A two-dimensional Kolmogorov–Smirnov test for crowded field source detection: \textit{ROSAT} sources in NGC 6397

S. A. Metchev\textsuperscript{1*} and J. E. Grindlay\textsuperscript{2}

\textsuperscript{1}Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
\textsuperscript{2}Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Accepted 2002 April 22. Received 2002 April 19; in original form 2001 November 28

ABSTRACT

We present a two-dimensional version of the classical one-dimensional Kolmogorov–Smirnov (KS) test, extending an earlier idea due to Peacock and an implementation proposed by Fasano and Franceschini. The two-dimensional KS test is used to optimize the goodness of fit in an iterative source-detection scheme for astronomical images. The method is applied to a \textit{ROSAT}\slash HRI X-ray image of the post-core-collapse globular cluster NGC 6397 to determine the most probable source distribution in the cluster core. Comparisons to other widely used source-detection methods, and to a \textit{Chandra} image of the same field, show that our iteration scheme is superior in measuring statistics-limited sources in severely crowded fields.

Key words: methods: data analysis – methods: statistical – globular clusters: individual: NGC 6397 – X-rays: stars.

1 INTRODUCTION

Deep X-ray imaging of crowded fields, even with increasing angular resolution and sensitivity (with \textit{Einstein} and \textit{ROSAT}, and now \textit{Chandra}), is invariably limited by the small number of source counts and by the relative size of the point spread function (PSF) compared to the angular separation between the objects. Determining the underlying source configuration in such a regime is often beyond the capabilities of conventional source-finding algorithms.

Classical X-ray source-detection methods are based on a sliding detection cell of a fixed size across the image, and calculating the signal-to-noise ratio (S/N) at each step. To find S/N, common detection algorithms for processing data from \textit{Einstein} and \textit{ROSAT} (implemented in IRAF\textsuperscript{1}/PROS\textsuperscript{2}) use either an average background (as determined from a source-free section of the image) or a local background (from a region around the detection cell). However, both methods fail to discern blended faint sources in crowded fields where the background is affected by overlapping PSFs. Source detection is somewhat improved by image deconvolution, e.g. with the Lucy–Richardson (LR) algorithm or with the maximum-entropy method (MEM), or by wavelet smoothing. Deconvolution algorithms provide higher positional sensitivity in moderately crowded fields, but suffer from such undesirable effects as noise amplification and ‘leakage’ (associating counts from fainter sources to brighter nearby ones). Wavelet detection implemented as task WAVDETECT in the \textit{Chandra} processing package\textsuperscript{3} does well in crowded fields, provided the sources are either sufficiently separated ($\gtrsim 3–5$ times the full width at half-maximum (FWHM), or $\gtrsim 3–5$ arcsec) or within $\sim 2–3$ FWHM and are similar in flux (Damiani et al. 1997; Freeman et al. 2001). When the PSFs are heavily blended (separation between the source centroids $\lesssim 1.5$ FWHM), individual sources cannot be distinguished and their relative fluxes cannot be measured.

A superior source-detection method is needed for severely crowded fields containing multiple faint sources, e.g. globular cluster cores (Hertz & Grindlay 1983), or nuclear bulges in external galaxies.

A powerful technique to compare statistics-limited samples is the Kolmogorov–Smirnov (KS) test, which, unlike its alternative – the Pearson $\chi^2$ test – does not require binning of the data. Unfortunately, the classical KS test is applicable only to one-dimensional distributions, and any attempts to convert a two-dimensional image to one dimension (e.g. by collapsing it on to a vector, or by azimuthal binning around a point) lead to unwanted loss of information and power. For some time now, a multidimensional version of the KS test has been known (Peacock 1983; Gosset 1987), which performs better than the $\chi^2$ test in the small-number statistics case (Fasano & Franceschini 1987, hereafter FF), and can be successfully applied in parameter point estimation in a manner similar to the widely used maximum-likelihood (ML) method. These properties of the multidimensional KS test make it viable for incorporation in source-detection methods.

In this paper, we revisit the characteristics of the KS test in two (and three) dimensions and examine its power in comparing different realizations of crowded low-S/N fields. As an application of the test, we devise an iterative source-modelling scheme that aims to

\textsuperscript{*}E-mail: metchev@astro.caltech.edu

\textsuperscript{1}Image Reduction and Analysis Facility; developed and maintained by the National Optical Astronomy Observatories.

\textsuperscript{2}Post-Reduction Off-line Software; developed and maintained by the Smithsonian Astrophysical Observatory.

\textsuperscript{3}Available at http://asc.harvard.edu/ciao
minimize the KS statistic in search of the optimum underlying data distribution in an image. Based on our Monte Carlo simulations, we find that our iterative algorithm is a powerful tool for faint object searches in crowded fields. We apply the algorithm to determine the faint X-ray source distribution in a deep ROSAT exposure of the post-core-collapse globular cluster NGC 6397, which has also been analysed with ML techniques by Verbunt & Johnston (2000). We compare the derived X-ray positions with those of Verbunt & Johnston and with our subsequent optical (HST, Taylor et al. 2001) and X-ray (Chandra, Grindlay et al. 2001a) identifications.

2 THE TWO-DIMENSIONAL KS TEST

2.1 Description

The classical one-dimensional (1D) KS test makes use of the probability distribution of the quantity $D_{KS}$, defined as the largest absolute difference between the cumulative frequency distributions of the parent population and that of an $n$-point sample extracted from it. Since $D_{KS}$ is approximately proportional to $1/\sqrt{n}$, one usually refers to the probability distribution of the quantity $Z_n = D_{KS}/\sqrt{n}$. For a given $n$, the values of $Z_n$ corresponding to a given significance level $SL$ (denoted as $Z_n,SL$) increase slightly with $n$. For large $n$, the integral probability distribution $P(Z_n) = 1 - SL$ approaches the asymptotic expression (Kendall & Stuart 1979)

$$P(Z_n) = 2 \sum_{k=1}^{\infty} (-1)^{k-1} \exp \left(-2k^2 Z_n \right),$$

which is satisfactory for $n \geq 80$. For the two-sample KS test, which compares distributions of different sizes ($n_1$ and $n_2$), the probability distribution $P(Z_n)$ remains unchanged provided that $n$ is set to $n_1 n_2 / (n_1 + n_2)$. The 1D nature of the test implies that it does not depend in any way on the shape of the parent distribution.

In a two-dimensional (2D) distribution, each data point is characterized by a pair of values, $(x, y)$. As with the 1D KS test, the maximum cumulative difference between two 2D distributions is found over the $(x, y)$-plane. In the case of distributions in more than one dimension, however, the procedure to cumulate the information on to the plane is not unique. FF made use of the total number of points in each of the four quadrants around a given point $(x_i, y_j)$, namely, the fraction of data points in the regions $(x < x_i, y < y_j)$, $(x < x_i, y > y_j)$, $(x > x_i, y < y_j)$, $(x > x_i, y > y_j)$. The 2D statistic $D_{KS}$ is defined as the maximum difference between data fractions in any two matching quadrants of the sample and of the parent population, ranging over all data points. The $Z_n$ statistic is defined similarly as in the 1D case, $Z_n = D_{KS}/\sqrt{n}$, where for a two-sample 2D KS test $n = n_1 n_2 / (n_1 + n_2)$.

Based on their Monte Carlo simulations, FF deduce that the 2D integral probability distribution $P(Z_n)$ depends solely on the correlation coefficient (CC) of the model distribution, i.e. that for a given CC the distribution of $Z_n$ in the 2D KS test is (nearly) independent of the shape of the model, as in the classical 1D KS test. FF also observe that, in the two-sample case, it is sufficient to take the average of the correlation coefficients $C_{11}$ and $C_{22}$ of the samples as an estimate of $CC$.

