Mon. Not. R. Astron. Soc. **368,** 497–510 (2006) doi:10.1111/j.1365-2966.2006.10125.x

Adaptive binning of X-ray data with weighted Voronoi tessellations

Steven Diehl⋆ and Thomas S. Statler⋆

Astrophysical Institute, Department of Physics and Astronomy, Ohio University, Athens, OH 45701, USA

Accepted 2006 January 30. Received 2006 January 24; in original form 2005 November 28

ABSTRACT

We present a technique to adaptively bin sparse data using weighted Voronoi tessellations (WVTs). WVT binning is a generalization of the Voronoi binning algorithm by Cappellari & Copin, developed for integral field spectroscopy. WVT binning is applicable to many types of data and creates unbiased binning structures with compact bins that do not lead the eye. We apply the algorithm to simulated data, as well as several X-ray data sets, to create adaptively binned intensity images, hardness ratio maps and temperature maps with constant signal-to-noise ratio per bin. We also illustrate the separation of diffuse gas emission from contributions of unresolved point sources in elliptical galaxies. We compare the performance of WVT binning with other adaptive binning and adaptive smoothing techniques. We find that the csmooth tool in CIAO versions 1.1–3.1 creates serious artefacts and advise against its use to interpret diffuse X-ray emission.

Key words: methods: data analysis – techniques: image processing – ISM: general – supernova remnants – galaxies: clusters: general – X-rays: galaxies.

1 INTRODUCTION

X-ray data are generally very sparse in nature. To deal with this problem, astronomers are often forced to either bin or smooth their data. The most commonly used techniques are simply binning to square blocks of a fixed size or convolving with a fixed kernel. However, due to the large dynamic range in many extended objects, ordinary binning and smoothing techniques are never able to capture structure on large scales without masking detail on smaller scales. This deficiency is the motivation for spatially adaptive algorithms.

With the advent of the two major X-ray satellites, *Chandra* and *XMM–Newton*, it is now possible to resolve fine morphological structures in extended X-ray emitting sources, such as galaxies, clusters or supernova remnants. This calls for new techniques to reliably extract spatial information. Sanders & Fabian (2001, hereafter SF01) were the first to answer with a two-dimensional adaptive binning algorithm, applicable to background-corrected intensity images and hardness ratio maps. However, this algorithm is restricted to a limited set of bin sizes, which prevents it from being fully adaptive and from adjusting its resolution so as to keep the signal-to-noise ratio (S/N) constant. This creates jumps in S/N of a factor of $\sim$2, which, along with its quadrilateral bin shapes, can lead the eye and suggest structure that is not there.

Motivated by the different problem of analysing two-dimensional optical integral field spectroscopic data, Cappellari & Copin (2003, hereafter CC03) developed an innovative adaptive binning technique using optimal centroidal Voronoi tessellations (CVTs). Their algorithm is able to smoothly adjust the bin size to the local S/N requirements and does not impose a Cartesian geometry on the image. Unfortunately, computing the optimal tessellation relies on an iterative scheme which can be used only with strictly positive, Poissonian or optimally weighted data whose S/N is guaranteed to add in quadrature. This prevents their full algorithm from being useful in even simple situations in X-ray astronomy, involving data corrected for exposure map effects or background, or in creating hardness ratio maps.

In this paper, we generalize the Voronoi binning technique of CC03 so that it can be used with any type of data. The generalized algorithm makes use of weighted Voronoi tessellations (WVT), and combines the virtues of both the techniques of CC03 and SF01. It is as robust as, and even more versatile than the code of SF01, yet retains the advantage of the flexible bin sizes of CC03. The algorithm produces smoothly varying binning structures that are geometrically unbiased and do not lead the eye.

In Section 2 of this paper, we review the two binning techniques of SF01 and CC03 in more detail, pointing out their advantages and drawbacks. In Section 3, we explain the functionality of the generalized WVT binning technique, and compare its performance to the two older algorithms in Section 4. Section 5 then demonstrates the utility of WVT binning in creating X-ray intensity images, hardness ratio maps and temperature maps, and in disentangling the diffuse gas emission in elliptical galaxies from the contribution of unresolved point sources. Finally, Section 6 quantitatively compares WVT binning to commonly used adaptive smoothing algorithms, before commenting on the availability of the code in Section 7 and ending with conclusions in Section 8.

⋆E-mail: diehl@phy.ohiou.edu (SD); statler@ohio.edu (TSS)

2 EXISTING ADAPTIVE BINNING ALGORITHMS

2.1 Quadtree binning

The pivotal work on spatial binning of sparse X-ray data is that of SF01, who produce surface brightness and colour maps for the analysis of X-ray cluster images. Their algorithm starts with the smallest possible bin size of 1×1 pixel and calculates the S/N for each bin. Each bin with an S/N higher than the user supplied minimal value $(\mathrm{S/N})_{\min}$ is marked as binned,[1] its pixel members are removed from the pixel list, and ignored for the rest of the binning process. In the next iteration, the unbinned pixels are rebinned with square bins of double the side length. The S/N of each bin is computed, and those exceeding the threshold are marked. This process is repeated until either all pixels are binned or the bin size exceeds the image size. Thus, the bins are generally square, with areas of 4^n pixel, except for regions at the transition between two binning levels. There, some pixels may have already been binned on a previous level and removed. The resulting bin shapes can be rectangular, L-shaped, or more complicated. Even non-contiguous bins are common.

Owing to its hierarchical structure, which resembles a quadratic tree commonly used in *N*-body simulations, we refer to this method as 'quadtree' binning. Although slightly different implementations are conceivable, we take the version of SF01 as representative. In Section 4.1, we make a rigorous quantitative comparison between this algorithm and WVT binning.

2.2 Voronoi binning

Motivated by the need to optimally bin integral field spectroscopic data, CC03 present a method to spatially bin two-dimensional images using Voronoi tessellations (VT). The goal is again to obtain a uniform S/N per bin over the entire image, while keeping each bin as compact as possible.

A VT is a partitioning of a region, defined by a set of points called the generators. Each point in the region, or in this case, each pixel in the image, is assigned to the generator to which it is closest. As a consequence of this scheme, the boundary between two adjacent bins is always the perpendicular bisector of the connecting line between the two generators (Fig. 1).

A subset of VTs, called CVTs, has the additional property that the generators coincide with the centroids of the bins. CVTs are meaningful when there is a density, ρ, defined over the region to be binned, and the generators are the density-weighted bin centroids. Since the centroids cannot be calculated before the bins themselves are constructed, it is necessary to construct a CVT by iteration. A helpful tool is the Lloyd algorithm (Lloyd 1982), which iteratively constructs CVTs with generators at each iteration taken as the centroids from the previous step. The Lloyd iterations have the desirable effect of moving generators into regions of higher density, thereby creating smaller bins. For a uniform density, this algorithm tends to create hexagonal lattice structures (Du, Faber & Gunzburger 1999).

In binning an image, one generally works with a signal S_k per resolution element k ('pixel' from now on)[2] and the associated noise per pixel, σ_k. One can compute the S/N of a bin $\mathcal{V}_i$ as

$$(\mathrm{S/N})_i = \frac{\sum_{k\in\mathcal{V}_i} S_k}{\sqrt{\sum_{k\in\mathcal{V}_i} \sigma_k^2}}. \quad (1)$$

For pure Poisson statistics, or certain forms of optimal weighting, the $(\mathrm{S/N})^2$ is additive (CC03). With this restriction, one can make use of a property of the Lloyd algorithm known as Gersho conjecture: applying the Lloyd algorithm to the square of the density tends to produce a configuration with equal mass per bin (Gersho 1979). CC03 exploit this conjecture by applying the Lloyd algorithm to the quantity $(\mathrm{S/N})^2$, thus producing a CVT with a constant S/N per bin.

In order for the Lloyd algorithm to converge, a good initial set of generators is necessary. CC03 solve this problem with a 'bin-accretion' algorithm. Starting from the pixel with the highest S/N in the input image, one grows a bin by accreting nearest neighbours until the bin reaches a minimum S/N or violates an imposed 'roundness' criterion. Then the next bin is started from the pixel closest to the weighted centroid of all previously binned pixels. This method is guaranteed to generate compact bins within the desired S/N range. Bins that do not meet both of these criteria are marked as 'bad' and their pixels are reassigned to the next closest bin. The centroids of the resulting bins are then used as the initial set of generators for the Lloyd algorithm.

By definition, a VT can produce neither gaps in the data nor non-contiguous bins. The tessellation adjusts to uneven boundaries smoothly, and the CC03 algorithm generally converges to a solution with small, spatially independent scatter around the target S/N. The bins are usually very compact, but can get more elongated or ragged close to the boundaries, or in regions with very strong S/N gradients. The principal drawback of the algorithm of CC03 lies in its applicability. The algorithm works only for data in which S/N adds in quadrature, as the iterative part is based on Gersho conjecture. This precludes the possibility of applying the code to background-corrected or exposure-corrected data, hardness ratios or other types of data where $(\mathrm{S/N})^2$ is not additive. The bin-accretion algorithm can be used without this restriction (Fabian et al. 2003; Sanders et al. 2004; Sanders, Fabian & Dunn 2005).

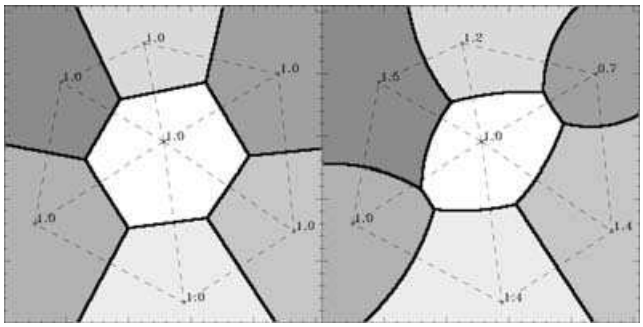

Figure 1. A normal VT (left-hand panel) and a WVT (right-hand panel) with identical bin generators z_i (crosses). The numbers attached to the bins are the associated scalelengths δ_i. The dashed lines connect neighbouring bin generators. Note how the bin boundaries are always perpendicular to them. For a normal VT, the bin boundaries are the perpendicular bisectors; for a WVT, the dashed lines are divided proportional to the respective scalelengths.

