td>2002-11-22</td>
<td>51866</td>
</tr>
<tr>
<td>0606320101</td>
<td>1721</td>
<td>2009-05-03</td>
<td>45717</td>
</tr>
<tr>
<td>0606320201</td>
<td>1722</td>
<td>2009-05-05</td>
<td>45645</td>
</tr>
<tr>
<td>0606320301</td>
<td>1724</td>
<td>2009-05-09</td>
<td>45548</td>
</tr>
<tr>
<td>0606320401</td>
<td>1725</td>
<td>2009-05-11</td>
<td>45447</td>
</tr>
<tr>
<td>0606321301</td>
<td>1727</td>
<td>2009-05-15</td>
<td>32644</td>
</tr>
<tr>
<td>0606321401</td>
<td>1728</td>
<td>2009-05-17</td>
<td>42433</td>
</tr>
<tr>
<td>0606321501</td>
<td>1729</td>
<td>2009-05-19</td>
<td>41813</td>
</tr>
<tr>
<td>0606321601</td>
<td>1730</td>
<td>2009-05-21</td>
<td>41936</td>
</tr>
<tr>
<td>0606321701</td>
<td>1733</td>
<td>2009-05-27</td>
<td>44919</td>
</tr>
<tr>
<td>0606321801</td>
<td>1734</td>
<td>2009-05-29</td>
<td>43726</td>
</tr>
<tr>
<td>0606321901</td>
<td>1736</td>
<td>2009-06-02</td>
<td>44946</td>
</tr>
<tr>
<td>0606322001</td>
<td>1737</td>
<td>2009-06-04</td>
<td>39756</td>
</tr>
<tr>
<td>0606322101</td>
<td>1739</td>
<td>2009-06-08</td>
<td>43545</td>
</tr>
<tr>
<td>0606322201</td>
<td>1740</td>
<td>2009-06-10</td>
<td>44453</td>
</tr>
<tr>
<td>0606322301</td>
<td>1743</td>
<td>2009-06-16</td>
<td>42717</td>
</tr>
</tbody>
</table>
time from the pn, for $\Delta t = 5\,\text{s}$ time bins). The outputs of this process were regularly sampled and uninterrupted, exposure-corrected time series for the source and background. The background time series was then subtracted from the corresponding source time series (after scaling by the ratio of the ‘good’ extraction region areas). Fig. 1 shows the resulting time series from the 15 separate 2009 observations.

3 POWER SPECTRUM

The power spectrum of the 2009 data was estimated using standard methods (van der Klis 1989). In particular, time series were extracted with binning $\Delta t = 5\,\text{s}$ and divided into uninterrupted 10-ks intervals. For each interval a periodogram was computed in units of relative power (Miyamoto et al. 1991; van der Klis 1997; Vaughan et al. 2003b), the Poisson noise level was subtracted and then the periodograms were averaged. Fig. 2 shows the resulting power spectrum estimate after rebinning in logarithmic frequency intervals. The model fitting was performed with the help of XSPEC v12.6.0 (Arnaud 1996) by using $\chi^2$ as the fit statistic; all data were binned to ensure that Gauss-normal statistics applied and errors quoted correspond to 90 per cent confidence limits (e.g. using a $\Delta \chi^2 = 2.71$ criterion for one parameter) unless otherwise stated.

The previous analysis of the power spectrum of NGC 4051 by McHardy et al. (2004) found a smoothly bending power-law continuum provided an adequate description of the data

$$\text{cont}(v) = \frac{N \nu^{\alpha_{\text{low}}}}{1 + (v/v_{\text{bend}})^{\nu_{\text{low}} - \nu_{\text{high}}}},$$

with indices of $\alpha_{\text{low}} \approx -1.1$ at low frequencies and $\alpha_{\text{high}} \approx -2$ above $v_{\text{bend}} \approx 8 \times 10^{-4}\,\text{Hz}$. This model provided an excellent fit to the 2009 power spectrum, as shown in Fig. 2, with the lower frequency index fixed at $\alpha_{\text{low}} = -1.1$ (since it is much better constrained by the RXTE data; McHardy et al. 2004). The best-fitting values for the free parameters of the model were $\alpha_{\text{high}} = -2.31 \pm 0.08$ and $v_{\text{bend}} = 2.3 \pm 1.0 \times 10^{-4}\,\text{Hz}$. The fit statistic was $\chi^2 = 31.19$ with 27 degrees of freedom (d.o.f.), giving a $p$-value of 0.26. Allowing $\alpha_{\text{low}}$ to vary provided no improvement in the fit ($\Delta \chi^2 < 2$) and the index itself was poorly constrained (with a 90 per cent confidence interval of $[-0.25, -1.78]$), and restricting $\alpha_{\text{low}}$ to vary only within the range $[-1.2, -0.9]$ the confidence interval determined by McHardy et al. (2004) from the RXTE data provided very little

---

1 The expected power density due to measurement uncertainties in this case is $P_\nu = 2\Delta T(\sigma^2)$, where $\Delta T$ is the sampling rate and $(\sigma^2)$ is the mean square error. See Vaughan et al. (2003b).

2 We have confirmed that the power spectral estimation and modelling were not significantly biased by sampling distortions – commonly known as ‘aliasing’ and ‘red noise leakage’ (see van der Klis 1989; Uttley et al. 2002; Vaughan et al. 2003a). Aliasing is minimal for individually binned data (van der Klis 1989), like those used here, and does not occur for the Poisson noise that dominates the highest frequencies probed. Leakage could in principle have been a problem, resulting from the finite duration of each observation, but in practice was not significant because the power spectrum was sufficiently flat ($\sim v^{-1}$) on time-scales comparable to the observation length. This was confirmed using an analysis discussed in footnote 4.

---

Figure 1. Time series of 0.2–10 keV EPIC pn count rate (background-subtracted) from each of the 15 observations taken in 2009. The bin size is $\Delta t = 100\,\text{s}$. The 15 separate $\sim 30–40$ ks observations, taken typically 2–4 d apart, have been concatenated in this plot, to allow easy comparison between them, with the end points marked by the dotted vertical line. The background count rate is also shown as a grey curve, but is almost always far weaker than the source. The labels above each section of data refer to the spacecraft revolutions during which the observations were made.

