# X-Ray Image Analysis

Francesco Damiani

*Istituto Nazionale di Astrofisica*
*Osservatorio Astronomico di Palermo*

# Outline

**Introduction:**

X-ray images in selected time- or energy ranges;
time- and energy-filtering of background;
common features of detectors: gaps and dithering;
spatially variable PSF/resolution; photon pile-up;
vignetting and exposure maps; particle detection and afterglows

**Historical perspective:**

from Einstein IPC to Chandra HRC/ACIS images;
going to extremely low background-per-resolution-element;
deep images and source confusion.

**Techniques:**

local (cell) and map detect; matched filters and wavelet detection;
analogies with optical techniques.

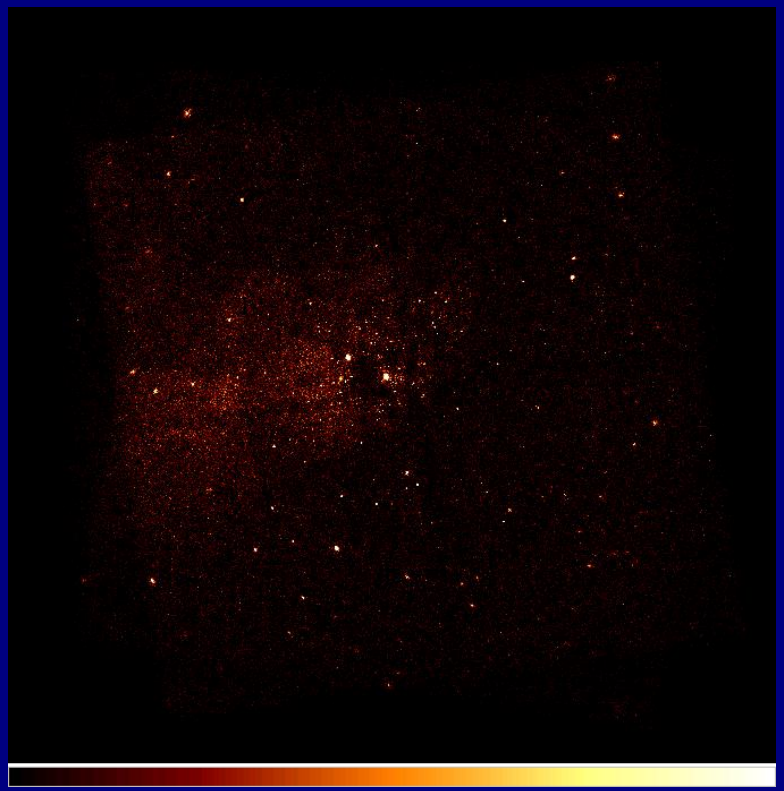
**Details of wavelet technique:**

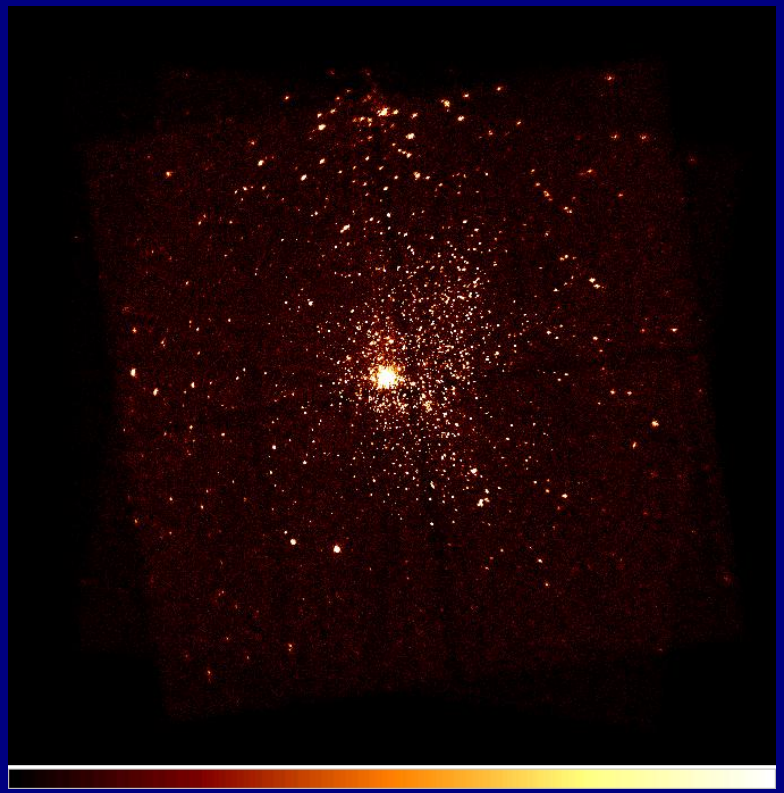
Multi-scale detection; edges and gaps; multi-scale background;
threshold calibrations and confidence levels for source existence;
count-rate and size estimation; errors; PSF corrections;
artifacts for bright sources.

## *X-ray images from binned position lists:*

- Opportunity to select time intervals or energy bands of astrophysical interest
- Ability to reject background “optimally” for your scientific aims
- Background filtering on time, energy, event type

M17: 0.5-1.0 keV (same Chandra/ACIS-I dataset!) M17: 1.5-7.0 keV

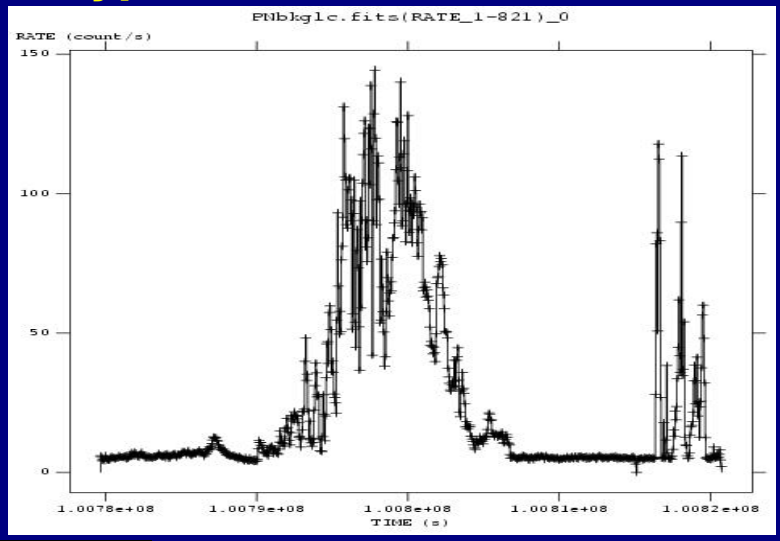


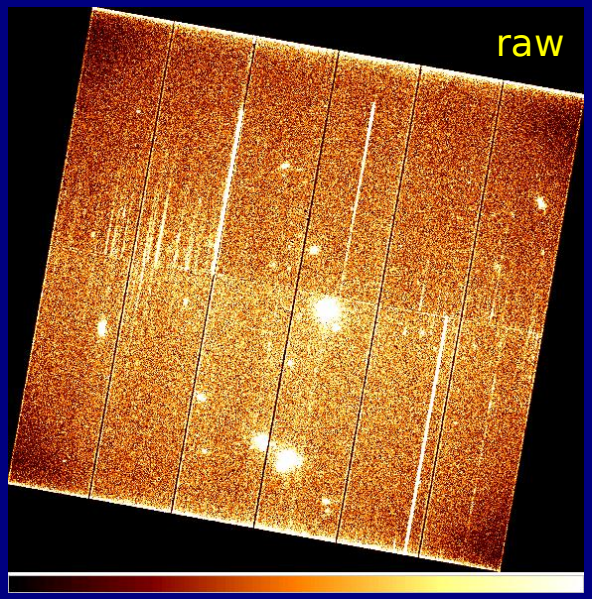


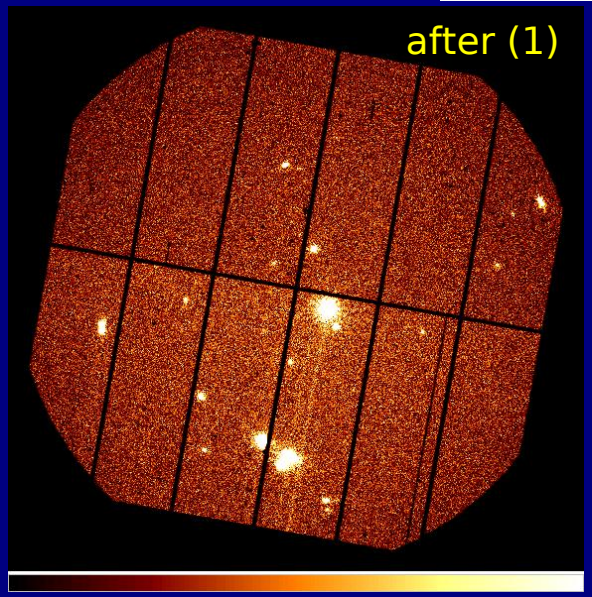
## Background filtering on time, energy, event type:

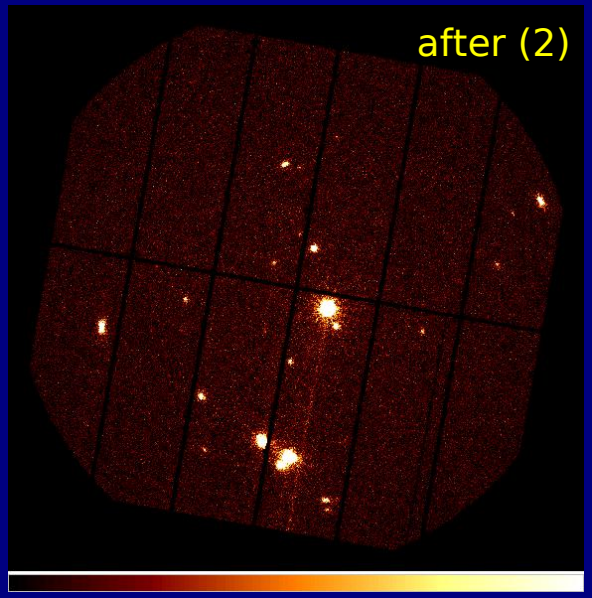
→ removal of:
(1) *detector noise, particles, bad pixel*
*(2) high-background time segments*

e.g. V410 Tau
XMM-Newton
EPIC-pn full-field
background lightcurve...



