DETAILED SPECTRAL ANALYSIS OF THE 260 ks XMM-NEWTON DATA OF 1E 1207.4–5209 AND SIGNIFICANCE OF A 2.1 keV ABSORPTION FEATURE

KAYA MORI,1 JAMES C. CHONKO,2 AND CHARLES J. HAILEY2

Received 2004 July 16; accepted 2005 June 8

ABSTRACT

We have reanalyzed the 260 ks XMM-NEWTON observation of 1E 1207.4–5209. There are several significant improvements over previous work. First, a much broader range of physically plausible spectral models was used. Second, we have used a more rigorous statistical analysis. The standard $F$-distribution was not employed, but rather the exact finite statistics $F$-distribution was determined by Monte Carlo simulations. This approach was motivated by the recent work of Protassov and coworkers and Freeman and coworkers. They demonstrated that the standard $F$-distribution is not even asymptotically correct when applied to assess the significance of additional absorption features in a spectrum. With our improved analysis we do not find a third and fourth spectral feature in 1E 1207.4–5209 but only the two broad absorption features previously reported. Two additional statistical tests, one line model independent and the other line model dependent, confirmed our modified $F$-test analysis. For all physically plausible continuum models in which the weak residuals are strong enough to fit, the residuals occur at the instrument Au M edge. As a sanity check we confirmed that the residuals are consistent in strength and position with the instrument Au M residuals observed in 3C 273.

Subject headings: methods: statistical — stars: neutron — X-rays: individual (1E 1207.4–5209)

Online material: color figure

1. INTRODUCTION

The thermal emission from isolated neutron stars (INSs) probes neutron star physics, since it is uncontaminated by emission from an accretion disk. In particular, many INSs (age $\sim 10^3$–$10^5$ yr) are good candidates because they are hot enough that the thermal emission is not obscured by nonthermal magnetospheric emission. Lines in thermal spectra are an excellent tool for probing the strong magnetic field and gravity of an INS.

Two broad absorption features were discovered in the INS 1E 1207 by the Chandra X-Ray Observatory (Sanwal et al. 2002) and later confirmed by XMM-NEWTON (Mereghetti et al. 2002). The discovery was followed by the detection of a single spectral feature in the X-ray thermal spectra of three nearby INSs (Haberl et al. 2003, 2004; van Kerkwijk et al. 2004). 1E 1207 still remains unique because it shows more than one spectral feature, as opposed to the other three INSs, which only show a single absorption feature.

The interpretation of features in neutron star (NS) thermal spectra is not straightforward because of effects such as magnetic field, gravity, and unknown surface composition. Soon after the discovery of features in the 1E 1207 spectrum, several different interpretations were proposed, such as helium atomic lines at $B \sim 10^{14}$ G (Sanwal et al. 2002), iron atomic lines at $B \sim 10^{12}$ G (Mereghetti et al. 2002), oxygen/Ne atomic lines at $B \sim 10^{12}$ G (Hailey & Mori 2002; Mori & Hailey 2003), and hydrogen molecule ion lines at $B \sim 10^{14}$ G (Turbine & López Vieyra 2004). Recently, Bignami et al. (2003, hereafter B03) proposed electron cyclotron lines based on the detection of two additional absorption features at 2.1 and 2.8 keV in the 260 ks XMM-NEWTON observations. They interpreted these four absorption lines as the fundamental and three harmonics of cyclotron lines. De Luca et al. (2004, hereafter DL04) reanalyzed the same XMM-NEWTON data and also found that the third and fourth features are significant.

In this paper we present our detailed analysis of the same 260 ks XMM-NEWTON observation and investigate the significance of the two additional absorption features at 2.1 and 2.8 keV.

2. DATA REDUCTION METHODOLOGY

XMM-NEWTON observed the INS 1E 1207 for two orbits starting on 2002 August 4. The pn camera was operated with a thin filter in the small-window mode. This allows accurate timing (6 ms time resolution) and minimizes pileup. The two MOS cameras were operated in full-frame mode with a thin filter. Our data reduction followed the latest XMM-NEWTON calibration document (Kirsch 2004), and we processed the data by running the pipeline directly from the opacity distribution function (ODF) files with SAS, version 5.4.1.3 We used canned response matrices for all the instruments and generated ancillary files by the SAS command arfgen. We also generated a response function using the SAS command rmfgen, but the difference in $\chi^2$ and fit parameters was tiny. Following Kirsch (2004), we generated a light curve of high-energy (>10 keV) single-pixel (PATTERN $= 0$) events on the full field of view and selected good time intervals using the criteria for count rates $< 1.0$ counts s$^{-1}$ (pn) and 0.35 counts s$^{-1}$ (MOS). The total exposure time for the pn and MOS after removing time intervals with high background is 156 and 231 ks, respectively.

We extracted the source photons from a 45" radius circle centered on the source for all cameras. There are sufficient photons in the pn camera to enable selection of only single events to achieve the best spectral resolution. Most of our results are based on this higher quality singles data; however, later we discuss some results from the doubles analysis. The MOS cameras have

\footnotesize

1 Canadian Institute for Theoretical Astrophysics, 60 Saint George Street, Toronto, ON M5S 3H8, Canada; kayam@cita.utoronto.ca.
2 Columbia Astrophysics Laboratory, 550 West 120th Street, New York, NY 10027; chonko@astro.columbia.edu, chuckh@astro.columbia.edu.

3 Since submission of the manuscript, we have confirmed that our results remain the same with the more recent SAS, version 6.0.
fewer photons, so we used all events in our analysis to improve the photon statistics.

Background regions were selected according to the recommendations of the XMM-Newton calibration team (Kirsch 2004). In the pn camera we chose a 45° radius circle background region the same distance from the readout node as the source region to ensure similar noise. An annular background region encircling the source was not used to avoid out-of-time events from the source. In the MOS cameras we used a circular region away from the source according to the calibration team’s recommendations. We also tried an annular region surrounding the source and found no significant deviations in the continuum fit parameters. The source is located in a supernova remnant, and nonuniform X-ray emission from the remnant could influence the NS spectra. To validate our choice of background region, we chose three different background regions for each instrument and confirmed that the spectral fit parameters of the source did not significantly change with our choices. As a result, source count rates in 0.2–5.0 keV are 1.546 ± 0.003 counts s⁻¹ (pn single+double), 0.346 ± 0.001 counts s⁻¹ (MOS1), and 0.353 ± 0.001 counts s⁻¹ (MOS2). Background count rates in the energy band are 0.089 ± 0.001 counts s⁻¹ (pn single+double), 0.0065 ± 0.0005 counts s⁻¹ (MOS1), and 0.0069 ± 0.0005 counts s⁻¹ (MOS2).