[1] The criterion of 'maximal fractional error' given by Sanders and Fabian is equivalent to a 'minimal S/N' threshold.

[2] For integral field spectroscopy, this is the averaged signal over a fixed wavelength interval.

3 ADAPTIVE BINNING WITH WVT

3.1 Introduction to WVT

As described above, a normal VT assigns each pixel k to the generator z_i to which it is closest; i.e. one finds the bin to which the pixel belongs by minimizing its distance to the generator $|x_k - z_i|$ over all bins i. In order to make this definition more flexible, we use a generalization known as a WVT (e.g. Møller 1994). In a WVT, each bin i has an associated scalelength δ_i in addition to its generator z_i, and a pixel k is assigned to the bin that minimizes $|x_k - z_i|/\delta_i$. One can picture δ_i as a factor that stretches or compresses the metric inside the bin i. An intuitive analogy is simultaneous crystal growth, with the generators representing the seeds and the scalelength representing the growth rate (Møller 1994). A WVT is completely described by its set of generators and scalelengths and can therefore be stored very efficiently.

Fig. 1 illustrates the appearance of a WVT (right-hand panel) with a simple example of bins with different relative scalelengths, and compares it to an unweighted VT with the same generators (left-hand panel). Note how the bin boundaries move closer to the generators with the smaller associated scalelengths, making their bins rounder and more compact. For WVTs, the boundaries b between two bins i and j always fulfil the equation $|b - z_i|/\delta_i = |b - z_j|/\delta_j$, implying that the ratio of the bin radii is equal to the ratio of scalelengths. This property is used below to modify the bin sizes by manipulating their relative scalelengths and letting the generator locations adjust. This replaces the Gersho prescription which changes the bin sizes by explicitly moving the generators.

3.2 Adaptive binning algorithm

In the following discussion, we assume that there is some general way to combine the signal and noise of various pixels to calculate the resulting S/N for the bins. We emphasize that the details of how the S/N is actually calculated are irrelevant for the functionality of the WVT binning algorithm.

Our algorithm creates a WVT, choosing the scalelengths δ_i such that the bins have a near-uniform S/N distribution, with the least possible scatter around the target signal-to-noise ratio $(\mathrm{S/N})_\mathrm{T}$. To find the appropriate scalelengths, it is useful to consider the quantity μ_i, defined by

$$\mu_i = \frac{(\mathrm{S/N})_i}{A_i}, \tag{2}$$

where A_i is the bin area and $(\mathrm{S/N})_i$ is the S/N ratio in bin i. The algorithm is aiming for a configuration where the bins have an S/N equal to the target value, $(\mathrm{S/N})_\mathrm{T}$. In this configuration, we should have, approximately,

$$\mu_i = \frac{(\mathrm{S/N})_\mathrm{T}}{q\,\delta_i^2}, \tag{3}$$

where q is a dimensionless constant that depends weakly on bin shape (for circular bins, $q = \pi$). Combining equations (2) and (3) gives a rule for setting the scalelength at each iteration:

$$\delta_i = \sqrt{\frac{(\mathrm{S/N})_\mathrm{T}}{q\,\mu_i}} = \sqrt{\frac{A_i}{q}\frac{(\mathrm{S/N})_\mathrm{T}}{(\mathrm{S/N})_\mathrm{i}}}. \tag{4}$$

Since the binning depends only on the ratio of the scalelengths, the value of q is unimportant. We show below that good results are obtained taking $q =$ constant for all bins, regardless of shape. This prescription replaces the Gersho–Lloyd procedure which, at each iteration, moves the generators to the $(\mathrm{S/N})^2$ weighted centroids. Such weighting is superfluous in our algorithm, so we adopt the geometric bin centres as the new generators.

The WVT binning procedure thus proceeds as follows:

(i) Start with an initial WVT.
(ii) For each bin i, evaluate the signal-to-noise ratio $(\mathrm{S/N})_i$, the area A_i and the geometric centres z_i.
(iii) Calculate the scalelength δ_i for each bin according to equation (4).
(iv) Reassign all pixels according to the new WVT with generators z_i and scalelengths δ_i.
(v) Return to step (ii) until the bins stop changing significantly.

A binning with constant S/N across the field is a natural stable fixed point of this iteration scheme, as it satisfies the relation $\delta_i/\delta_j = \sqrt{A_i/A_j}$.

As in ordinary Voronoi binning, this algorithm requires a good set of initial generators. We adopt the solution of 'bin accretion' given by CC03 with a few modifications for speed, relaxing some acceptance criteria to suit sparse and not strictly positive data (e.g. background-subtracted X-ray images). We also employ a soft lower S/N boundary for accepting bins, in which the S/N has no longer to be larger than the fixed target S/N. Instead, accretion terminates if the addition of another pixel would increase the scatter around the target S/N. This modification keeps the average S/N closer to the target value.

4 PERFORMANCE

4.1 Comparison with Quadtree

While the simplicity of the quadtree binning algorithm makes it easy to understand and apply, there are several disadvantages, which we illustrate in this section using simulated X-ray data. A suitable model for the surface brightness profiles $I(r)$ of a galaxy or galaxy cluster is the circular β model (Sarazin 1988):

$$I(r) = I_0 \left[1 + \left(\frac{r}{r_\mathrm{c}}\right)^2\right]^{0.5-3\beta} + I_\mathrm{Bg}, \tag{5}$$

where I_0 is the central surface brightness, r_c is the core radius, β is a slope parameter and I_Bg is an additive, flat background. We adopt the same parameters used by SF01: the image size is set to 512×512 pixels, r_c is 128 pixels, I_0 is 100 cts pix^{-1}, β is 0.67 and I_Bg is 20 cts pix^{-1}. The simulated X-ray image is obtained by populating the image with counts according to a Poisson distribution. For the quadtree algorithm, a minimum S/N limit of 14 ($\sim 20/\sqrt{2}$) produces an average S/N of 19.92 in the test image and is therefore chosen as the equivalent to a target S/N of 20 for the comparison with the WVT algorithm.

The results of the quadtree and WVT binning algorithms are shown on the left- and right-hand panels of Fig. 2, respectively. The middle panels show the full images; the upper panels zoom in on a small region to emphasize the differences between the binning structures. Note that the quadtree algorithm regularly forms non-contiguous bins and can leave single pixels or small sets of pixels 'stranded'. In a few cases, these pixels can be directly picked out as isolated dark spots in the image, since a larger binning level usually also corresponds to a lower average flux per bin. This effect occurs predominantly in regions where neighbouring pixels have already been binned on a previous binning level. SF01 offer two ways to deal with this problem. The first is to handle isolated sets of pixels of a non-contiguous bin separately, violating the minimum

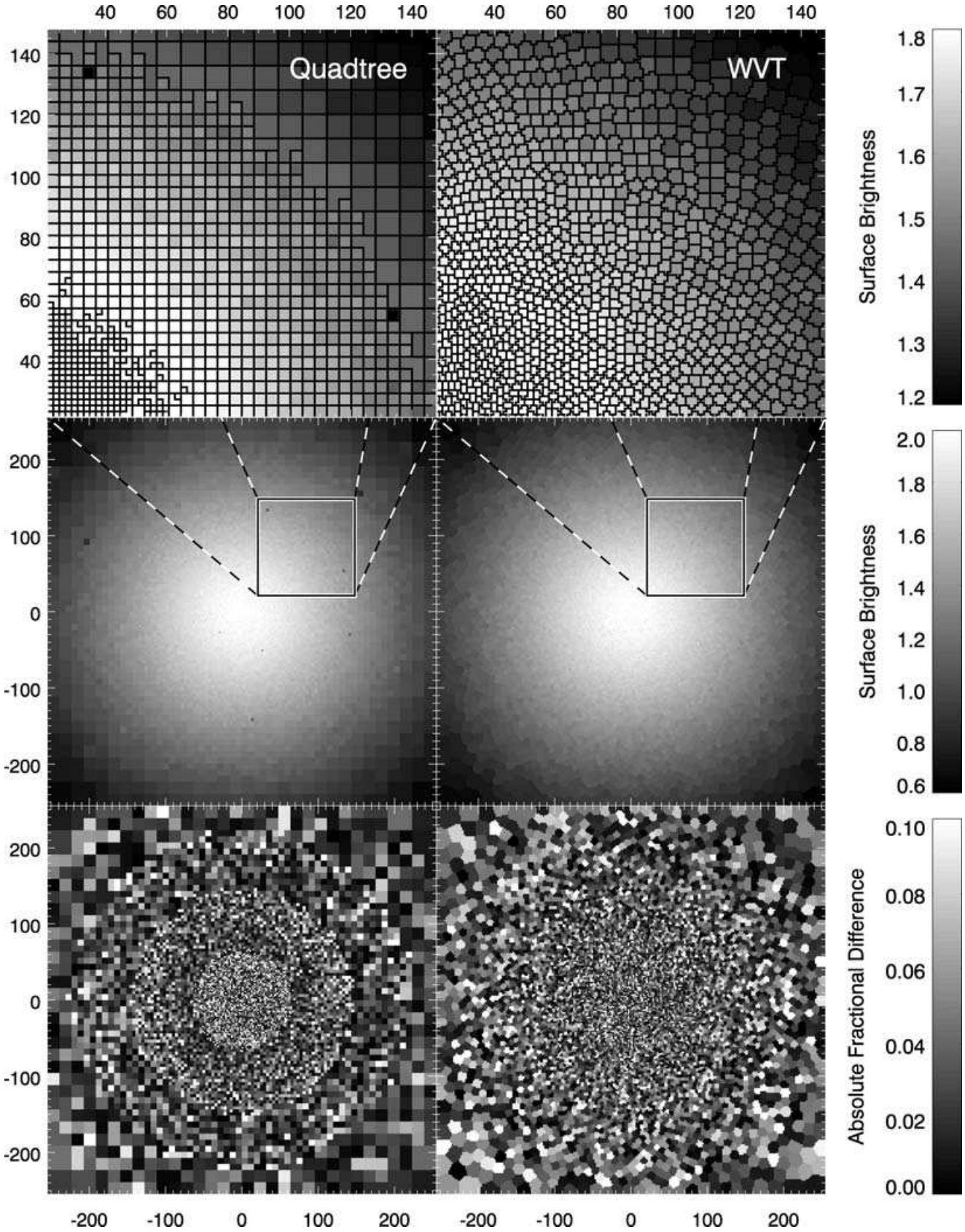

Figure 2. Comparison between Quadtree (left-hand column) and WVT binning (right-hand column); middle panels: logarithmically scaled, binned intensity images. The square indicates the region of the zoom-in shown in the upper panels. Each bin has been outlined to emphasize the difference in the binning structure. Note the darker 'stranded' bins on the left-hand panels; lower panels: absolute fractional difference between the model surface brightness and the binned simulated data.