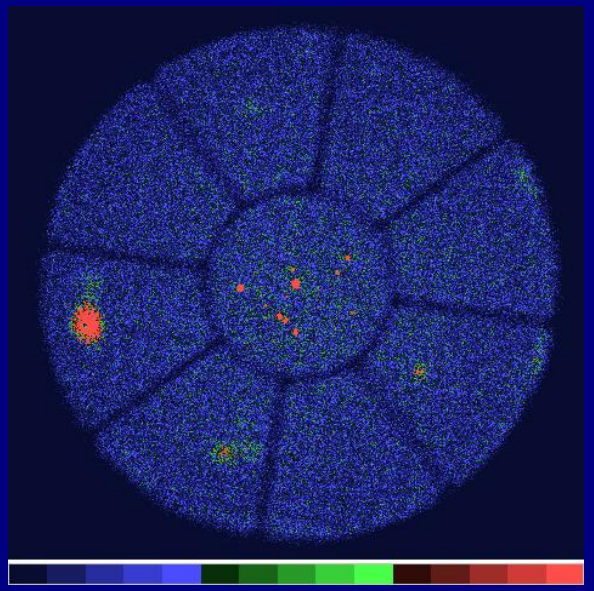




...and images,
with increasing degrees of filtering

## Detector shadows/support ribs

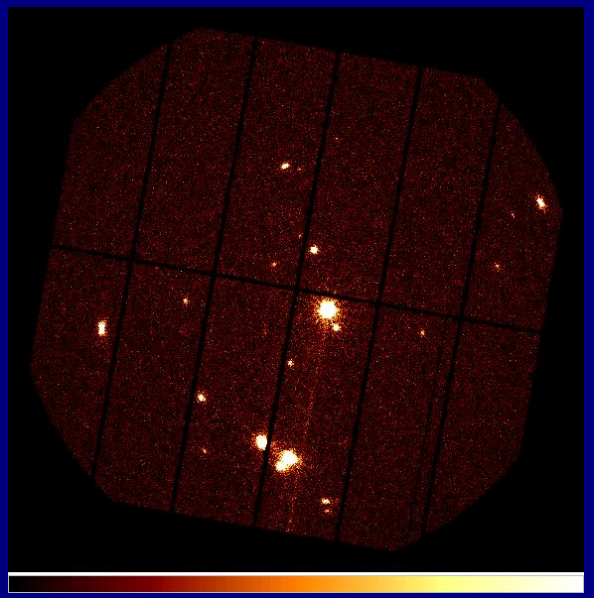
Example: ROSAT PSPC – modulated by "wobbling", to obtain only partially underexposed regions under the ribs (and to prevent bright point sources from "burning" the detector...):

→ *reduced information under ribs*



Example: XMM EPIC/pn – no wobbling, zero exposure between CCD chips:

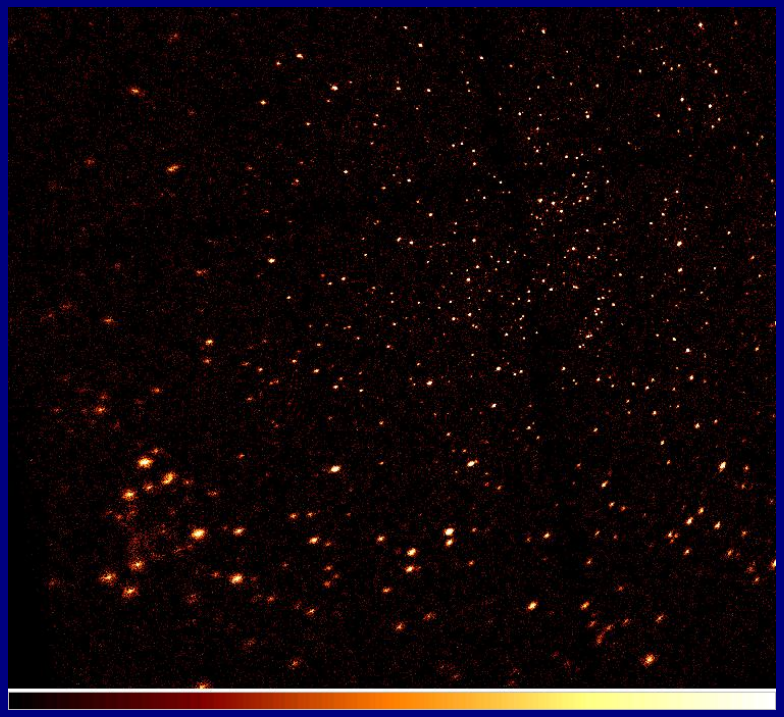
→ *no information in gaps*



***In both cases a problem for reliable source detection!***

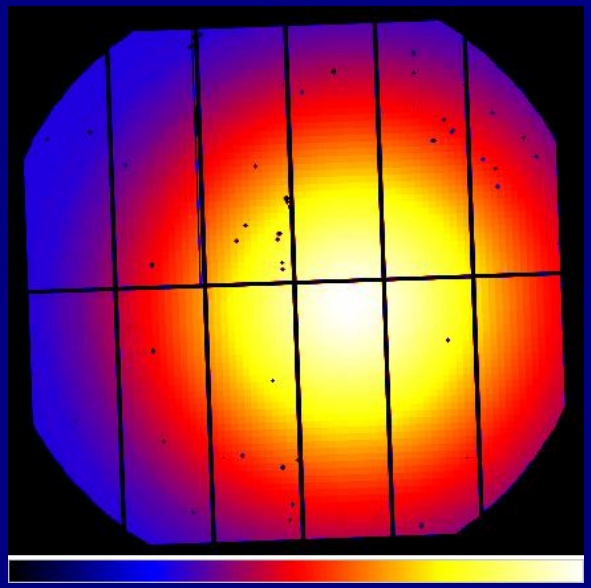
**Mirror+detector PSF:**

- Strong variation with off-axis distance (e.g. a factor of ~10 for Chandra ACIS-I):
- Choice of detection cell/scale size is very important
- *Main factor* to determine off-axis limit sensitivity (reduced source/background contrast)



**Telescope vignetting:**

- Information stored in exposure map (shown for XMM EPIC-pn):
- Slight energy dependence of decrease
- Minor factor for decreased off-axis sensitivity



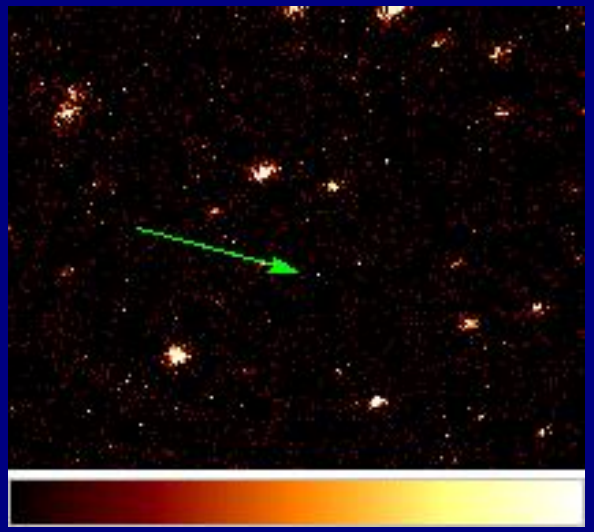
## Other disturbances

- Photon pile-up: *more than one* X-ray photon/pixel/CCD frame

⇒ detected as *one* event, with wrong shape (i.e. pattern or grade), and wrong (cumulated) energy. Kind of “saturated”, unusable pixels, in the core of very bright point sources (e.g. $\Theta^1$ Ori):

- Out-of-time events (in CCD images of bright sources): true X-ray events, recorded *while the CCD is being read*, and assigned to a wrong position along the same CCD column.