The spectra were rebinned such that each bin contained at least 40 counts and analyzed using XSPEC, version 11.3.1 in the 0.3–4.0 keV energy range. Modest oversampling of the energy resolution kernel was used. Our number of bins and degrees of freedom are comparable to those of other analyses of this source. We adopted the energy band between 0.3 and 4.0 keV for the pn instrument. For the MOS instruments we selected the 0.5–4.0 keV band because the MOS is not well calibrated at lower energies. Continuum fit parameters for the pn and MOS instruments were in good agreement.

The long XMM-Newton observation permits splitting the full observation in order to search for time variability. We split the observation into four approximately equal time intervals and found no deviations between the four observations. They yielded statistically consistent continuum fit parameters. We conclude that there was no time dependence during the total observation, and therefore we use the data from the entire observation for all subsequent analysis.

3. ANALYSIS OF DATA USING CONTINUUM MODELS

3.1. Introduction

In this section we consider fits to the global 1E 1207 spectra using a variety of physically motivated spectral models. Most of these models are not compatible with the data and are rejected. For those models that provide good fits to the data, we determine the relevant fitting parameters. The issue of properly fitting additional spectral features is a complex one, and we postpone addressing it until § 4. Here we only set the stage by providing information on the best-fit continuum models.

3.2. Atmosphere Models for Neutron Star Thermal Spectra

The fitted continuum models must reflect the full range of plausible INS atmospheric conditions. The surface of an INS can be any element from hydrogen to iron, since only ~10⁻¹⁹ M☉ of material produces an optically thick atmosphere (Romani 1987). The atmosphere distorts the spectral energy distribution (SED) from Planckian, as will the presence of a magnetic field. For B = 0, hydrogen atmosphere models have the hardest spectrum due to free-free absorption. INS spectra soften and show a proliferation of lines and edges at higher energies, becoming closer to blackbody (Romani 1987). For increasing B and atomic number there is significant spectral softening of the SED (Ho & Lai 2003). Even for hydrogen, electron binding effects can amount to an appreciable fraction of the INS surface kT at a high enough B-field, leading to incomplete ionization and softening of the SED due to opacity from bound species (Ho et al. 2003). From the above considerations we conclude that realistic atmospheric models for 1E 1207 will have an SED somewhere between a blackbody and a fully ionized hydrogen model, and our subsequent fitting reflects this range of SEDs.

3.3. Description of Model Fits to the Data
(No Third or Fourth Line Assumed)

In the following few sections we describe our fit to the data without the assumption of a third or fourth line in the spectrum. Fits with such additional lines are discussed in § 4. Therefore, our “baseline” models always consist of the two universally accepted lines in the 1E 1207 spectrum, which we model as Gaussian absorption lines at 0.7 and 1.4 keV.

For each continuum component we adopt one of the following three physically plausible models: blackbody (BB), magnetized hydrogen atmosphere (HA), or power law (PL). The magnetized hydrogen atmosphere model was constructed assuming a fully ionized hydrogen atmosphere at B = 10¹² G, close to the dipole field strength as measured from the spin-down parameters (Pavlov et al. 1995). We also tried B-fields as low as B = 0 (Zavlin et al. 1996), which did not alter the continuum parameters significantly. Higher B-fields (≥4.4 × 10¹³ G) just give results closer to the blackbody case (Ho & Lai 2003).

In our first fits to the spectrum we used one continuum component along with the two absorption lines. For all instruments the one-component models did not yield acceptable χ² values (χ² > 1.3) and showed significant residuals above ~2 keV. This suggested the need for another continuum component.

Given two continuum components and two absorption lines, there are three types of continuum models, defined as below. They consist of two continuum components (C1 and C2) and two absorption lines (L1 and L2). Here C1/L1 and C2/L2 refer to the lower and higher energy spectral component, respectively. The three classes of models considered are as follows (with the notation C* L meaning line L resides on continuum component C):

Model I : ‚ C1L1 + C2 + L2,
Model II : ‚ C1L1 + C2L2,
Model III : ‚ (C1 + C2)L1 + L2.

For L1 and L2 the Gaussian absorption line was modeled as

\[ F(E) \propto \exp \left\{ -\tau \exp \left[ -(E - E_0)^2 / 2\tau^2 \right] \right\}, \]

where τ and w refer to the line depth and width, respectively. We fitted other absorption-line profiles, such as a Lorentzian, but the results were similar. While the exact shape of the absorption features may be asymmetric or have substructure from blended lines (Mori & Hailey 2003), we obtained excellent fits without invoking asymmetric profiles. Thus, we chose simple Gaussian line shapes. When we fitted photoabsorption edges to the two absorption features at 0.7 and 1.4 keV, the χ² value was not as good as for the Gaussian lines in any of the continuum models.

3.3.1. Physical Meaning of Models I, II, and III

Interpretation of spectral features in NS thermal spectra is not straightforward, since the line parameters depend on various
NS parameters, such as surface element, ionization state, magnetic field strength, and gravitational redshift (Hailey & Mori 2002; Mori & Hailey 2003). Therefore, it is important whether the observed spectral features originate from the same region or layer.

The presence of two continuum components complicates our interpretation of the observed absorption features. We assume that the two continuum components originate from different regions on the NS surface. For instance, C1 is emitted from a large area on the surface and C2 is emitted from a hot polar cap (DL04).

Model I assumes that L1 and L2 are from the same region as the emission of C1 and model II assumes that they are from different regions. Model III assumes that L1 and L2 are from a layer above the two regions emitting continuum photons. Several physical models predicting a layer made of electron-positron or electron-ion pairs a few NS radii above the NS surface have been proposed (Dermer & Sturle 1991; Wang et al. 1998; Ruderman 2003). Such a layer may become optically thick and modify the thermal spectra from the NS surface. In addition, absorption in model III can take place anywhere between the NS and the observer (e.g., magnetosphere, supernova remnant, interstellar medium, or materials on the XMM-Newton mirrors and cameras).

3.4. Results of Continuum Model Fitting to XMM-Newton Data

We searched for continuum models that are consistent with the data using models I, II, and III. The set of models we considered is extensive, including all permutations of blackbody (BB), magnetized hydrogen atmosphere (HA), and power law (PL) models for C1 and C2.