An important distinction between the Peacock (1983) and the Fasano & Franceschini (1987) versions of the 2D KS test was pointed out to us by the referee, which makes the latter generally less stringent. FF restrict the search for the maximum cumulative difference $D_{KS}$ to loci harbouring a data point, thereby often missing the location of the true maximum difference, which is almost always found for $(x < x_i, y < y_j)$, where $i$ and $j$ are two different data points. Nevertheless, the maximum cumulative difference computed in such a way will have a tendency to vary in the same manner as the true maximum difference. Thus the FF statistic is probably well-behaved, at least as long as the genuine parent population distribution and the assumed one are not too different (Gosset 1987). The latter situation is indeed expected when comparing distributions of point spread functions in two images. The advantage of FF’s approach is speed: order $n$, instead of $n^2$. The disadvantage is its approximate nature and that the $D_{KS}$ statistic is sensitive to the correlation coefficient $CC$ of the distributions, requiring its inclusion as a free parameter in the reference tables. Our Monte Carlo experiments below take into account both factors.

2.2 Monte Carlo experiments

Following the procedure in FF, we used the 2D KS test computer code provided in Press et al. (1997) to run our own Monte Carlo experiments. We studied the $Z_n$ statistic by means of a Monte Carlo procedure using a uniform distribution (CC = 0) within the unit square as a parent population. The analysis comprised cases with number of points $n$ per sample ranging from $n = 5$ to $n = 50000$. For any given $n$ we produced a large number of simulations (from 100 000 for $n = 5$ to 1000 for $n = 50000$), enabling us to construct the integral probability distribution $P(Z_n)$ with sufficient accuracy. Values of $Z_n,SL$ for uniform samples of all tested sample sizes $n$ are listed in Table 1.

In Fig. 1 presents a comparison between our and FF’s results for the critical values $Z_{n,SL}$ as a function of $n$. Both sets of Monte Carlo simulations show similar tendencies in the behaviour of the $Z_n$ statistic, and are indistinguishable from each other for $n \geq 500$ within the statistical uncertainties (≈5 per cent for high $n$, because of the limited number of simulations). There is, however, a marked inconsistency between the two sets of data in the low- SL, small-$n$ part of the graph, where at the 30 per cent significance level, the values of $Z_{n,SL}$ differ by a factor of ~1.15. The difference is highly significant, given the fact that both FF’s and our results for small $n$ are based on 100 000 simulations. We attribute this discrepancy to a detail in the implementation of the 2D KS test: in particular, to whether the data point around which the $D_{KS}$ statistic is computed is included in one of the quadrants or not. This discrepancy

<table>
<thead>
<tr>
<th>$n^a$</th>
<th>No. of simul.</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>95</th>
<th>99</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>100000</td>
<td>0.76</td>
<td>0.82</td>
<td>0.88</td>
<td>0.95</td>
<td>1.03</td>
<td>1.13</td>
<td>1.27</td>
<td>1.38</td>
<td>1.62</td>
</tr>
<tr>
<td>10</td>
<td>100000</td>
<td>0.79</td>
<td>0.85</td>
<td>0.92</td>
<td>0.98</td>
<td>1.06</td>
<td>1.15</td>
<td>1.29</td>
<td>1.41</td>
<td>1.65</td>
</tr>
<tr>
<td>20</td>
<td>100000</td>
<td>0.85</td>
<td>0.91</td>
<td>0.97</td>
<td>1.03</td>
<td>1.10</td>
<td>1.20</td>
<td>1.33</td>
<td>1.45</td>
<td>1.69</td>
</tr>
<tr>
<td>50</td>
<td>100000</td>
<td>0.91</td>
<td>0.97</td>
<td>1.04</td>
<td>1.10</td>
<td>1.18</td>
<td>1.27</td>
<td>1.41</td>
<td>1.53</td>
<td>1.78</td>
</tr>
<tr>
<td>100</td>
<td>100000</td>
<td>0.96</td>
<td>1.02</td>
<td>1.08</td>
<td>1.15</td>
<td>1.22</td>
<td>1.32</td>
<td>1.45</td>
<td>1.57</td>
<td>1.81</td>
</tr>
<tr>
<td>200</td>
<td>50000</td>
<td>1.01</td>
<td>1.07</td>
<td>1.13</td>
<td>1.19</td>
<td>1.27</td>
<td>1.37</td>
<td>1.51</td>
<td>1.62</td>
<td>1.86</td>
</tr>
<tr>
<td>500</td>
<td>50000</td>
<td>1.06</td>
<td>1.12</td>
<td>1.18</td>
<td>1.25</td>
<td>1.33</td>
<td>1.43</td>
<td>1.57</td>
<td>1.68</td>
<td>1.91</td>
</tr>
<tr>
<td>1000</td>
<td>50000</td>
<td>1.09</td>
<td>1.15</td>
<td>1.21</td>
<td>1.28</td>
<td>1.35</td>
<td>1.45</td>
<td>1.60</td>
<td>1.73</td>
<td>1.98</td>
</tr>
<tr>
<td>2000</td>
<td>50000</td>
<td>1.11</td>
<td>1.17</td>
<td>1.24</td>
<td>1.30</td>
<td>1.38</td>
<td>1.47</td>
<td>1.61</td>
<td>1.74</td>
<td>1.99</td>
</tr>
<tr>
<td>5000</td>
<td>50000</td>
<td>1.14</td>
<td>1.21</td>
<td>1.27</td>
<td>1.34</td>
<td>1.42</td>
<td>1.50</td>
<td>1.64</td>
<td>1.74</td>
<td>1.96</td>
</tr>
<tr>
<td>10000</td>
<td>50000</td>
<td>1.15</td>
<td>1.22</td>
<td>1.28</td>
<td>1.35</td>
<td>1.43</td>
<td>1.54</td>
<td>1.69</td>
<td>1.82</td>
<td>2.13</td>
</tr>
<tr>
<td>20000</td>
<td>50000</td>
<td>1.14</td>
<td>1.21</td>
<td>1.29</td>
<td>1.37</td>
<td>1.45</td>
<td>1.56</td>
<td>1.74</td>
<td>1.89</td>
<td>2.20</td>
</tr>
<tr>
<td>50000</td>
<td>50000</td>
<td>1.16</td>
<td>1.24</td>
<td>1.32</td>
<td>1.40</td>
<td>1.50</td>
<td>1.60</td>
<td>1.78</td>
<td>1.91</td>
<td>2.19</td>
</tr>
</tbody>
</table>

$^a$Size of sample.

$^b$Significance level (SL = 1 - $P(Z_n)$).

© 2002 RAS, MNRAS 335, 73-83
broadly not suitable when the PSFs of the individual sources are heavily blended (source separation is ≲0.5 FWHM). Classical X-ray source-detection methods fail to distinguish the individual objects. Crowded field optical photometry tools (e.g. DAOPHOT, Stetson 1987, 1991) are also not suitable for the small-number Poisson statistics of X-ray images. Even the recently introduced detection algorithm based on wavelet transforms (Damiani et al. 1997; Freeman et al. 2001, and references therein) implemented in the Chandra X-ray data analysis package (CIAO),4 does not produce adequate results in the regime of severe source confusion and small number of counts per source.

Below we describe an implementation of the 2D KS test to astronomical images. Provided that the PSF of the (unresolved) sources in a crowded field is known and constant over the image, by comparing the image to a simulation of a proposed source distribution, we can obtain the KS probability \( P(Z_{0}) \) that the image and the simulation represent the same source configuration. The obtained probability can be used as a measure of the accuracy of both the positions and the intensities of the proposed sources, as well as an indication of the necessity for additional sources to account for the photon distribution. Because of the high sensitivity of the 2D KS test, we expect that it should be able to discern positional discrepancies equal to a fraction of the FWHM of the PSF.