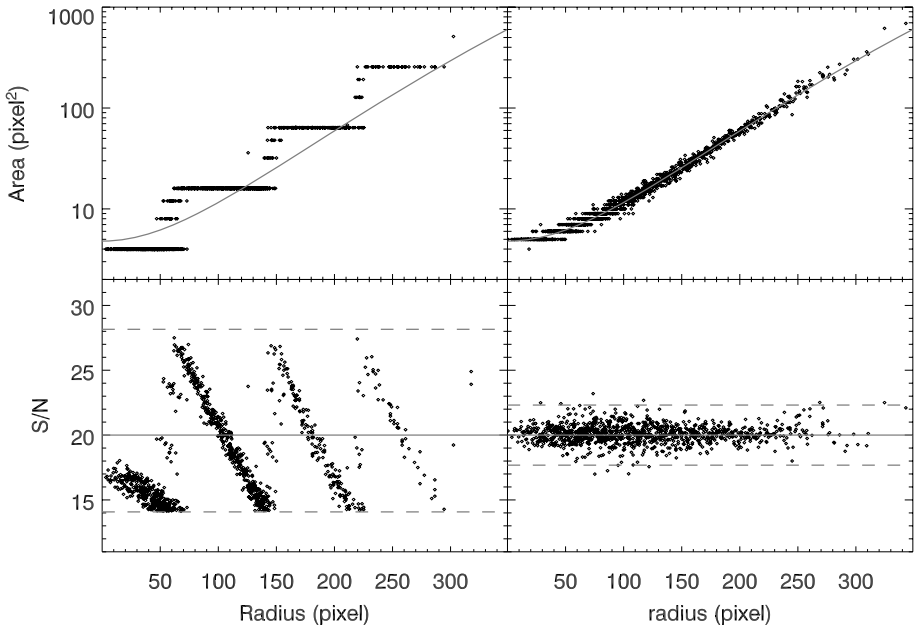

Figure 3. Upper panels: area per bin versus radius for quadtree (left-hand panels) and WVT (right-hand panels), the solid line indicates the theoretical prediction to produce a constant S/N of 20 per bin for our test model. Lower panels: corresponding S/N per bin; note the jumps in S/N due to the discrete bin sizes for the quadtree binning, which is completely absent in WVT; the solid line shows the target S/N, the dashed lines indicate the natural scatter of ∼2 in quadtree and the 3σ rms scatter in WVT.

S/N criterion as the bin is being split up. The alternative is to redistribute the pixels to an adjacent neighbouring bin. In the latter case, the S/N of the neighbouring bin will be elevated, which can lead to an increased scatter in S/N. At the same time, one sacrifices resolution, as the effective number of bins is decreased. Throughout the remainder of this discussion, we do not enforce contiguous bins in the quadtree algorithm for simplicity. We simply note that this problem is absent in the WVT algorithm, which can easily be made to enforce contiguous bins.

The main problem with the quadtree algorithm lies in its small set of discrete bin sizes. Except in small transition regions, where the bin shapes are not square, the bin area is restricted to values of 4^n pixels. This discontinuous distribution of bin sizes is visible in the binned image, and illustrated in the upper panels of Fig. 3, which show the radial dependence of the bin areas for quadtree (left-hand panels) and WVT (right-hand panels) binning. The solid line indicates the optimal, theoretical bin size needed to produce the target S/N. The discrete steps in the quadtree bin area translate into an inhomogeneous S/N distribution, shown in the lower left-hand panel of Fig. 3. Each sharp increase in S/N corresponds to a sudden jump in bin size, which decreases the local resolution beyond the requirements of the target S/N. These jumps in S/N are easily visible as circular rings in the fractional difference image in the bottom left-hand panel of Fig. 2, showing that the binning algorithm can create spurious structure. In contrast, WVT binning allows bins to adjust their size smoothly in single pixel steps, which results in a reduced scatter around the target S/N and the removal of misleading spurious features (bottom right-hand panels of Figs 2 and 3).

The spatially correlated fractional error distribution resulting from quadtree binning can be particularly misleading when the bin value is decoupled from the actual S/N distribution. A good example is a hardness ratio map. Here, the signal is given by a flux ratio of two independent bandpasses, whereas the S/N is determined by the total flux of both bands. Fig. 4 show quadtree- and WVT-binned hardness ratio maps of the Perseus cluster. The eye identifies two concentric rings in the quadtree-binned map (left-hand panel) at around 50 and 150 arcsec. These features are imprints of the discontinuous jumps in bin size, and are completely absent in the WVT-binned map (right-hand panel). WVT-binned hardness ratio maps are described in more detail in Section 5.2.

4.2 Comparison with VT

As the WVT algorithm generalizes the method of CC03, which was designed for optical integral field spectroscopic data, it is natural to test the code on the same type of data. We apply our WVT binning algorithm to the test data provided by CC03, in their online code release. The test data consist of a list of coordinates and signal and noise values of the wavelength-integrated spectra of a SAURON (Bacon et al. 2001) observation of NGC 2273.

Fig. 5 compares the results of both algorithms for a target S/N of 50. Both yield consistent results with a comparable scatter around the target S/N of only ∼6 per cent. The only noticeable difference lies in the compactness of the individual bins, especially close to the border and in regions of strong gradients. While the code of CC03 tends to generate strongly elongated shapes in these cases, the WVT bins stay consistently rounder. To quantitatively measure roundness, we introduce the average bin radius $R^{\rm av}$ and the effective bin radius $R^{\rm eff}$:

$$R_i^{\rm av} = \frac{1}{A_i} \sum_{j \in \mathcal{V}_i} R_j, \tag{6}$$

$$R_i^{\rm eff} = \sqrt{\frac{A_i}{\pi}}. \tag{7}$$

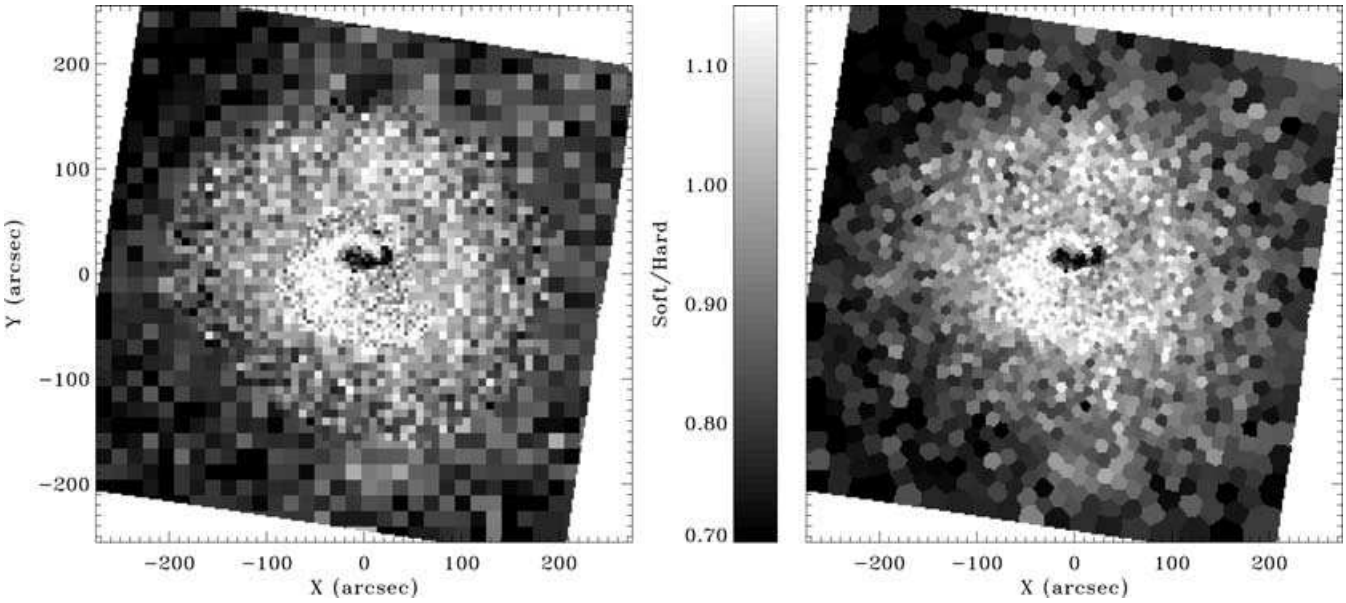

Figure 4. Comparison of quadtree (left-hand panel) and WVT binning (right-hand panel); both panels show adaptively binned hardness ratio maps of the core of the Perseus cluster, with dark colours indicating regions of higher temperature and/or lower photoelectric absorption. Note how quadtree binning leads the eye into identifying two ring structures, due to the strong jumps in the S/N where the bin area suddenly quadruples.

