A Brief Intro to the Chandra Mission

Jonathan McDowell
The Chandra X-ray Observatory

Launched over 20 years ago 23 July 1999

A revolution in X-ray astronomy
and astronomy in general
What is Chandra?

NASA Chandra X-ray Observatory – operated by Chandra X-ray Center (CXC) at the Center for Astrophysics (Cambridge, MA) in collaboration with MIT

The greatest X-ray telescope ever built!

Orbits the Earth to be above the atmosphere

Goes 1/3 of the way to the Moon

every 64 hours (2 ½ days)

Observing only while outside radiation belts

Chandra takes superbly sharp images:

with good spectral resolution too!
X-ray Telescopes are different

Chandra’s mirrors are almost cylinders

X-rays don’t reflect off a normal mirror – they get absorbed.
Only by striking a mirror at a glancing angle, about 1°, do X-rays reflect.
Focal length 10 m
4 mirror pairs - perfectly aligned (not quite)
The Chandra spacecraft

10 meters (32 ½ ft) from mirror to detector, 1.2 meters (4 ft) across mirror

...but focuses X-rays onto a spot only 25 microns across
Chandra Instruments and Gratings:

**ACIS**  Advanced CCD Imaging Spectrometer
- Most Chandra observations use ACIS
- 10 x 1 Mpix chips, use any 4 (caveats)
- Sub arcsecond imaging on axis
- Each chip is 8 x 8 arcmin
- Spectrum in each pixel, 0.5 (ish) to 7 (ish) keV

**HRC-I and HRC-S**  High Resolution Camera
- Microchannel plate, highest spatial and timing res but no spectral res

**HETG**  High Energy Transmission Grating
- Disperses spectrum onto ACIS
- Two gratings, HEG and MEG, at an angle to each other
- Order sorting possible via CCD energy discrimination

**LETG**  Low Energy Transmission Grating
- Disperses spectrum onto HRC-S or ACIS
- Single dispersed spectrum
- Orders superimposed
Chandra science center
Smithsonian Observatory, at Harvard (Cambridge, MA)

Chandra mission control
Near MIT in Cambridge, MA

DSN control at Jet Propulsion Lab Pasadena, CA
Main CIAO software release December each year (typically)

Proposals due in March each year

Archive available all year round!! cxc.harvard.edu
Introduction to X-ray Data Analysis

Jonathan McDowell
Introduction to X-ray Data Analysis

- X-ray astronomy is different ..... 
  - Problem 1: Photon counting with small number statistics
  - Problem 2: Spectral line spread function is often broad and messy - forced to foward-folding approach
  - Problem 3: Bands are very broad, so energy (wavelength) dependence more obvious (e.g. in PSF)
  - Problem 4: Different optics - PSF degrades rapidly off axis
  - Problem 5: The telescope is not pointing steadily like, say, HST - it's moving back and forth across the source.
  - But:
  - Advantage: We have more information on each photon (position, energy, arrival time)
Imaging data has limited energy resolution and modeling can only be done via 'forward folding' spectral analysis (a theoretical model is fitted to the data until the best fit is found).

In Poisson statistics regime because of the small number of photons

2 decades of photon energy (0.1 to 10 keV)
Complexities in X-Ray and Chandra Data Analysis

In Poisson statistic regime because of the small number of photons, imaging data has limited energy resolution and modeling can only be done via 'forward folding' spectral analysis (a theoretical model is fitted to the data until the best fit is found).

2 decades of photon energy (0.1 to 10 keV)
Complexities in X-Ray and Chandra Data Analysis

Every aspect of the observation varies with:
- energy
- position
- time

(e.g. image sharpness, sensitivity, instrumental energy scale)

The Chandra PSF

Q0836+7104 predicted vs. observed

Pileup Effect

Specific instrumental effects eg.
readout streak
pileup - two or more photons detected as single event
Basics of CIAO

- Data files are in FITS format (usually binary tables, not just images)
  FITS: Flexible Image Transport System (1979-present)
- CIAO can also operate on ASCII files in many cases
- All (well, almost all) CIAO tools that want an input file can accept a CIAO Data Model "virtual file"
  e.g instead of     evt.fits
take                 “evt.fits[energy=300:1000,sky=circle(4096,4096,20)]”

Each file (dataset) is made up of sections called 'blocks' (HDUs for FITS fans)
Blocks can be tables or images

Key tools:
  dmcopy  infile outfile
  dmclist infile opt=blocks,cols,data

  ahelp dmclist → help for tool dmclist
  plist dmclist → list parameters for dmclist

Key applications:
  Sherpa  - fitting
  ds9 – imaging and analysis
  (also Prism, now within ds9: file explorer)
The Event File

- In optical astronomy, the primary data set is an image. In radio interferometry, it's a visibility array.

