The Chandra Source Catalog 2.0: Estimating Source Fluxes

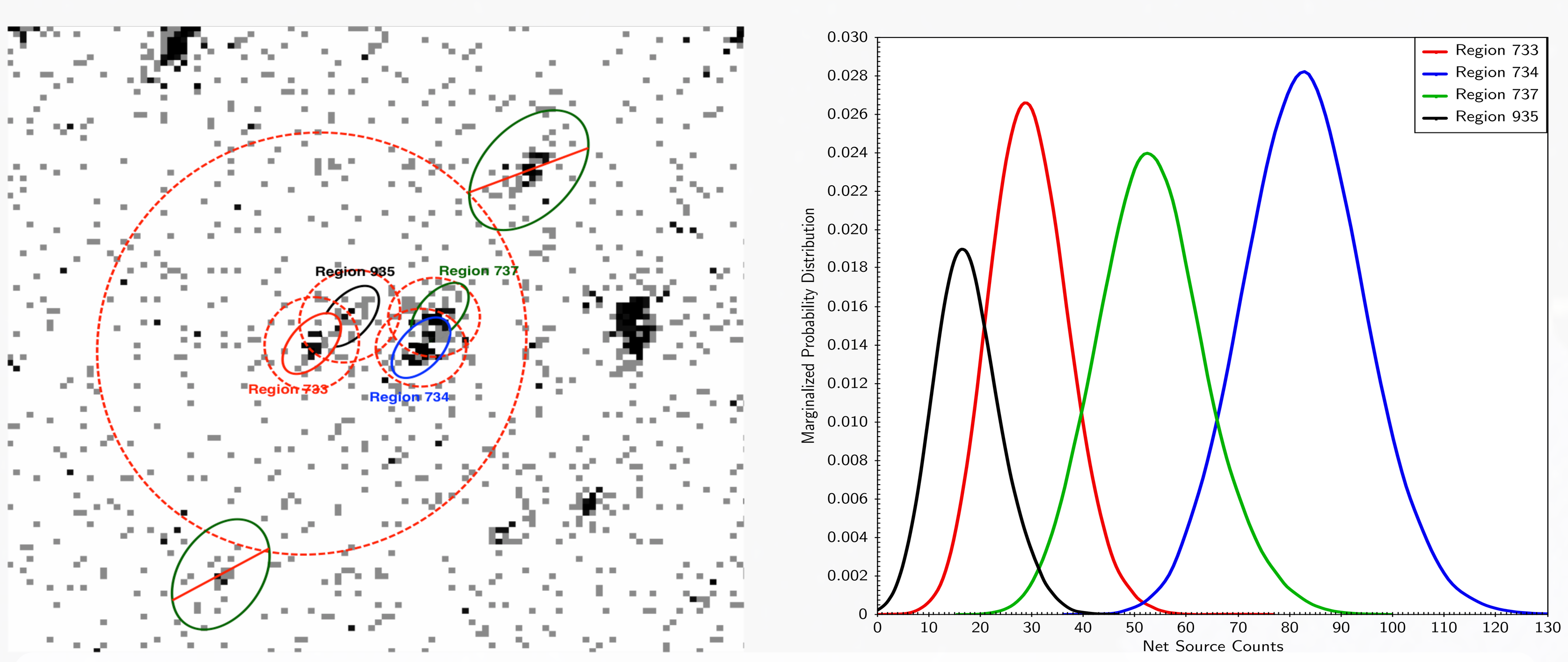
Francis A. Primini[1], Christopher Allen[1], Joseph B. Miller[1], Craig S. Anderson[1], Jamie A. Budynkiewicz[1], Douglas Burke[1], Judy C. Chen[1], Francesca Civano[1], Raffaele D'Abrusco[1], Stephen M. Doe[2], Ian N. Evans[1], Janet D. Evans[1], Giuseppina Fabbiano[1], Danny G. Gibbs II[1], Kenny J. Glotfelty[1], Dale E. Graessle[1], John D. Grier[1], Roger M. Hain[1], Diane M. Hall[3], Peter N. Harbo[1], John C. Houck[1], Jennifer Lauer[1], Omar Laurino[1], Nicholas Lee[1], J. Rafael Martinez-Galarza[1], Michael L. McCollough[1], Jonathan C. McDowell[1], Warren McLaughlin[1], Douglas L. Morgan[1], Amy E. Mossman[1], Dan T. Nguyen[1], Joy S. Nichols[1], Michael A. Nowak[4], Charles Paxson[1], David A. Plummer[1], Arnold H. Rots[1], Aneta Siemiginowska[1], Beth A. Sundheim[1], Michael S. Tibbetts[1], David W. Van Stone[1], Panagoula Zografou[1]