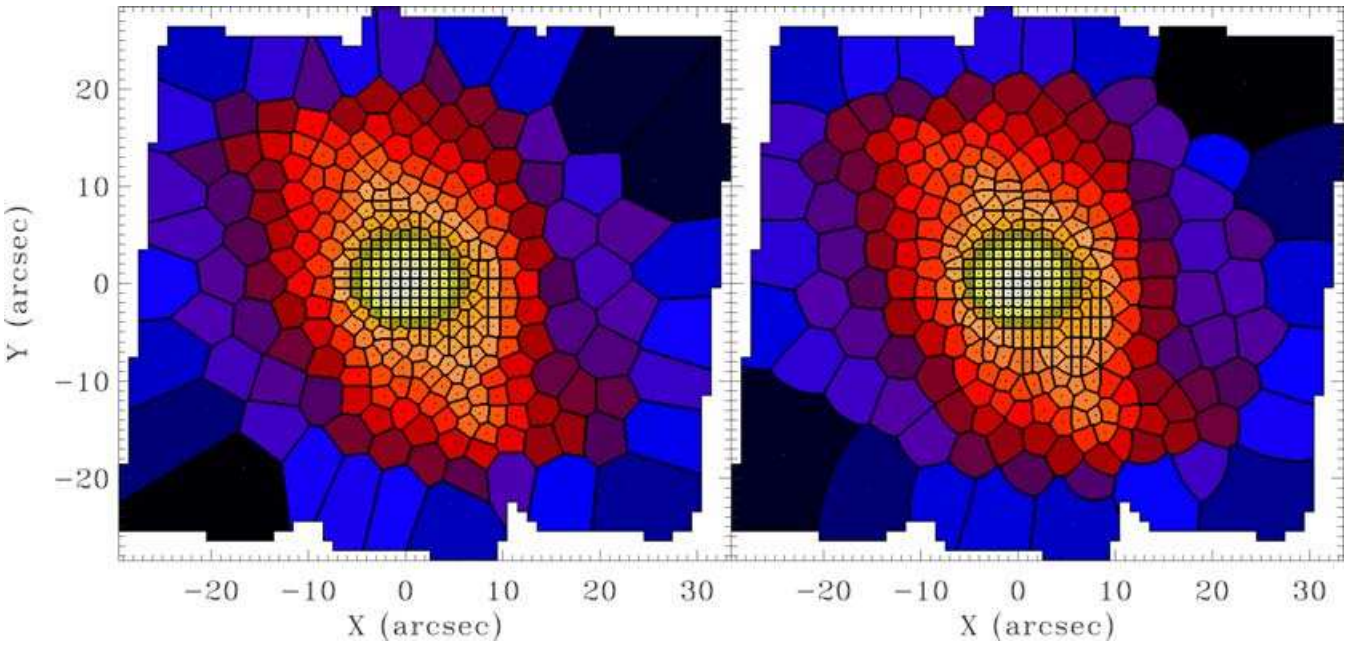

Figure 5. Final result after application of the Voronoi binning code (left-hand panel) and the WVT algorithm (right-hand panel) of CC03 to the SAURON data of NGC 2273. After completion of the binning process, the bins have been projected on to a finer grid to make it easier to identify differences in the shape of the bins.

The more compact the bin, the smaller is the ratio $R^{\rm av}/R^{\rm eff}$. Its minimum at a value of $2/3$ represents a perfectly circular bin. Fig. 6 shows $R^{\rm av}/R^{\rm eff}$ as a function of the distance from the galaxy centre and confirms that the WVT algorithm (filled circles) produces more compact bins without edge effects. At a distance of 20 arcsec, Voronoi bins given by CC03 (open circles) get more elongated, as shown by the jump in $R_{\rm av}/R_{\rm eff}$. This is mainly due to the use of a weighted bin centroid in CC03, which pushes the generators toward the bright end of the bin, elongating the bins in the opposite direction.

5 APPLICATIONS TO X-RAY DATA

5.1 Intensity maps

The most common application for adaptive binning in X-ray astronomy is intensity binning. As discussed above, Voronoi binning algorithm of CC03 is valid only for purely Poissonian data. This would correspond to raw counts for a perfectly flat detector response without any background. However, real X-ray data are not as simple. Many X-ray faint targets have surface brightness values comparable

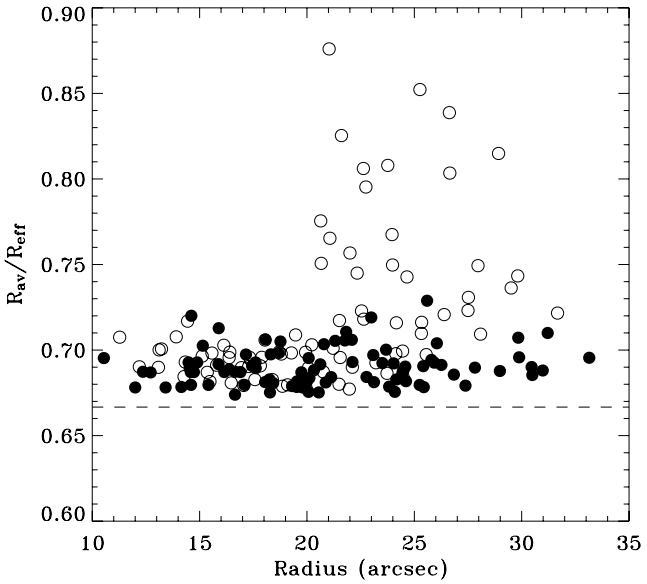

Figure 6. Comparison of bin compactness between Voronoi binning (open circles) and WVT binning (filled circles). The dashed line indicates the theoretical limit for a perfect circle.

to the background, which itself may be spatially dependent. In addition, real observations exhibit strong variations in effective area per pixel, due to chip gaps, node boundaries, or partly overlapping multiple exposures. The effective area of an observation E_k per pixel k is saved in an 'exposure map', which together with the effective exposure time τ, can be used to convert raw counts C_k per pixel into a flux F_k with physical units[3] of photons s^{-1} cm^{-2} $arcsec^{-2}$:

$$F_k = \frac{C_k}{E_k\,\tau} - B_k. \tag{8}$$

Here, B_k is the background flux per pixel. The variance in the same pixel can be expressed as

$$\sigma^2_{F_k} = \frac{C_k}{E_k^2\,\tau^2} + \sigma^2_{B_k}, \tag{9}$$

where σ_{B_k} denotes the uncertainty that is attached to the background value. The prescription for combining these quantities to produce an S/N per bin is given by equation (1). For the hypothetical case where $\tau = 1$, $E_k = 1$ and $F_{\mathrm{Bg}_k} = 0$, the signal F_k reduces to pure counts and the binning scheme will converge to a solution with a constant number of counts per bin.

We use the well-known 50-ks *Chandra* observation of Cassiopeia A (Hwang, Holt & Petre 2000) to demonstrate the power of adaptive binning for X-ray images. The lower left-hand panel of Fig. 7 shows the unbinned, exposure map-corrected counts image for the full exposure. In the panel directly above, we restrict the data to only 1 ks of exposure time. We bin this image with the WVT algorithm to a target S/N of 5 (upper right-hand panel). A comparison with the full 50-ks exposure shows that WVT binning successfully reduces the noise, bringing out the large-scale features in the outer parts of the image, while keeping the appropriate resolution in the better-exposed filamentary features. Even in this short exposure, one is able to pick out the central neutron star in the WVT-binned image.

[3] Alternatively, one can multiply the photon counts with their detected energy to get units of erg s^{-1} cm^{-2} $arcsec^{-2}$; see http://cxc.harvard.edu/ciao/download/doc/expmap_intro.ps for more details on exposure maps.

The image in the lower right-hand panel of Fig. 7 shows the full 50-ks image, adaptively binned to an S/N of 20, to demonstrate the applicability of WVT binning to a different S/N regime.

5.2 Hardness ratio maps

Another useful tool in X-ray astronomy is the hardness ratio (or 'colour') map. The hardness ratio H^{AB} can generally be defined as the quotient between the fluxes F_A and F_B in two different bands A and B, summed over all pixels of the bin $\mathcal{V}_i$:

$$H_i^{AB} = \frac{\sum_{k\in\mathcal{V}_i} F_{A,k}}{\sum_{k\in\mathcal{V}_i} F_{B,k}}. \tag{10}$$

The associated error can be expressed in terms of the noise in the individual bands:

$$\sigma_{H_i^{AB}} = H_i \sqrt{\frac{\sum_{k\in\mathcal{V}_i} \sigma^2_{A,k}}{\sum_{k\in\mathcal{V}_i} {F_{A,k}}^2} + \frac{\sum_{k\in\mathcal{V}_i} \sigma^2_{B,k}}{\sum_{k\in\mathcal{V}_i} {F_{B,k}}^2}}. \tag{11}$$

Depending on the choice of energy bands, a hardness ratio map can be used as a diagnostic for any spectrally identifiable properties, such as temperature gradients or photoelectric absorption features (e.g. SF01). A general discussion about the physical interpretation of these maps, an appropriate choice of bands, and a generalization to n different bands can be found in SF01 or Fabian et al. (2000).

We use the well-known 25-ks *Chandra* observation of the Perseus cluster (Fabian et al. 2000) to demonstrate a WVT-binned hardness ratio map. The right-hand panel of Fig. 4 shows a colour map for the Perseus cluster, in which bright colours indicate regions of lower temperature and/or lower photoelectric absorption. Our choice of bands (A: 0.3–1.2 keV, B: 1.2–5 keV) shows both effects for illustrative purposes; in principle, a different choice is able to separate these two properties. The sharp dark feature close to the centre is due to the photoelectric absorption 'shadow' of an infalling dwarf galaxy in the line of sight (Fabian et al. 2000). To the north-east of the centre, one can pick out a giant radio cavity, with cooler rims surrounding it. The smooth colour gradient toward the centre also supports a cooling flow model (see e.g. Sarazin 1988) and the 'swirl' of the bright emission has been interpreted as a sign for angular momentum of the infalling gas (Fabian et al. 2000). Note that the WVT colour map does not contain the spurious circular features present in its quadtree counterpart.