All the two-component thermal models (and combinations of BB and HA) fitted the data well, yielding $\chi^2$/dof very close to unity (Tables 1 and 2). For brevity we do not consider mixed cases (BB+HA) because tests showed that in assessing absorption-line significances ($\chi^2$) they always produced results intermediate between the pure BB and pure HA cases. Continuum models with PL components are ruled out because they do not adequately fit

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Parameter & \multicolumn{2}{c|}{MOSI III} & \multicolumn{2}{c|}{MOSII III} & \multicolumn{2}{c|}{MOSII III} \\
 & (BB+BB) & (HA+HA) & (BB+BB) & (HA+HA) & (BB+BB) & (HA+HA) \\
\hline
$N_{\text{H}}$ (10$^{22}$ cm$^{-2}$) & 0.05$^{+0.10}_{-0.04}$ & 0.176$^{+0.025}_{-0.017}$ & 0.076$^{+0.074}_{-0.051}$ & 0.182$^{+0.011}_{-0.010}$ & 0.123$^{+0.018}_{-0.020}$ & 0.204$^{+0.021}_{-0.030}$ \\
$kT_1^\alpha$ (keV) & 2.0$^{+0.3}_{-0.2}$ & 2.01$^{+0.027}_{-0.016}$ & 2.0$^{+0.3}_{-0.2}$ & 2.01$^{+0.027}_{-0.016}$ & 2.0$^{+0.3}_{-0.2}$ & 2.01$^{+0.027}_{-0.016}$ \\
$R_1$ (km) & 3.4$^{+1.2}_{-1.0}$ & 20.3$^{+2.7}_{-3.7}$ & 3.4$^{+1.2}_{-1.0}$ & 20.3$^{+2.7}_{-3.7}$ & 3.4$^{+1.2}_{-1.0}$ & 20.3$^{+2.7}_{-3.7}$ \\
$kT_2^\alpha$ (keV) & 0.374$^{+0.074}_{-0.060}$ & 0.41$^{+0.11}_{-0.06}$ & 0.373$^{+0.056}_{-0.039}$ & 0.41$^{+0.11}_{-0.06}$ & 0.373$^{+0.056}_{-0.039}$ & 0.41$^{+0.11}_{-0.06}$ \\
$R_2$ (km) & 0.39$^{+0.23}_{-0.22}$ & 0.49$^{+0.46}_{-0.38}$ & 0.37$^{+0.18}_{-0.16}$ & 0.43$^{+0.27}_{-0.24}$ & 0.37$^{+0.18}_{-0.16}$ & 0.43$^{+0.27}_{-0.24}$ \\
$E_1$ (keV) & 0.70$^{+0.003}_{-0.003}$ & 0.72$^{+0.013}_{-0.013}$ & 0.71$^{+0.014}_{-0.014}$ & 0.72$^{+0.013}_{-0.013}$ & 0.71$^{+0.014}_{-0.014}$ & 0.72$^{+0.013}_{-0.013}$ \\
$E_2$ (keV) & 1.418$^{+0.009}_{-0.010}$ & 1.417$^{+0.010}_{-0.010}$ & 1.403$^{+0.010}_{-0.010}$ & 1.402$^{+0.009}_{-0.010}$ & 1.403$^{+0.010}_{-0.010}$ & 1.402$^{+0.009}_{-0.010}$ \\
$\chi^2$/dof & 0.85 & 0.80 & 1.20 & 1.18 & 139 & 137 \\
\hline
\end{tabular}
\caption{Best-Fit Continuum Parameters to MOS Data}
\end{table}
the data, leaving significant residuals above 2 keV. Models I, II, and III with thermal continuum components are all acceptable.

Figure 1 shows the spectra and residuals for the best fit to model I (BB+BB) using the pn single/double events and MOS1/ MOS2 data. We also fitted the data with a third spectral line, calculated $\chi^2$, and evaluated the difference of $\chi^2$ between the third line case and the no third line case. This is a key parameter in our subsequent statistical tests.

We show our best-fit continuum parameters and 90% confidence levels for six different models in the energy range from 0.3 to 4.0 keV for the pn in Table 1. All the errors quoted are statistical errors, and we did not include any systematic error. We calculated 90% confidence level errors using $S = S_{\text{min}} + 2.706$ for each parameter while allowing the other parameters to vary freely (Lampton et al. 1976, hereafter LMB76). These models yield excellent fits to the data with $\chi^2$ values for the pn slightly less than unity. We found that Gaussian line profiles yield a statistically good fit to the data. The pn blackbody models I, II, and III all give consistent values for these parameters.

We note that for some of our models the line depths (especially for the 1.4 keV line) are not well constrained. This is caused by the presence of two continuum components. If the absorption feature is fixed to reside on one of the continuum components, then the level of that component becomes very important. If the depth of the absorption feature is close to the continuum level, the residuals cannot be adequately fitted, rendering the line depth poorly constrained. It is possible for both of the broad absorption features to lie on the higher temperature continuum component ($C1 + C2L1L2$). However, the fit was not acceptable, leaving significant deviations from the data.

As an aside we note that we took an additional step to confirm the robustness of our continuum parameter fitting and $\chi^2$ determination. We examined a range of degrees of freedom in the pn analysis spanning those previously reported in the literature and consistent with obtaining normal statistics and sensible energy resolution kernel sampling. Both our continuum parameter determination and quality of fit were insensitive to this. However, we emphasize the extremely important technical point that we selected the actual bin size a priori, not by adjusting the size empirically to obtain the lowest $\chi^2$. It is well known that this would result in a fit statistic that is not $\chi^2$ distributed (Eadie et al. 1983). It is important for the subsequent statistical analysis that the conditions for a distribution-free statistic are met, even in the finite-N limit.

3.5. Comparison with and Improvements to Previous Work

In several previous investigations (DL04; B03) the data were fitted with two blackbody components and two absorption lines. In both cases $\chi^2$ significantly greater than 1 were obtained, requiring the inclusion of additional absorption lines to fit some higher energy residuals. Other work did not require such extra lines when fitting either blackbody continua (Sanwal et al. 2002; Hailey & Mori 2002; Mereghetti et al. 2002) or magnetized HA models (Mereghetti et al. 2002).

In order to more thoroughly investigate these deviations in fit quality using just continuum models, we have exploited, in
addition to more up-to-date response matrices and a more time-
consuming study of background effects, a much broader set of
continuum models (models I and II in addition to model III). In
the case of model III, for which our comparison with previous
work is direct, we find agreement at the 90% confidence level for
all derived parameters for both MOS and pn data.

In regard the overall fits, we find acceptable $\chi^2$ ($\sim 1-1.3$) for
all three classes of models, i.e., our fits do not require additional
spectral features. However, introduction of extra lines does
improve the overall $\chi^2$ very slightly. Assessing the necessity of
such additional features is a matter of great statistical delicacy,
which we discuss fully in $\S$ 4.