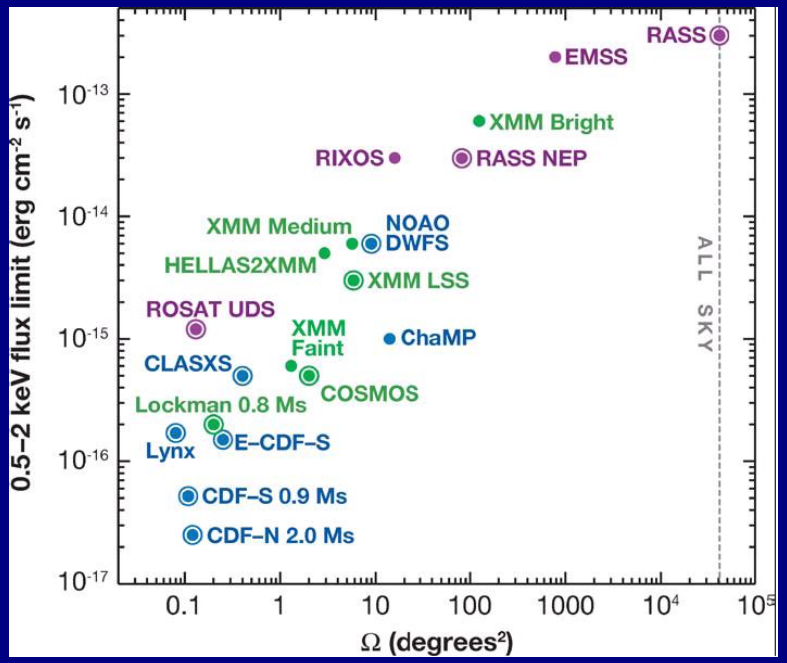
- Particle detection, and “afterglow” events: CCD pixels are hit by energetic particles, and generated electrons are read over many (2-7) readout frames, mimicking faint sources (often smaller than the PSF), with impulsive variability:



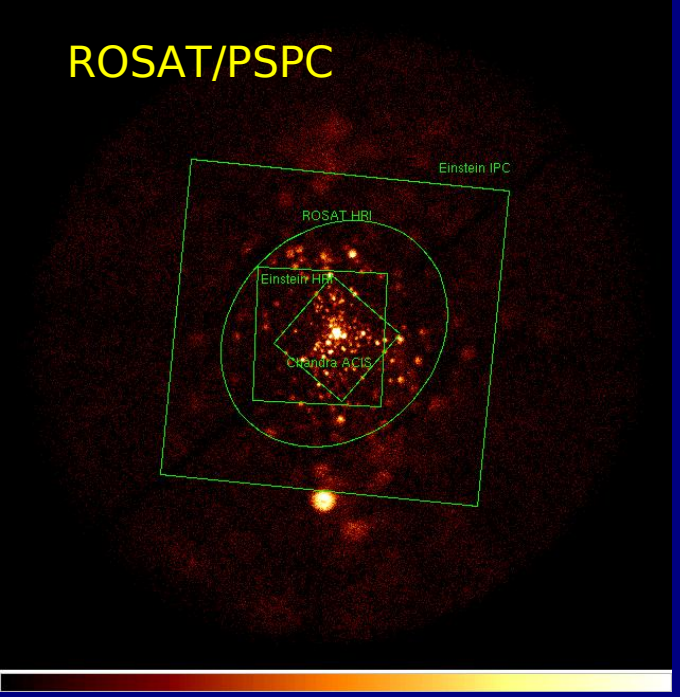
## Historical perspective: 30 years of cluster imaging

- **Einstein Observatory IPC** (Imaging Proportional Counter) and **HRI** (High Resolution Imager) – 1978-1981
- **ROSAT PSPC** (Position Sensitive Proportional Counter) and **HRI** (High Resolution Imager) – 1990-1999
- **XMM-Newton EPIC** (European Photon Imaging Camera), with **MOS1**, **MOS2** and **pn** detectors – 2000-today
- **Chandra HRC** (High Resolution Camera) and **ACIS** (Advanced CCD Imaging Spectrometer) – 2000-today
- Other satellites (e.g. **ASCA** – 1993-2000) observed individual bright, young stars (but *PSF too wide* for proper cluster imaging)

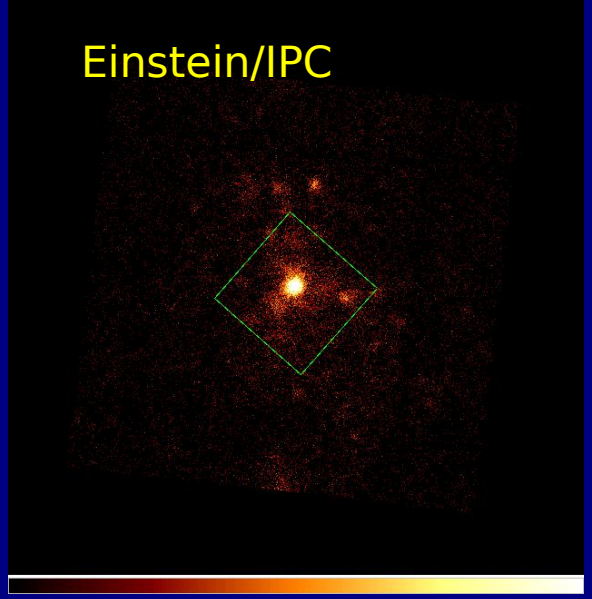
A noticeable sensitivity improvement with time: four orders of magnitude!



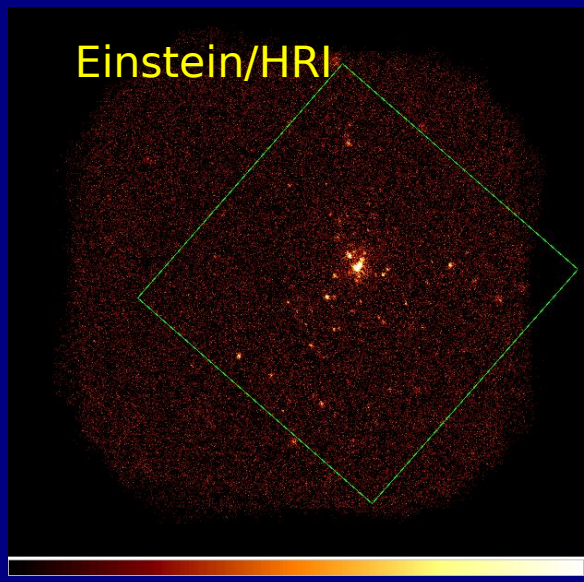
ROSAT/PSPC



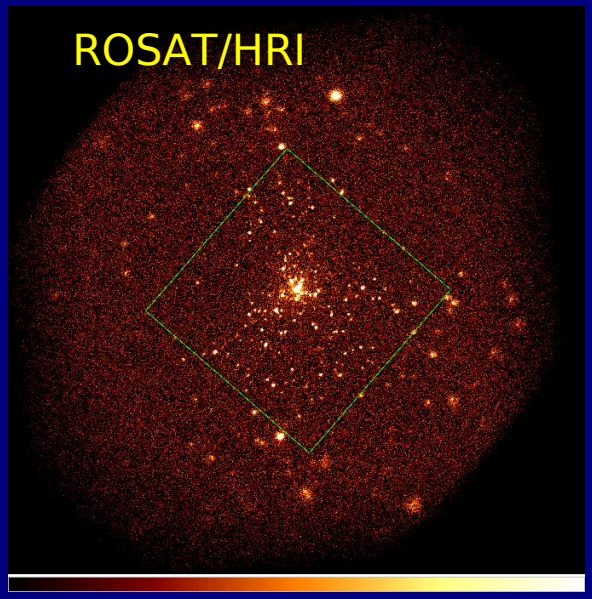

Einstein/IPC



Einstein/HRI

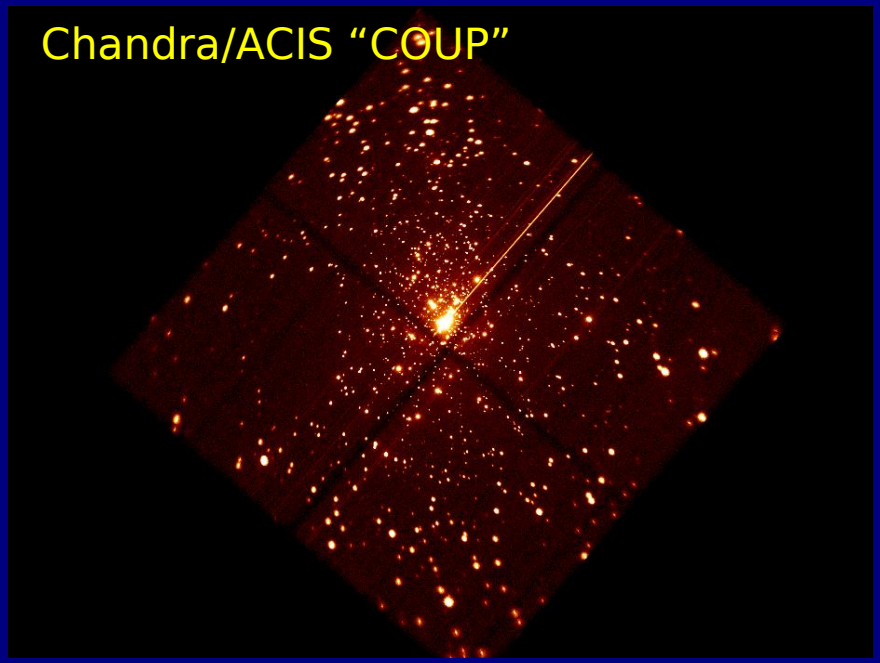


ROSAT/HRI



A few views of Orion Trapezium, with increasing detail

Chandra/ACIS “COUP”