5.3 Maps of temperature (or other spectral parameters)

Hardness ratio maps are often insufficient to disentangle the spectral components of extended sources. This requires a detailed spectral analysis of multiple regions within the field of view. Current state-of-the-art techniques to generate maps of temperature or other spectral parameters usually specify a regular grid of points, within which circular or square regions are 'grown' until they reach a minimum number of counts for the spectral analysis (e.g. Nulsen et al. 2002; O'Sullivan, Vrtilek & Kempner 2005). One can extract a spectrum and create response files for each region and feed them into an X-ray spectral fitting package such as XSPEC,[4] ISIS[5] or SHERPA.[6] Just as in adaptive smoothing (see also Section 6), the measurements in this 'adaptively accreted' temperature map are not independent of

[4] http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/

[5] Interactive Spectral Interpretation System, http://space.mit.edu/ASC/ISIS/.

[6] http://cxc.harvard.edu/sherpa/

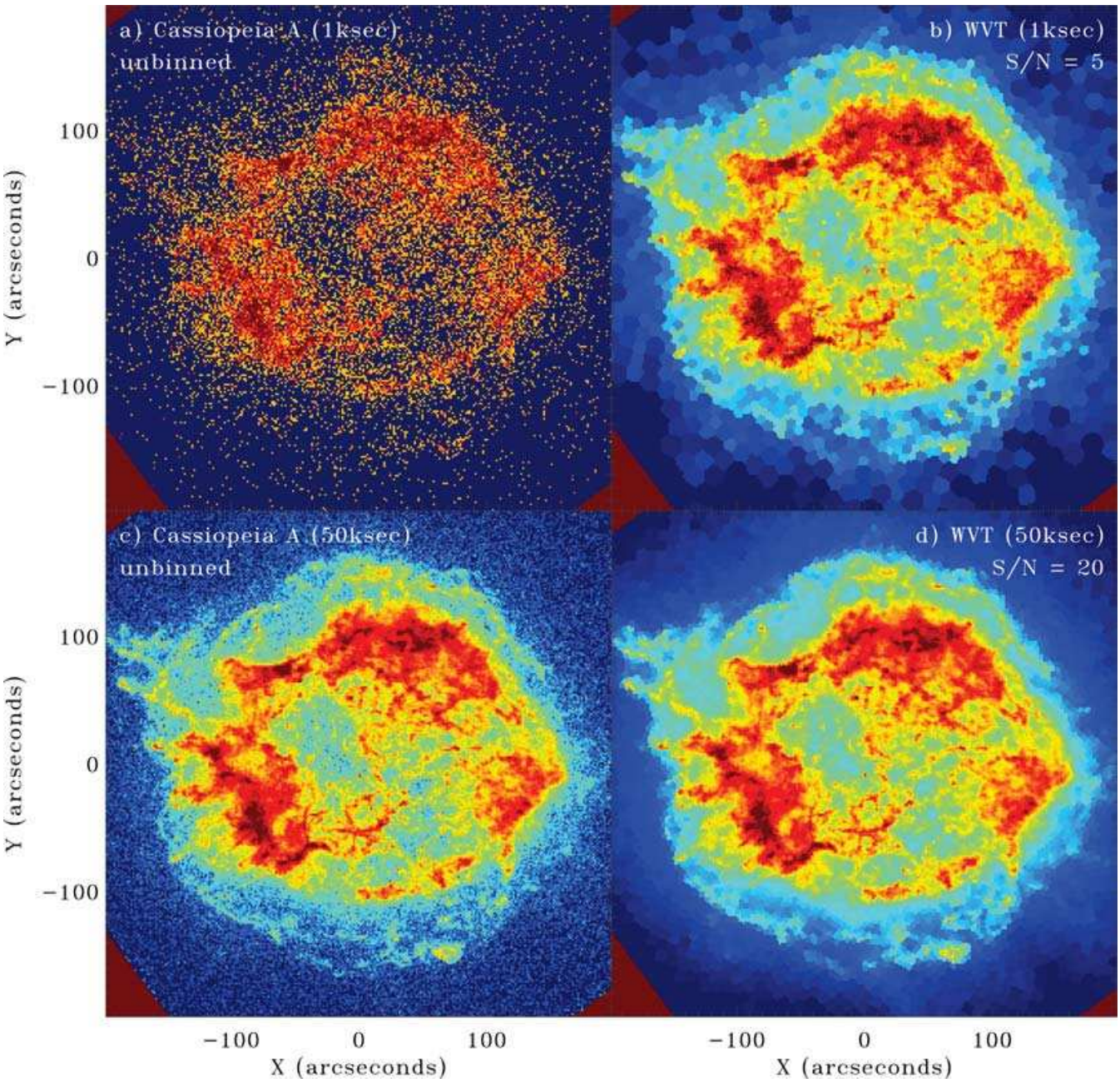

Figure 7. Top: flux-calibrated *Chandra* image of Cassiopeia A with an exposure time of 1 ks (left-hand panels) and the same data adaptively binned with WVT to a target S/N of 5 (right-hand panels). Bottom: Cassiopeia A with the full exposure of 50 ks (left-hand panels) and the same image binned to an S/N of 20 (right-hand panels).

each other. The user is burdened with the task of deciding which features are actually resolved and thus believable, i.e. which features are larger than the extraction regions. WVT binning, on the other hand, has the power to automatically divide the field into unbiased, independent bins with constant source counts per bin, while keeping the individual bins as compact as possible. In addition, this reduces automatically the number of time-consuming spectral fits from the total number of pixels to the number of independent bins.

To demonstrate this capability, we adaptively bin the 46-ks *Chandra* observation of NGC 4636 (Jones et al. 2002) to 900 counts per bin. We then extract a spectrum for each bin and fit it with an absorbed, single-temperature APEC[7] model. The left-hand panel of Fig. 8 shows the resulting temperature map, with the corresponding relative error distribution on the right-hand panel. This figure is directly comparable to fig. 2(a) of O'Sullivan et al. (2005). They interpret the asymmetric temperature distribution as the result of hotter gas surrounding the cool core of NGC 4636, which is penetrated by a 'plume' of gas extending to the south-west. Inside of this plume sits a concave, rising radio bubble, surrounded by cool rims. The

[7] Astrophysical Plasma Emission Code, http://cxc.harvard.edu/atomdb/sources_apec.html.

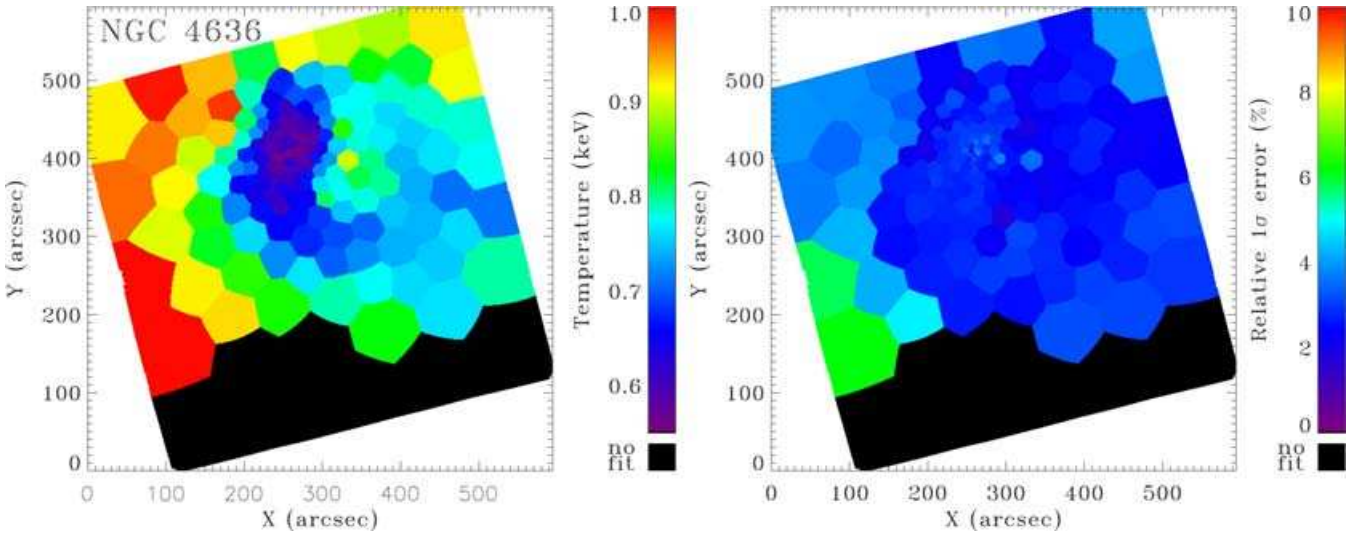

Figure 8. Temperature map of NGC 4636. Temperatures are scaled linearly from 0.55 to 1.0 keV, as indicated by the colour bar. The relative 1σ uncertainties, in per cent, of the fitted values are shown in right-hand panel. The error distribution is not completely uniform in this case owing to a combination of different levels in background contribution, degradation of the instrument response for large off-axis angles, and differences in the spectral shape for varying temperatures.

higher temperature inside this cavity is interpreted as a projection effect, as the bubble pushes the cooler gas away and thus increases the relative contribution of the hotter surrounding gas along the line of sight. The WVT-binned temperature map shows the same large-scale features as the O'Sullivan et al. (2005) map, with no ambiguity as to their statistical significance.

5.4 Isolation of different components

In many astronomical sources, the observed emission comes from multiple overlapping components. Good examples are normal elliptical galaxies, where the diffuse X-ray emission is made up of contributions from interstellar gas and low mass X-ray binaries (LMXBs). Because of their spectral differences, hot gas and LMXBs contribute differently to the soft- and hard-band images. Diehl & Statler (2005) and Diehl & Statler (in preparation) show how this fact can be exploited to recover the gas emission alone. Let $F_{S,k}$ and $F_{H,k}$ represent the background-subtracted soft and hard images in each pixel k. We can express both as linear combinations of the unresolved point source emission P_k, the gas emission G_k, and their respective softness ratios γ and δ:

$$F_{S,k} = \gamma P_k + \delta G_k, \tag{12}$$

$$F_{H,k} = (1-\gamma)P_k + (1-\delta)G_k. \tag{13}$$

Thus, the uncontaminated gas image and its associated noise can be expressed as

$$G_k = \frac{1-\gamma}{\delta-\gamma}\left[F_{S,k} - \left(\frac{\gamma}{1-\gamma}\right) F_{H,k}\right]; \tag{14}$$

$$\sigma_{G,k} = \frac{1-\gamma}{\delta-\gamma}\sqrt{\sigma^2_{S,k} + \left(\frac{\gamma}{1-\gamma}\right)^2 \sigma^2_{H,k}}. \tag{15}$$

Diehl & Statler (2005) discuss the determination of the constants γ and δ with spectral models.