### 3.1 Implementation of the two-dimensional KS test

#### 3.1.1 A two-dimensional versus a three-dimensional KS test

Astronomical images (e.g. from charge-coupled devices, CCDs) have three dimensions: \( x \) and \( y \) pixel coordinates, and pixel intensity. Although the spatial distribution of photons that strike the detector is two-dimensional (two photons never fall at exactly the same position), they are binned by the detector into integer bins corresponding to the digital pixel size. Thus the incoming 2D distribution of photons with real-valued coordinates is transformed into a three-dimensional (3D) one with integer coordinates. It is possible to apply the 2D KS test to an integer-valued 3D distribution by simply recalculating the two-dimensional \( D_{KS} \) statistic for every count in a given pixel. However, since the KS test is designed to deal with real-valued distributions, any point in the \((x, y)\)-plane that has more than one count at exactly the same real-valued position (which would be the case if a pixel contained more than one count) becomes disproportionately significant, and distorts the value of \( D_{KS} \).

Therefore, to apply a 2D KS test to 3D images, it is first necessary to ‘un-bin’ the images by spreading the counts in each pixel over the area of the pixel. Since no information is preserved about the exact impact locations of the photons on the detector, we introduce a random shift (between \(-0.5\) and \(+0.5\) pixel) in the integer-valued coordinates of each count. Every count in the image is thus assigned a unique position (within the precision limits of the computer), and the integer-valued 3D coordinates are converted to real-valued 2D coordinates. Since this routine distributes the data on a subpixel scale, we refer to it as ‘subpixelization’.

Unfortunately, subpixelization incurs an undesirable effect, resulting from the randomness with which the counts are moved around within the pixels. In particular, two subpixelized versions of the same image are never the same. Therefore, KS testing of different subpixelizations of the same two images will produce different results for \( Z_{0} \) every time. Our experience is that \( Z_{0} \) varies by about 2–4 per cent between runs; the corresponding variations in the significance level \( SL \) may be as high as ±8 per cent for values of \( Z_{0} \) near SL = 50 per cent (Table 1). To obtain a mean value for the KS probability \( P(Z_{0}) \) (equal to \( 1 - SL \)) with which two images represent the same parent distribution, it is necessary to run the 2D KS test multiple times. The error in the KS probability can be estimated from the variations in \( Z_{0} \) among the different runs.

The above method may appear contrived and unnecessary, when instead of applying a 2D KS test, by following the generalization in FF, a 3D KS test can be implemented. A 3D test does not have the undesirable uncertainties associated with subpixelization, and gives the exact KS probability that two images represent the same parent distribution. FF report that the 3D test exhibits greater power in rejecting wrong hypotheses than a three-fold 2D test along each of the planes \((x, y), (y, z)\) and \((z, x)\). However, our experiments with the tests on simulated ROSAT images show that the 2D KS test applied to subpixelized images is more powerful than its 3D counterpart (with accordingly generated look-up tables) applied to the original (not subpixelized) images. KS tests performed on simulations of an 880-count five-source distribution with centre-to-centre distances of 1.1–2.1 times the PSF size (5 arcsec for ROSAT/HRI) show that the 2D test on subpixelized images can distinguish individual source position shifts as small as 2–3 arcsec (depending on the relative source locations and intensity) at the \( \geq 97 \) per cent significance level. The 3D test finds such shifts insignificant: it gives a probability of \( \geq 70 \) per cent that the simulations represent the same parent distribution.

The higher power of the subpixelized 2D test is explained by the manner in which pixels are weighted. Subpixelization ensures that all photons are assigned equal weights: a desirable effect, since pixels are weighted proportionally to their intensity. The 3D KS test, on the other hand, assigns equal weights to all pixels, so single-count pixels (often from background) are as significant as pixels.
with multiple counts (denoting sources). Hence, the 2D test on subpixelized images is more sensitive to variations in the source distribution than its 3D counterpart.

We also note that subpixelized images represent more realistically the incoming photon distribution. Repeated 2D KS tests on independent subpixelizations produce an estimate of the random error in the KS probability induced by photon binning.

### 3.1.2 One-sample versus two-sample two-dimensional KS test

When comparing an image to a proposed model distribution of sources (using an analytical or a fitted PSF), perfect information is available about the shape of the model distribution. It is therefore appropriate to use a one-sample KS test to compare the image to the model distribution. However, the analytical form of a fitted PSF is often very complex. The added complication of having several nearby sources with overlapping PSFs (as in a crowded field) makes it computationally very tedious to calculate the fraction of the analytic model distribution in the quadrants around every count in the image (needed to compute the $D_{KS}$ statistic; Section 2.1).

To avoid lengthy 2D surface integrals, the two-sample KS test can be used instead, to compare the image to a simulation generated from the proposed model. The random deviations in the representation of the model can be decreased if a ‘bright’ (high number of counts $n$) simulation is used, that follows the model closely. The fractional deviation of the simulated versus expected counts per pixel will decrease as $1/\sqrt{n_{\text{tmp}}}$ for large $n_{\text{tmp}}$, where $n_{\text{tmp}}$ is the number of counts per pixel in the model image. The running time of the two-sample 2D KS test is, however, proportional to the square of the sum $n_1 + n_2$ of counts in the samples compared, and is impracticable to use the two-sample KS test on simulations containing a high number of counts.

To take advantage of the higher power of the one-sample test (given perfect information about the shape of the model distribution), and to avoid long running time, we compare the image (containing $n$ counts) to a bright simulation ($n_{\text{mod}} = k \times n$ counts; $k \sim 50$) of the proposed model, using the one-sample test. In essence, we use Monte Carlo integration for the model, the assumption being that a bright simulation can be made to represent the model with sufficient accuracy, and simple summing of the counts in the quadrants can be substituted for analytic integration of the PSF. We therefore do not expect the properties of this pseudo one-sample KS test to be significantly different from those of the FF one-sample KS test.

In particular, we assume that the new test is still distribution-free for a fixed correlation coefficient $CC$ of the bright simulation (hereafter referred to as the ‘model’), and that its properties vary slowly with $CC$ (as observed in FF for their 2D KS test). Clearly, our test will converge to the FF one-sample test for $k \rightarrow \infty$, but will run faster than the latter for moderate-sized $k$, since lengthy analytical integrations are substituted with Monte Carlo integration. The pseudo one-sample test [running time proportional to $n(n + n_{\text{mod}}) = (k + 1)n^2$] is also $k + 1$ times faster than the two-sample test [running time proportional to $(n + n_{\text{mod}})^2 = (k + 1)n^2$]. Thus, for moderate-sized $k$, the pseudo test approximates the power of the one-sample KS test, and is faster than both the one-sample and the two-sample KS tests for a general model distribution.

Table 2 presents a comparison of the performance of the two versions of the test, using the same five-source setup as in Section 3.1.1, with one source shifted by only 0.4 FWHM – 2s. These are due to the fact that the small-number statistical uncertainties in the two samples are not averaged out in the one-sample test, whereas in the two-sample test they are. For $k \geq 20$, however, the values for $Z_{n_1}$ become self-consistent (as well as consistent with those from the two-sample test), and there is little power to be gained from increasing $k$.

