X-ray Stacking

Tom Aldcroft   SAO/CXC

Motivation for stacking

- Multiwavelength analysis is a major theme in modern astrophysics
- Many types of astrophysical sources emit (at some level) over a wide range spanning radio to gamma rays
- Detection efficiency of a source class (e.g. obscured active galactic nuclei) is often best in a particular band
- Observations in other bands may result in non-detections
- How can we get the most information from the non-detections?
  - Individual SEDs – upper limits constrain physical models
  - Survival analysis for a sample
  - Stacking – mean properties of sample

- Chandra X-ray data (faint point sources) are photon-limited with low background => stacking in X-rays is very effective

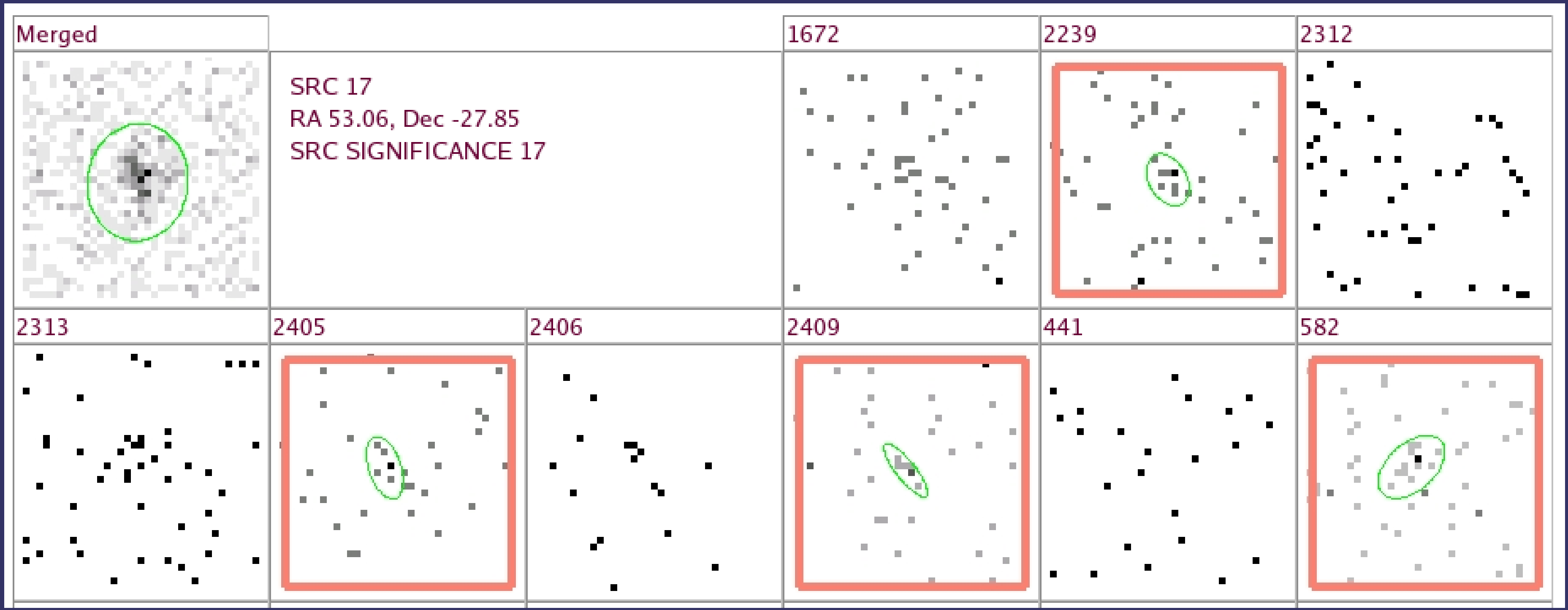
Stacking basics: Make an image

- Define source positions from input catalog

EXAMPLES:

  - 1376 (2308) optically-selected unobscured QSOs from SDSS
  - ~1600 (2150) IR-selected sources from Spitzer (obscured AGN)