Fig. 9 demonstrates the isolation of the gas emission using WVT binning, in a simulated observation. We assume gas and LMXB sources with very different spatial distributions for purposes of illustration. We adopt an elliptical de Vaucouleurs profile for the LMXBs (top left-hand panel), and a β model, with a nearly orthogonal major axis, for the gas (top right-right panel) and simulate hard- and soft-band images. The bottom left-hand panel of Fig. 9 shows the full-band emission, which is nearly round. Applying WVT adaptive binning (lower right-hand panel) to the gas image (equation 14), we are able to reconstruct the true shape of the diffuse gas emission very accurately.

6 ADAPTIVE BINNING VERSUS ADAPTIVE SMOOTHING

6.1 Adaptive smoothing in X-ray astronomy

Two adaptive smoothing algorithms are in widespread use due to their inclusion in the main data analysis systems of *Chandra* and *XMM–Newton*. Confusingly, both algorithms are named asmooth. The most widely known asmooth algorithm was invented by Ebeling, White & Rangarajan (2006) and implemented in IDL. The IDL implementation has not been publicly available prior to 2006. An early C++ version has been distributed as part of the *Chandra* Interactive Analysis of Observations (CIAO) software since 1999 under the name of csmooth. For clarity, we will refer to both versions. The second asmooth algorithm is the adaptive smoothing tool of the *XMM* Science Analysis System (*XMM*-SAS), which we call *XMM*-asmooth. Although the output of any of these algorithms is generally not used for quantitative analyses, they have become the primary tools to create 'pretty pictures' for papers, talks and press releases. Adaptive smoothing algorithms are thus instrumental in forming and influencing the perceptions of the broader astronomical community and the public.

In adaptive smoothing, the size of the smoothing kernel changes over the field of view to create a constant S/N per pixel in the output image. It is worth emphasizing that the number of independent measurements is equally decreased in adaptively smoothed and adaptively binned images of the same target S/N. Thus, one does

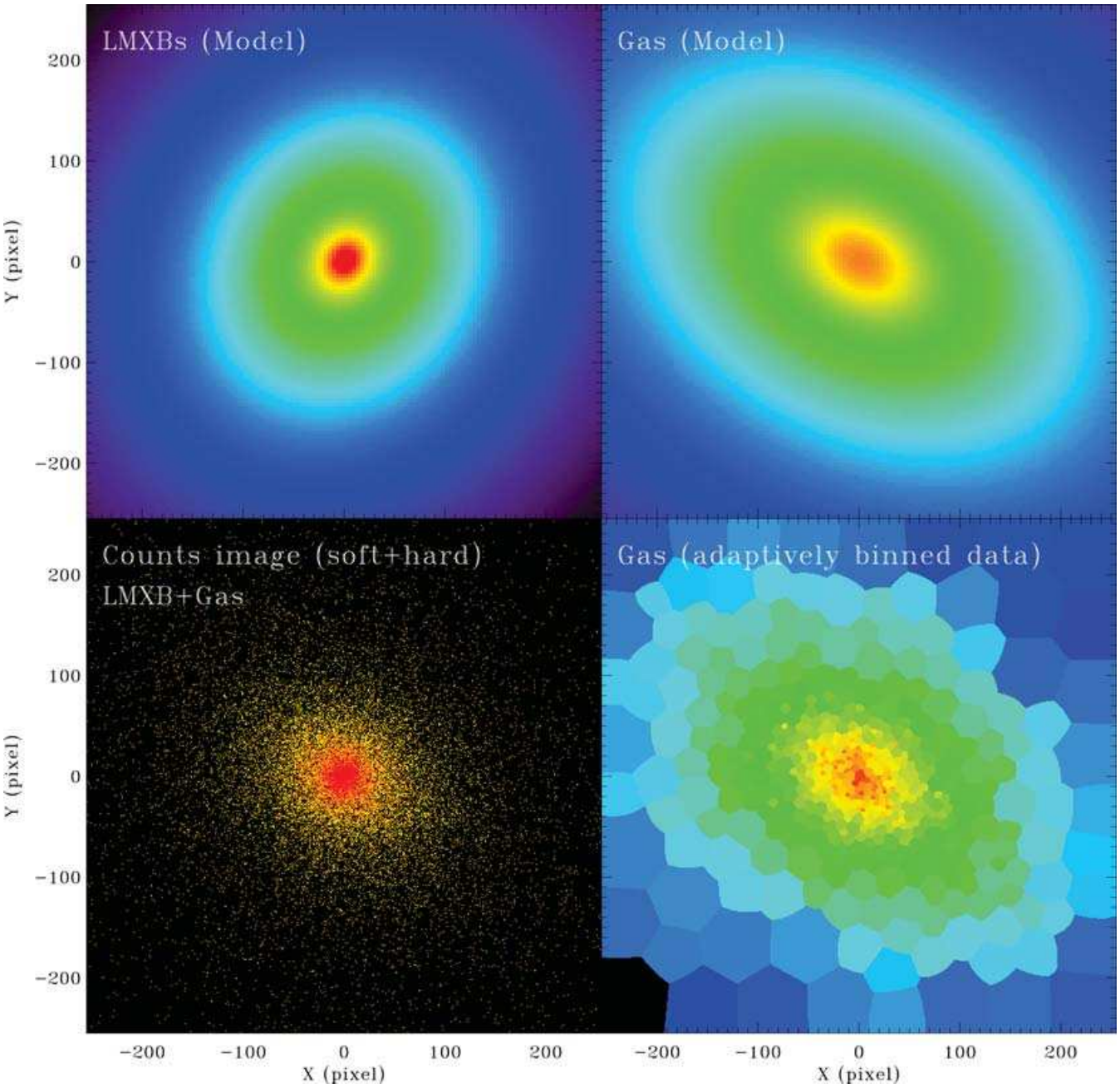

Figure 9. Upper left-hand panel: model surface brightness distribution for unresolved point sources; upper right-hand panel: model surface brightness distribution for the hot, isothermal gas; lower left-hand panel: simulated Poisson image for the full band, including contributions of both point sources and gas; lower right-hand panel: reconstruction of the gas surface brightness distribution with the help of WVT binning.

not gain any additional spatial information by smoothing rather than binning.[8]

In this section, we give some cautionary advice on the interpretation of adaptively smoothed images. To illustrate, we take a β-model surface brightness distribution (equation 5) with $I_0 = 10\,\mathrm{cts\,pix^{-1}}$, $\beta = 0.67$ and $r_c = 64\,\mathrm{pix}$, with a background of $2\,\mathrm{cts\,pix^{-1}}$, shown in the upper left-hand panel of Fig. 10. We simulate a counts image, which we adaptively smooth or bin to a target S/N of 5.

[8] An exception to this statement is when there is filamentary structure narrower than the local bin size, just above the detection threshold.

6.2 Comparison with *XMM*-asmooth

The *XMM*-asmooth algorithm is thoroughly described in the *XMM*-SAS 6.0.0 user manual.[9] The basic idea is to increase the size of the smoothing kernel for each pixel until the pixel can 'accrete' enough signal to meet the S/N requirement. Thus, each pixel has a scale associated with it, which determines the size of the convolution kernel that contributes to the smoothed flux value at this point. In WVT binning, each *bin* has a scale associated with it. In both cases, the scale is determined from the local S/N distribution. We