Figure 2. Power spectrum estimate (histogram) for the broad-band (0.2–10 keV) EPIC pn data, computed using 10-ks intervals at $\Delta t = 5$ bins, with the best-fitting bending power-law continuum model (solid curve). Also shown (grey circles) are two low-frequency power density estimates computed using 28-ks intervals. These lower frequency data were not used in the fitting (see the text for details).
change in the confidence interval for $\nu_{\text{bend}}$ compared to the fixed $\alpha_{\text{low}} = -1.1$ model. There is no obvious indication, in either the overall fit statistic or the residual plot, of additional power spectral components. We discuss this point in more detail in Section 5.

A sharply broken power-law model also provided a good fit, with $\chi^2 = 37.52$ for 27 d.o.f. ($p = 0.086$), slightly worse than the smoothly bending model. The index below the break was again fixed to $\alpha_{\text{low}} = -1.1$, and the best-fitting break frequency was $\nu_{\text{break}} = 2.5 \pm 0.6 \times 10^{-4}$ Hz, and a high-frequency index of $\alpha_{\text{high}} = -2.18 \pm 0.05$. These parameters, and the overall fit quality, are very similar to those of the bending power-law model, indicating that the data are rather insensitive to the detailed shape of the change from flat to steep index. By contrast, the power spectrum is less well explained in terms of an exponentially cut-off power law, which gives $\chi^2 = 44.03$ for 26 d.o.f. ($p = 0.015$), but only by allowing $\alpha_{\text{low}} \approx -1.8$ which is inconsistent with the long-term RXTE results (McHardy et al. 2004). Assuming $\alpha_{\text{low}} = 1.1$ gave an unacceptable fit ($\chi^2 = 217.81$ for 27 d.o.f.; $p \ll 0.001$).

The model fitting was repeated using a Bayesian scheme by treating the likelihood function as $\sim \exp(-\chi^2/2)$ and assigning simple prior distributions to the parameters. See Vaughan (2010), and references therein, for a brief introduction to these ideas. Uniform priors were assigned to $\alpha_{\text{high}}$ and the overall normalization, and to log $\nu_{\text{bend}}$ [equivalent to a $p(\nu) = 1/\nu$ Jeffreys’ prior], although in practice using uniform or Jeffreys’ prior made no significant difference to the results. As expected, the parameter values at the posterior mode were almost identical to those giving the best fit using the classical min $\chi^2$ method. But having the problem restated in Bayesian terms allowed for an exploration of the posterior distribution of the parameters, $p(\theta|x^{\text{obs}})$, with $\theta_{\nu}$ and $x^{\text{obs}}$ being the parameters specifying the continuum and the observed data, respectively, using a Markov chain Monte Carlo (MCMC) scheme to randomly draw sets of parameters from the posterior.3 The marginal density for the bend frequency, $p(\nu_{\text{bend}}|x^{\text{obs}})$, as computed using the MCMC data, is shown in Fig. 3; the 90 per cent credible interval on $\nu_{\text{bend}}$ was $[1.3, 3.2] \times 10^{-4}$ Hz.

The limited durations of the individual 2009 observations make it difficult to directly probe frequencies lower than $10^{-4}$ Hz. But for completeness, the power spectrum was also estimated using longer uninterrupted intervals (28-ks duration, of which there were 15 in 2009, one for each observation). However, the estimates of power density at the lowest frequencies were not normally distributed in this case, due to the small number of estimates contributing to the average (Papadakis & Lawrence 1993). These data were not included in the power spectral fitting, but the lowest frequency points are shown in Fig. 2 and are clearly consistent with an extrapolation of the best-fitting model.

As a test for a possible high-frequency cut-off in the power spectrum, the bending power-law model was modified by including an exponential cut-off (the highest cut model in XSPEC). This provided only a very small improvement in the overall fit statistic ($\Delta\chi^2 = 4.1$ for two additional free parameters) with only poorly constrained parameters (cut-off frequency $\nu_{\text{cut}} \leq 8 \times 10^{-3}$ and e-folding frequency $\nu_{\text{fold}} = 2 \pm 1 \times 10^{-2}$). The power spectrum estimated from the high flux data (see the next section) has a slightly higher signal-to-noise ratio above 10 mHz, but again shows no significant difference in the quality of the fit between the bending power laws with and without an exponential cut-off at high frequencies.

### 3.1 Energy dependence of the power spectrum

The dependence of the power spectrum on photon energy was investigated by dividing the data into three broad energy bands – 0.2–0.7, 0.7–2 and 2–10 eV – and estimating and fitting the power spectrum for each of these. A simultaneous fit to all these spectra using the same continuum model (equation 1) gave a rather poor fit ($\chi^2 = 118.9$ with 85 d.o.f.; $p = 0.009$) with the parameters tied between the three spectra. But allowing to be different for each energy band, while keeping $\nu_{\text{bend}}$ tied between them, gave a much better fit ($\chi^2 = 98.5$ with 83 d.o.f.; $p = 0.12$); the improvement is significant in an $F$-test, with $p < 0.001$, indicating that the high-frequency power spectrum is energy dependent. The slopes estimated for the low-, medium- and high-energy bands were $2.31 \pm 0.07, 2.24 \pm 0.09$ and $2.04 \pm 0.10$, respectively. The flattening of the power spectrum at higher energies is very similar to that detected by McHardy et al. (2004) from the 2001 data, and previously observed in several other Seyfert galaxies (e.g. Nandra & Papadakis 2001; Vaughan & Fabian 2003).

Allowing the bend frequency $\nu_{\text{bend}}$ to be different for each energy band, but with $\alpha_{\text{high}}$ tied between them, provided a considerably worse fit ($\chi^2 = 108.0$ with 83 d.o.f.) than the energy-dependent $\alpha_{\text{high}}$ model. Furthermore, allowing both $\nu_{\text{bend}}$ and $\alpha_{\text{high}}$ to differ between energy bands did not improve significantly on the variable $\alpha_{\text{high}}$, $\chi^2 = 95.7$ with 81 d.o.f.). In this case, the best-fitting $\nu_{\text{bend}}$ values for the low- and medium-energy bands were very similar while the hard band gave a somewhat lower value, but with a much larger confidence interval. We conclude that there is strong evidence for a change in the high-frequency slope with energy, but the data do not resolve any difference in the location of the bend.