Despite the similar behaviour of the FF KS test and our pseudo one-sample test, the reference tables for the former (tables A1–A5 in FF: Table 1 herein) cannot be used for the latter, since the $Z$ distributions are different. We therefore ran Monte Carlo simulations to create a separate look-up table for the new test. Table 3 is based on comparing simulated source distributions with $n_1 \approx 850$ counts to $k \approx 50$ times brighter simulations (‘models’, $n_2 \approx 45000$) of the same source distributions. Sources were simulated on a $128 \times 128$ pixel field using the analytic 5-arcsec (10-pixel) HRI PSF. We also ran Monte Carlo simulations with higher $n_1$ ($n_1 \approx 1800$ against $n_2 \approx 45000$ count models for the purpose of interpolation. Although in this case the value of $k$ is lower ($k = 25$), we choose to set the value of $n_2 = 45000$ as the standard, as it is a

<table>
<thead>
<tr>
<th>$k$</th>
<th>$n_1$</th>
<th>Pseudo one-sample test $Z_{n_1}$</th>
<th>P($&gt;Z_{n_1}$)</th>
<th>$h^b$</th>
<th>Two-sample test $Z_{n_2}$</th>
<th>P($&gt;Z_{n_2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>880</td>
<td>$1.78 \pm 0.07$</td>
<td>$440$</td>
<td>$1.30 \pm 0.04$</td>
<td>$34 \pm 5$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>880</td>
<td>$1.72 \pm 0.05$</td>
<td>$587$</td>
<td>$1.44 \pm 0.03$</td>
<td>$19 \pm 3$</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>880</td>
<td>$1.56 \pm 0.04$</td>
<td>$733$</td>
<td>$1.48 \pm 0.04$</td>
<td>$16 \pm 4$</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>880</td>
<td>$1.88 \pm 0.05$</td>
<td>$800$</td>
<td>$1.79 \pm 0.03$</td>
<td>$3 \pm 1$</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>880</td>
<td>$1.59 \pm 0.06$ $\lesssim 4 \text{ per cent}^d$</td>
<td>$838$</td>
<td>$1.64 \pm 0.03$</td>
<td>$8 \pm 2$</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>880</td>
<td>$1.62 \pm 0.04$ $\lesssim 2 \text{ per cent}^d$</td>
<td>$863$</td>
<td>$1.65 \pm 0.05$</td>
<td>$8 \pm 3$</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>880</td>
<td>$1.62 \pm 0.06$</td>
<td>$871$</td>
<td>$1.68^c$</td>
<td>$\sim 7$</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>880</td>
<td>$1.66 \pm 0.06$</td>
<td>$876$</td>
<td>$1.74^c$</td>
<td>$\sim 6$</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Errors determined from KS comparisons of 10 different subpixelizations of the same simulations. Systematic uncertainties associated with representing a continuous model by a discrete distribution are not included.

$^b$For the two-sample test $h = \left(\frac{n_1 n_2}{n_1 + n_2}\right) = \frac{k n_1^2}{n_1 + n_2}$.

$^c$Obtained from Table 1 for $CC = 0.10 \approx 0.0$ of the two samples.

$^d$Interpolated from lines 1 ($k \approx 50$) and 2 ($k = 25$) of Table 3 ($CC = 0.10 \approx 0.12$).

$^e$Only one comparison was performed due to the longer running time.
measure of the precision of the Monte Carlo integration. High values of the sample correlation coefficient CC were not pursued, since astronomical images rarely have highly correlated photon distributions (especially in crowded fields, but also given random background). Thus constructed, Table 3 should be applicable to most low-background, low-S/N imaging cases (especially X-ray).

Using Table 3 for the values of $Z_{n,SL}$, the results for the KS probability $P(Z_n)$ from the pseudo one-sample KS test will be consistent with those from the FF KS test using the FF tables. Hereafter, unless explicitly stated otherwise, we shall refer to our pseudo one-sample 2D KS test as ‘the 2D KS test’, or simply as ‘the KS test’.

### 3.2 Application of the two-dimensional KS test to simulated images

After establishing that the 2D KS test can be successfully applied to images, it is important to determine the sensitivity of the test with respect to deviations from the proposed model. Here we test the responsiveness of the 2D KS test with respect to changes in the parameters of individual sources in the model distribution. We look for the minimum deviation in a single parameter (position or intensity of a source) that enables the test to tell the distributions apart at the $\geq 99$ per cent significance level. As a trial source distribution we use the one simulated in Fig. 2(a), whose parameters are listed in Table 4 (sources X1–X5), and compare it to models that have one parameter changed.

It is worth noting here that, given the significant overlap of PSFs in a crowded field such as the one in Fig. 2(a), the sensitivity of the test with respect to changes in the distribution need not be isotropic. It depends on the relative positioning and brightness of the sources. We therefore test several possible scenarios of such changes. The scenarios shown in Fig. 2 and described in its caption correspond to the limiting cases, in which the KS test can distinguish the distributions at the $\geq 99$ per cent level.

After performing KS tests between the simulation in Fig. 2(a) and the models in Figs 2(b)–(d), we establish that the positional sensitivity of the KS test in crowded fields depends on the relative source intensity, and on the direction in which a source is allowed to move with respect to the crowded region. The sensitivity to moving a source is higher for brighter sources (3–4 pixel for source X2) than for the fainter ones (4–9 pixel for source X1), and is generally (although not conclusively) poorer when the source is moved towards the region of crowding (9 pixel for source X1, 3 pixel for X2) as opposed to when it is moved away (4 pixel for X1 and X2). The brightness sensitivity of the KS test is also dependent on the relative source intensity in a crowded field: brighter sources can vary by a smaller percentage (~45 per cent for X2) than fainter sources (~70 per cent for X1). For a source that is sufficiently far away (~3 FWHM) from the crowded region (source X5) the sensitivity of the KS test is greater and nearly position-independent. For source X5 we set the limits at a 2-pixel shift in any direction [$P(Z_n) \approx 1$ per cent] or a 30 per cent change in intensity [$P(Z_n) < 1$ per cent].

The above-determined sensitivity limits are based on varying only one parameter (position or intensity) of a single source, while keeping all other parameters fixed. This is the approach used to determine the 95 per cent ($\approx 2\sigma$) confidence limits (Table 6) on the positions and intensities of the detected sources in our NGC 6397 ROSAT image (Section 5).

### 4 THE 2D KS TEST IN PARAMETER POINT ESTIMATION FOR SOURCE DETECTION

#### 4.1 Algorithm

An iterative source-fitting algorithm was devised that aims to minimize the $Z_n$ statistic, thus maximizing the probability that an image and a simulation represent the same parent distribution of sources. The final simulation that results from this algorithm will contain the
best estimate for the number, positions and intensities of the sources in
the image, subject to limitations arising from the sensitivity of the
test. The iterative procedure steps through the following algorithm:

(i) An initial guess of the source distribution is made. This can be
a source at the location of the brightest pixel (thus starting with a one-
source configuration) or a guess with \( N_{\text{initial}} > 1 \) number of sources.
Both initial guesses will produce the same results for a distribution
with \( N_{\text{initial}} \geq N_{\text{initial}} \). Sources. The PSF is fitted to a single unresolved
source in an uncrowded part of the image so that aspect or other
image systematics are included.

(ii) A bright simulation (also known as a ‘model’; see Section
3.1.2) based on the current guess for the source distribution is created
(using the fitted PSF) and normalized to the image intensity. The
normalized model and the image are smoothed with a Gaussian
function to roughly match the \textit{ROSAT}/HRI resolution. The residual
between the two is then formed.

(iii) The next guess for the source distribution is obtained by
moving the source positions against the steepest gradient in the
residual, and by adjusting the intensities, so as to decrease the max-
imum deviation from zero (\( D_{\text{max}} \)) in the smoothed residual. We
indeed observe that \( D_{\text{max}} \) is strongly correlated to the value of \( Z_n \)
correlation is 0.93; Fig. 3), and hence (for constant CC) to the KS
probability \( P(>Z_n) \).