## Chandra ACIS and HRC images:

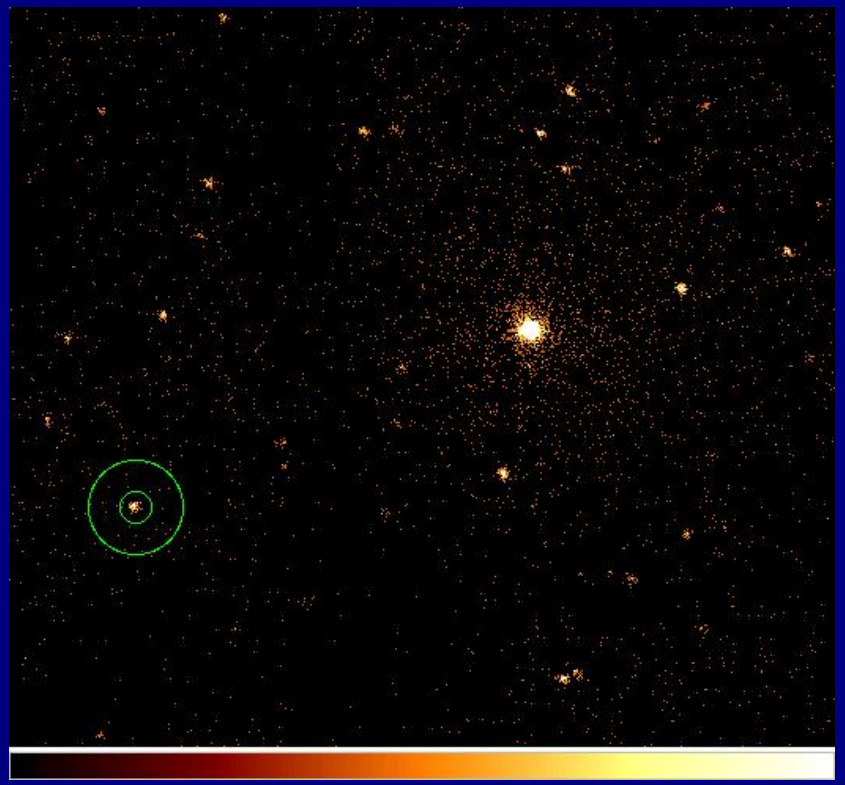
- *Best spatial resolution ever in X-rays:* on-axis PSF FWHM = 0.5”
- *Very low background* (detector + sky): 1-2 · $10^{-6}$ counts/sec/sq.arcsec

  ⇒ in a resolution element of 1”x1” and 100 ksec exposure, one has a typical background of 0.1-0.2 counts!

  ⇒ very few source photons needed to have a *statistically significant* detection

- However: number of resolution elements is *huge* ⇒ threshold cannot be set too low
- And: not straightforward to actually measure such a low background locally

Example: 2'x2', center of ACIS FOV,
100-ksec exposure,
inner circle radius = 2.5” (full PSF inside)
outer circle radius = 7.5” ⇒ poor background measurement ⇒ ill-defined threshold



## Chandra deep imaging:

- Thanks to very high spatial resolution and low background, exposures lasting ~ 1-2 Msec can be usefully made, e.g. COUP in Orion, Galactic Center, Chandra Deep Fields North and South, Antennae (extragalactic)...

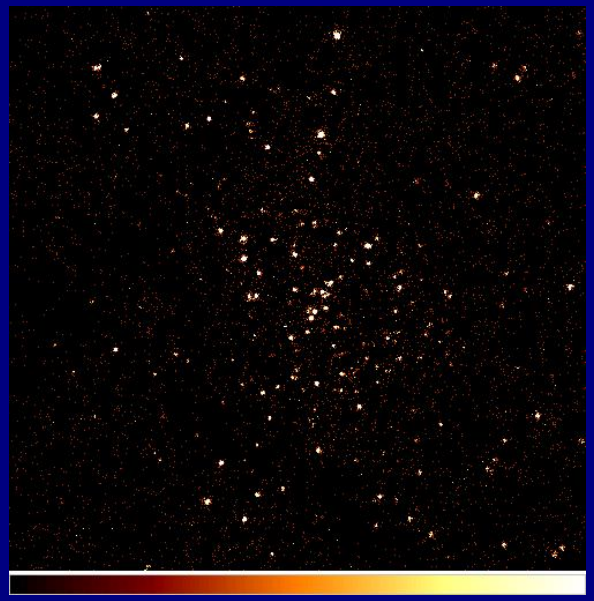
- Only with the longest exposure times, a background-limited regime is reached at field center

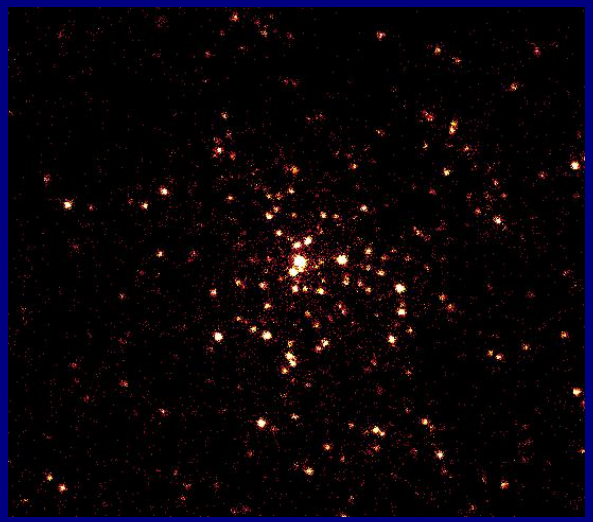
- Most important practical limit is source confusion. This is reached sooner (i.e. already with shorter exposures) in dense clusters than in “average” fields such as those of CDFS, CDFN.
  ⇒ young, dense clusters are a *real challenge* for source detection methods!

Examples:

(left): core of Mon R2 (800 pc), 4'x4':

(right): core of Tr 14 (~2 kpc), 2'x2':





**Techniques for X-ray source detection:**

- **Sliding-cell** (with local or mapped background)
- **Maximum-likelihood**
- **Wavelets**

Advantage of Local sliding-cell: simpler (to implement and understand!), but too much w.r.t. the image complexity...

Wavelets enable very efficient detection of sources, since useful information in current X-ray images is inherently multi-scale (non-uniform PSF, diffuse sources).

Different flavors of wavelet detection methods are available for the same instrument datasets, e.g. *wavdetect* (developed at Chicago/CfA), and *pwdetect* (developed at INAF/OAPA). Of this latter, there is also a version working on XMM data, called *pwxdetect*.

*Pw(x)detect* has been used in major X-ray projects, both stellar and extragalactic: COUP, XEST, DROXO, COSMOS...

**Basics of sliding-cell method:**

- A square cell ($n$ x $n$ pixels) is defined as the source photons extraction region (let the number of counts in it be = $S$).
- A concentric, larger square is used to extract the source background (with total counts = $B$).
- The net source counts $N = S - B \times A_B/A_S$ (with $A_S$, $A_B$ being areas of S and B regions, respectively) are computed.
- As a refinement, the PSF knowledge permits to compute the fraction of total source counts falling in S, and also those residual source counts falling in B, and therefore corrective coefficients to $S$ and $B$.
- The signal/noise ratio SNR is computed at the cell position.
- The detection cell is slided across the image in steps of 1/3 its size, a map of SNR is computed, and local maxima (sources) are found.
- The value of Image background used to define thresholds for detection is derived either from the local “annulus” around the source position, or from a pre-determined background spatial map.
- Optionally, the procedure is repeated with different cell sizes $n$.

## What are wavelets?

Many answers possible...

Shortly, wavelets allow a “mixed” representation of data, in terms of “conjugate” variables (e.g. time and frequency, or space and spatial frequency).

Fourier analysis tells us that each such pair of quantities cannot be *simultaneously* determined (an “Heisenberg uncertainty principle” in data analysis), *with arbitrarily high accuracy*.

It is nevertheless common experience that we e.g. hear sounds having *fairly definite* frequencies over *finite time intervals*.
We do not need waiting forever before perceiving a definite frequency!

Analogous quantities in spatial domain are *positions* and *sizes* (or generally shapes) of objects in images: a wavelet transform (unlike Fourier transforms) permits both properties to be studied at the same time.

Since a wavelet transform (WT) is able to “know” the source size/shape, there is no need to either assume a PSF, nor to set a fixed extraction cell size.
Large PSF variations across the FOV are easily handled by WT. Source apparent sizes are automatically derived in the detection process.

**A little mathematics for WTs:** *(I'll give you a lot more if you want...)*

A WT is computed as a *convolution* of the original image *f(x,y)* with a suitable function *g(x,y,a)*, called *generating wavelet*.

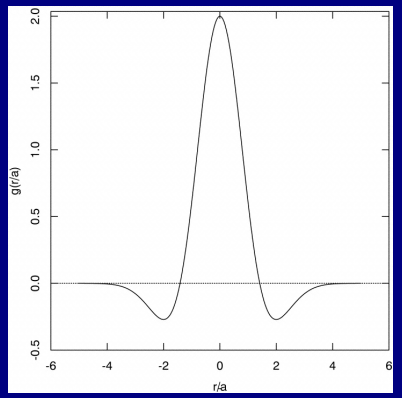
$$w(x,\ y,\ a) = \iint g\left(\frac{x - x'}{a}, \frac{y - y'}{a}\right) f(x',\ y')dx'\,dy'$$

Here *a* is the "scale" parameter (assumed to be the same for both *x* and *y*). Therefore, the WT is a function of both space *and* scale, *w(x,y,a)*.

The *g(x,y,a)* must be bounded, square-integrable in spatial domain, must have zero mean and compact support.

A commonly used form for *g(x,y,a)* is the so-called *Mexican-Hat*:

$$g\left(\frac{x}{a}, \frac{y}{a}\right) \equiv g\left(\frac{r}{a}\right) = \left(2 - \frac{r^2}{a^2}\right) e^{-r^2/2a^2} \quad (r^2 = x^2 + y^2)$$




...which despites tending rapidly to zero for increasing *r*, has no compact support and is not "mathematically" a true wavelet! But is enough so for our purposes...

In addition, the Mexican-Hat generating wavelet has the property of "mimicking" a typical "gaussian-like" PSF: as such the WT is a kind of "matched filter", extracting useful information from an image by a proper weighting. Its analogues in optical data analysis are gaussian-weighted photometric extraction, or optimal extraction for fiber spectroscopy.