[9] http://xmm.vilspa.esa.es/sas/current/doc/asmooth/index.html

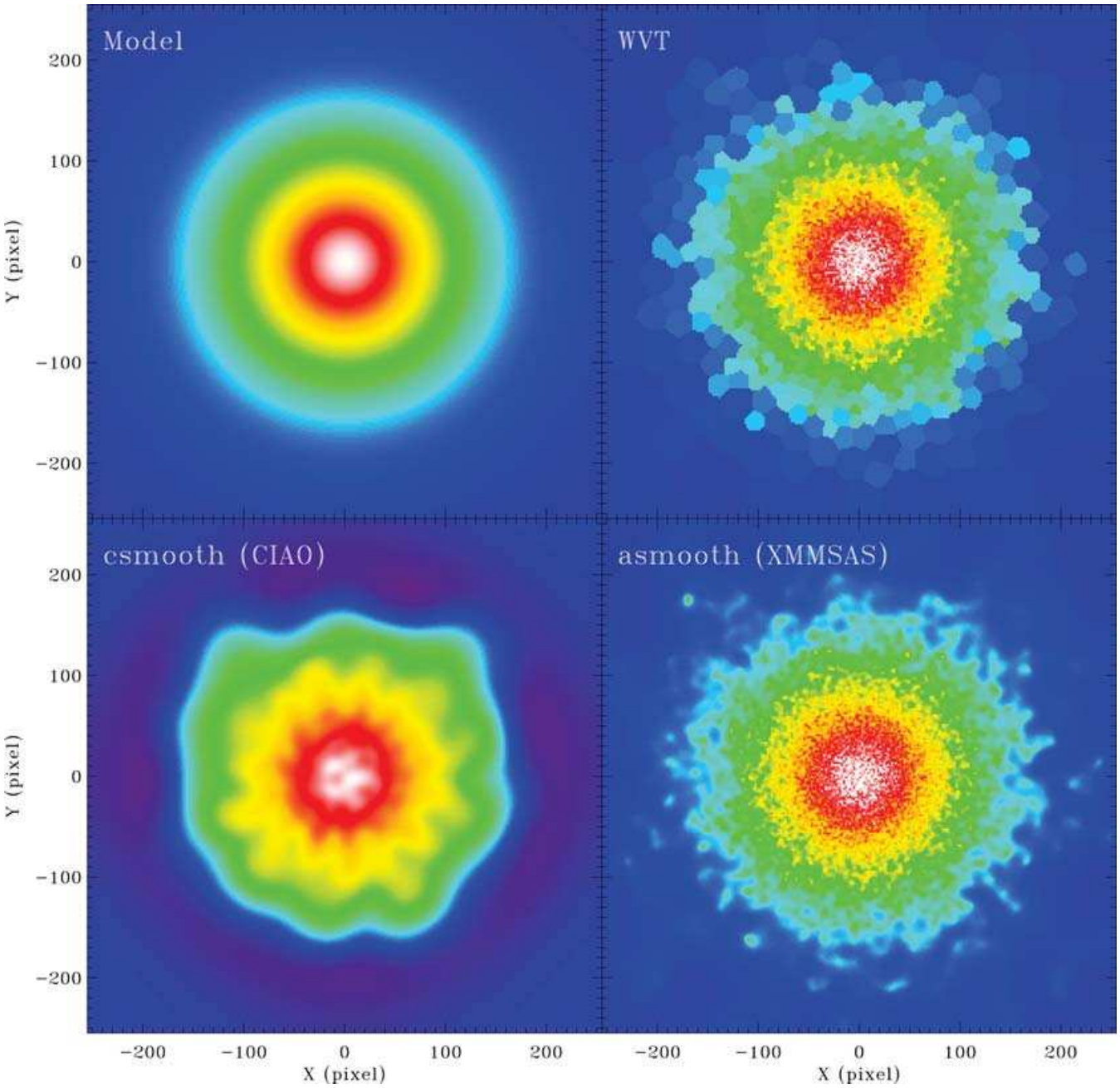

Figure 10. A comparison of WVT binning with adaptive smoothing. Upper left-hand panel: model surface brightness distribution. The other three panels show the simulated counts image, adaptively binned image with WVT binning (upper right-hand panel), adaptively smoothed with the CIAO tool csmooth (lower left-hand panel) and *XMM*-asmooth (lower right-hand panel). In the csmoothed image, note the radial 'fingers', the annulus of deficient emission (deep purple) and the boundary effects in the corners.

choose default Gaussian kernel of *XMM*-asmooth, as it is the most commonly used. Fig. 10 shows the results of applying WVT binning (upper right-hand panel) and *XMM*-asmooth (lower right-hand panel) to the same test model. Both algorithms are able to reproduce the underlying surface brightness distribution. Fig. 11 compares the distributions of relative errors for *XMM*-asmooth (long dashed line) and WVT binning (solid line). Both distributions are consistent with the constant targeted S/N value of 5, but the error distribution of *XMM*-asmooth is not as regular as that of WVT and is skewed slightly toward higher fluxes.

The skewed error distribution is a result of the tendency of *XMM*-asmooth to preferentially misidentify high flux pixels over low flux pixels as real features. We find that the *XMM*-asmooth algorithm tends to build 'bridges', connecting nearby, independent noise peaks and making them appear as linear filamentary structures. Examples can be seen in the outer parts of Fig. 10. A cleaner illustration is shown in Fig. 12. Here, we have simulated a flat-field image, with a vertical gradient in S/N. The figure compares the results of *XMM*-asmooth (left-hand panel) and WVT binning (right-hand panel). The smoothed image shows a wealth of spurious linear features that

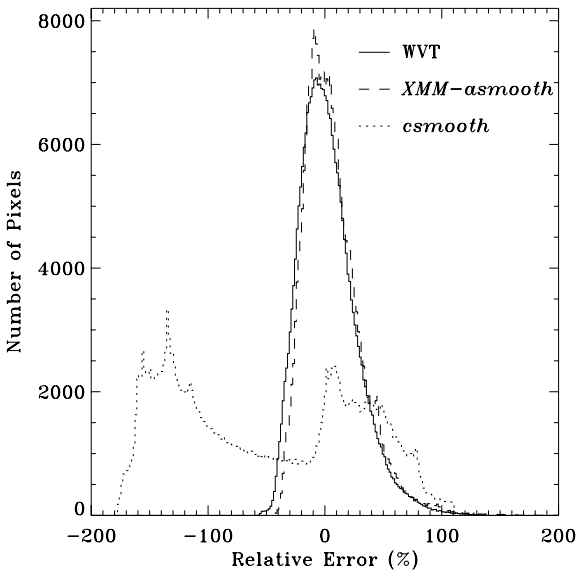

Figure 11. Histogram of relative errors, compared to the model surface brightness, for the example of Fig. 10. The WVT binning (solid line) and *XMM*-asmooth (dashed line) histograms are consistent with the target S/N value of 5 (i.e. they approximate a Gaussian with a width of 20 per cent). Note that the adaptively smoothed image is *not* a statistically better representation of the true surface brightness. The histogram of csmooth results (dotted line) is very irregular with a wide range of positive and especially negative errors, demonstrating the failure of this algorithm.

strongly lead the eye, suggesting filaments and cavities. The binned image, on the other hand, looks to the eye like a featureless but noisy flat-field. One can easily see that all of the apparent structure in the smoothed image happens at scales slightly smaller than the WVT bin sizes, and is therefore not statistically significant despite the identical target S/N of 5 given as input.

In conclusion, if an adaptively smoothed image is necessary, we urge that it be published only in conjunction with its smoothing scale map or an equivalent WVT-binned image.

6.3 A cautionary note on CIAO tool csmooth

More than half of all *Chandra* press release images of the diffuse emission from galaxies and clusters of galaxies are generated with the CIAO tool csmooth. We demonstrate here that this algorithm creates very serious artefacts for images of diffuse emission and should be used only with extreme caution. Similar concerns were also previously raised by SF01.

The csmooth algorithm (Ebeling et al. 2006) first calculates a set of smoothing kernel sizes, ranging from the size of a single pixel to that of the entire image. Starting with the smallest kernel, all pixels with a sufficient S/N within the kernel to match the target S/N requirements are convolved and added to the output image. These pixels are then removed from the input image, so they make no contribution at larger scales. The algorithm then picks the next larger kernel and starts over with the remaining pixels. This continues until no more pixels are left in the image, or the kernel size reaches its maximum. The final csmoothed image is the sum of all these individually convolved slices, and the flux from each pixel is spread over a different area, according to its smoothing scale.

Because a different smoothing kernel is assigned to each pixel in the input image, each pixel in the output image is the sum of many convolutions of different parts of the input image with different kernels. This is fundamentally distinct from the *XMM*-asmooth algorithm, where the kernel is assigned to the pixel in the output image, whose flux is then a result of a single convolution using a single kernel. In other words, *XMM*-asmooth *collects* flux, while csmooth *distributes* flux. These procedures are identical only for pure convolution with a fixed kernel. When the kernel is variable, the csmooth algorithm has the effect of moving flux from low surface brightness regions into high surface brightness regions. This

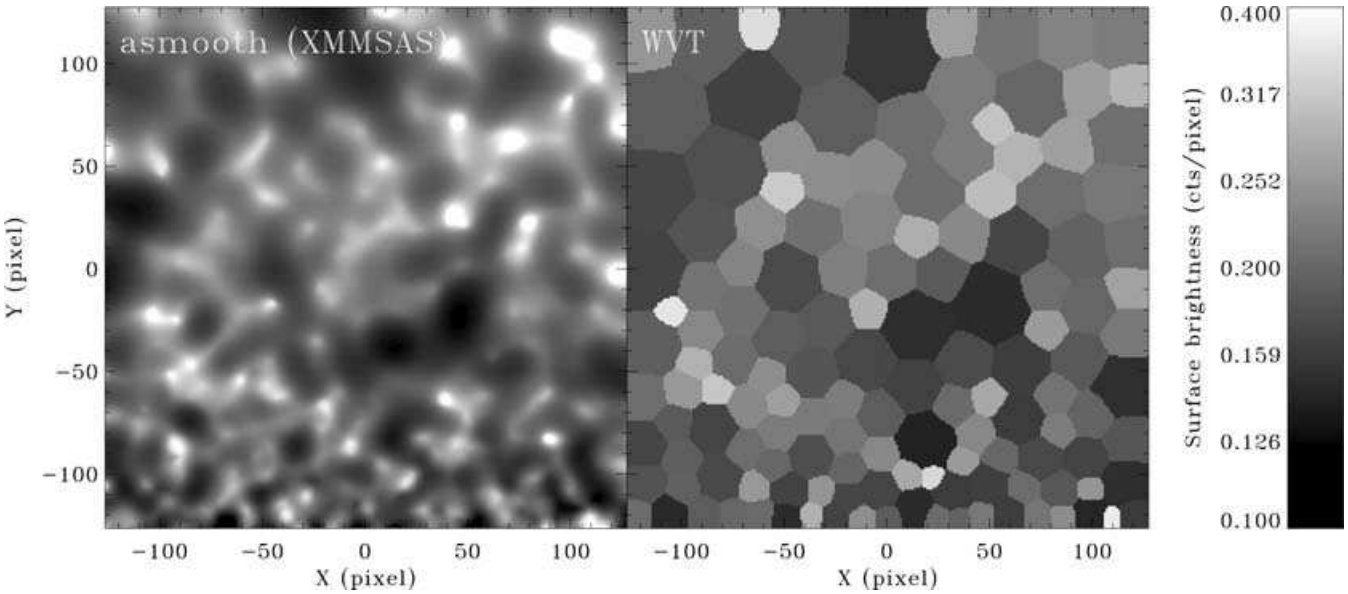

Figure 12. Comparison between *XMM*-asmooth (left-hand panel) and WVT binning (right-hand panel): the simulated counts data was derived from a flat Poisson distribution with a count rate of 1 cts pix^{-1}, with a spatially variable background 'ramp' increasing linearly from 0 cts pix^{-1} at the bottom to 5 cts pix^{-1} at the top of the image.