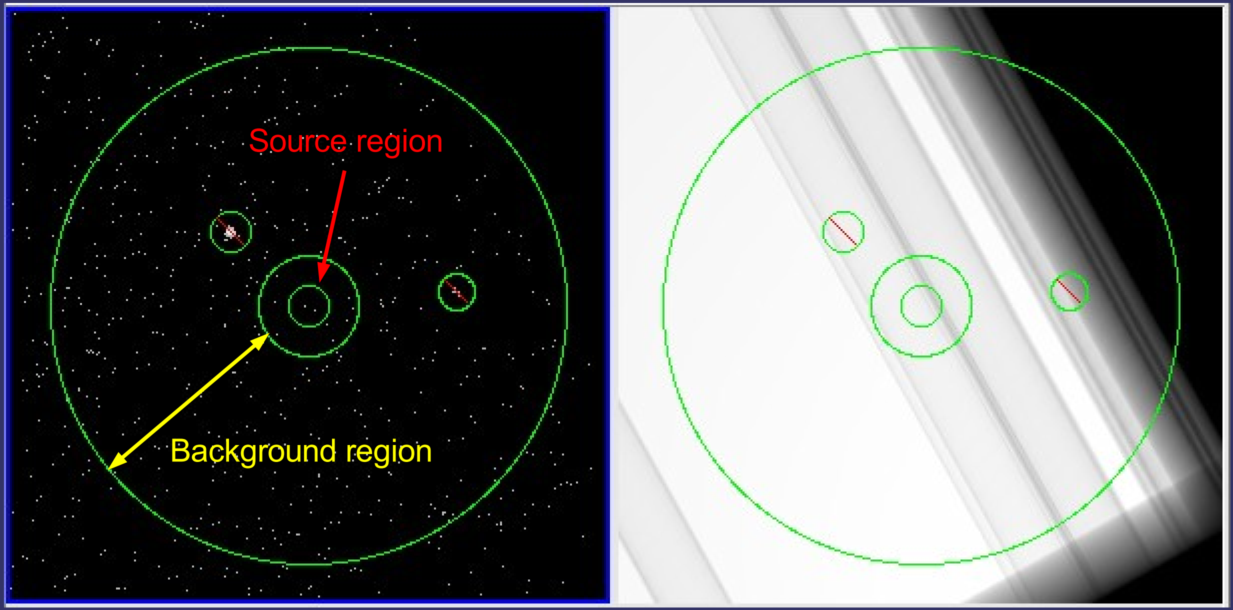
- Extract image cutouts in X-ray data at each source position
- Co-add images for sources that are *not detected* in X-ray



Source, background and exposure map

- At the position of source *i* in input catalog define source region S_i based on local PSF and background region B_i
- Exclude sources that are known from X-ray or input catalogs
- Exposure map $E_{S,i}$ is effective area (cm^2) $\times$ exposure time (sec) at each pixel