Using this generating wavelet, the WT will be:

- Zero for a flat background (as for any *g(x,y,a)*)
- Zero for a uniform-gradient background
- Very sensitive to 2nd derivative of data *f(x,y)* (in fact, is its smoothed laplacian):

$$w(x, y, a) = \iint \left(2 - \frac{r^2}{a^2}\right) e^{-r^2/2a^2} f(x', y') dx' dy'$$

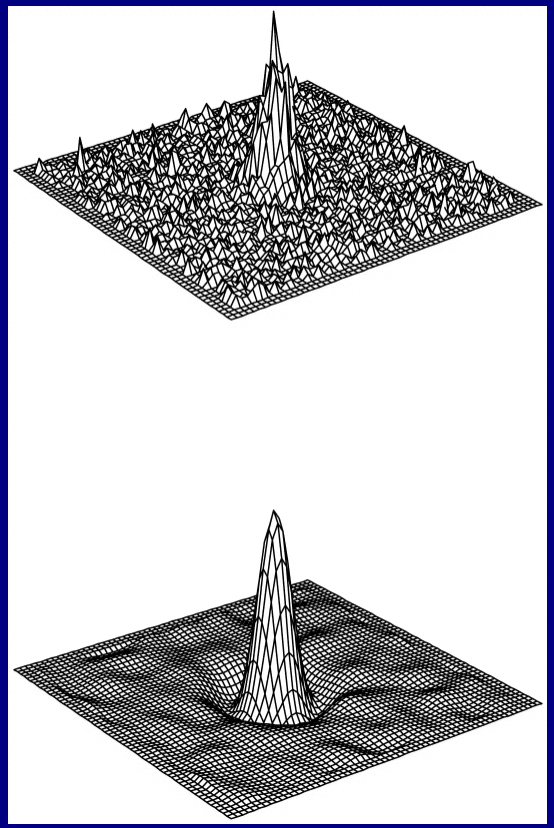
$$= -a^2 \iint e^{-r^2/2a^2} \left(\frac{\partial^2 f}{\partial x'^2} + \frac{\partial^2 f}{\partial y'^2}\right) dx' dy'$$

...and any image edge has an *horribly large* 2nd derivative! → handle with caution

## WT of a gaussian source plus flat background:

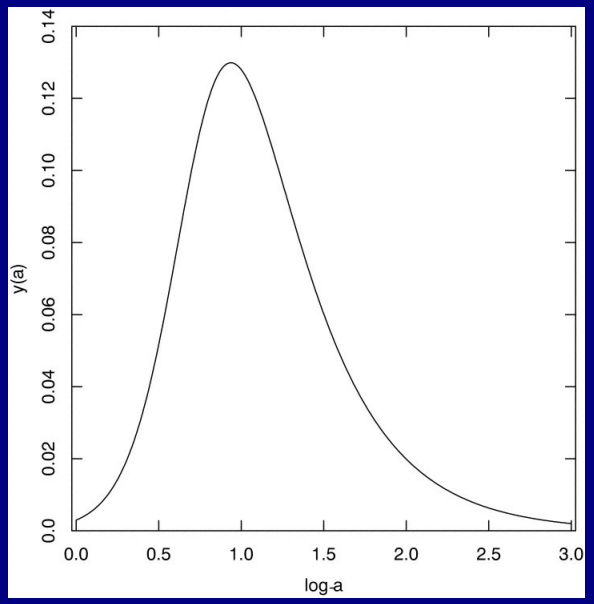
$$w(r,\,a) = \frac{N_{\rm src}}{(1+\sigma_{\rm src}^2/a^2)^2}\left(2-\frac{r^2}{a^2+\sigma_{\rm src}^2}\right)e^{-r^2/2(a^2+\sigma_{\rm src}^2)}$$

i.e., as a function of *space* *(x,y)*:



...and as a function of *scale* *a*, at spatial peak (*r=0*):

$$w_{\rm peak}(a) = \frac{2N_{\rm src}}{(1+\sigma_{\rm src}^2/a^2)^2}\,.$$



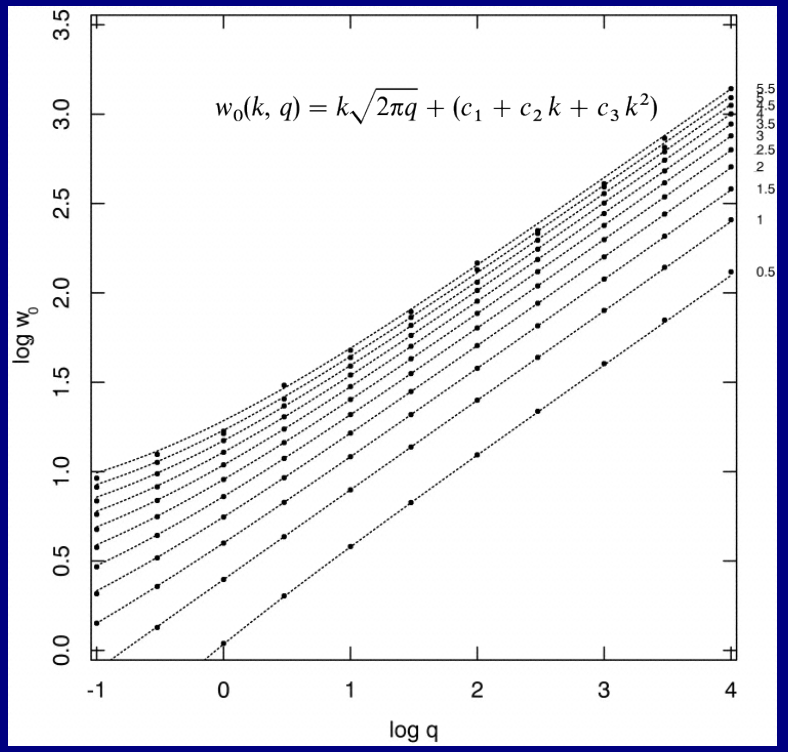

Also useful is the function $y(a)=w_{peak}(a)/a$

peaking at $a_{\rm max} = \sqrt{3}\,\sigma_{\rm src}$

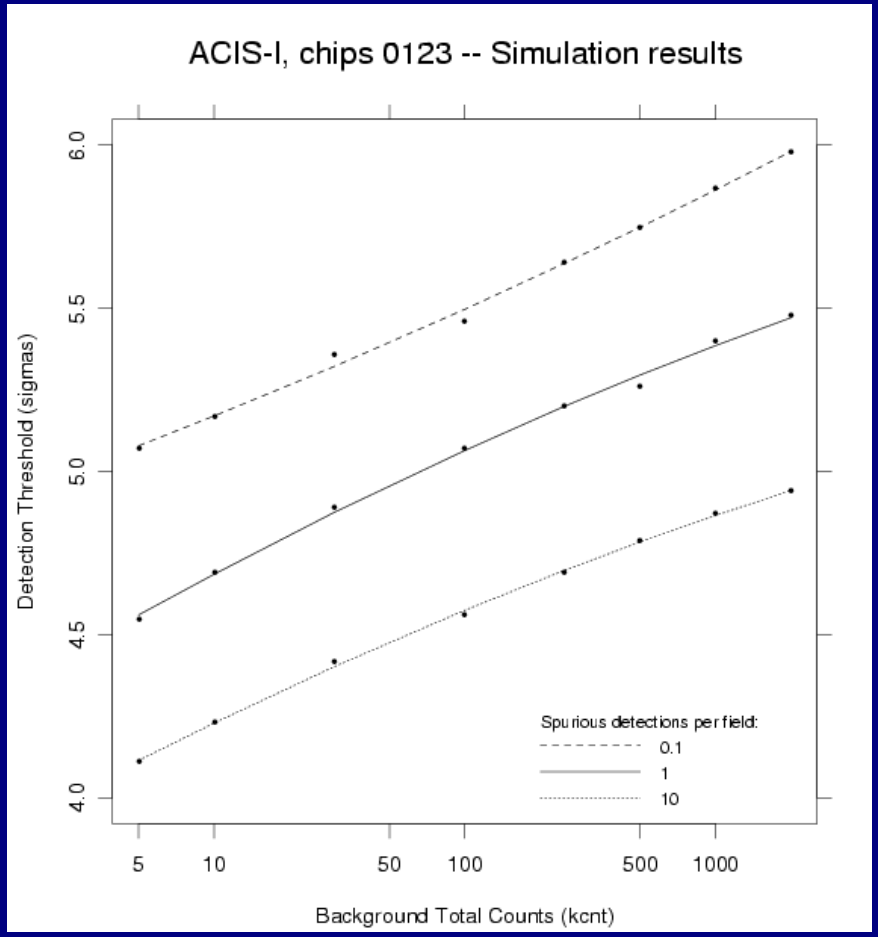
→ A range of scales near $a_{max}$ must be explored to find source properties