To facilitate a direct comparison with previous work, we have
used the results of Table 2 in DL04, which indicates the best
continuum model fits (without two additional lines) as giving
$\chi^2 < 1.8$, 1.3, and 1.6 for the pn, MOS1, and MOS2, respec-
tively. While the MOS fits of DL04 are not as good as ours (for
any of our three models), the discrepancy is not really that large;
DL04 only marginally required extra lines in the MOS data.
Only the pn fits appear to have a large discrepancy. One must be
very cautious in assuming that this discrepancy arises from dif-
fences in processing of the data. Such an assumption is belied
by the agreement in overall signal and background count rates in
the MOS and pn, in the MOS and pn fitting parameters, and even
in the overall $\chi^2$ for MOS1/MOS2 data between all groups.
The problem is that in the case of finite count statistics, there can be
substantial deviations in $\chi^2$ from its ideal, distribution-free theo-
retical form (Eadie et al. 1983).

Simulations support this assertion. We generated 1000 sets of
bootstrapped data based on the 1E 1207 data set, fitting this data
set with continuum model III and then comparing the theoretical
and observed $\chi^2$ distributions. It was found that $\chi^2$ of 1.6 or
greater were found more than an order of magnitude more times
than predicted by the standard $\chi^2$ distribution. Thus, there is a
nonnegligible probability (10%) of obtaining a deceivingly
large $\chi^2$ using proper fitting procedures. The difference in $\chi^2$
for the pn data may be nothing more than artifacts of finite-N sta-
tistics. Our results are more immune to this because we fitted so
many test cases. Moreover, such a large $\chi^2$ is not problematic
per se. Rather, what is required is an appropriate finite-N statis-
tical methodology applicable to each reduced data set. An approp-
iate approach has recently been developed, and we present it
and apply it to the 1E 1207 data set in $\S$ 4.3.

4. STATISTICAL TESTS FOR EXTRA
SPECTRAL FEATURES

While it may naively appear that assessing the existence of
and need for an extra spectral feature in a continuum spectrum is
straightforward, this is not true. We present three approaches for
assessing the need for extra spectral features. These include a
goodness-of-fit test for lines in a continuum, direct spectral fit-
ting and significance assessment, and the modified $F$-test.

We have chosen to apply the $F$-test because it has been used in
previous work on 1E 1207. Despite its great familiarity to astro-
physicists, the use of the $F$-test to determine the need for extra
spectral features is extremely problematic. To effect this analysis
we exploit the recent work of Protasov et al. (2002, hereafter
P02) and Freeman et al. (1999, hereafter F99). These authors
have developed approaches that surmount the difficulties in the
standard $F$-test and permit a much more reliable and accurate
determination of feature significance than is possible with the
“standard” $F$-test. Our analysis exactly follows that of P02, so
we emphasize P02 in our brief discussion of theoretical and
practical particulars in $\S$ 4.1. The reader is directed to the above
references for a more thorough treatment. In $\S$ 4.2 and 4.3 we
consider the other methods for attacking this problem.

4.1. F-Test for Detecting Extra Model Components

In the $F$-test and its closely related cousin the likelihood ratio
test (LRT), a test statistic is formed from the distribution function
$f(x|\theta)$ governing the probability of counts $x_i$ in bin $i$ under a set
of parameters $\theta = (\theta_1, \ldots , \theta_p)$. The test statistic is a functional
containing the $f(x|\theta)$ with $\theta$ free to assume any value and also
$f(x|\theta)$, where some parameters are constrained by the values
that are required by the null hypothesis. To be concrete, in the
case of the $F$-statistic the statistic $F = (\chi^2(x, \theta) - \chi^2(x, \theta_0))/\chi^2_0$
(Bevington 1969; P02).

The $F$-test is extremely powerful because it is “distribution
free.” That is, in the asymptotic (large count) limit and under the
null hypothesis (specifically, in our case, $\tau = 0$), it does not
depend on the initial probability distribution, which is the
underlying distribution $f(x|\theta_1, \ldots , \theta_p)$. Rather, the test statistics
only depend on the reference distribution (in our case, the
$F$-distribution). The asymptotic behavior effectively decouples
details such as the shape of the continuum and the particulars of
the line profile and strength from the way the test statistic is
distributed asymptotically. This is why the test is so powerful and
broadly applicable.

P02 and F99 made the crucial observation that there are cer-
tain regularity and topological conditions that must be met for
distribution-free statistics. The regularity conditions are related
to the integrals and derivatives of $f(x|\theta_1, \ldots , \theta_p)$ (see Serfling
[1980] and Roe [1992] for a precise statement) and are of no
interest to us, since these conditions are weak and thus almost
always satisfied in astrophysics modeling. Of profound impor-
tance are the topological conditions (P02). These are (1) the
model under test must be nested, that is, the parameter under test
in the constrained model must be a subset of that parameter in
the unconstrained model; and (2) the parameter under test in the
constrained model cannot be on the boundary of the possible
values of the parameter under the unconstrained test.

Evaluating these topological conditions for equation (1) we
see that condition 1 is met; the constrained model has $\tau = 0$, and
thus the parameterization in the constrained model is a subset of
that in the unconstrained model. However, condition 2 is vio-
lated: $\tau = 0$ is on the boundary of the possible values that $\tau$ can
assume under the unconstrained model. When condition 2 is
violated, the $F$-statistic does not even asymptotically approach
the $F$-distribution. P02 illustrate, through simulations for an ab-
sorption line on a continuum, that the reference distribution over-
estimates the strength of the line by almost an order of magnitude.
P02 properly note that in general nothing can be said about the
significance of a line with the standard $F$-test, because we do not
even in principle know the asymptotic reference distribution. In
this regard, for instance, the XSPEC users’ manual specifically
advises against the use of the $F$-test in assessing the necessity for
extra lines in a spectrum.

4.1.1. Analysis of 1E 1207 Line Significance Using Modified F-Test

Following P02, we seek to more accurately assess the signifi-
cance of the alleged third and fourth lines in 1E 1207 by employ-
ing the $F$-statistic, which we now recognize is not $F$-distributed.
The $F$-statistic is actually a less obvious choice than the LRT, but
it facilitates direct comparison with previous work. Since an approp-
ate asymptotic distribution does not exist, even in principle,
what is now required is the establishment of a reference
distribution for finite statistics. P02 and F99 explain in detail how
to do this. We have determined the reference distribution using
the posterior predictive $p$-value methodology (PPPM) of P02, and the reader is referred there for details. The idea behind PPPM is straightforward, although the computer resources required to implement it are substantial. We only outline the procedure here. Under the null hypothesis, we can use data to generate fake data from which we can determine the relevant reference distribution. However, this standard bootstrapping technique is not correct, since we do not know the "true" values of the parameters to use in the simulation. The failure of the topological condition means we lose conditioning properties that would make it meaningful to replace parameters by best estimators. The solution, as pointed out by P02, is to use parameter values in the simulation that are likely, given the observed data. This simply requires a statistical distribution of parameter values that can be sampled and incorporated into the simulation that generates the distribution of the statistic. P02 advocate a Bayesian approach for determining the probability distribution of each parameter, given the data. PPPM is a specific Bayesian implementation. In effect, it is a fancy statistical bootstrap with a method for explicitly incorporating parameter-fitting errors. The entire approach is completely straightforward to implement in a Monte Carlo by brute force, although clever and efficient non–brute force methods for implementing the Monte Carlo exist (van Dyk et al. 2001).