### 4 RMS–FLUX RELATION

The 15 observations shown in Fig. 1 clearly show very different mean levels and amplitudes of variability, with noticeably stronger variations occurring preferentially during brighter intervals (e.g. rev1730) and the lowest flux intervals being extremely stable (e.g. rev1739). This behaviour is exactly as expected for

---

3 Specifically, five chains of a length of $5 \times 10^4$ were generated (after removing an equal length “burn-in” period), using a Metropolis–Hastings algorithm with a multivariate Cauchy proposal distribution. Convergence of the chains was confirmed using the $R$ statistic (Gelman & Rubin 1992; Gelman et al. 2004). The results from each chain were then combined to produce a large sample of posterior draws.
sources that obey the rms–flux relation (Uttley & McHardy 2001; Uttley et al. 2005).

In order to test this explicitly and in detail, the power spectrum was estimated in different flux bins. As before, the time series were divided into uninterrupted 10-ks intervals, the periodograms computed and the contribution of the Poisson noise subtracted, but using absolute units of power (i.e. without normalizing to the squared mean rate; see Vaughan et al. 2003b). The intervals were sorted according to their mean flux and the periodograms were then averaged in flux bins to provide estimates of the power spectrum as a function of mean flux. Fig. 4 shows the power spectral estimates when the data were divided into three flux bins. There is clearly a large difference in the overall amplitude (in absolute units) but little change in shape, as revealed by the ratio of bright to faint power spectra. Direct fitting of the high and low flux power spectra gives the same result: the best-fitting values for \( \alpha \) and \( \nu \) estimated for the high flux data fall within the 90 per cent confidence intervals derived from the low flux data. The same effect was seen in Cygnus X-1 by Uttley & McHardy (2001) (see their fig. 2) and is consistent with there being a single rms–flux relation independent of the frequencies used to compute the rms.

The detailed rms–flux relation was then examined explicitly by dividing the data into 1-ks uninterrupted intervals, computing a periodogram for each (again, in absolute units), subtracting the Poisson noise from each and then averaging the results in several flux bins\(^4\) \((F_i)\). The 1–10 MHz rms, \( \sigma_i \), was then estimated for each flux bin by integrating the corresponding power spectrum estimate over this frequency range. The resulting rms–flux data are shown in Fig. 5. (The uncertainties on the rms estimates were calculated based on the variance of the sum of the averaged periodograms, which are themselves calculated based on the \( \chi^2 \) distribution of periodogram ordinates. The method is explained in more detail in Heil, Vaughan & Uttley, in preparation.)

The dependence of rms on flux is clearly very close to linear over the full flux range, as previously noted in this source and accreting black holes generally (e.g. Uttley & McHardy 2001; Uttley et al. 2005; Heil, Vaughan & Uttley, in preparation). The best-fitting linear function \(- \sigma_i = k (F_i - C)\), where \( k \) is the gradient and \( C \) is the offset on the flux axis – gave a rather high fit statistic \((\chi^2 = 30.05\) for 12 d.o.f.) but the residuals show little structure. The best-fitting linear function has a negative offset in the rms axis, equivalent to a positive offset on the flux axis. The energy spectrum of this offset was computed by dividing the pn data into different energy bands (14 logarithmically spaced) and for each one calculating the rms–flux data and fitting with a linear model. The spectrum of the offset \( C(E) \) is shown in Fig. 6, along with some representative pn spectra from bright (rev1730), intermediate (rev1743) and faint (rev1739) flux observations. The spectra were ‘fluxed’ for display purposes, that is, they have been normalized to flux density units by dividing out the effective area curve (see Appendix A for details). The spectrum of the offsets from the rms–flux relation is clearly soft: it is significantly positive in the lower energy bands, decreasing until consistent with zero intercept above \( \sim 2\) keV. Most interestingly, it has a very similar shape and strength to the faint flux spectrum below 2 keV.

5 SEARCH FOR QUASI-PERIODIC OSCILLATIONS

The power spectral fitting discussed above provided no clear evidence for narrow features, i.e. QPOs. Such features are usually described in terms of Lorentzian ‘lines’ seen in the power spectrum in addition to the continuum. In order to assess the evidence for a QPO more rigorously, we defined two competing hypotheses: the null hypothesis \( H_0 \) of a continuum power spectrum and the alternative hypothesis \( H_1 \) which includes an additional Lorentzian component:

\[
H_0 : P(\nu) = \text{cont}(\nu, \theta_C)
\]

\[
H_1 : P(\nu) = \text{cont}(\nu, \theta_C) + \text{Lorentzian}(\nu, \theta_L)
\]

\[
\text{Lorentzian}(\nu, \theta_L) = \frac{P_0}{\pi} \frac{\nu - \nu_0}{(\nu - \nu_0)^2 + \gamma^2}
\]

\( P_0 \) is the peak power, \( \gamma \) is the Lorentzian width and \( \nu_0 \) is the frequency of the Lorentzian.

\(^4\) We have checked that the power spectra estimated using short time series intervals (e.g. 1 ks) are not significantly biased by leakage (see Uttley et al. 2002, and references therein) in two different ways. First, we compared the average power spectrum estimated from the real data using 1-, 10- and 28-ks intervals and confirming that they are all consistent. Secondly, we compared the average power spectrum by simulating data (using the best-fitting bending power-law model) with the same sampling pattern as the real data and estimating the bias between the power spectral estimate and the model.
\begin{equation}
H_1 : P(v) = \text{cont}(v, \theta_C) + \text{qpo}(v, \theta_Q),
\end{equation}

where the continuum model is defined above (equation 1) and the QPO will be described by a Lorentzian profile:

\begin{equation}
\text{qpo}(v, \theta_Q) = \frac{2R^2Qv/\pi}{v_Q^2 + 4(v - v_Q)^2}.
\end{equation}

The parameters of the continuum are \( \theta_C = \{ \alpha_{\text{high}}, \nu_{\text{high}}, N \} \) and the QPO parameters are \( \theta_Q = \{ v_Q, Q, R \} \), where \( v_Q \) is the centroid frequency, \( Q \) is the width parameter \( Q = v_Q/\Delta v \), where \( \Delta v \) is the full width at half-maximum) and \( R \) is the amplitude parameter (equal to the total rms in the limit of high \( Q \)). See Pottschi et al. (2003) for details of this parametrization. In the present case we have assumed two different values, namely \( Q = 5 \) and \( 10 \), which are fairly typical values for HF QPOs in GBHs (e.g. Remillard et al. 2002, 2003). Using these two representative values rather than allowing \( Q \) to be a free parameter made the QPO search and upper limit calculations (discussed later) much more efficient.