## Image background and WT noise

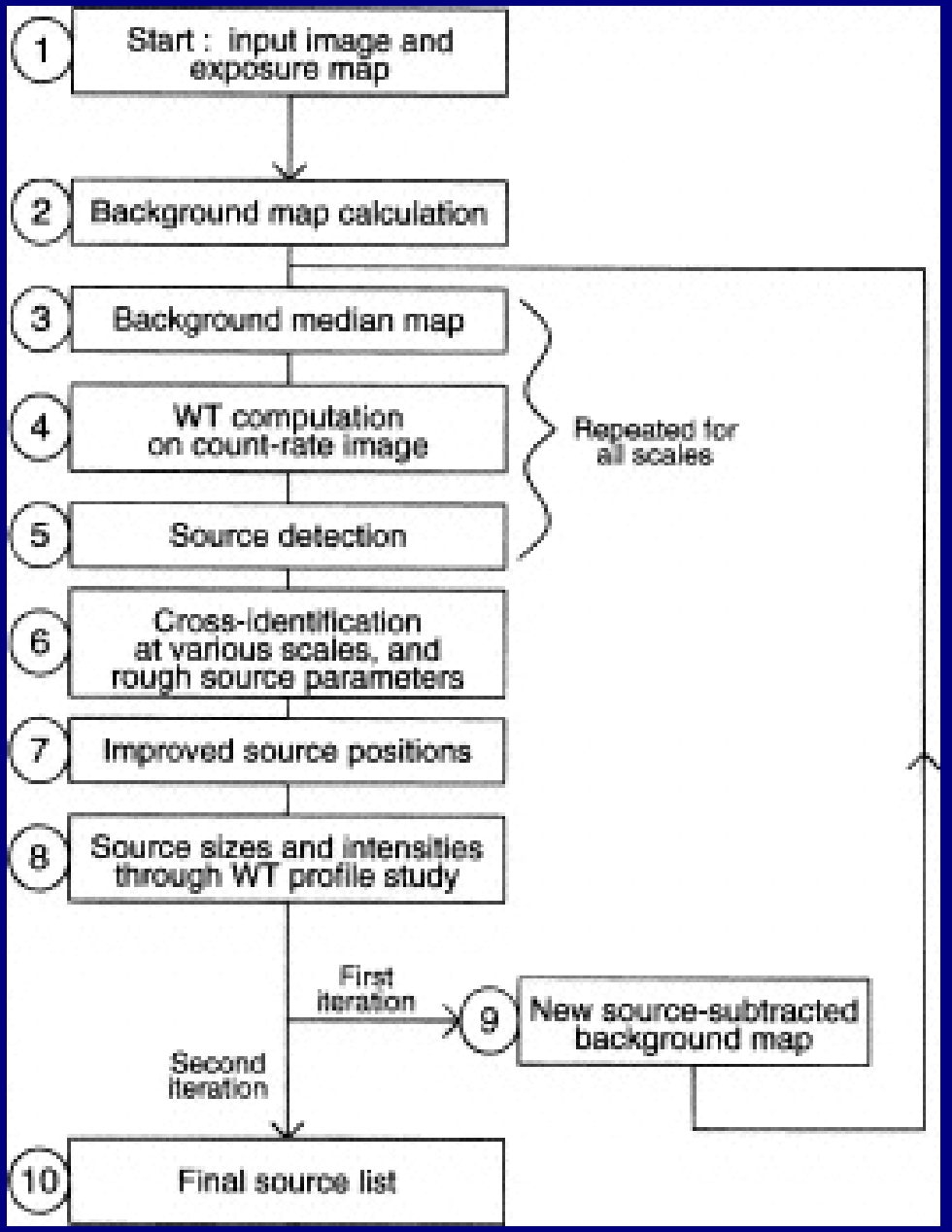
- The WT of a flat background has zero *expected value,* but non-zero *fluctuations*
- Such fluctuations translate into thresholds in *WT space*, to avoid spurious detections with a desired confidence level
- The probability distribution of these WT fluctuations is highly non-Gaussian (and non-analytic) for a low-intensity background – best modeled through simulations
- The relevant parameter is defined as $q = bg \times a^2$, i.e. the number of background counts per squared scale
- This gives a “local” confidence value (n-σ significance) for a detection, but...



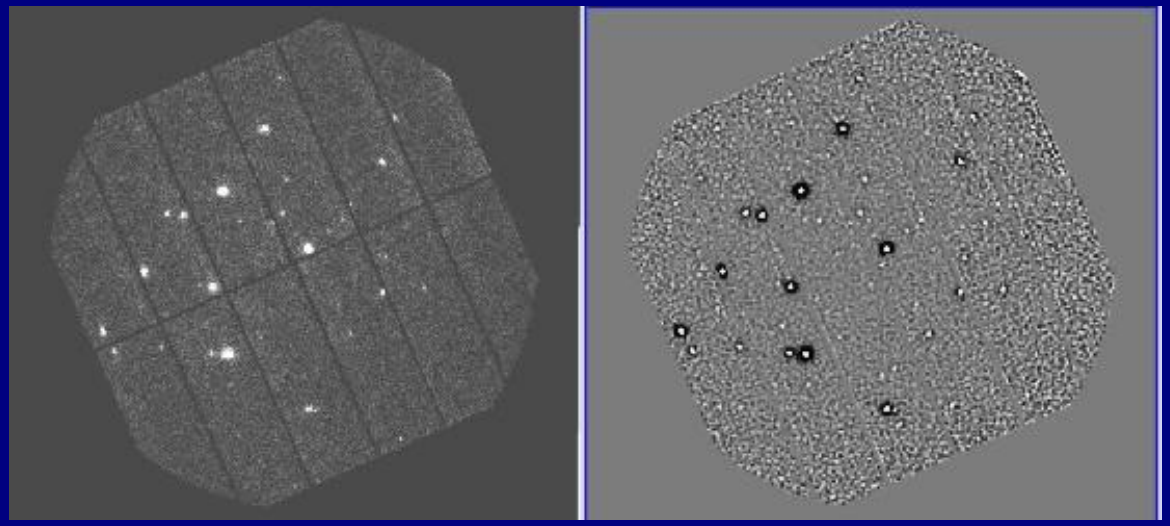
- ...don't forget the often very high number of resolution elements (~ 1 million) in high-resolution images!
- → we better speak of the *number of full-field spurious detections* (calibrated via simulations, for various total background values), e.g.:

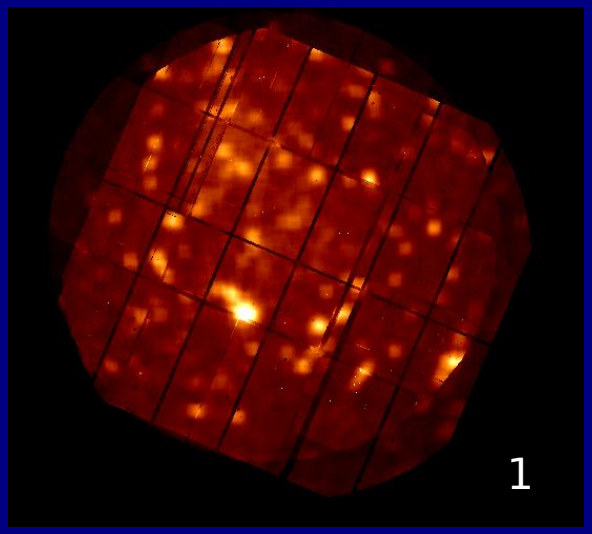


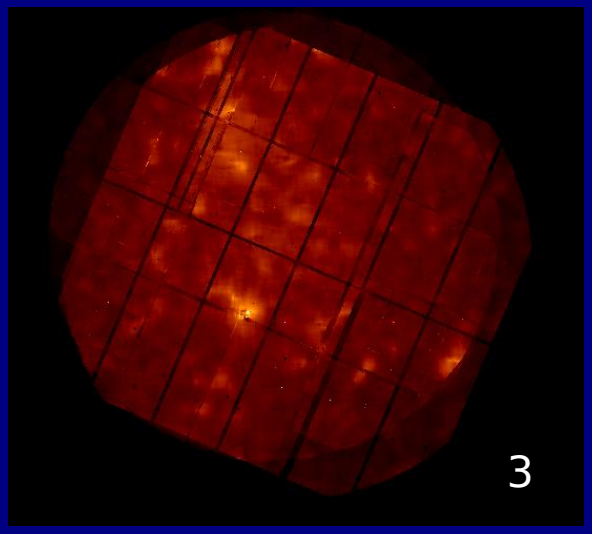
**A block diagram** for *pwdetect* (and *pwxdetect*) algorithm:



An example: **EPIC pn image and its WT** at a particular scale *a*:
N.B. the image is interpolated over gaps to avoid strong background discontinuities!






**Background computation:**

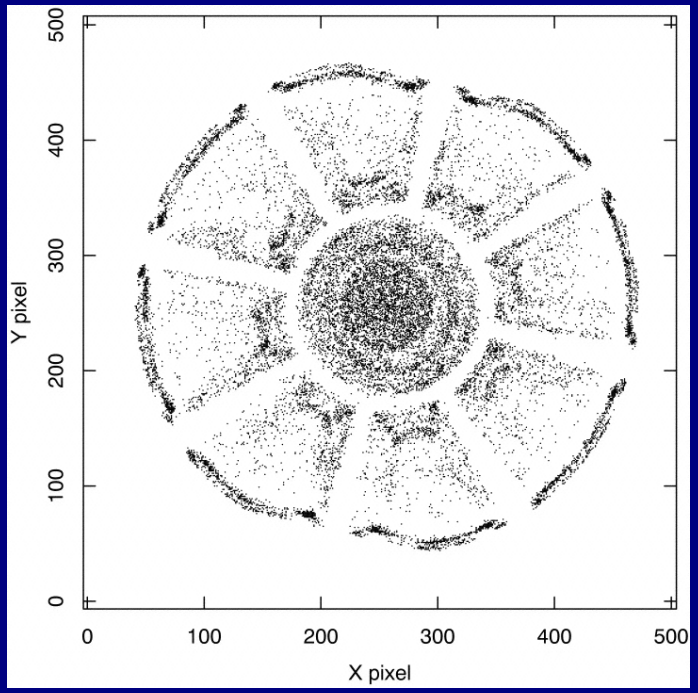
1. adaptive smoothing (with Gaussian $\sigma = \sigma_{PSF}$)
2. Median filter (over cells of size ~ scale *a*)
3. After first pass of detection, interpolate over source positions (using a “swiss-cheese” mask), also usable to study diffuse emission
4. Repeat step 2.
5. For very low background (Chandra), a Poisson distribution is fit to the local image histogram