(iv) Repeat steps (ii) and (iii) until \( D_{\text{max}} \) is minimized. Compare
the image to the model with the 2D KS test and, if necessary, further
minimize \( Z_n \) by applying small changes (e.g. single-pixel shifts) to
the model (since the nature of the relation between \( D_{\text{max}} \) and \( Z_n \)
is not established rigorously). The final simulation will contain the
best guess of the positions and intensities of the assumed sources.
The image is compared to the model using the 2D KS test.

(v) Steps (i) through (iv) are run for a fixed number of sources
(guessed in step (i)). If the final KS probability of similarity is not
satisfactory (e.g. not \( \geq 5 \) per cent), a new source is added at the
location of the highest residual, and the algorithm is repeated from
step (ii).

(vi) If the addition of the last source did not incur a decrease in
the \( Z_n \) statistic larger than its uncertainty (\( \pm 2 \) to \( \pm 4 \) per cent),
the last added source is considered marginal, and the previous best
guess for the number, positions and intensities is taken as the final one.

As noted in step (i) an initial guess with \( N_{\text{initial}} > 1 \) number of
sources can also be fed into the algorithm. Such a guess can be
made either from visual inspection of the image, or after applying
a deconvolution algorithm. We found that Lucy–Richardson (LR)
deconvolution (see Section 6) gives good initial estimates of the
positions of the individual sources. However, since deconvolution
can introduce spurious sources, the initial guess of the number of
sources should be conservative.

4.2 Performance

The above procedure has not been automated, and therefore (be-
cause of subjectivity in ‘guessing’ a simulated source distribution
after having created it) we have not performed tests to determine
explicitly its efficiency in detecting sources. We quote the ability of
the 2D KS test to detect small changes in the positions (within \( \sim 0.2 \)
FWHM) and/or intensities of individual sources (Section 3.2) as an
indication of the power of the iterative algorithm. Nevertheless, we
have devised a method to test the confidence with which a certain
number of sources can be claimed in a given photon distribution.
The method takes our best guess for the source distribution in the
image with a given number of sources (e.g. \( N \)), and compares a
model of it to a faint simulation of our best guess with one source
fewer (\( N - 1 \)). In this way we can test in what fraction \( P_{N-1,N} \)
of the cases our proposed model (with \( N \) sources) can describe a source
distribution with \( N - 1 \) sources. In other words, we test for the sign-
ificance (\( 1 - P_{N-1,N} \)) of the addition the \( N \)th source; \( P_{N-1,N} \) is
thus its false-detection probability. If this comparison is performed
many times (of the order of the number of Monte Carlo simulations
done for each row in Table 3), a \( Z_{n,N}^{-1,N} \) curve for the two guesses
is recovered. The latter can then be compared to a \( Z_{n,N}^{1,N} \) curve,
obtained in a similar fashion comparing \( N \)-source simulations to an
\( N \)-source model.

Example \( Z_{5,N}^{1,N} \) and \( Z_{880,N}^{N-1,N} \) curves for \( N = 5 \) and \( n = 880 \) are
shown in Fig. 4. The overlap of the two curves gives the fraction of
Monte Carlo simulations in which a best-fitting distribution with
\( N - 1 \) sources produces an image that has the same KS similarity
to the model as that of a best-fitting distribution with \( N \) sources.
The ratio of the overlap area to the area of either of the \( Z_n \) curves
(assuming they are both normalized to the same area) is the desired
fraction \( P_{N-1,N} \).

To investigate the dependence of the overlap area \( P_{N-1,N} \) on the
width of the PSF, we ran KS test simulations built with an 8-arcsec

![Figure 3](image-url) **Figure 3.** Values of \( Z_n \) versus the maximum deviation from zero, \( D_{\text{max}} \), in
the smoothed residual. The error bars in \( Z_n \) represent the standard deviation
of \( Z_n \) due to 10 independent subpixelizations of the compared simulations.
The statistics \( Z_n \) and \( D_{\text{max}} \) are highly correlated (correlation coefficient
0.93), and we therefore use a minimum in \( D_{\text{max}} \) as an indication of being
near a minimum in \( Z_n \) [and hence, for constant CC, near a maximum in
\( P(>Z_n) \)].

![Figure 4](image-url) **Figure 4.** A comparison between a \( Z_{880}^{2.5} \) and a \( Z_{880}^{4.5} \) curve obtained using the
5-arcsec predicted HRI PSF. The \( Z_{880}^{4.5} \) curve is obtained from KS tests be-
tween a five-source model and 10000 realizations of a four-source simulation
that best represents the five-source model. The overlap area \( P_{4.5} \) determines
the false-detection probability of the fifth source. Here \( P_{4.5} = 0.23 \).
Gaussian PSF. The result is that for a wider PSF the $Z_n^{N-1,N}$ curve is narrower and is shifted towards smaller $Z_n$ (the model and the simulations look more alike). The $Z_n^{N,N}$ curve, however, is not affected. The overall effect is that the false-detection probability $P_{N,n}^{N,N}$ of the $n$th source increases (from 0.23 to 0.43 for $N = 5$; Table 5).

The above technique can be generalized to produce $P_{N-1,N}$ for arbitrary integers $i < N$ and $j$. An application of $P_{N-1,N}$ is discussed in Section 7.2.

5 APPLICATION TO A DEEP ROSAT IMAGE OF NGC 6397

The iterative source-modelling procedure (Section 4.1) was applied to our 75-ks ROSAT/HRI exposure (1995 March) of the core region of the post-core-collapse globular cluster NGC 6397 (Fig. 5). Standard aspect correction routines (Harris, Silverman & Hasinger 1998a; Harris et al. 1998b; Harris 1999a, b) were applied to the image to improve the S/N ratio. After the aspect corrections the PSF improved from 10.3 $\times$ 8.3 arcsec$^2$ to 8.3 $\times$ 7.9 arcsec$^2$, as measured from the shape of a background point-source quasi-stellar object (QSO) at 3.7 arcmin off-axis (source ‘D’ in Cool et al. 1993). The obtained size of the PSF was still much worse than the predicted 5 arcsec. This effect is not due to the known deterioration of the ROSAT PSF with increasing off-axis angle beyond $\sim$5 arcmin, since the QSO is only 3.7 arcmin from the centre of the field. Residual (unknown) aspect errors are present, and in the analysis below we use a fitted PSF instead of the nominal one.

We analyse the central 128-pixel (64-arcsec) square region of the image, containing 980 counts. Model simulations were created using an analytical PSF fitted to the QSO with the IRAF/DAPHOT routines PSF and ADDSTAR. The PSF comprised a FWHM $\approx$8-arcsec Gaussian core and Lorentzian wings, where the core and the wings could be tilted along different directions in the image. In determining the false-detection probabilities (from the overlap of the $Z_n$ curves), however, for faster iteration we used a symmetrical 8-arcsec Gaussian PSF, meaning the $Z_n$ distributions based on the fitted PSF and on the Gaussian PSF are expected to be indistinguishable, since the test is distribution-free.

Table 5 presents results for the KS statistics of the best-fitting models for a given number of sources. The error in the values of $Z_n$ and $P(>Z_n)$ is the 1σ uncertainty due to subpixelization, as determined from KS comparisons between the model and 10 independent subpixelizations of the same image. It is an estimate of the error in the mean of the $Z_n$ distribution of comparisons between the image and the N-source model.