X-ray photons for source *i* Exposure map

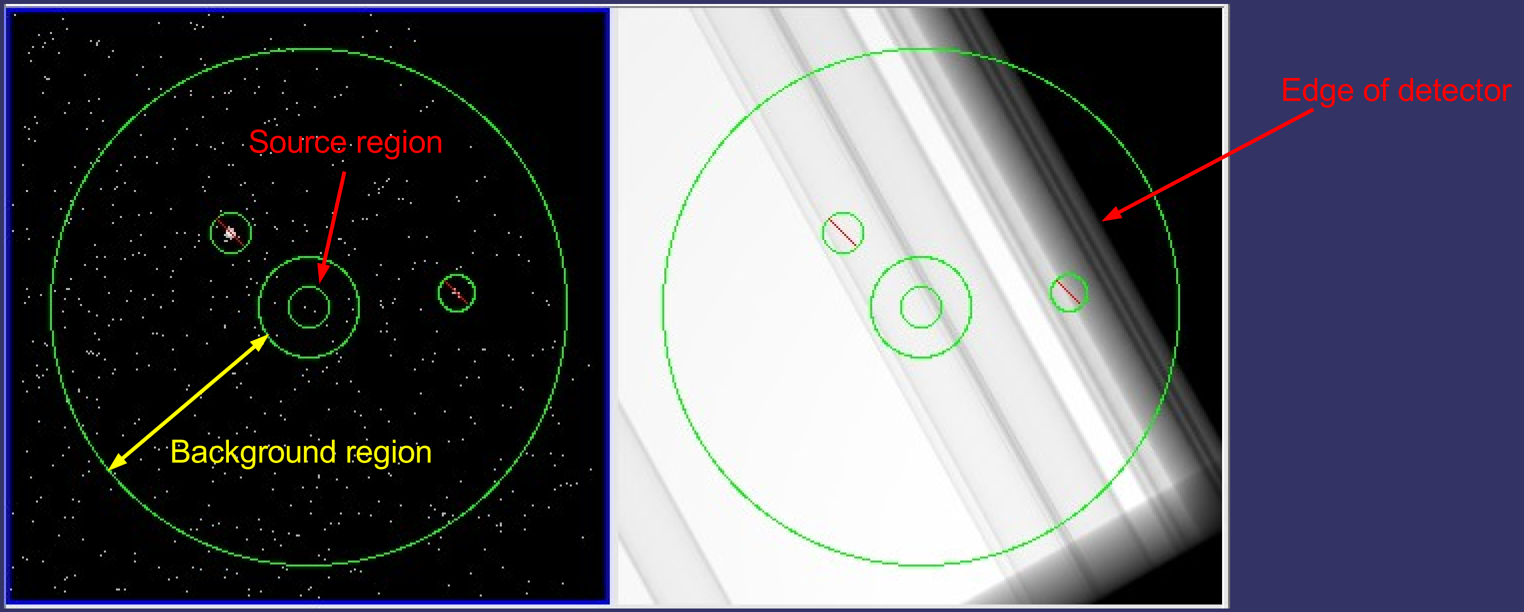


Stacking basics: Net counts and flux

- Net counts *C* and mean flux *f*

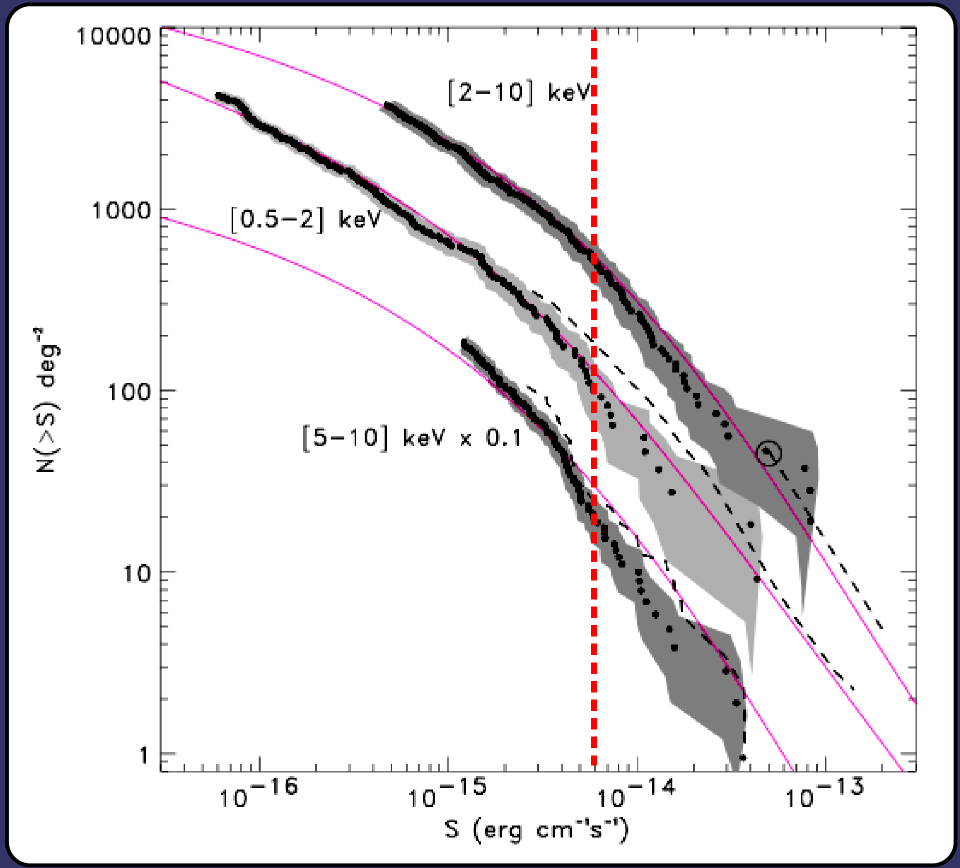
$$C = \sum_i \left(S_i - B_i \frac{\sum_{pix} E_{S,i}}{\sum_{pix} E_{B,i}} \right) \qquad \bar{f} = \frac{C}{\sum_i \bar{E}_{S,i}}$$