4.1.2. Details of the Line Search Procedure

We simulated our spectra with parameters determined by the Bayes formula. These spectra are folded through the instrumental response using XSPEC. The actual line search was done by performing a blind search for absorption lines between 0.3 and 4.0 keV. The statistic $F$ was calculated as $F = (\chi_2^2 - \chi_1^2)/\chi_1^2$ (Bevington 1969; P02), where $F$ is obtained from the fit to a continuum model (subscript c) and continuum model plus absorption line (subscript u). A blind search for spectral features in fake spectra in XSPEC is dependent on the initial values and paths (Rutledge et al. 2002). In order to reduce this effect, we started the line search from four different line energies. We may not find all the significant absorption lines when more than four are present in the simulated spectrum. However, such cases, which entail a large $F$-statistic, occur very rarely. To confirm this, we performed Monte Carlo simulations with an increased number of initial starting energies and found that the number of occurrences of $F$ larger than 5 did not change. In searching for a line from four starting locations, we found that sometimes XSPEC fitted the same feature more than once. We corrected for such double-counted events a posteriori.

The distribution of the statistic $F = \Delta\chi^2/\chi^2$ is shown in Figure 2. This distribution was determined for each instrument. In formulating this distribution we recall (§ 4.1) that this corresponds to the distribution of $F$ under the null hypothesis, i.e., that there is no line. This is what is shown in Figure 2, along with the classical $F$-distribution. The value of $F$ for six models is plotted, along with a significance line. We conclude that there are no grounds for rejecting the null hypothesis (no third line). The results are always in the 1–3 \( \sigma \) range and only reach the higher level in a few isolated instances involving blackbody models and pn singles data.

Our results differ from previous work (B03; DL04) for two reasons. First, our $F$-statistic is about 50% smaller due to our better fits to the continuum. Second, the use of the correct finite distribution for the $F$-statistic drops the significance of an extra absorption line by a huge amount, as happened in the likelihood simulations of P02. The net result is a reduction of more than 6 orders of magnitude in the significance of the strongest line found in the pn data.

Further support for our conclusions is presented below, where essentially the same result is arrived at by independent statistical tests.

4.2. Goodness-of-Fit Test for Lines in Smooth Spectra

This type of test is described in numerous papers, but we mainly follow the arguments and notation of Eadie et al. (1983). We test the null hypothesis ($H_0$) that there is no deviation of counts with respect to the continuum level against the alternative ($H_1$) that there is such a deviation. We define a test statistic $T$, and unless $T$ exceeds some critical value $T_c$ then we accept $H_0$. Under $H_0$, $T$ has some distribution $P(T)$. Then $T_c$ can be implicitly defined by the formula $\alpha = \int_{T_c}^{\infty} P(T) dT$, where $\alpha$ is defined by the 4 \( \sigma \) tail of a normal, standard distribution. This procedure ensures that the null hypothesis will not be rejected unless a 4 \( \sigma \) burden of proof is met. This test is binary—unless $T$ exceeds $T_c$, we accept $H_0$. This result can be further quantified by calculating $1 - C = \int_{c}^{\infty} P(x) dx$. Here $1 - C$ can be converted into units of $\sigma$ using the same method as employed for $\alpha$ and provides a rough measure of the significance of the absorption feature.

We need only define our test statistic $T$ and its probability distribution under $H_0$. We are searching for a deficit of counts in a region of the spectrum beyond that predicted by our continuum model. We assume that the region of interest is defined by $K$ energy bins around 2.1 keV. Define $n_i$ as observed counts in bin $i$;
structured additionally show (Appendix) that a new variable can be con-
how close (far) the measured
This is shown in Table 4 and provides a direct comparison of
the random variable
The natural test statistic is
null hypothesis
alternatives defined by

This can be expressed, under suitable conditions (Appendix), as
This test statistic is composed of random variables
This test statistic is composed of random variables

\[ T = \sum_{i=1}^{K} \frac{(n_{i} - \hat{b}_{i})^2}{\hat{b}_{i} + \sigma_i^2}. \]  

4.2.1. Results Applied to 1E 1207 Data

For all the models we considered, the observed \( T \) was much
less than \( T_c \), so we accept \( H_0 \)—i.e., there is no line. Table 3 lists
the observed \( T \) and \( T_c \) for various models. Note \( T_c \) changes
generally because of its dependence on the continuum level and
fit variances within that model. Table 3 also indicates 1 - C,
which gives some feel for the “distance” of \( T \) from \( T_c \) in probability
space. Because \( K \) is sufficiently large in this analysis, we formed
the random variable \( y \) from the observed \( T \), as mentioned above.
This is shown in Table 4 and provides a direct comparison of
how close (far) the measured \( y \) is from \( y_c \), and the “center” of the
distribution \( P(y) \) in units of \( \sigma \).

The formulation above is correct in searching a region where
there is a priori reason to expect a line. There is a complication in
how exactly to determine the \( K \) bins of interest (other than that they are around 2.1 keV). If the spectral feature is narrow and the
region of the \( K \) bins broad, then the feature can be washed out.
This is also true in the converse case. To be conservative, we
checked varying-sized regions of interest around 2.1 keV. The
results were not significantly changed for regions comparable to
a few times the energy resolution (from less than to slightly
greater than the widths reported in B03 and DL04).

\[ s_i = n_i - \hat{b}_i, \] where \( b_i \) is the true continuum in bin \( i \) and \( \hat{b}_i \) is the
estimated continuum in bin \( i \). We can then formally construct the
null hypothesis \( H_0 \), there is no line (\( s_i = 0 \) for all \( K \)), against all
alternatives defined by \( H_1 \). This is the classic goodness-of-fit test.
The natural test statistic is \( T = \sum_{i=1}^{K} (n_{i} - \hat{b}_{i})^2/\{\text{var}(n_{i} - \hat{b}_{i})\} \).

Note—Models I and II show very similar results to model III.