Following the logic of step (vi) in the KS probability maximization algorithm (Section 4.1), we conclude that four sources are insufficient to represent the image conclusively, since the addition of a fifth source decreases the $Z_n$ statistic by more (17 per cent) than the 3.5 per cent error in the $Z_n$ statistic for four sources. However, the addition of a sixth source is not justified, since the decrease (0.04) in $Z_n$ is smaller than the error (0.05). We therefore claim that five sources are sufficient, and that at least four sources are necessary (at the 1$\sigma$ level) to account for the observed photon distribution in the ROSAT image. The source centroids for the optimal source configurations with four, five and six sources are shown in Fig. 6. The number of counts per source for the five- and six-source cases are the same as listed in Table 4.

Source X6 is sufficiently faint and detached from the central group (X1–X4) that its addition did not necessitate any changes in the prior five-source configuration. It is 11 arcsec away from the closest source (X1), and 4.5 times dimmer than the faintest one (also X1; Table 4). The false-detection probability (determined as $P_{5,6}$) for source X6 is 90 per cent.

Source X1,4 in the four-source best-fitting solution falls approximately in the middle between sources X1 and X4 of the five-source solution, which is consistent with them having comparable intensities and being fainter than X2 and X3 (Table 4).

The derived optimal number of five sources in our 75-ks image of NGC 6397 is consistent with the earlier suggestion (Cool et al. 1993) that at least four X-ray sources (X1, X2, X3 and X5) are present in an 18-ks exposure of the same region (found by visual inspection of the peaks in the image, and confirmed by a one-dimensional azimuthal KS test around the source centroids). The locations of the detected sources are also consistent (up to a $\approx$2-arcsec systematic offset) with the positions of known optical cataclysmic variables

<table>
<thead>
<tr>
<th>$N$</th>
<th>CC of model</th>
<th>$Z_n$</th>
<th>KS prob $P(&gt;Z_n)$ (per cent)</th>
<th>$P_N^b$ (per cent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.21</td>
<td>1.88 ± 0.04</td>
<td>&lt;1</td>
<td>1.5</td>
</tr>
<tr>
<td>4</td>
<td>0.22</td>
<td>1.48 ± 0.05</td>
<td>4 ± 1</td>
<td>43</td>
</tr>
<tr>
<td>5</td>
<td>0.19</td>
<td>1.23 ± 0.05</td>
<td>22 ± 6</td>
<td>90$^a$</td>
</tr>
<tr>
<td>6</td>
<td>0.21</td>
<td>1.19 ± 0.04</td>
<td>26 ± 6</td>
<td>(100)</td>
</tr>
</tbody>
</table>

$^a$For $N = 980$ counts in the NGC 6397 image.

$^b$Probability that a best-fitting model with $N$ sources and the ROSAT image represent the same parent distribution; $1 - P_N^b$ is the significance of adding an ($N + 1$)th source. $P_N^b$ is calculated as $P_{N,5}$ (Section 7.2).

$^c$Estimated as $P_{5,6}$, i.e. the probability that a five-source best-fitting distribution can represent a six-source one; $1 - P_{5,6} = 10$ per cent is the significance of adding a sixth source.

Figure 5. A 75-ks ROSAT/HRI exposure of the central region of NGC 6397, with the detected sources marked. The X-ray image is smoothed with a 2D $\sigma = 2$ arcsec Gaussian. The cluster centre is at (α, δ) = (17:40:41.3, –53:40:25) (Djorgovski & Meylan 1993). The conversion from pixel to celestial coordinates is accurate to within 1 arcsec.
CV1 and CV4 are unresolved in the ∼expected variation (data; Cool et al. 1995, 1998). Such an offset is well within the HST et al. 1995; Edmonds et al. 1999) and/or faint (− offset of the X-ray emission from X4. Its subsequent confirmation as a CV candidate (Hα-emission object) in our follow-up deep ROSAT data. The larger detect cells fail to find more than three sources in the central region of NGC 6397, whereas the 4-arcsec detect cell size is too small for use with our PSF (FWHM ≈ 8 arcsec), and produces an unjustified high number of individual sources. The local detect algorithm (task LDETECT) calculates S/N around each pixel, using the local background (in a region between 1.5 and 2.5 arcsec from the source) as an estimate of the noise. As a result it does not handle crowded fields adequately, and cannot distinguish blended sources. Even the two smallest detection cells (6 × 6 arcsec² and 9 × 9 arcsec²) do not find more than three sources in the image in Fig. 5.

Wavelet detect

This algorithm based on the wavelet transform has only recently been applied to imaging astronomy (Freeman et al. 2001; Damiani et al. 1997, and references therein), and has been demonstrated to out-perform other source-detection algorithms in low-S/N fields. We used an implementation of the wavelet detect based on the Marr wavelet, or the ‘Mexican hat’ function, coded in the WAVDETECT task in the Chandra DETECT 1.0 package. The algorithm is most sensitive to structures of size approximately equal to the width of the Mexican hat function. Running WAVDETECT on our NGC 6397 image (Fig. 5) with transforms of width ≤ 8 arcsec produced only the same three sources already found by the sliding-cell algorithms. This was not unexpected, since in simulated images for the Chandra High Resolution Camera (FWHM = 0.5 arcsec), WAVDETECT is unable to discern point sources less than 2 FWHM apart.

Image deconvolution

There exist a number of widely used image deconvolution algorithms that are applicable to moderately crowded fields. After vicinity comes most likely from both CV1 and CV4), and weigh the measured shifts inversely to the square of the positional uncertainties of the sources along x and y. The 95 per cent confidence radius of the source coordinates corresponds to the maximum shift of a single source from its listed position (keeping all intensities and other source positions constant) that maintains the KS probability above 5 per cent.

6 COMPARISON TO OTHER SOURCE-DETECTION METHODS

To get an idea of the superior performance of our source-modelling scheme, we compared it to established source-detection algorithms, such as the classical ‘sliding-cell’ detect, the wavelet detect, the IRAF/DAOPHOT PSF-fitting task ALLSTAR and ML analysis. We also used deconvolution routines on the image to determine possible source locations. Below we discuss briefly each of these alternatives.

Sliding-cell detect

The sliding-cell detect algorithm is based on S/N calculation and was not expected to perform well in a crowded low-S/N field. Indeed, the two versions of this algorithm in the IRAF/PROD package (tasks IMDETECT and LDETECT in the XSPATIAL package) fail to produce the expected number of X-ray sources in the cluster. IMDETECT uses a constant average background for the entire image, and a variable detect cell size (squares with sides from 4 to 24 arcsec) to search for sources. The larger detect cells fail to find more than three sources in the central region of NGC 6397, whereas the 4-arcsec detect cell size is too small for use with our PSF (FWHM ≈ 8 arcsec), and produces an unjustified high number of individual sources. The local detect algorithm (task LDETECT) calculates S/N around each pixel, using the local background (in a region between 1.5 and 2.5 arcsec from the source) as an estimate of the noise. As a result it does not handle crowded fields adequately, and cannot distinguish blended sources. Even the two smallest detection cells (6 × 6 arcsec² and 9 × 9 arcsec²) do not find more than three sources in the image in Fig. 5.

Wavelet detect

This algorithm based on the wavelet transform has only recently been applied to imaging astronomy (Freeman et al. 2001; Damiani et al. 1997, and references therein), and has been demonstrated to out-perform other source-detection algorithms in low-S/N fields. We used an implementation of the wavelet detect based on the Marr wavelet, or the ‘Mexican hat’ function, coded in the WAVDETECT task in the Chandra DETECT 1.0 package. The algorithm is most sensitive to structures of size approximately equal to the width of the Mexican hat function. Running WAVDETECT on our NGC 6397 image (Fig. 5) with transforms of width ≤ 8 arcsec produced only the same three sources already found by the sliding-cell algorithms. This was not unexpected, since in simulated images for the Chandra High Resolution Camera (FWHM = 0.5 arcsec), WAVDETECT is unable to discern point sources less than 2 FWHM apart.