[1]Smithsonian Astrophysical Observatory [2] formerly Smithsonian Astrophysical Observatory [3]Northrop Grumman Mission Systems [4]MIT Kavli Institute for Astrophysics and Space Research

The Second Chandra Source Catalog (CSC2.0) will provide information on approximately 316,000 point or compact extended x-ray sources, derived from over 10,000 ACIS and HRC-I imaging observations available in the public archive at the end of 2014. As in the previous catalog release (CSC1.1), fluxes for these sources will be determined separately from source detection, using a Bayesian formalism that accounts for background, spatial resolution effects, and contamination from nearby sources. However, the CSC2.0 procedure differs from that used in CSC1.1 in three important aspects. First, for sources in crowded regions in which photometric apertures overlap, fluxes are determined jointly, using an extension of the CSC1.1 algorithm, as discussed in Primini & Kashyap (2014ApJ...796…24P). Second, an MCMC procedure is used to estimate marginalized posterior probability distributions for source fluxes. Finally, for sources observed in multiple observations, a Bayesian Blocks algorithm (Scargle, et al. 2013ApJ...764..167S) is used to group observations into blocks of constant source flux.

1. BAYESIAN ESTIMATION OF FLUXES

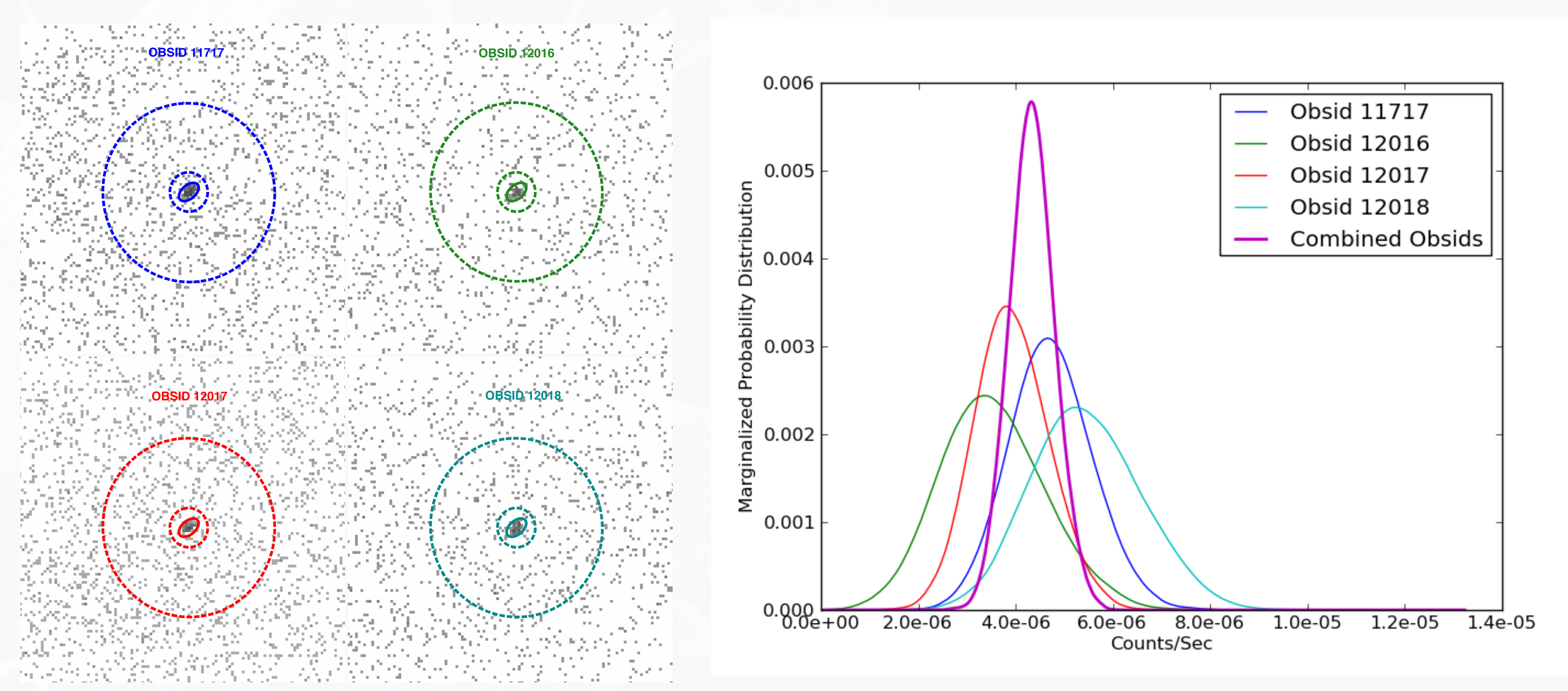
- Analyze sources with overlapping apertures, near-by sources, and background simultaneously.
- Assign counts and area in overlap region to aperture of brighter source.
- Assume counts in source and background apertures $\{C_i, B\}$ governed by Poisson Distributions, i.e.,
  - $Likelihood\ L(C_i|\mu_i) = \mu_i^{C_i} e^{-\mu_i}/C_i!$ where
  - $\mu_i = expected\ counts\ in\ aperture\ i = \sum_j f_{ij} \cdot T_i \cdot s_j + \Omega_i \cdot T_i \cdot b$,
  - $\{s_i, b\}$ = model source, background fluxes,
  - f_{ij} = fraction of normalized point spread function for source j in aperture i,
  - T_i = exposure in aperture i,
  - Ω_i = area in aperture i.
- Joint posterior probability distribution for $\{s_i, b\}$ given by
  - $P(s_i \cdots b|C_i \cdots B) \sim P(\mu_b) L(B|\mu_b) \times \prod_i P(\mu_i) L(C_i|\mu_i)$.
  - Assume non-negative, but otherwise non-informative priors for $P(\mu)$.
- Use *sherpa* (cxc.cfa.harvard.edu/sherpa) to optimize $P(s_i \cdots b|C_i \cdots B)$.
- Use *MCMC* (hea-www.harvard.edu/AstroStat/pyBLoCXS) to sample posterior probability. distributions to estimate errors in parameters $\{s_i, b\}$:
  - Approximate marginalized posterior probability distributions (MPDFs) for $\{s_i, b\}$ with smoothed distributions of draws with respect to $\{s_i, b\}$ separately.
  - Report modes and 68% percentiles of MPDFs as flux values and error bounds.