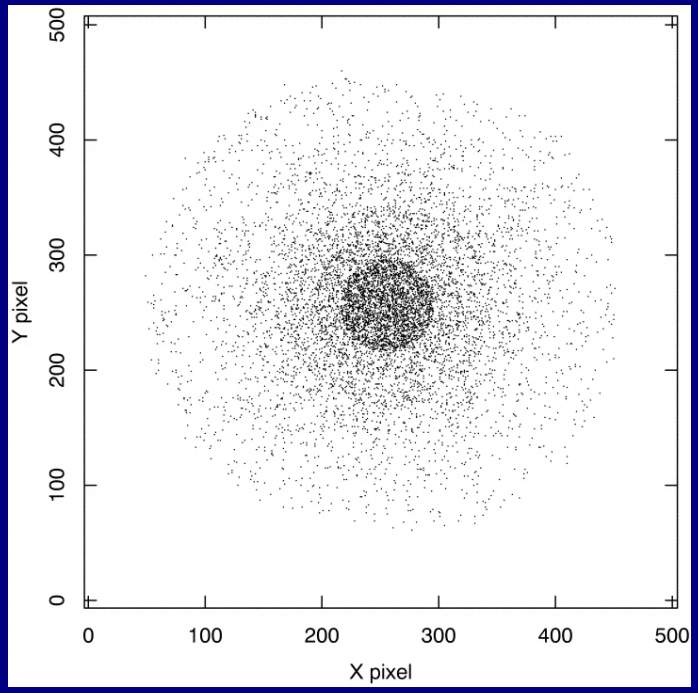
## WT of image gaps and edges

WT detection on simulated background images + *instrument exposure variations*

On *raw* ROSAT/PSPC count image:
lots of *spurious detections* near gaps and edge

On *exposure corrected* image:
spurious detections *uniformly distributed*
(~ density of resolution elements)



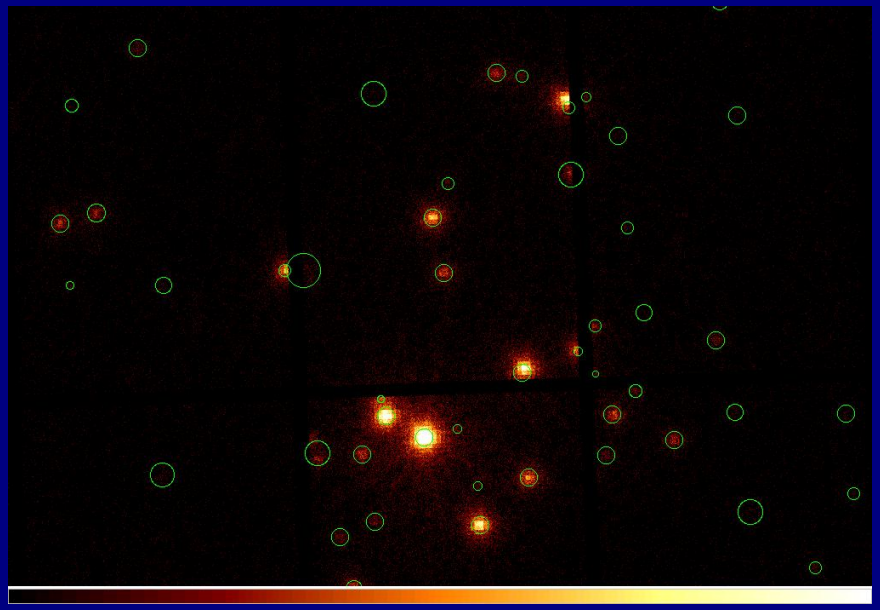




**Use of exposure map is mandatory to deal with these nonuniformities.**

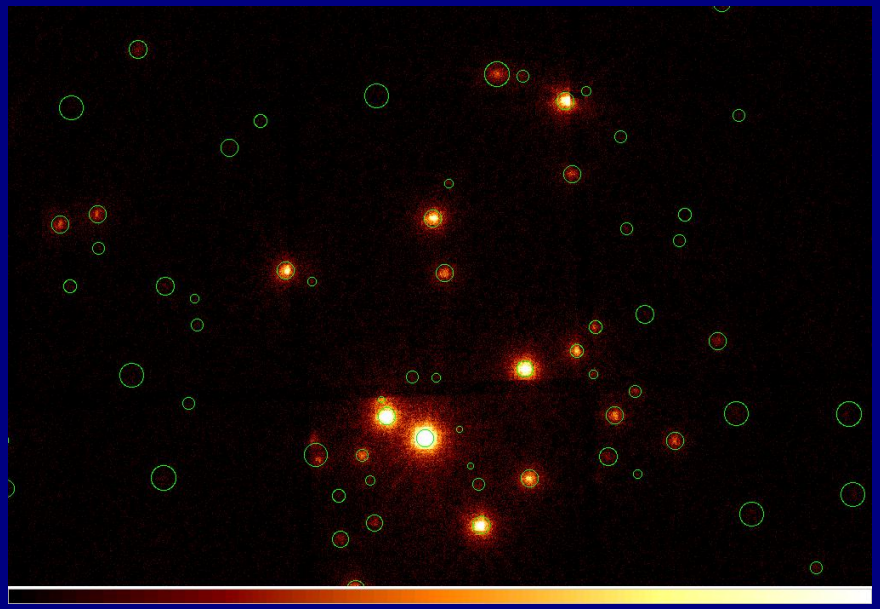
The WT is computed on count-rate images, and its statistics suitably modified.

## Real sources near gaps

Completely *black gaps* (XMM EPIC) are hard to deal with (example: σ Ori cluster):



...but *combining EPIC MOS and pn* datasets before detection helps to find good source positions near gaps!



**Source properties:**

- Detection (maximum) significance and corresponding scale
- Position + error
- Count rate + error
- Apparent size + error
- Background count rate

Detection significance is determined independently of assumptions on source shape. All other parameters are derived from fitting the “theoretical” $w_{peak}(a)$ profile for a Gaussian source to that found in the WT.
Error computation is highly non-trivial.

Source count-rates and sizes are simultaneously determined, under the only hypothesis of a Gaussian source shape.
Corrections for a non-Gaussian instrument PSF are later applied to count rates.

- At very low background levels, common to Chandra images, sources may be reliably detected with very few counts (4-6), at high confidence levels (e.g. 5 σ, or 99.99994%), but have count-rates uncertain even by 50%!

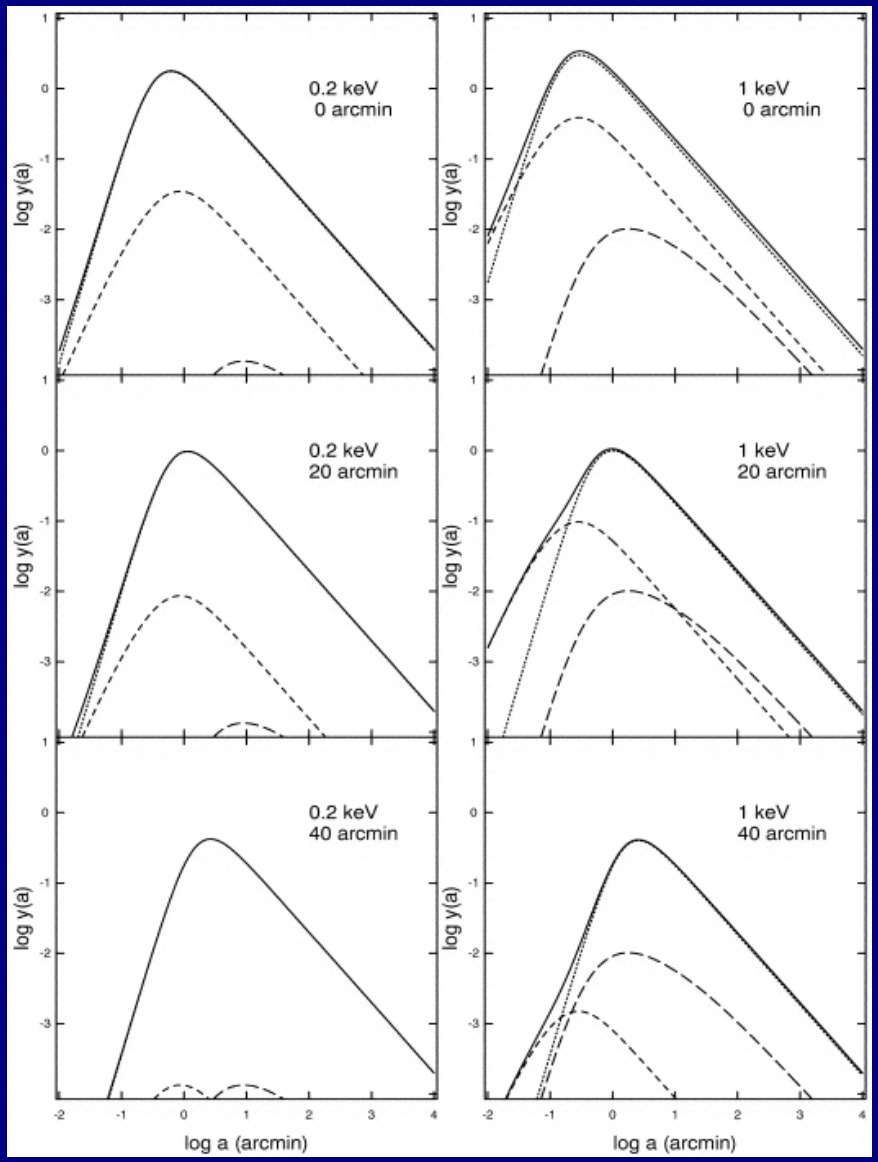
→ **Probability of existence** and **accuracy of count-rates** (fluxes) are *two different problems*

## Example of WT of ROSAT/PSPC PSF

as modeled by Hasinger et al. (1993): near the WT peak, the Gaussian component is dominant...