### Table 3

**Goodness-of-Fit Parameters of Third Line Test**

| Parameter | pn Single | pn Double | MOS1 | MOS2 |
|-----------|-----------|-----------|------|------|
| \( T_c \) | 112.8 | 71.38 | 65.26 | 65.26 |
| \( T \) | 85.90 | 43.42 | 24.79 | 41.03 |
| 1 - C & | 1.7 x 10^-2 | 5.4 x 10^-2 | 5.3 x 10^-1 | 3.1 x 10^-2 |

### Table 4

**Summary Statistics for Third Feature Significance**

| Model | \( P(F) \) Analysis \( \sigma \) | Goodness-of-Fit Test \( \sigma \) | EW Analysis \( \sigma \) |
|-------|-----------------|-----------------|-----------------|
| I (BB+L3+BB) | 1.77 | 2.13 | b |
| I (BB+BB+L3+BB) | 3.29 | 2.13 | 2.07 |
| I (HA+L3+HA) | 2.81 | 1.18 | 1.69 |
| I (HA+HA+L3+L3) | 2.84 | 1.18 | 1.96 |
| II (BB+BB+L3) | 3.27 | 2.22 | 2.18 |
| II (HA+HA+L3) | 2.55 | 1.69 | 1.85 |
| III (BB+BB+L3) | 3.26 | 2.17 | 2.14 |
| III (HA+HA+L3) | 3.12 | 1.83 | 1.88 |

Note—Models I and II show very similar results to model III.

In such a blind search, one must correct for the fact that there are
N possible regions of interest (\( N \) is the energy band divided by
the width of a single region of interest) in which to find a
statistical fluctuation of a given level. Very roughly this increases
1 - C by the same factor, which is \( \sim 6 \). A more detailed treatment
can determine the exact factor (Eadie et al. 1983). We have
neither corrected for this factor in this section nor reduced the
significances accordingly in the next section. Thus, it should be
noted that the actual significances are approximately 6 times less
than those indicated in Table 3.

4.3. Significance of the Absorption Lines by Equivalent Width

The final and perhaps most straightforward method for deducing
an absorption feature’s existence is to simply assume it
exists and estimate the significance of the detection. In this
section we evaluate the significance of the residuals around 2.1 keV
by fitting a Gaussian absorption profile and then calculate the
equivalent width (EW) and error bars. This particular absorption-
line profile gives excellent fits to the data (\( \chi^2 \sim 1 \)), so no
improvement would be had by alternate line shapes. We confirmed
that other line shapes did not statistically significantly affect either
the EW or the line position. Of critical importance here is a proper
calculation of the error on the line fit. This is a standard nonlinear
\( \chi^2 \) fitting problem of the type discussed by LMB76. We follow
their approach exactly. The spectrum is fitted allowing all the
continuum parameters to vary freely, and the line depth and width are stepped through. The minimum value of $\chi^2$ is jointly estimated for the line depth and width. A contour plot of the 68% confidence interval for line depth and width is then calculated using the prescription of LMB76. Using the 68% ($\tau, \omega$) contour we integrate to find the EW extrema, and the best-fit ($\tau, \omega$) provide the best estimator of the EW. We held the line centroid at its best-fit we integrate to find the EW extremum, and the best-fit ($\tau, \omega$) provide the best estimator of the EW. We held the line centroid at its best-fit

The results of this analysis are shown in Table 4 for all the relevant models. In no case is the statistical significance of the line detection larger than $\sim 2\sigma$. As noted in § 4.2, this is a firm upper limit, uncorrected for a blind search. The strongest residuals are for pn data, and the EW for these cases are shown in Figure 4. Except for the blackbody models, the MOS1/MOS2 residuals were essentially nonexistent, so we gave up attempting to fit them. In some cases in Table 4 the absorption line resided on a low-temperature blackbody component that was so weak that no EW could be determined. We note for thoroughness (although we think it obvious) that one must calculate the EW using both the continuum model, and residuals are comparable to or smaller than the energy resolution, indicating that they are due to statistical fluctuation. This is borne out by statistical analysis of the residuals using the approach of § 4.1. Figure 5 shows the typical phase-resolved spectrum fitted with continuum model III (BB+BB) and residuals. We fixed $N_H$ to the fitted value (1.32 $\times 10^{20}$ cm$^{-2}$ from the phase-averaged pn spectral analysis). Our results did not change when we let $N_H$ vary in spectral fitting.

Around 2.1 keV, only a few spectral bins deviate from the continuum model, and residuals are comparable to or smaller than the energy resolution, indicating that they are due to statistical fluctuation. This is borne out by statistical analysis of the residuals using the approach of § 4.1. Figure 5 shows the results for the statistical tests on the phase-resolved data. There are no spectral features around 2.1 keV.

Table 4. The level of significance of the third spectral feature and measures of its significance are consistent with each other. In only a few isolated cases of blackbody continua does the line rise even to the $3\sigma$ level (and only for the $F$-test), much less something firmly indicative of a line. Since the fourth spectral feature is much, much weaker than the third feature, we were unable to analyze it in detail. In § 6 we present plausibility arguments that the statistically insignificant spectral residuals in the pn data are due to instrumental effects.

5. PHASE-RESOLVED SPECTRAL ANALYSIS

While no evidence of a third spectral feature was found in the time-averaged data, for thoroughness we decided to repeat our analyses for phase-resolved data. First we searched for the spin period using the pn data after correcting photon arrival times to the barycenter. The pn was operated in small-window mode with time resolution of $\sim 6$ ms. The MOS data do not have sufficient time resolution for spin-period determination. We used XRONOS, version 5.20 for our timing analysis. Using the epoch-folding method we found significant pulsation at $P = 424.1307 \pm 0.0001$ ms, consistent with the results of B03 and DL04. After tagging photons with the best-fit spin period, we reduced spectra in four spin phases. Spectral bins were binned to have at least 40 counts in each bin. Figure 5 shows a typical phase-resolved spectrum fitted with continuum model III (BB+BB) and residuals. We fixed $N_H$ to the fitted value (1.32 $\times 10^{20}$ cm$^{-2}$ from the phase-averaged pn spectral analysis). Our results did not change when we let $N_H$ vary in spectral fitting.

6. PHYSICAL SIGNIFICANCE OF THE RESIDUALS IN THE CONTINUUM SPECTRUM

We have demonstrated that there is no third line in the 1E 1207 spectrum. Thus, any attempt to explain the origin of $\leq 1–3\sigma$ residuals should be considered suspect. We have several ground rules in this section. We explicitly avoid detailed fitting of line shapes and in-depth searches for specific instrumental origins of the residual. It would certainly be futile, given the strength of the

---

**Table 4**

| Model          | Equivalent Width (in electron volts) |
|----------------|--------------------------------------|
| BB + BB        | 100                                 |
| BB + BB + BB   | 200                                 |
| BB + BB + BB   | 300                                 |
| BB + BB + BB   | 400                                 |
| BB + BB + BB   | 500                                 |
| BB + BB + BB   | 600                                 |
| BB + BB + BB   | 700                                 |
| BB + BB + BB   | 800                                 |
| BB + BB + BB   | 900                                 |
| BB + BB + BB   | 1000                                |

---

**Figure 3**

The $\tau$-$\omega$ contour plot for a third line (pn, MOS1, and MOS2). The contours were calculated at the 68% confidence level. We fitted model III (BB+BB)/L3 to the data. The plus signs show the best-fit parameters at 90% confidence of DL04 (solid contour: pn; dashed contour: MOS1; dotted contour: MOS2).