One of the simplest ways to select between these two competing models is to assess the improvement in the \( \chi^2 \) fit statistic upon the addition of the QPO, i.e. the difference between the minimum \( \chi^2 \) for \( H_0 \) and \( H_1 \). For a given data set, a large reduction (improvement) in the fit statistic between \( H_0 \) and \( H_1 \) is taken to be evidence in favour of a QPO. For the 2009 data, and the full 0.2–10 keV energy band, the reduction in the \( \chi^2 \) statistic was only 6.85 (assuming \( Q = 10 \), and a very similar value assuming \( Q = 5 \), with best-fitting parameters \( v_Q = 3.8 \pm 0.3 \) mHz and \( R = 1.2 \pm 0.5 \) per cent. The modest improvement in the fit falls short of our threshold for a significant detection, as discussed below. The QPO search was repeated for the three energy bands discussed above, with very similar results. The 2–10 keV band gave the largest improvement in the fit upon the addition of a QPO (\( \Delta \chi^2 = 10.94 \) assuming \( Q = 5 \)), but this still falls short of a detection in the \( \alpha = 0.01 \) test.

The lack of detection may itself be an interesting result if the observation was sufficiently sensitive, and to assess this requires a calculation of the upper limit on the strength of a QPO. The first step is to define a precise detection procedure and then calibrate its statistical ‘power’ as a function of the QPO strength. This would allow us to determine the weakest QPOs that would be detected with reasonable probability.

### 5.1 Detection procedure

The search for QPOs in binned power spectral data is no different from the search for ‘lines’ in any other one-dimensional spectrum and suffers from the same statistical and computational challenges. See Freeman et al. (1999), Protassov et al. (2002) and Park, van Dyk & Siemiginowska (2008) for elaboration of the issues.

The difference in the minimum \( \chi^2 \) statistic for each model was used as a test statistic. For a data set \( x \), we compute

\begin{equation}
T(x) = \chi^2(x, H_0) - \chi^2(x, H_1).
\end{equation}

The nuisance parameters – specifying the continuum and the QPO – are eliminated by minimization, i.e. \( \chi^2(x, H_0) = \min_{\theta_C, \theta_Q} \{ \chi^2(x, H_0, \theta_C) \} \). This test statistic is closely related to the likelihood ratio test, as has been discussed by Protassov et al. (2002). A large value of the test statistic (i.e. a large improvement in the fit upon adding a QPO to the model) is taken as evidence for a QPO.

The critical value of \( T \) for detection, \( T_{\text{crit}} \), was calculated by a Monte Carlo method as discussed in Appendix B. In the present case, we found \( T_{\text{crit}} = 12.28 \) for \( \alpha = 0.01 \) (this is for \( Q = 10 \); using \( Q = 5 \) gave 13.05 which is practically the same). In other words, assuming the continuum model to be true [with parameters represented by the posterior distribution \( p(\theta_C|x_{\text{obs}}) \) derived above from the real data], and no QPO, a reduction in the \( \chi^2 \) fit statistic of \( T \geq 12.28 \) will occur with probability \( \alpha = 0.01 \) (and have a posterior predictive \( p \)-value of \( \leq 0.01 \)).

### 5.2 Upper limit procedure

In order to define an upper limit on the strength of possible QPOs we use the definition of an upper limit in terms of the power\(^5\) \( \beta \) of a detection procedure, as discussed very clearly in the recent paper by Kashyap et al. (2010). A QPO of a given frequency and strength \( (v_Q, R) \) has a probability of detection \( \beta(v_Q, R) \) using the above method (for fixed \( \alpha \)).

We used a Monte Carlo method, as discussed in Appendix B, to estimate the detection probability \( \beta(v_Q, R) \) for a range of different QPO locations and strengths \( (v_Q \) and \( R \) values). The estimate is simply the fraction of simulations (for each \( v_Q \) and \( R \) values) that gave a significant detection defined by \( T(x_{\text{obs}}) \geq T_{\text{crit}} \). Contours of constant

\(^5\) Here we follow the notation used by Kashyap et al. (2010) and use \( \beta \) to denote the power of the test, i.e. the probability of meeting the detection criterion assuming that the effect is real (e.g. there is a QPO present).
law model ($\chi^2 = 26.22$ for 27 d.o.f., $p = 0.51$) had a slightly higher bend frequency ($\nu \approx 5.5 \times 10^{-4}$ Hz with a 90 per cent confidence interval of [3.0, 9.0] $\times 10^{-4}$ Hz and a credible interval of [2.7, 8.4] $\times 10^{-4}$ Hz) and high-frequency index ($\alpha_{\text{high}} = -2.48 \pm 0.16$) compared to the 2009 data. The marginal distributions of $v_{\text{bend}}$ from both data sets are shown in Fig. 3. These appear to suggest that there was a slight change in the power spectral shape between the two sets of observations. It is for this reason that we have not combined the 2001 and 2009 data to produce a composite power spectrum.

7 DISCUSSION AND CONCLUSIONS

We have presented an analysis of the rapid X-ray flux variability of the low-mass Seyfert 1 galaxy NGC 4051 covering the range from $10^{-4}$ to $\sim 2 \times 10^{-3}$ Hz, based on a series of 15 XMM–Newton observations taken during 2009. This is among the best X-ray power spectra for any active galaxy, and the large range of fluxes sampled during the observations has allowed for an examination of the flux dependence of the variability. Below we discuss the results and, where possible, relate them to the known behaviour of GBHs (for which we use a fiducial mass of $M_{BH} = 10 M_\odot$).

7.1 Power spectral shape in relation to other sources

The power spectrum of the variations is well described by a bending power-law model that bends from an index of $\alpha_{\text{low}} \approx -1.1$ to $\alpha_{\text{high}} \approx -2.3$ around $v_{\text{bend}} \sim 2 \times 10^{-3}$ Hz. This is fairly consistent with the McHardy et al. (2006) relation:

$$\log(T_b) = 2.1 \log(M_\odot) - 0.98 \log(L_{44}) - 2.32,$$

where $T_b$ is the break time-scale in days ($v_{\text{bend}} = 1.16 \times 10^{-5}/T_b$ Hz), $M_\odot = M_{BH} / 10 M_\odot$ and $L_{44} = L_{44} \times 10^{44}$ erg s$^{-1}$ is the bolometric luminosity. The average $2 - 10$ keV X-ray flux during the 2009 observations was $F_{2-10} \approx 1.2 \times 10^{-11}$ erg s$^{-1}$ cm$^{-2}$ which gives $L_{2-10} = 3.5 \times 10^{44}$ erg s$^{-1}$ at an assumed distance of 15.7 Mpc (Russell 2002), and $L_{\text{bol}} \sim 10^{43}$ (assuming a $L_{\text{bol}}/L_{2-10} \sim 27$; Elvis et al. 1994). The black hole mass is thought to be $M_{BH} \sim 1.7 \pm 0.5 \times 10^4 M_\odot$ (Denney et al. 2009); therefore, $M_{\text{bol}} \sim L_{\text{bol}}/L_{\text{bol}} \sim 0.05$. Together these give a predicted bend frequency of $8 \times 10^{-3}$ Hz ($T_b \sim 0.14$ d), within a factor of $\sim 3$ of the observed value. In fact, the estimated $v_{\text{bend}}$ from the 2009 data is closer to the value expected from the McHardy et al. (2006) relation than that estimated by McHardy et al. (2004), when calculated using the revised black hole mass estimate (Denney et al. 2009, 2010).