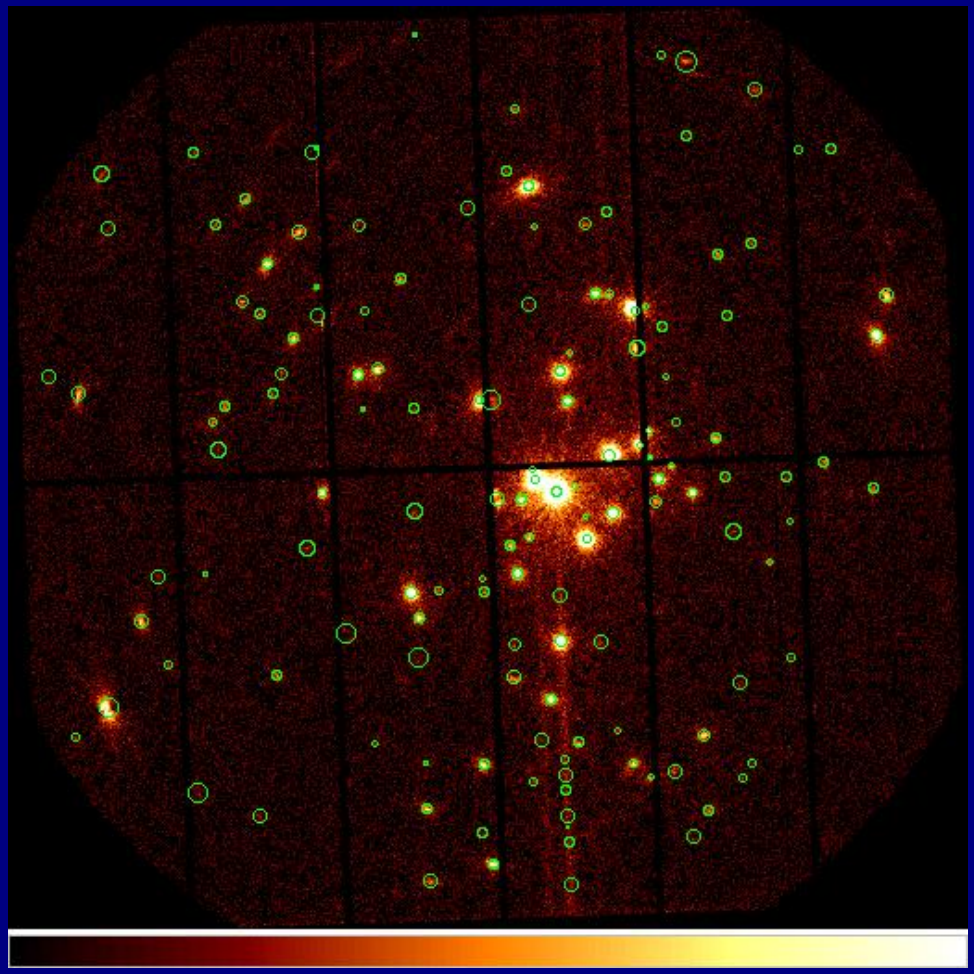
...but this is not so for all instruments, and accurate calibrations must be (and have been) made in every case!

After detection, you may also choose to compute fluxes using more "instrument-PSF-optimized" tools, like e.g. *acis_extract.*



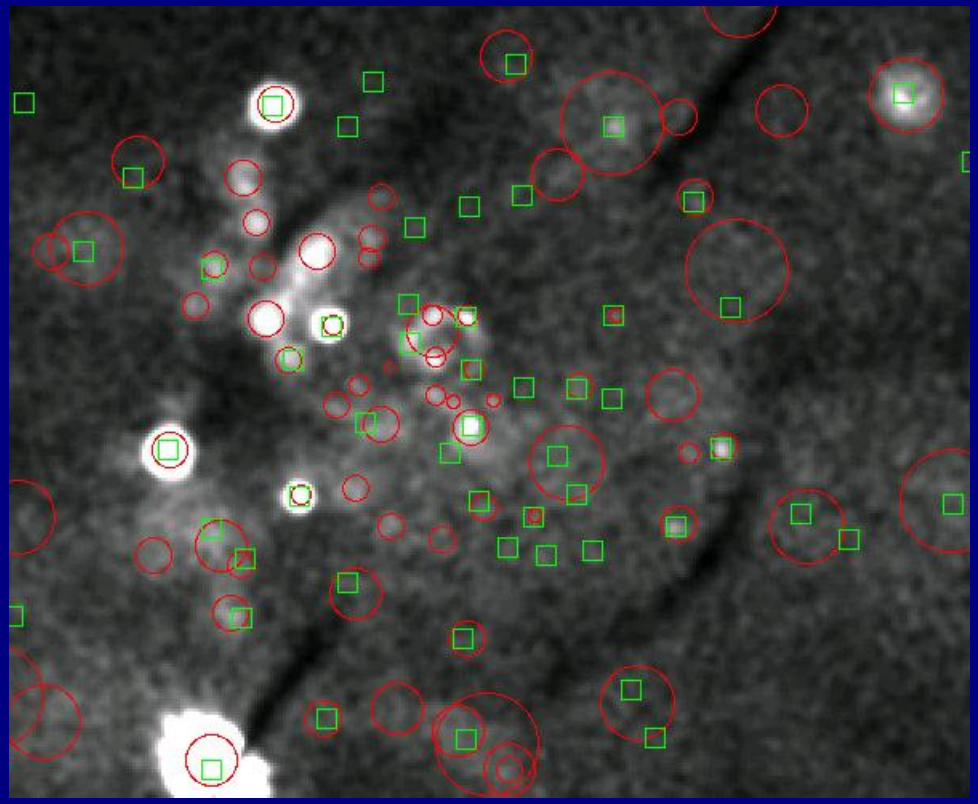
**Residual problems:**

- Spurious detections along out-of-time event trails for bright sources
- Spurious detections in the PSF wings around very bright point sources
- Misbehaviour for highly distorted non-radially symmetrical PSF, near outer image border



## Some comparison

Wavelets (red circles) vs. another method (green squares), ROSAT/PSPC data on SN1987a in LMC:



Wavelets perform better in a crowded field, and near obscuring ribs!

For more information see:


Damiani,F. et al. 1997, ApJ, 483, 350 *(wavelets)*
Damiani,F. et al. 1997, ApJ, 483, 370
Harnden,F.R.,jr. et al. 1984, SAO special report 393 *(sliding-cell)*
Cruddace, R., et al. 1988, in Astronomy from Large Databases, ESO *(max-likelihood)*

# Sample *pwdetect* session

(...is not as terrible as you maybe expect!)

```
prompt> ./pwdetect_1.3.2.lnx
 Thu Jul 26 19:03:22 2001

                    Welcome to PWDETECT V1.2
               Palermo Wavelet Detect for CHANDRA data

Input file name (event_list.fits):
Final threshold significance for detection (5.2):
Maximum wavelet scale (arcsec) [powers of sqrt(2)] (16):
Detection at Maximum Scale only (y/n) (n):
Additional edge margin (arcsec) (7):
Relative exposure threshold (0-1, recommended: 0.9) (0.9):
ACIS single chip mode flag (0=Full ACIS-I, 1=Single Chip) (0):
ACIS single chip Identifier (0 to 9) (0):
Output source list filename root (det):
Output background filename root (bkgmap):
Output Log file name (det_log):
Live show (1=yes, 0=no) ? (0):
Exposure map file name (exposure_map_1.0keV.fits):

 Reading exposure map:  exposure_map_1.0keV.fits ...done

 initial threshold =  4.00 sigmas
 final threshold   =  5.20 sigmas
 Detector is ACIS-I
   ... reading data
   ... successfully read       20295 events
 performing detection at 11 scales,
 from 0.50 to  16.00 arcsec, in logarithmic steps by a factor sqrt(2)
```

*Input data*

← event file to analyze
← detection threshold
← maximum detection scale

} optional parameters

} output (root) file names

← exposure map

*Pwdetect reads data*

(continues...)

```
computing background image... done
    51  sources at scale =  0.50 arcsec
    69  sources at scale =  0.71 arcsec
    69  sources at scale =  1.00 arcsec
    98  sources at scale =  1.41 arcsec
   113  sources at scale =  2.00 arcsec
   143  sources at scale =  2.83 arcsec
   143  sources at scale =  4.00 arcsec
   118  sources at scale =  5.66 arcsec
   105  sources at scale =  8.00 arcsec
    86  sources at scale = 11.31 arcsec
    63  sources at scale = 16.00 arcsec
computing best source properties... done
computing background image update... done
    51  sources at scale =  0.50 arcsec
    69  sources at scale =  0.71 arcsec
    69  sources at scale =  1.00 arcsec
   100  sources at scale =  1.41 arcsec
   113  sources at scale =  2.00 arcsec
   146  sources at scale =  2.83 arcsec
   148  sources at scale =  4.00 arcsec
   121  sources at scale =  5.66 arcsec
   106  sources at scale =  8.00 arcsec
    91  sources at scale = 11.31 arcsec
    66  sources at scale = 16.00 arcsec
computing best source properties... done

 Total number of detected sources =         105

 Nominal effective area (cm^2) =   286.6157

Thu Jul 26 19:10:46 2001
total elapsed time:         444 sec
prompt>
```

Background map computation.
Source detection at each scale...

...their matching, and fit to *y(a).*
New background map with sources removed.
Source detection using new background...

...and final source properties:
significance, position, count rate, size, ...

All done!