Image deconvolution

There exist a number of widely used image deconvolution algorithms that are applicable to moderately crowded fields. After
Table 6. X-ray sources detected in NGC 6397.

<table>
<thead>
<tr>
<th>Source</th>
<th>RA (J2000)</th>
<th>Dec. (J2000)</th>
<th>95 per cent conf. radius</th>
<th>Count rate ($k_{s^{-1}}$)</th>
<th>$L_X$ (erg s$^{-1}$)</th>
<th>Optical counterpart</th>
<th>Chandra offset (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>17:40:41.46</td>
<td>-53:40:18.4</td>
<td>2.5 arcsec</td>
<td>1.2 ± 0.5</td>
<td>$3.3 \times 10^{31}$</td>
<td>CV1, CV4</td>
<td>-0.2</td>
</tr>
<tr>
<td>X2</td>
<td>17:40:42.48</td>
<td>-53:40:18.0</td>
<td>1.5 arcsec</td>
<td>2.7 ± 0.6</td>
<td>$7.4 \times 10^{31}$</td>
<td>CV3</td>
<td>0.7</td>
</tr>
<tr>
<td>X3</td>
<td>17:40:42.32</td>
<td>-53:40:27.4</td>
<td>1.5 × 2.5 arcsec$^2$</td>
<td>2.5 ± 0.6</td>
<td>$6.9 \times 10^{31}$</td>
<td>U18, CV2</td>
<td>0.6 ± 1.8</td>
</tr>
<tr>
<td>X4</td>
<td>17:40:41.53</td>
<td>-53:40:26.4</td>
<td>2.0 arcsec</td>
<td>1.3 ± 0.5</td>
<td>$3.6 \times 10^{31}$</td>
<td>CV5</td>
<td>0.7 ± 0.1</td>
</tr>
<tr>
<td>X5</td>
<td>17:40:41.36</td>
<td>-53:40:02.0</td>
<td>1.2 arcsec</td>
<td>3.0 ± 0.6</td>
<td>$8.2 \times 10^{31}$</td>
<td>U24</td>
<td>0.0</td>
</tr>
<tr>
<td>X6</td>
<td>17:40:40.4</td>
<td>-53:40:18</td>
<td>$\sim$</td>
<td>0.4</td>
<td>$&lt;1.1 \times 10^{31}$</td>
<td>U43</td>
<td>0.8 ± 2.8</td>
</tr>
</tbody>
</table>

$^a$Unabsorbed luminosities listed in the 0.5–2.5 keV band, for a cluster distance of 2.2 kpc, and column density of $1.0 \times 10^{21}$ cm$^{-2}$. For best-fitting (Chandra) column densities and bremsstrahlung spectra for individual sources, see Grindlay et al. (2001a).

$^b$U18, U24 and U43 are Grindlay et al. (2001a) Chandra IDs. U18 also identified as either a BY Dra or MSP by Grindlay et al. (2001a), and U43 identified as a BY Dra binary by Taylor et al. (2001). CV2 first identified as H0 object by Cool et al. (1995).

$^c$Offset between the given positions (subscript R) and the ones listed in Grindlay et al. (2001a, subscript C). A bore-sight offset ($\Delta_{R-C}$ = 4.9 pixel and $\Delta_{R-C}$ = 4.6 pixel) has been applied to match the positions of our best-constrained source (X5) and its Chandra counterpart (U24). When the emission from two Chandra/HST sources corresponds to a single ROSAT source, the latter has been associated with the mean position of the Chandra/HST sources.

$^d$The 2σ confidence radius of the position of X6 is infinite, because the KS probability that the model and the image represent the same parent distribution is always above 5 per cent, regardless of the source location.

Comparing results from the IRAF implementations of the maximum entropy method, the Lucy–Richardson (LR) algorithm (both applicable primarily to optical images) and from CLEAN (used mostly in radio imaging), we found that LR deconvolution (Richardson 1972; Lucy 1974) most reliably discerns the five-source distribution found by our iterative source-modelling scheme (Section 5). The positions of the peaks in the deconvolved image are in excellent agreement (to within ±1 pixel = ±0.5 arcsec) with the KS best-fitting source positions, which exemplifies the usefulness of LR deconvolution in analysing crowded fields. Unfortunately, the LR method does not provide a measure of the goodness of fit of these positions and of the significance of the peaks in the reconstructed image. These need to be determined separately with a multi-source fitting routine (since the field is crowded, such as DAOPHOT/ALLSTAR, or the current (2D KS) iterative method. Furthermore, the obtained intensities of the deconvolved sources are in much poorer agreement with the ones from the 2D KS best-fitting model. Nevertheless, LR deconvolution does give an indication for the existence of more than three sources (which could not be determined with the source-searching methods). The LR method thus provides a very good initial guess for the source configuration, which can be input to iterative source-modelling algorithms.

**DAOPHOT/ALLSTAR**

The DAOPHOT package is designed for the analysis of crowded optical images, and as such it assumes that the images are in the Gaussian statistics (high number of counts per pixel) regime. Thus, strictly speaking, the package is inapplicable to data governed by Poisson statistics, such as most X-ray images (including our NGC 6397 image, containing ≤3 counts per pixel), because it severely underestimates random errors. However, until recently DAOPHOT was the only widely available software for reduction of crowded fields, and it has been suggested (Cool et al. 1993) that it can be useful for analysing crowded X-ray fields.

Our experience with ALLSTAR is that it is heavily dependent on several loosely defined parameters, which, in regimes of severe source confusion and low S/N ratio as in our NGC 6397 image (Fig. 5), critically determine the performance of the task. We found that different combinations of the values of the parameters and of the initial guess for the source distribution produced different final results, in which the number of detected sources in the NGC 6397 image varied from two to five. By judiciously adjusting its parameters, ALLSTAR can be made to detect five sources; however, that combination is not favoured statistically over other combinations with fewer sources. In the case when ALLSTAR detects five sources, the obtained positions and intensities are such that the KS probability of similarity with the ROSAT image is <1 per cent ($Z_{n} = 2.2$).

**Maximum likelihood**

Given our method of optimization – minimizing the maximum residual $D_{\text{max}}$ (albeit we then further minimize the KS statistic $Z_{n}$) – ML analysis would be expected to produce a similar fit. This is indeed the approach of Verbunt & Johnston (2000) in analysing the same ROSAT field. The results for the five detected sources (model 1 in Verbunt & Johnston; X1–X5 in this paper) agree well; in addition, our analysis suggests the possible presence of the faint source X6. We choose to employ a 2D KS test to assess the goodness of fit instead, banking on its sensitivity to diffuse distributions. As pointed out by the referee, it is a good test for the location of smeared objects, but it is rather insensitive to their width. Via KS, a source may be deduced to be unresolved, despite having broader profile, which can frequently be the case in Poisson noise-limited images.

We have thus demonstrated that, under conditions of severe source confusion and low S/N, our source-detection method based on a 2D KS test works better than other available techniques. We attribute its performance to the fact that our approach uses the actual PSF in searching for sources (sliding-cell and wavelet detect algorithms do not), that no information is lost to binning (as in the Pearson $\chi^2$ test, used in ALLSTAR), and that it is more sensitive to broad emission than other tests (e.g. ML).

We have not made a comparison of our method against the Pixon deconvolution method (Pina & Puetter 1993). Our method was originally intended to enhance sensitivity for crowded point-source detection; the Pixon method also shows good results for the detection of low surface brightness features.