Four source apertures (solid ellipses) in the image on the left, are analyzed together with the background (region bounded by dashed ellipses). MCMC draws are used to characterize marginalized posterior probability distributions for each source, shown by the colored curves on the right.

2. COMBINING OBSERVATIONS

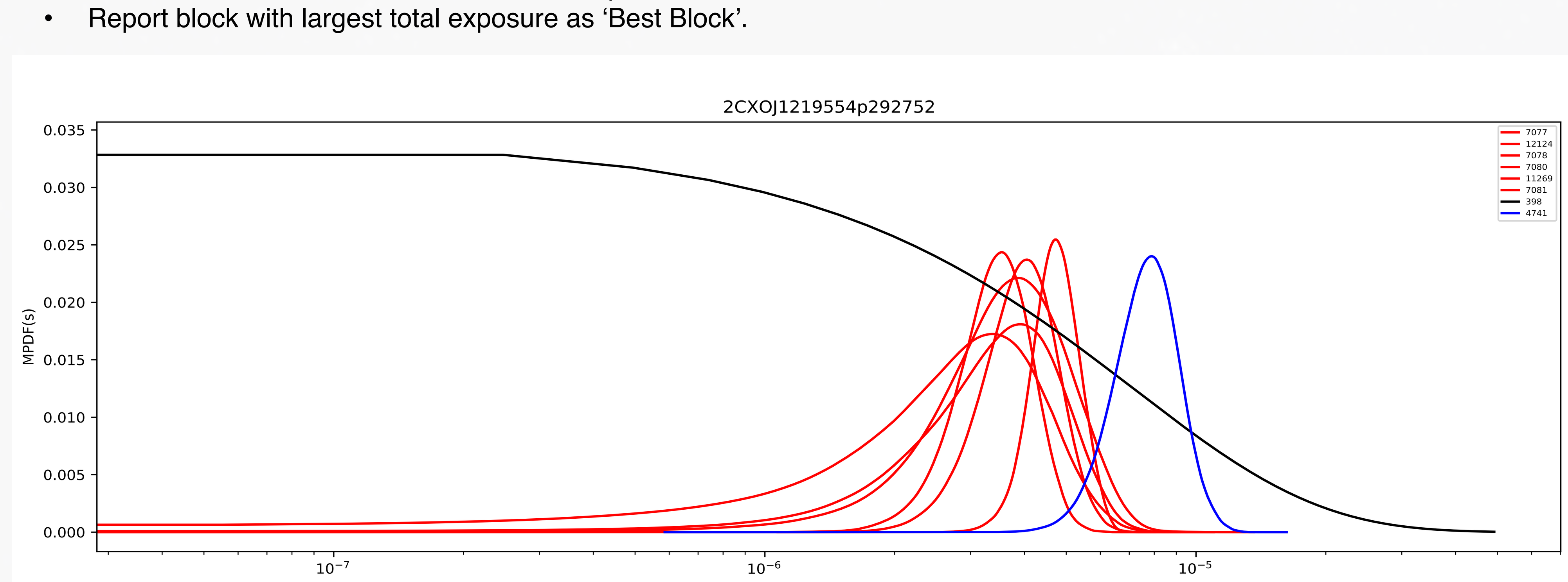
- Estimate single flux for source, combining aperture data from multiple observations.
- Joint posterior probability distribution $P(s, b^1 \cdots b^n|C^1 \cdots C^n, B^1 \cdots B^n)$ is the product of those described in Box 1 for each observation.
- Model fluxes are now $\{s, b^1 \cdots b^n\}$ for n observations (i.e., separate background fluxes are allowed for each observation, but only a single source flux).
- Again use *sherpa* to optimize $P(s, b^1 \cdots b^n|C^1 \cdots C^n, B^1 \cdots B^n)$ and $MCMC$ to estimate errors.