The high-frequency power spectrum above the bend rarely provides strong, unambiguous signatures of the accretion state of GBH systems, which tend to differ more at lower frequencies where features such as low-frequency QPOs, flat-topped noise and broad Lorentzians occur (see e.g. McClintock & Remillard 2006; van der Klis 2006; Belloni 2010). McHardy et al. (2004) suggested that NGC 4051 may be analogous to a soft state (particularly like that seen in Cygnus X-1) on the basis of the low-frequency noise (a $v^{-1}$ spectrum extending over two decades or more), the high-frequency slope ($\alpha_{\text{high}} < -2$) and the value of $v_{\text{bend}}$ (which tends to be higher in the soft state, making the ratio $T_b/M_{BH}$ smaller).

Note that McHardy et al. (2006) defined $T_b$ as the break frequency estimated by fitting power spectra with a sharply breaking power-law model. This difference is of little consequence to the present discussion as we obtained very similar estimates for the characteristic time-scale using smoothly bending and sharply broken models (Section 3).

6. LONG-TERM STATIONARITY

The power spectral estimation and fitting described in Section 3 were repeated for the 2001 data. Fig. 8 shows a comparison of the 2001 and 2009 power spectra which are in broad agreement in shape and overall (fractional) amplitude. The best-fitting bending power-

---

Figure 7. Upper limit on the strength of a QPO, assumed to have a Lorentzian spectrum, defined for an $\alpha = 0.01$ test. The solid curve was calculated assuming a $Q = 10$ QPO and the dotted curve was calculated assuming $Q = 5$. In each case, the three curves show (from bottom to top) the upper limits corresponding $\beta = 0.5$, 0.9 and $>0.995$ powers. The detection power at each point in the plane was computed using 200 Monte Carlo simulations of data with a QPO present and recording the fraction of simulations that gave a detection (using a $\Delta X^2 \geq 12.28$ detection criterion for $Q = 10$). See Section 5 and Appendix B for details.

β on the $v_0$–$R$ plane define the upper limit on the strength $R$ of a QPO as a function of frequency $v_0$. Fig. 7 shows the contours for $\beta = 0.5, 0.9$ and $>0.995$ assuming $Q = 10$ (solid curves) and $Q = 5$ (dotted curves). The middle of each set of curves represents the weakest QPO that could be present (at each frequency) and have a probability of $>0.9$ of being detected. This limit extends below $R = 5$ per cent for $v \lesssim 3 \times 10^{-4}$ Hz and below $R = 2$ per cent for $v \gtrsim 3 \times 10^{-3}$ Hz. The shape and amplitude of this curve are similar to those obtained using the simpler $\Delta X^2$ confidence contour method used previously by Vaughan & Uttley (2005).

Figure 8. Comparison of 2009 and 2001 power spectrum estimates. Upper panel: the estimated power spectra computed for the 2001 observation (blue histogram) and the combined 2009 observations (black; same as Fig. 2). Lower panel: ratio of the two estimates.
For the 2009 data, we found a very similar high-frequency shape and it is therefore consistent with this identification. Ignoring any dependence and simply scaling frequencies inversely with mass \(v_{\text{break}} \propto 1/M_{\text{BH}}\), the bend frequency for NGC 4051 corresponds to \(\sim 34\) Hz in a typical GBH. This is close to the \(v_{\text{break}} \approx 23\) Hz estimated from a soft state observation of Cygnus X-1 by McHardy et al. (2004).

The state analogy is supported by radio observations. At the highest available resolution, the source is resolved into a core and two oppositely directed sources separated from the core by \(\sim 0.4\) arcsec (Ulvestad & Wilson 1984; Kukula et al. 1995; Ho & Ulvestad 2001; Giroletti & Panessa 2009). This structure is similar to the core and hotspots commonly seen in Fanaroff–Riley II radio galaxies, although no radio jet has been detected in NGC 4051. Jones et al. (2011) recently used radio monitoring observations and found no evidence for strong variability of the core flux. Despite the presence of these radio sources, Jones et al. (2011) showed that NGC 4051 falls below the ‘Fundamental Plane’ connecting the radio and X-ray luminosities to black hole mass in radio-loud AGN and hard-state GBHs (Merloni et al. 2003; Falcke et al. 2004; Körding et al. 2006), i.e. it is less radio-loud than would be expected for a hard state or radio-loud system (see fig. 14 of Jones et al. 2011). This is consistent with the ‘soft state’ interpretation of NGC 4051 mentioned above.

The \(\text{XMM–Newton}\) data reveal the power spectrum of NGC 4051 up to frequencies as high as \(\gtrsim 10^{-2}\) Hz (see Figs 2 and 4), a factor of \(\sim 40\) above the bend frequency \(v_{\text{break}}\). Assuming a black hole mass as above, the gravitational radius is \(r_g = GM/c^2 \approx 2.5 \times 10^6\) km, and the period of the ISCO should be \(\sim 100-800\) s, depending on the dimensionless spin parameter, \(j\) (see e.g. van der Klis 2006, and references therein). The power spectrum of NGC 4051 above \(v_{\text{break}}\) shows no indication of deviating from a power law up to the highest frequencies (but see Section 7.3) and in particular shows no signature of the ISCO period, which might be expected to produce an excess of variability power due to Doppler boosting of ‘hotspots’ (Syunyaev 1973; Revnivtsev, Gilfanov & Churazov 2000). There is no sign of the additional power spectral break seen at very high frequencies in the hard state of Cygnus X-1 by Revnivtsev et al. (2000). Indeed, assuming linear scaling of time-scales with black hole mass, we detect power at frequencies corresponding to \(\sim 1.7\) kHz for a GBH, far above the fastest variations yet detected from these sources (Revnivtsev et al. 2000; Strohmayer 2001).