**Figure 4**

Equivalent width (in electron volts) of a third line at 2.1 keV and 68% confidence level.
residuals, and we are unlikely to offer deep insights into possibly subtle instrumental effects. Finally, we concentrate on motivating our suspicions using pn data, since the residuals are marginally stronger and thus likely to be more fruitful hunting ground.

6.1. Analysis of Residual Line Centroids

The analysis of § 4.3 yields the centroid of the weak residuals for each of the models described in § 3. A plausible hypothesis is that the residuals are associated with the Au M edge, which occurs in the pn and MOS instruments. To pursue this hypothesis we need to focus particular attention on the line centroids associated with model III. If the residuals arise in the Au M5 edge of the instrument, then model III clearly corresponds to the only physically correct model for the data. So if this hypothesis is true, we would expect to see residuals around the Au M5 energy.

Figure 6 shows the line centroids from fits to the third line using model III for pn single data, pn double data, and combined pn single and double data. We also show the B03 and DL04 90% confidence interval band for their line centroid determination. The dotted line is the theoretical position of the Au M5 edge. The striking feature of this plot is that the pn single data, which is without question the most reliable data because it rejects split events, occurs at the position of the Au M5 edge. Moreover, the potential origin of the lower residual energy found in previous work is suggested by the fact that the pn doubles (split events) residuals occur at lower energy. Our combined pn singles/doubles data set is in agreement with the results of B03 and DL04. Thus, we conclude that the pn split events have reduced the apparent energy of the residuals, while the pristine pn singles events are at the expected energy for an instrumental effect. One would be tempted to say that this lower energy in the split events is expected. This would get into a discussion about how well split events are reconstructed in EPIC data and is far too detailed for us. We simply note that the highest quality event data show residuals at the Au M5 edge, and the lowest quality data show it at much lower energy. We also show in Figure 6 the position of the residuals in 3C 273, using a fitting procedure described in the next section. The 3C 273 residuals, which are most certainly associated with the Au M5 edge, define the "fit" position of the Au M5 edge and are a more relevant metric than the theoretical position of the edge. Nevertheless, the position of the Au M feature in 3C 273 is consistent at the 90% level with the theoretical Au M edge and also with the position of the residuals in the pn singles data. A better line shape approximation would no doubt move the 3C 273 result downward, but this would require detailed understanding of the instrument response. Our results are meant only to be suggestive, since, after all, the residuals are statistically insignificant.

6.2. Consistency Test between pn and MOS Data

In DL04 a large inconsistency (~5 σ) between the significance of the lines detected in the pn and MOS data was reported, despite fit parameters being in statistical agreement. They explained this as being due to low-energy calibration uncertainties in the MOS. This observation motivated us to perform a consistency check on our data set. Because we used a MOS energy cut slightly higher in energy, it would reduce low-energy calibration uncertainties.
and thus yield better consistency between the MOS and pn data sets. This is what we have found, as described below.

We set up a Monte Carlo simulation to test whether the third line residual detected by the pn could have appeared in the MOS data as a deficit with a significance comparable to that observed. We used the III (BB+BB)/L3 model fit to the data because it showed the largest significance of a third feature, therefore the largest discrepancy between the pn and the MOS. Hereafter we describe our formalism using the pn single and MOS1 cases as an example: (1) we fit the model III (BB+BB)/L3 to the pn single data and calculate best-fit parameters, (2) we fold the model with variance in both the continuum and line parameters through the MOS1 detector response and simulate a spectrum with the same exposure time and spectral bins as our analysis of the real data, and (3) we evaluate $\Delta \chi^2$ by fitting a simulated spectrum with and without a third line. We repeat this procedure 1000 times and compute the distribution of $\Delta \chi^2$. Figure 7 shows the results from the consistency test between the pn and MOS1 (left) and the pn and MOS2 (right).

We calculated the fraction of simulated $\Delta \chi^2$ points smaller than or equal to the measured $\Delta \chi^2 (S)$, i.e., $f(\Delta \chi^2 \leq S)$. The results are $f(\Delta \chi^2 \leq S) = 2.9 \times 10^{-2}$ (MOS1) and $3.1 \times 10^{-2}$ (MOS2). The pn and MOS1 results are just consistent at the 1.90 $\sigma$ level and at the 1.86 $\sigma$ level between the pn and MOS2. We conclude that the behavior of the pn and MOS instruments is internally consistent with regard to fitting of the lines, as we showed previously was the case for the continuum fits.

### 6.3. Strength of the Residuals in the Au M Region

There are no statistically significant additional spectral features in the 1E 1207 spectrum. For thoroughness we indicate just how weak the residuals are that we have uncovered at the Au M edge. We do this by considering the residuals relative to 3C 273, where the same instrument residuals are well established (Kirsch 2004).

We extracted PATTERN = 0–4 events from a 40$''$ radius region centered on the source. We used a double power-law continuum model to fit the pn singles data in the 0.3–6.0 keV energy range. We did not fit the Au M region but instead fitted our model III with blackbody spectra for 1E 1207 and a double power-law continuum for 3C 273 and plotted the ratio of the residuals to the model spectrum (Fig. 8). We note that the residuals are $\sim 6\%$–$7\%$ in the quasar data and $\sim 10\%$–$15\%$ in the NS data. The NS data has larger error bars.

We directly calculated the EW of the 2.2 keV residuals without a line fit and find them to be $\sim 40$ eV. This is consistent with the EW that we determined by line fitting. A similar calculation (as a sanity check) gave us 8 eV for the EW residuals at 2.2 keV in 3C 273, in agreement with the measurements of the XMM-Newton team (Kirsch 2004). Of course the much higher noise level in the 1E 1207 data renders the residuals insignificant. In Figure 8 we show the ratio between the continuum model and the data for 1E 1207. The same plot is shown for 3C 273. We see there is nothing extraordinary in the NS data. The fluctuations are uniform.
through the entire spectrum. In 3C 273 the Au M edge is readily apparent because of the better counting statistics.