X-ray photons for source *i* Exposure map



Issues

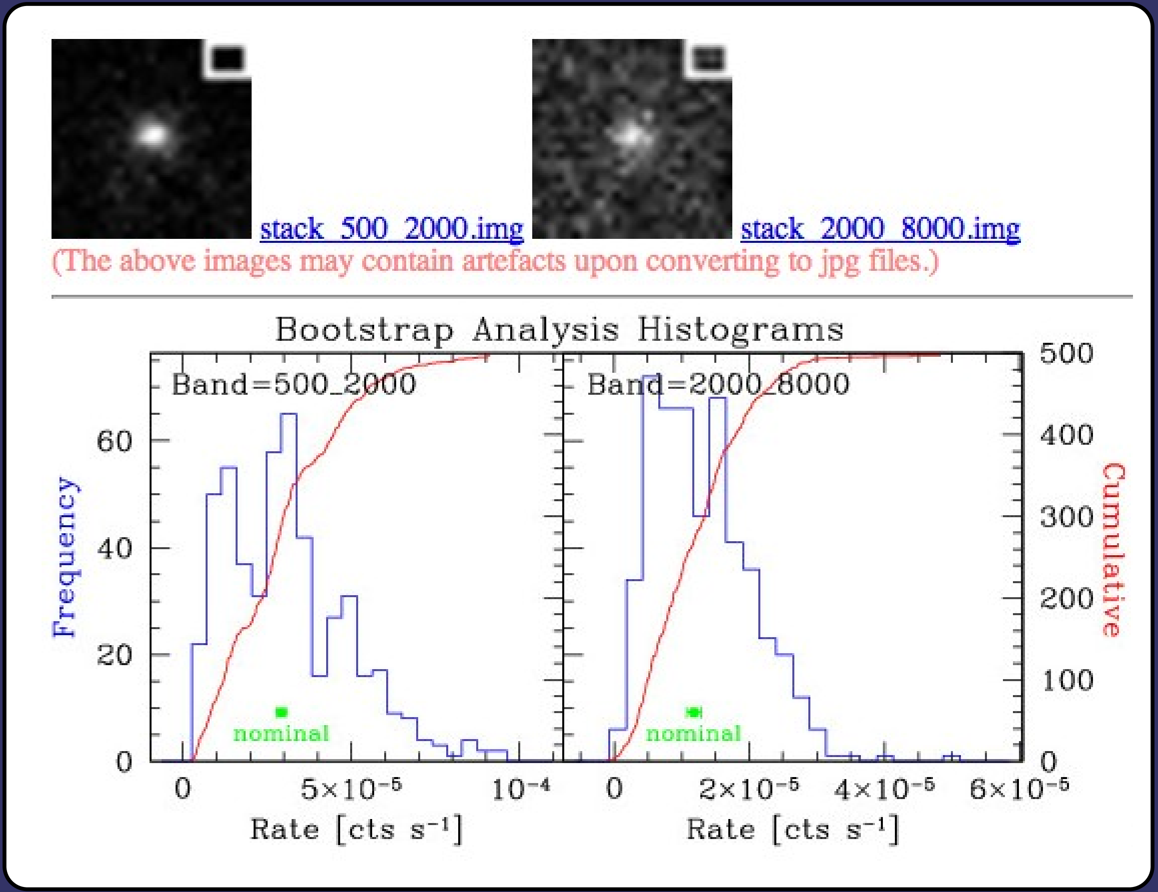
- The sample of X-ray "non-detections" depends on detect method and draws a hard distinction between sources straddling the detection limit
- Are a few sources dominating the signal?