7 Above the break the power spectrum of NGC 4051 is sufficiently steep, \(\alpha_{\text{high}} < -2\), such that the integrated power in the power spectra of both \(L_X\) and \(dL_X/dt\) converges at high frequencies. The rate of change of the luminosity may be limited by the radiative efficiency of accretion (Fabian 1979). The power spectrum is therefore not required to become steeper at even higher frequencies.

7.2 Stationarity and the existence of ‘states’

The location of the power spectrum bend frequency is lower than that reported by McHardy et al. (2004) from their analysis of the 2001 \(\text{XMM–Newton}\) data. The 90 per cent confidence intervals from these two analyses are \([5, 12] \times 10^{-4}\) mHz and \([1.4, 3.3] \times 10^{-4}\) mHz – do not overlap, although some of these may be due to the interval reported by McHardy et al. (2004) being an underestimate (see Mueller & Madejski 2009). The difference between the two results may in part be due to differences in the data reduction and power spectrum estimation techniques; the reanalysis of the 2001 data presented here (Section 6) gives a lower \(v_{\text{break}}\) estimate and shows the difference between the two observations to be only moderately statistically significant (compare also the Bayesian posterior distributions of Fig. 3). If this is a real effect, it is different from the expected scaling of \(v_{\text{break}}\) with the accretion rate in the central region implied by the McHardy et al. (2006) relation, since the X-ray luminosity in the month leading up to the 2009 observation (based on the \(\text{RXTE}\) monitoring) was \(\sim 20\) per cent higher than in the month leading up to the 2001 observations, but the bend frequency was \(\sim 60\) per cent lower.

The separation of \(\sim 8\) yr between the two observations corresponds to a separation of only \(\sim 30\) min for a 10-M\(_{\odot}\) GBH; on these time-scales, GBH power spectra are approximately stationary except during very rapid state transitions. By the same linear scaling, each of the \(\sim 40\)-ks \(\text{XMM–Newton}\) observations is equivalent to \(\sim 0.2\) s, and the 45-d span of the observations covers the equivalent of only \(\sim 20\) s for a typical GBH. Unless NGC 4051 is drastically more non-stationary than most GBHs, we might therefore expect all the \(\text{XMM–Newton}\) observations to be representative of almost exactly the same state and variability properties.

During both the 2001 and 2009 observations the variability does, to a very good approximation, follow a single rms–flux relation over the full flux range (see Fig. 5) and the shape of the power spectrum remains roughly the same at different flux levels (see Fig. 4). The lack of variability during low flux intervals (e.g. rev1739) and the extreme variability during high flux intervals (rev1730) are opposite extremes of this relation. These apparently different behaviours are therefore not the result of distinct variability states, despite their very different energy spectra (see Fig. 6). (See also Uttley et al. 2003.) The variations in energy spectral shape are also continuous, as discussed by Vaughan et al. (in preparation).

7.3 Absence of QPOs

We found no strong evidence (using a \(\alpha = 0.01\) test) for QPOs (additional, narrow components) in the 2009 power spectrum, and we place limits of \(R \lesssim 2\) per cent rms on the strength of a QPO in the \(\gtrsim 3 \times 10^{-3}\) Hz range and \(\lesssim 5-10\) per cent over the \(10^{-4}-3 \times 10^{-3}\) Hz range. The \(10^{-3}-0.1\) Hz bandpass over which these data are sensitive to weak QPOs corresponds to \(\gtrsim 170\) Hz for a GBH, where HF QPOs are often found, and the upper limit of \(R \sim 2\) per cent on the QPO strength is comparable to the strength of typical HF QPOs (Remillard et al. 2003). Such HF QPOs have been detected in GBHs only in intermediate states (particularly the soft intermediate state; Belloni 2010), but are not always present (at least not at detectable levels), and when detected may be strongly energy dependent. The lack of detection in NGC 4051 therefore does not constitute strong evidence against an analogous state in this AGN.

Despite not being formally significant in an \(\alpha = 0.01\) test, the best-fitting narrow Lorentzian QPO component has a frequency of \(v_{Q} = 4 \times 10^{-3}\) Hz and rms \(R \sim 1\) per cent. This frequency is very close to that found previously by Vaughan & Uttley (2005) from fitting the power spectrum of the 2001 data. This may be a coincidence in the random sampling fluctuations around a smooth continuum power spectrum. But it is around this frequency that the signal-to-noise ratio in the data is highest, and so model fitting will

---

7 Not in the sense normally applied to GBHs where different states are often characterized by distinct power spectral shapes and amplitudes as well as different energy spectra. See van der Klis (2006), Remillard & McClintock (2006) and McClintock & Remillard (2006).
be most sensitive to small (~20 per cent) deviations from the smooth power-law model. The high-frequency power law may be only an approximation to a more complex continuum power spectrum, the structure of which is on the edge of detectability in these data.

It is well known that HF QPOs in GBHs can be strongly energy dependent, and frequently stronger at higher energies, and in fact we did find that the higher energy band (2–10 keV) gave slightly stronger evidence for a possible QPO. But this could also indicate that this energy band contains a stronger contribution from a second, variable emission component with a different power spectrum from that which dominates the softer energy bands.

7.4 Quasi-constant emission component

The flux offsets in the energy-resolved rms–flux relations, C(E), reveal the spectrum of a quasi-constant component (QC) of the emission. A simple model comprising a single power law (Γ = 3.06 ± 0.13) modified by Galactic absorption (Elvis, Wilkes & Lockman 1989) gave a very good fit, with χ^2 = 10.65 for 18 d.o.f.: A Comptonizing plasma model (COMPTT) with optical depth τ = 1 and a seed photon energy of ∼50 eV, which has an approximately power-law shape, also gave a good fit (χ^2 = 9.45 for 17 d.o.f.). The simplest interpretation is that this represents an emission component that does not vary at all during the 2009 observations, emission that remains if the highly variable emission component(s) drops to zero flux. We note that the Chandra data taken during an extended low flux period in 2001 and the XMM–Newton observation taken during a low flux period in 2002 were previously interpreted in terms of an unresolved (∼100 pc), non-varying soft spectral component and a strongly variable, higher spectral component (Uttley et al. 2003, 2004). However, the QC emission is not required to be absolutely constant over the observations; it is only required to have a far lower (absolute) amplitude (at least 1 mHz) and be weakly correlated (or uncorrelated) with the highly variable component that dominates brighter flux spectra (e.g. rev1730) and drives the rms–flux relation.