The same source is observed in four separate observations in the same stack, and exhibits a range of fluxes. The MPDF from the combined data has a much narrower distribution, as expected.

3. WHICH OBSERVATIONS TO COMBINE?

- Many x-ray sources are variable in either overall intensity, spectral shape, or both.
- Combining data from all observations may not result in optimum dataset for further analysis, such as spectral fitting.
- Use Bayesian Blocks algorithm of Scargle et al. (2013ApJ…764..167S) to group multiple $MPDF_j(s)$ into a set of blocks $\{B_i\}$ for which constant flux is consistent with MPDFs in each block, within observational errors.
- Prior Probability of getting N_{blocks} blocks: $P(N_{blocks}) \sim \gamma^{N_{blocks}}$, $N_{blocks} \le N_{obs}$, $0 < \gamma \le 1$.
- Probability of getting particular block B_i of observations with $\{MPDF_j\}: F(B_i|MPDF_j \in B_i) = \int_0^\infty ds \left[\prod_{j|MPDF_j \in B_i} MPDF_j(s)\right]$.
- Probability of getting set of blocks: $P(\{B_i\}|MPDF_j) = P(N_{blocks}) \prod_{i=1}^{N_{blocks}} F(B_i|MPDF_j \in B_i)$.
- Maximize $\log[P(\{B_i\}|MPDF_j)]$.
- Combine observations in each block to compute block fluxes.
- Report block with largest total exposure as 'Best Block'.



Eight separate b-band observations of source 2CXOJ121955.4+292752 reveal a range of fluxes. Our analysis groups those observations shown in red into a single block.

4. UNDER THE HOOD

- DIFFERENT FLUX UNITS:
  - Net counts, count rate, photon flux, and energy flux, in 6 energy bands (5 for ACIS observations, one for HRC observations).
  - Units for T_i (see Box 1) determine the flux units ($T_i \times s_i$ has units of counts).
- TWO APERTURE SIZES:
  - 'ECF90' (90% of the point-spread function) and a typically larger 'source region', derived from source detection process.
  - Comparison of fluxes in two apertures illuminates possible source extent.
- BAYESIAN BLOCKS ANALYSIS:
  - Observations sorted in both time and flux. Time-ordered blocks include observations that are sequential in time. Flux-ordered blocks can include non-time-sequential observations.
  - Performed in the s, m, h, and b bands for ACIS observations. Final ACIS blocks are determined from the intersection of band-specific blocks.
  - HRC blocks determined separately.
- *SHERPA*:
  - Optimize $P(s_i \cdots b|C_i \cdots B)$ by fitting expected counts in source or background apertures to observed counts, with model parameters being source and background fluxes.
  - *Monte Carlo* optimization technique used.
- UPPER LIMITS:
  - Determined when the best-fit flux is ≤ 0, or for when lower confidence bound is 0.
  - Determined for sources not detected in stack but detected in overlapping stacks.
  - MPDFs typically exponential rather than Gamma distributions.

CHANDRA
SOURCE CATALOG

This work has been supported by NASA under contract NAS 8-03060 to the Smithsonian Astrophysical Observatory for operation of the Chandra X-ray Center.