Rosati et al. ApJS 139 (2002)

Issues

- Estimate distribution in *N* or *f* by random resampling of input sources
- Question – what is this really telling us?



http://saturn.phys.cmu.edu/cstack
User: guest Password: guest

Advanced stacking: ChaMP + SDSS

- Simple stacking tools calculate net counts or flux assuming homogenous data sets - exposure and responses are fairly uniform.
- The Chandra Multiwavelength Project (ChaMP) is a very large X-ray survey using archival data spanning 6 years: GOOD
- The ChaMP has with a wide range in exposure and responses: BAD
- The goal is to properly account for these complications to explore AGN and galaxy physics and evolution.
- The Sloan Digital Sky Survey (SDSS) is a massive optical survey of the sky which provides > 100,000 spectroscopically confirmed quasars with redshifts (and much much more).
- There is a 30 deg^2 overlap between ChaMP and SDSS.

Advanced stacking: Luminosity for a heterogenous sample

- Since we have redshift (distance) estimates for the input SDSS samples we want to infer luminosity rather than flux.

$$
L(\mathcal{E}_{\rm rest}) = \begin{array}{lll}
 & \textit{Dependencies} & \textit{Units} \\
C(\mathcal{E}_{\rm obs}) & & \text{counts} \\
\times \text{ApertureCorrection} & (\text{Position}) & \text{counts} \\
\times \text{ExposureMap}^{-1} & (\text{Time, position, detector, etc}) & \text{photons/cm}^2\text{/sec} \\
\times \text{EnergyCorrectionFactor} & (\mathcal{E}_{\rm obs}, \text{spectrum}, \text{z}) & \text{erg/cm}^2\text{/sec} \\
\times (4\pi d_l^2)^{-1} & (z, \text{cosmology}) & \text{erg/sec}[\mathcal{E}_{\rm obs}/(1+\text{z})] \\
\times \text{Kcorrection} & (z, \mathcal{E}_{\rm obs}, \mathcal{E}_{\rm rest}, \text{spectrum}) & \text{erg/sec}
\end{array}
$$

$$
\bar{L} = \frac{1}{N} \sum_i C_i \frac{A_{\rm corr,i} \times \mathrm{ECF_i} \times \mathrm{K_{corr,i}}}{\bar{E}_{S,i}\, 4\pi d_{l,i}^2}
$$

Mean of L values

$$
\bar{L} = \frac{1}{\sum_i \bar{E}_{S,i}} \sum_i C_i \frac{A_{\rm corr,i} \times \mathrm{ECF_i} \times \mathrm{K_{corr,i}}}{4\pi d_{l,i}^2}
$$

Weight by exposure as per flux calculation

Is this even “stacking” any more?

Discussion!

- Now the real question: Some mean luminosity will be generated but how do we calculate the confidence intervals?

$$C_i = S_i - B_i \frac{\sum_{pix} E_{S,i}}{\sum_{pix} E_{B,i}}$$

~0 to 5 cts

~5 to 25 cts

$$\bar{L} = \frac{1}{\sum_i \bar{E}_{S,i}} \sum_i C_i \frac{A_{\mathrm{corr,i}} \times \mathrm{ECF_i} \times \mathrm{K_{corr,i}}}{4\pi d_{l,i}^2}$$

- Would it be useful to plot / analyze histograms of L_i? This avoids the (somewhat) arbitrary distinction between detected and non-detected sources.