7 DISCUSSION

7.1 Applicability to distributions with unknown parameters

Rigorously, the presented look-up table, Table 3 (generated by comparing simulations of models with a priori known parameters), is
not applicable when comparing an image of an unknown source distribution to a simulation with known parameters. Lilliefors (1967) investigates this situation for the case of the 1D KS test and sampling from a distribution with unknown mean and variance (‘the Lilliefors test for normality’). He finds that the standard 1D KS test table is too conservative, i.e. with an appropriately generated look-up table (via Monte Carlo simulations), one can reject the null hypothesis that a sampled distribution is normal at a higher significance level than with the standard table.

The implications of this to our case are not known, and have not been investigated. Speculatively extrapolating Lilliefors’ conclusion, the 2D KS test for comparing an unknown to a known distribution should be, if anything, more powerful than presented. This would increase the significance of source X6, making its association with the suggested BY Dra variable more likely.

The advantage of an ML approach here would be that likelihood ratios between different models do not suffer from such problems.

7.2 Significance of the detections

The developed detection significance test for additional sources in Section 4.2 may seem subjective, since prior knowledge is needed about the $Z_{M,N}^{M,N}$ curve (where $M$ is the number of sources in the image). Naturally, this information is not available when working with an astronomical image representing an unknown source distribution, where $M$ is a sought parameter. However, owing to the (nearly) distribution-free character of the pseudo one-sample 2D KS test (Section 3.1.2), all that is needed is the correlation coefficient CC of the counts in the image, which is readily available (CC = 0.17 for the ROSAT image in Fig. 5). Provided that the best-fitting model with $N$ sources represents the image reasonably well (KS probability $\geq 1$ per cent), the $Z_{M,N}^{M,N}$ curve will be indistinguishable from the $Z_{M,N}^{M,N}$ curve of the image, since the correlation coefficients of the $N$-source model and of the ($M$-source) image will be very similar. Indeed, in our case the best-fitting five-source model has $CC = 0.19$, which, given the slow dependence of the $Z_{n}$ distribution on CC, well approximates the $Z_{n}$ distribution for $CC = 0.17$ (the correlation coefficient of the counts in the ROSAT image).

More general than the false-detection probability is the fraction $P_{k}^{N}$ of cases in which the observed image can be represented by a best-fitting model containing $K < N \leq M$ sources, where $K$ does not necessarily equal $M - 1$. Here $N$ is our best guess for the number of sources in the image, and $M$ is the actual (unknown) number of sources. The quantity $1 - P_{k}^{N}$ is the significance level at which we can reject the hypothesis that the image contains only $K$ sources. To determine $P_{k}^{N}$ using the 2D KS test, we need to compare multiple images of the same field to a single model simulation with $K$ sources (Section 3.1.2). Naturally, this cannot be done, since there rarely exist multiple available images of the same field. However, $P_{k}^{N}$ is well approximated by the quantity $P_{k}^{N}$, which is, as discussed above, $Z_{M,N}^{M,N}$ describes well the $Z_{M,N}^{M,N}$ distribution of the image. This is the value listed in Table 5 (using $N = 5$) for the probability that the ROSAT image can be fitted with fewer than five sources. For $K = N - 1$, in $P_{k}^{N}$, we recover the false-detection probability for the $N$th source, as already discussed in Section 4.2.

7.3 Detected sources

The positions and the count rates of the detected sources are in excellent agreement with model I (based on 1995 ROSAT data) of Verbunt & Johnston (2000) in their maximum-likelihood analysis of the same HRI field. A Chandra image of the core region of NGC 6793 (Grindlay et al. 2001a) reveals a greater complexity of sources (Fig. 6). The source ‘doubles’ CV1 and CV4, as well as CV2 and U18, are too close ($\sim 2.5$ arcsec $\approx 0.3$ FWHM) to be distinguished as separate sources in the ROSAT/HRI image, and are represented as blended sources X1 and X4, respectively. The remainder of the sources marked with open squares are too faint ($L_{X} < 10^{31}$ erg s$^{-1}$; Grindlay et al. 2001a) to be detected given the crowedness of the field. None the less, there is a clear one-to-one correspondence between the brightest ($L_{X} > 10^{31}$ erg s$^{-1}$) Chandra sources, and the ones detected in the ROSAT/HRI image using the 2D KS technique.

Although in their model IV Verbunt & Johnston 2000 predict the existence of separate X-ray counterparts to sources CV2 and U18, that model is fitted to 1991 ROSAT/HRI data when CV2 was more prominent in X-rays relative to U18 (hence could be more accurately centroided; cf. fig. 1 in Cool et al. 1993), and three of the sources in the model have fixed positions. On the other hand, in our 2D KS test iterative analysis we have not used any fixed parameters. Moreover, the KS test suggests the existence of source U43, detected (albeit inconclusively, and offset by $\sim 4.5$ arcsec from its Chandra position) as source X6, for which there exists no X-ray identification prior to the Chandra results of Grindlay et al. (2001a). Although the source is fainter ($\log L_{X} = 29.4$; Grindlay et al. 2001a) than other undetected sources in the complex X1–X4, the source must have been $\geq 10$ times brighter to have been detectable with ROSAT (indeed, BY Dra binaries were discovered in globular clusters as faint and flaring X-ray sources, Grindlay et al. 2001b).

8 CONCLUSION

We have developed an application of the 2D KS test (Peacock 1983; Fasano & Franceschini 1987) in a source-detection algorithm for astronomical images. By employing the ‘subpixelization’ technique on 3D astronomical images, we show that the 2D KS test has greater power than the 3D KS test – the intuitive choice for such images.

We use Monte Carlo integration to determine the cumulative values of the proposed model distribution in all four quadrants around each count and, recognizing the deviations that this incurs from the derived $Z_{n}$ distributions, we provide our own reference tables for estimating the KS probability.

We devise an iterative source-modelling routine that employs the KS probability as a goodness-of-fit estimator, and can be used to find the optimum number, positions and intensities of blended sources. We then apply the iteration scheme to a deep (75 ks) ROSAT/HRI exposure of the core region of NGC 6397 and find five blended sources, as well as a possible sixth one. The locations of the five brightest (and possible sixth) X-ray sources match closely (within the positional error bars) the locations of probable CVs and BY Dra systems, discovered with HST (Cool et al. 1993, 1995; Taylor et al. 2001), and confirmed with Chandra (Grindlay et al. 2001a). The sixth source, X6, is a marginal detection with the 2D KS technique and is likely identified with the much fainter Chandra source, U43 (Grindlay et al. 2001a), which is in turn identified with a BY Dra binary, PC-4 (Taylor et al. 2001).

Comparisons to other source-detection schemes (sliding-cell, wavelet detect, DAOPHOT/ALLSTAR, LR deconvolution and ML techniques) applied to the same image demonstrate the superior power of our method in heavily crowded fields with low signal-to-noise ratio. The example with the ROSAT/HRI deep field indicates that the proposed iterative source-modelling scheme can find applications in small-number statistics high-energy imaging, e.g. in deep exposures of globular clusters and extragalactic nuclear regions with Chandra.
where the size of the PSF is often comparable to the angular separation between the objects.

ACKNOWLEDGMENTS

We thank our collaborators Adrienne Cool and Peter Edmonds for numerous discussions. This research was partially supported by NASA/LTSA grant NAG5-3256 and by HST grant GO-06742.

REFERENCES


Harris D. E., 1999a, AAS HEAD Meeting No. 31, Paper 17.12, The ROSAT/HRI: The Last Hurrah! (a.k.a. How to get better images)
Harris D. E., 1999b, The HRI Bug for Aspect Time; at sao-ftp.harvard.edu//pub/rosat/dewob/asptime.doc
Harris D. E., Silverman J. D., Hasinger G., 1998a, Spatial Corrections of ROSAT/HRI Observations from Wobble-related Aspect Problems. US ROSAT Science Data Center/SAO

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