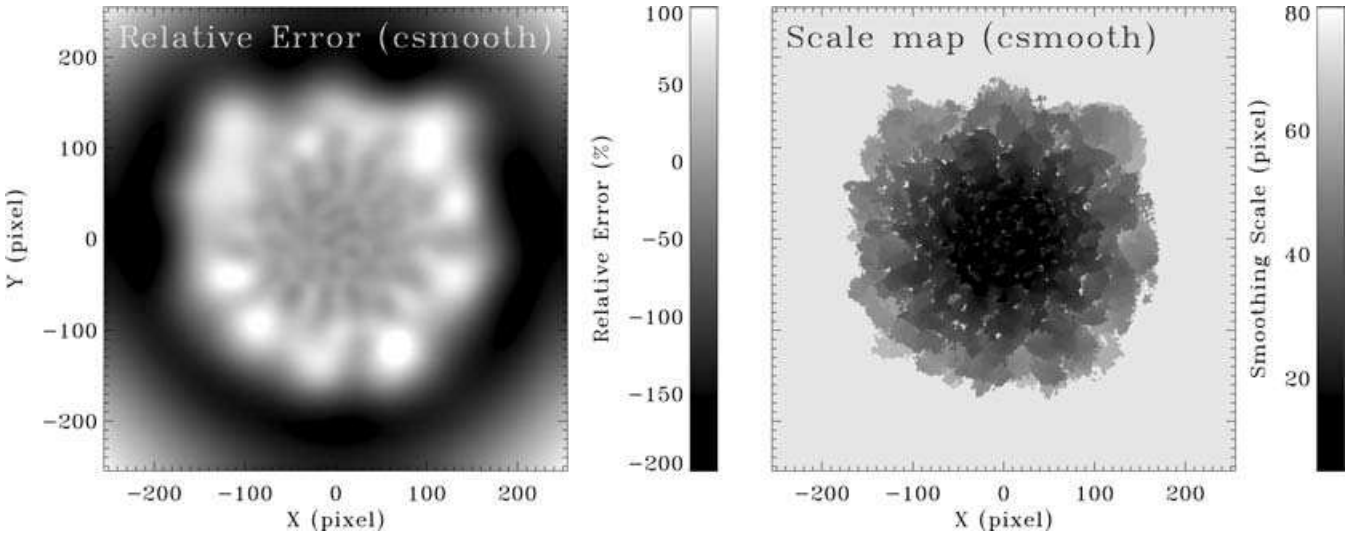

Figure 13. Left-hand panel: relative error distribution for the *csmoothed* image in the lower right-hand panel of Fig. 10; note the large-scale spatially dependent error distribution which leads to the identification of spurious features. Right-hand panel: the csmooth scale map, indicating the spatial distribution of the smoothing kernel size. Note how smaller kernel sizes translate into stronger positive deviations in the error distribution; in particular, note the correspondence between the transition to the largest smoothing kernel and the sharp edge in the error distribution that leads to the annulus of deficient emission.

has previously been noted by SF01 in the adaptively smoothed image of the Perseus cluster, where they find the flux per pixel in the X-ray cavities dropping to half the value derived from raw data. This behaviour is also particularly destructive in regions of relatively flat emission, where csmooth will move flux into the high-flux tail of the noise distribution, creating spurious emission features in the smoothed image. A good example is given in the lower left-hand panel of Fig. 10, where csmooth obviously produces spurious radial features.[10]

An additional problem affecting csmooth is its definition of S/N. In this paper, we define the S/N as the quotient of the background-corrected signal divided by the total noise in each bin. Unfortunately, up to CIAO 3.1, the option of csmooth to supply an external background map to calculate the S/N distribution in the equivalent way does not function correctly. Instead, the only available option is to compute the background from a local annulus surrounding the smoothing kernel. This local significance approach is valid for point sources on top of diffuse emission, but not for the diffuse emission itself. For extended sources, the local 'background' region includes significant amounts of the diffuse emission, resulting in strong S/N variations across the field and an overestimate of required smoothing scales. Fig. 13 shows the spatial error distribution (left-hand panel), as well as the scale map, indicating the distribution of smoothing kernel sizes. Due to computing the S/N locally, csmooth will *always* employ the largest available smoothing kernel in the outer parts of an image, no matter what target S/N is chosen, or how high the flux at the boundaries is. In our example, the kernel reaches its maximum size at a radius of about 150 pixel, and starts to disperse the flux over a large area. This results in an annulus of depressed emission at larger radii (deep purple colours in Fig. 10 and dark regions in Fig. 13). The missing flux from this annulus accumulates in adjacent regions with smaller smoothing scales, producing a relatively sharp surface brightness edge at the transition point. The relative errors in these two regions range from +100 per cent to −200 per cent (see Figs 11 and 13), indicating the magnitude of this effect.

[10] This image looks very similar to the claimed radial 'finger' structures, seen in the csmoothed image of NGC 4649 (Randall, Sarazin & Irwin 2004). The noticable strength of smoothing artefacts actually increases with an increase in target S/N in csmooth. This may be caused by the asymmetric Poissonian error distribution, which will be taken into account in a future csmooth release (H. Ebeling private communication).

7 AVAILABILITY OF THE CODE

WVT binning is implemented in IDL 5.6,[11] and publicly available under http://www.phy.ohiou.edu/∼diehl/WVT. It has also been submitted to the *Chandra* contributed software website.[12] An extensive manual and download instructions are provided on the website, along with multiple examples for its usage.

The published version is fully functional and applicable to two-dimensional images as well as one-dimensional pixel lists to generate intensity maps and colour maps. We also provide csh and SHERPA s-lang scripts to generate temperature maps from adaptively binned images.

The code is not limited to be used with X-ray data or integral field spectroscopy. To adapt the code to handle completely different types of problems, it is only necessary to adjust two external functions which calculate the S/N per bin, without any modifications to the main program. Our manual includes detailed instructions on how to achieve this.

8 CONCLUSIONS

We have presented a generalization of the Voronoi adaptive binning technique by Cappellari & Copin (2003), broadly applicable to X-ray and other data. The generalized algorithm exploits the properties of WVT, rather than the overly restrictive Gersho conjecture. WVT binning is applicable to any type of data as long as there is a way to robustly calculate the S/N, and the S/N distribution changes smoothly over the size of a bin. We have demonstrated the capabilities of WVT binning on exposure- and background-corrected X-ray

[11] http://www.rsinc.com

[12] http://cxc.harvard.edu/cont-soft/soft-exchange.html

intensity images, colour and temperature maps, and in isolating the diffuse gas emission in elliptical galaxies.

WVT binning overcomes the shortcomings of both Voronoi and quadtree binning, the latter of which is in growing use in X-ray astronomy. Sanders et al. (2005) have recently published results using a 'contour binning' algorithm. In this new adaptive binning algorithm, the bin boundaries follow the isophotes of an adaptively smoothed image. We are unable to make a rigorous quantitative comparison with this technique, as the details are still unpublished. However, the maps that are published suggest that this new binning algorithms creates very irregular and elongated bins, which lead the eye and introduce a shell-like appearance in their binned X-ray cluster images of Perseus. In anticipation of their paper, we would like to re-emphasize that our WVT binning produces an unbiased distribution of compact bins, and does not lead the eye.

We have also demonstrated the pitfalls of adaptive smoothing, and regretfully advise against the use of the CIAO tool csmooth for images of extended diffuse emission, as it creates very serious artefacts. If an adaptive smoothing technique has to be used, we recommend the *XMM*-SAS tool *XMM*-asmooth instead. However, we urge that adaptively smoothed images be published only in conjunction with the smoothing scale map or an equivalent WVT-binned image to facilitate the identification of real structures.

ACKNOWLEDGMENTS

We would like to thank Michele Cappellari and Yannick Copin for several extremely helpful discussions and for making their Voronoi binning code publicly available. Our implementation of WVT binning is inspired by and based on their original algorithm. We would also like to thank Harald Ebeling for many useful discussions regarding the use of csmooth. Support for this work was provided by the National Aeronautics and Space Administration (NASA) through *Chandra* Awards G01-2094X and AR3-4011X, issued by the *Chandra X-ray Observatory* Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-39073.

REFERENCES

Bacon R. et al., 2001, MNRAS, 326, 23
Cappellari M., Copin Y., 2003, MNRAS, 342, 345 (CC03)
Diehl S., Statler T. S., 2005, ApJ, 633, L21
Du Q., Faber V., Gunzburger M., 1999, SIAM Rev., 41, 637
Ebeling H., White D. A., Rangarajan F. V. N., 2006, MNRAS, in press (doi:10.1111/j.1365-2966.2006.10135.x) (astro-ph/0601306)
Fabian A. C. et al., 2000, MNRAS, 318, L65
Fabian A. C., Sanders J. S., Allen S. W., Crawford C. S., Iwasawa L., Johnstone R. M., Schmidt R. W., Taylor G. B., 2003, MNRAS, 344, L48
Gersho A., 1979, IEEE Trans. Inform. Theory, 44, 373
Hwang U., Holt S. S., Petre R., 2000, ApJ, 537, L119
Jones C., Forman W., Vikhlinin A., Markevitch M., David L., Warmflash A., Murray S., Nulsen P. E. J., 2002, ApJ, 567, L115
Lloyd S., 1982, IEEE Trans. Inform. Theory, 28, 129
Møller J., 1994, Lectures on Random Voronoi Tessellations. Springer-Verlag, New York
Nulsen P. E. J., David L. P., McNamara B. R., Jones C., Forman W. R., Wise M., 2002, ApJ, 568, 163
O'Sullivan E., Vrtilek J. M., Kempner J. C., 2005, ApJ, 624, L77
Randall S. W., Sarazin C. L., Irwin J. A., 2004, ApJ, 600, 729
Sanders J. S., Fabian A. C., 2001, MNRAS, 325, 178 (SF01)
Sanders J. S., Fabian A. C., Allen S. W., Schmidt R. W., 2004, MNRAS, 349, 952
Sanders J. S., Fabian A. C., Dunn R. J. H., 2005, MNRAS, 360, 133
Sarazin C. L., 1988, Cambridge Astrophysics Series, X-Ray Emission from Clusters of Galaxies. Cambridge Univ. Press, Cambridge

This paper has been typeset from a TEX/LATEX file prepared by the author.