The spectral shape and absolute level of this component are similar to the lowest flux spectrum taken during 2009 (rev1739) below ∼2 keV, as shown in Fig. 6. This strongly suggests that most of the soft X-ray continuum emission observed during low flux periods is due to the soft QC component. This may be the tail of the spectrum of inverse-Compton scattering of optical–UV continuum photons in shock-heated gas where the ionized outflow collides with the interstellar medium of the host galaxy, as recently predicted to occur in AGN by King (2010). Indeed, the velocity–ionization structure of the outflow as resolved by the RGS data provides further evidence to support this scenario (Pounds & Vaughan 2011).

ACKNOWLEDGMENTS

We thank the referee, Iossif Papadakis, for a thorough and constructive report. PU is supported by an STFC Advanced Fellowship, and funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement number ITN 215215 ‘Black Hole Universe’. This research has made use of NASA’s Astrophysics Data System Bibliographic Services and the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This paper is based on observations obtained with XMM–Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and the USA (NASA).

REFERENCES

Gelman A., Meng X.-L., Stern H. S., 1996, Statistica Sinica, 6, 733
Gierliński M., Middleton M., Ward M., Done C., 2008a, Nat, 455, 369
McHardy I. M., Koerding E., Knigge C., Uttley P., Fender R. P., 2006, Nat, 444, 730

© 2011 The Authors, MNRAS 413, 2489–2499
Terashima Y. et al., 2009, PASJ, 61, 299

APPENDIX A: CONSTRUCTION OF ‘FLUXED’ SPECTRA

X-ray spectra are conventionally given in terms of count rates per detector channel. These ‘raw’ spectra are of limited use for data visualization because of the distorting effect of (i) the energy-dependent efficiency of the telescope and detector (described by the effective area function) and (ii) blurring by the line spread function of the instrument (modelled by the redistribution matrix). The true source spectrum \( S(E) \) (in units of photon s\(^{-1}\) cm\(^{-2}\) keV\(^{-1}\)) is related to the observed count rate in channel \( i \) by a Fredholm integral equation:

\[
C(i) = \int R(i, E) A(E) S(E) dE,
\]

where \( A(E) \) is the instrument effective area (cm\(^2\)) and \( R(i, E) \) gives the probability that a photon of energy \( E \) will be recorded in channel \( i \) (i.e. the redistribution function). Here the observed spectrum \( C(i) \) is assumed to be background-subtracted, and the instrumental response defined by \( R(i, E)A(E) \) is assumed to be linear, that is, independent of the value of \( S(E) \) which is a good approximation so long as pile-up is negligible. This integral equation can be approximated using a discrete formulation in which the energy range is divided into \( N \) bins \( E_i \):

\[
C_i = \sum_{j=1}^{N} R_{ij} A_j S_j.
\]

where \( R_{ij} \) is the average and \( A_j \) and \( S_j \) are the definite integrals of their continuous counterparts, over the energy range of each bin \( E_j \rightarrow E_{j+1} \).

In general, methods that attempt to invert the above equations to recover \( S(E) \) from an observed spectrum \( C_i \) are unreliable and one must turn to the ‘forward fitting’ method to build a model of the spectrum (see Arnaud 1996). However, it is often useful to view even an approximate ‘fluxed’ spectrum, largely free from the distorting effects of the instrument efficiency, if only to give a better visual representation of the underlying source spectrum. Nowak (2005) discussed a simple method to accomplish this in which the observed count spectrum is normalized by an effective area curve that has been blurred using the redistribution function:

\[
\tilde{S}_i = \frac{C_i}{\sum_{j=1}^{N} R_{ij} A_j}.
\]

This method matches the blurring of the effective area to the blurring of the observed spectrum to give an estimate of the spectrum in true flux units (e.g. photon s\(^{-1}\) cm\(^{-2}\)) that is, to a large extent, free from the distorting effects of the energy-dependent effective area. This process is essentially equivalent to calibrating the sensitivity using an observation of an ideal standard star (with a perfectly flat spectrum).\(^9\) The resulting spectrum can be easily converted to conventional flux density units (e.g. erg s\(^{-1}\) cm\(^{-2}\) keV\(^{-1}\)).

It is crucial that the effective area function \( A_j \) be blurred by the redistribution matrix. The reason is that the effective area function may contain abrupt changes in efficiency (e.g. near photoelectric edges in the detector-mirror system, such as the K-edges of O and Si or the M-edge of Au) that can be far sharper than the resolving power of the detector, which means that normalizing the observed (blurred) spectrum by the exact (not blurred) effective area introduces very strong spurious features into the spectrum near the edges. Blurring the effective area curve matches the resolution to that of the data and suppresses this effect. The result is a spectrum with the energy dependence of the detector/mirror efficiency removed (to a reasonable approximation), but which remains at the detector spectral resolution. We emphasize that this is not, even approximately, a deconvolved spectrum. It has the advantage over the popular ‘unfolding’ method (in xspec) of being independent of the spectral model and therefore providing a more objective visualization ‘fluxed’ spectrum.

\(^9\)This can be done trivially within xspec by defining a constant model spectrum \( M(E) = 1.0 \) (e.g. a power law with index 0 and normalization 1) and plotting the ‘unfolded’ spectrum. For the special case of a constant model, this is equivalent to equation (A3). The ‘unfolded’ plot shows the model in flux units (=1) multiplied by the ratio of the observed spectrum to the ‘folded’ model, i.e. \( M(E) \times C_i / \sum_{j} R_{ij} A_j M_j = C_i / \sum_{j} R_{ij} A_j \). The SAS task efluxer can also perform this transformation.

© 2011 The Authors, MNRAS 413, 2489–2499
**APPENDIX B: QPO DETECTION AND UPPER LIMITS**

This appendix presents first an outline of the method used to search for QPOs in the power spectral data and secondly how upper limits were derived based on the detection method.