We note that a common error in estimating residual strength is to fit absorption lines and a continuum and then to use the resulting continuum without lines as a basis for comparison with the residuals. This procedure grossly overestimates the continuum and thus the strength of any absorption features with respect to it. The correct procedure of comparing the dips with respect to the best-fit continuum markedly reduces the significance of the residuals. With this proper procedure our results here are consistent with those we obtained in § 4.

7. CONCLUSIONS

1. Using a wider range of physically plausible INS models and careful background subtraction, we have obtained continuum fits for 1E 1207 that effectively rule out the need for additional harmonics of the first two spectral absorption features. Our derived continuum parameters for both the pn and MOS data are consistent with those found by previous investigators.

2. To explore the significance of the weak residuals found at higher energies in a few cases, we have implemented an improved F-statistic analysis for additional weak lines in 1E 1207. It corrects for systematic overestimation of absorption-line strengths using the standard F-test, as previously reported in the literature. Calculating the correct, finite statistics reference distribution for the F-statistic, we find that lines at harmonics of the two firmly established lines are statistically insignificant.

3. We performed three completely independent statistical analyses to test for the presence of third and fourth spectral features. All tests give consistent results, and all tests indicate that there is no third or fourth line.

4. While there are no statistically significant spectral features at higher energies, we explicitly fitted the spectral residuals around 2.1 keV. Using the continuum model favored by previous investigators, and for which we get excellent continuum fits, we find that the pn single spectrum residuals occur at the position of the instrumental Au M edge. The residuals appear at the same energy as those obtained by fitting the known Au M residuals in 3C 273 with an identical procedure. We further show that pn doubles events yield a much lower energy for the residuals, as does the combined pn singles/doubles data set. This is highly suggestive of an instrumental origin for the insignificant residuals. Indeed, in all eight physically plausible INS models we considered, the residuals always appear at the Au M edge.

5. We have demonstrated that the strength of the weak instrumental residuals in the 1E 1207 spectrum is consistent with the stronger instrumental residuals observed in the 3C 273 spectrum.

We thank the referee for a very careful reading of the manuscript and for suggesting changes that strengthened our analysis. We also thank the editor for numerous suggestions of style and for helping make the paper more concise and focused. J. C. acknowledges financial support from NASA’s Graduate Student Researchers Program NGRS-50392. The authors thank Maurice Leutenegger for assistance and advice on various data-processing issues.

APPENDIX

TESTING FOR A SPECTRAL LINE IN A CONTINUUM

Some important results used in § 4.2 are summarized here. The arguments of Eadie et al. (1983) are closely followed.

Consider searching a region of interest (ROI) consisting of \( K \) bins for a deficit of counts over the expected continuum. The total spectrum has \( N \) bins. Define \( b_i \) as the true continuum, \( \hat{b}_i \) as the estimated continuum, and \( n_i \) as the observed counts in bin \( i \). We can define the difference of counts in bin \( i \) as \( s_i = n_i - b_i \).

The search for a residual in the \( K \) bins of the ROI can be formulated as a goodness-of-fit test, and under the null hypothesis \( H_0 \) there are no residuals, and \( E(s_i) = E(n_i) - \hat{b}_i = b_i - \hat{b}_i \approx b_i - b_i = 0 \), where \( E \) is the expectation (mean) operator. A reasonable test statistic is \( T = \sum_{i=1}^{K} (n_i - \hat{b}_i)^2 / [V(n_i - \hat{b}_i)] \), where \( V(.) \) is the variance operator. This statistic is a function of the random variables \( z_i = (n_i - \hat{b}_i)(V(n_i - \hat{b}_i))^{-1/2} \).

The \( z_i \) have the following properties if \( H_0 \) is true: \( E(z_i) = 0 \), and \( V(z_i) = V(n_i - \hat{b}_i) = V(n_i) + V(\hat{b}_i) - \text{cov}(n_i, \hat{b}_i) = V(\hat{b}_i) + V(\hat{b}_i) = \hat{b}_i + \sigma^2_i \), where \( \text{cov}(.) \) is the covariance operator. The trick is that the \( \hat{b}_i \) should be estimated using the \( N - K \) bins outside the ROI, for then \( n_i \) and \( \hat{b}_i \) are uncorrelated and their covariance vanishes. The continuum is simply fitted without using the \( K \) bins in the ROI. Also note that \( V(\hat{b}_i) \equiv \sigma_i^2 \) can be directly obtained from the covariance matrix of the continuum fit. Thus, under \( H_0 \),

\[
T = \sum_{i=1}^{K} \frac{(n_i - \hat{b}_i)^2}{b_i + \sigma_i^2} = \sum_{i=1}^{K} z_i^2, \tag{A1}
\]

\( E(z_i) = 0 \), and \( V(z_i) = 1 \). Thus, the \( z_i \) are standard, normal variables (for \( s_i \) not too small), and \( P(T) \) is a \( \chi^2 \) distribution with \( K \) dof.

For \( K \geq 30 \) we can simplify this result by defining a new statistic \( y = (T - K)(2K)^{-1/2} \). Since the mean and variance of the \( \chi^2 \) distribution are \( K \) and \( 2K \), respectively, \( y \) has a standard, normal distribution. Then as described in the text we can test \( H_0 \) by simply calculating \( y \); and if \( y < y_c = 4 \) we accept \( H_0 \).

REFERENCES

Bevington, H. A. 1969, Data Reduction and Error Analysis for the Physical Sciences (New York: McGraw-Hill)
Bignami, G. F., Caraveo, P. A., De Luca, A., & Mereghetti, S. 2003, Nature, 423, 725 (B03)
De Luca, A., Mereghetti, S., Caraveo, P. A., Moroni, M., Mignani, R. P., & Bignami, G. F. 2004, A&A, 418, 625 (DL04)
Dermer, C. D., & Sturme, S. J. 1991, ApJ, 382, L23
Eadie, W. T., Driijard, D., & James, F. E. 1983, Statistical Methods in Experimental Physics (Amsterdam: North-Holland)
Freeman, P. E., Lamb, D. Q., Wang, J. C. L., Wasserman, I., Loredo, T. J., Fenimore, E. E., Murakami, T., & Yoshida, A. 1999, ApJ, 524, 772 (F99)
Giacani, E. B., Dubner, G. M., Green, A. J., Goss, W. M., & Gaensler, B. M. 2000, AJ, 119, 281
Haberl, F., Schwope, A. D., Hambranyan, V., Hasinger, G., & Motch, C. 2003, A&A, 403, L19
Haberl, F., Zavlin, V. E., Trümper, J., & Burwitz, V. 2004, A&A, 419, 1077
Hailey, C. J., & Mori, K. 2002, ApJ, 578, L133
Ho, W. C. G., & Lai, D. 2003, MNRAS, 338, 233