**B1 Calibrating the QPO detection procedure**

An $\alpha$-threshold detection procedure uses a test statistic $T$ to quantify evidence for a detection and is calibrated such that a detection occurs when $T(x) > T_{\text{crit}}$ for some threshold value $T_{\text{crit}}$ chosen to have a small probability $\alpha$ of occurring by chance when $H_0$ is true. Specifically, the detection threshold is chosen such that $\Pr(T > T_{\text{crit}}|H_0) \leq \alpha$, where $\alpha$ is the chance of a ‘false detection’ (a ‘Type I error’ when $H_0$ is true). Commonly used values are $\alpha = 0.05$ and 0.01. But in order to calculate $T_{\text{crit}}$ we need to know the reference distribution of the test statistic on the assumption that the null hypothesis is true, $p(T|H_0)$. In general, this may be estimated by calculating the distribution of the test statistic from a large number of Monte Carlo simulations of data generated assuming $H_0$ and distributed as expected for realistic data.

The above would be valid if the null hypothesis was simple, that is, did not contain any unknown (free) parameters. In the present case, the null hypothesis does involve parameters estimated using the observed data, $\theta_C$. We account for the uncertainties in their values by using random parameters drawn from the posterior distribution to generate the simulated data, using the MCMC discussed in Section 3. This gives the distribution of the statistic known as the posterior predictive distribution (see Rubin 1984; Meng 1994; Gelman, Meng & Stern 1996; Protassov et al. 2002; Gelman et al. 2004; Vaughan 2010):

$$p(T|H_0, x^{\text{obs}}) = \int p(T|H_0, \theta_C)p(\theta_C|x^{\text{obs}})d\theta_C. \quad (B1)$$

In order to determine $T_{\text{crit}}$, this distribution was mapped out using Monte Carlo simulations as follows.

1. Repeat for $k = 1, 2, \ldots, N$ simulated data sets.
   - (i) Draw a set of parameter values from the posterior $\theta_C^k$.
   - (ii) Simulate data from $H_0$ with parameters $\theta_C^k$: $x^k$.
   - (iii) Fit $H_0$ to get $\chi^2(x^k,H_0)$.
   - (iv) Fit $H_1$ to get $\chi^2(x^k,H_1)$.
   - (v) Compute $T^k$.
2. Use the distribution of $T^k$ to define $T_{\text{crit}}$ for which $\Pr(T > T_{\text{crit}}|H_0) \leq \alpha$.

We used $N = 3000$ simulations to calibrate $T_{\text{crit}}$ and found that the critical value for an $\alpha = 0.01$ test was $T_{\text{crit}} = 12.28$. (This was calculated using a $Q = 10$ QPO, but using $Q = 5$ gave a very similar result for the data and models discussed in this paper.) In order to ensure that the global minimum was found in each case (this can be tricky for $H_1$ which may produce multiple minima in $\chi^2$), we used the robust automated fitting scheme discussed by Hurkett et al. (2008, section 3.2.2)

**B2 Upper limit procedure**

We wish to determine the upper limit $U$ on the QPO strength $R$, which is the smallest value of $R$ such that the probability of detecting the QPO (assuming that it is present, i.e. $H_1$) is $\geq \beta_{\text{min}}$ for some specific $\beta_{\text{min}}$. See Kashyap et al. (2010) for a very clear introduction to the meaning and calculation of upper limits in astronomy.

In the present case, we are searching for a QPO of an unknown location (frequency) but the upper limit may vary as a function of this location – because the ‘background’ continuum level and the uncertainty on the spectrum vary as a function of frequency – so we must calculate the limit for a set of different QPO locations $\nu_Q$. The probability of a detection, $\beta$, which is $1 - (\text{probability of ‘Type II error’})$, is

$$\beta(\nu_Q, R) = \Pr(T \geq T_{\text{crit}}|H_1, \nu_Q, R) \quad (B2)$$

and the upper limit $U(\nu_Q)$ is the smallest value of $R$ which satisfies

$$\beta(\nu_Q, R) \geq \beta_{\text{min}} \quad (B3)$$

for some chosen minimum power $\beta_{\text{min}}$.

Given the $T_{\text{crit}}$ defined above, we can compute the upper limit using Monte Carlo simulations of realistic data under $H_1$. These can be used to map out the distribution $p(T|H_1, \nu_Q, R)$ for different $\nu_Q$ and $R$ values and hence find $U(\nu_Q)$ that satisfies the above equations. In these terms, the upper limit is the strength of the weakest QPO that would have a reasonable probability $\beta \geq \beta_{\text{min}}$ for some specific $\beta_{\text{min}}$ such as 0.5 or 0.9.

This method will, in general, be sensitive to the parameters of the continuum model, $\theta_C$, which of course we do not know exactly. But, as before, we can average over the posterior distribution of these as derived from the data $p(\theta_C|x^{\text{obs}})$:

$$p(T|H_1, \nu_Q, R) = \int p(T|H_1, \nu_Q, R, \theta_C)p(\theta_C|x^{\text{obs}})d\theta_C. \quad (B4)$$

Again, this can be done by taking values from the posterior using the MCMC discussed in Section 3.

The calculation of $\beta$ for a given pair of QPO parameters ($\nu_Q$ and $R$) was performed using a Monte Carlo scheme as follows.

(1) Loop over $k = 1, 2, \ldots, N$ simulated data sets.
   - (i) Draw a set of parameter values from the posterior $\theta_C^k$.
   - (ii) Simulate data from $H_1$ with continuum parameters $\theta_C^k$ and QPO parameters ($\nu_Q, R$): $x^k$.
   - (iii) Fit $H_0$ to get $\chi^2(x^k, H_0)$.
   - (iv) Fit $H_1$ to get $\chi^2(x^k, H_1)$.
   - (v) Compute $T^k$.
2. Use distribution of $T^k$ to find $\beta(\nu_Q, R) = \Pr(T > T_{\text{crit}} | H_1, \nu_Q, R)$.

This process is repeated over a grid of QPO parameters, $\nu_Q$ and $R$, and at each frequency $\nu_Q$ the smallest $R$ for which $\beta(\nu_Q, R) \geq \beta_{\text{min}}$ is the upper limit $U(\nu_Q)$. For the present analysis we used 30 logarithmically spaced values of the QPO location $\nu_Q$, 40 logarithmically spaced values of the QPO strength $R$ and, at each pair generated $N = 200$ simulations. This is a sufficient number of simulations to accurately define contours of e.g. $\beta = 0.9$ on the $\nu_Q$–$R$ plane.

The detection procedure was calibrated using simulated data generated assuming $H_0$. But to evaluate $\beta$, and hence the upper limit, the data were simulated assuming $H_1$. The former is concerned with the chance of spurious detection assuming that no QPO is present; the latter is concerned with the chance of detection when a QPO is present.

This paper has been typeset from a \LaTeX file prepared by the author.