- In X-ray astronomy, the primary data set is an event list - a table of (putative) photons
  - Our software makes it easy to generate an image from the event list, so it's easy to forget that's what you have. But making the image loses information.
  - First cut way of thinking about the event list: it's a 4-dimensional array of x, y, time, energy. But most pixels are empty (we don't have many photons!) so it's more compact to just list the non-empty ones.
  - Complication: we actually have many more parameters for each photon, not just 4.
Inside the event list

<table>
<thead>
<tr>
<th>CoNo</th>
<th>Name</th>
<th>Unit</th>
<th>Type</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>time</td>
<td>s</td>
<td>Real18</td>
<td>154361559.6127999964:154436827.4158589973 S/C TT corresponding to mid-exposure</td>
</tr>
<tr>
<td>2</td>
<td>ccd_id</td>
<td></td>
<td>Int2</td>
<td>0:9 CCD reporting event</td>
</tr>
<tr>
<td>3</td>
<td>node_id</td>
<td></td>
<td>Int2</td>
<td>0:3 CCD serial readout amplifier node</td>
</tr>
<tr>
<td>4</td>
<td>expno</td>
<td></td>
<td>Int4</td>
<td>0:2147483647 Exposure number of CCD frame containing event</td>
</tr>
<tr>
<td>5</td>
<td>chip(chipx,chipy)</td>
<td>pixel</td>
<td>Int2</td>
<td>1:1024 Chip coords</td>
</tr>
<tr>
<td>6</td>
<td>tdet(tdet,t_dety)</td>
<td>pixel</td>
<td>Int2</td>
<td>1:8192 ACIS tiled detector coordinates</td>
</tr>
<tr>
<td>7</td>
<td>det(detx,dey)</td>
<td>pixel</td>
<td>Int4</td>
<td>0.50: 8192.50 ACIS detector coordinates</td>
</tr>
<tr>
<td>8</td>
<td>sky(x,y)</td>
<td>pixel</td>
<td>Int4</td>
<td>0.50: 8192.50 sky coordinates</td>
</tr>
<tr>
<td>9</td>
<td>pha</td>
<td>adu</td>
<td>Int4</td>
<td>0:36855 total pulse height of event</td>
</tr>
<tr>
<td>10</td>
<td>pha_ro</td>
<td>adu</td>
<td>Int4</td>
<td>0:36855 total read-out pulse height of event</td>
</tr>
<tr>
<td>11</td>
<td>energy</td>
<td>eV</td>
<td>Real14</td>
<td>0:100000.0 nominal energy of event (eV)</td>
</tr>
<tr>
<td>12</td>
<td>pi</td>
<td>chan</td>
<td>Int4</td>
<td>1:1024 pulse invariant energy of event event grade, flight system</td>
</tr>
<tr>
<td>13</td>
<td>fllgrade</td>
<td></td>
<td>Int2</td>
<td>0:255 binned event grade</td>
</tr>
<tr>
<td>14</td>
<td>grade</td>
<td></td>
<td>Int2</td>
<td>0:7 event status bits</td>
</tr>
<tr>
<td>15</td>
<td>status[4]</td>
<td></td>
<td>Bit(4)</td>
<td></td>
</tr>
</tbody>
</table>
Energy slices through an event list, 0.1 - 10 keV
Level 1 Event List - Calibrated but Dirty

Node boundaries

Lots of background

Bad columns

Source!

Bad pixels
Level 2 event list - cleaned and filtered

Energy filter 300-7000 eV removes background but not signal
Grade filter removes cosmic ray events etc
Good time filter removes times of high background, poor data quality

Sources fuzzy far off axis (PSF big)

Beware chip gaps!

More sources!
During an observation, Chandra's optical axis describes this 'dither pattern' on the sky, (Problem 5), smearing the image of a point source. The RA, Dec, roll angle of the telescope versus time is called the 'aspect solution'; the asol1.fits file provides this for each observation.

We record the motion of the guide stars in the star tracker so that we can calculate RA and Dec for EACH PHOTON and so reconstruct the image.
Chandra aspect-corrected data

This is what you get after calibration but before cleaning the data. Note the sharp point sources near the center.
In instrument space, the photons are spread out over 20 arcsec and have bad columns going through them - so be careful of the effective exposure time. If you didn't dither, you could lose the source entirely if it landed on a bad pixel.
Spatial Response: EXPOSURE MAP

The Exposure Map, \( E(\Delta h, \lambda, \hat{p}) \), retains spatial information at the expense of spectral. It has units of \([\text{cm}^2 \text{ counts photons}^{-1}]\).

\[
\int d\lambda \ S(\lambda, \hat{p}) \approx \frac{C(\Delta h, \hat{p})}{E(\Delta h, \lambda, \hat{p})}
\]

\( C \) is the observed counts per spatial bin in a pulse-height bin. \( S \) is the source flux, with units of \([\text{phot cm}^{-2}\text{s}^{-1}\text{Å}^{-1}]\).

**Instrument Map** – efficiency calibration information, band integrated. (create with `mkinstmap`)

\[=\text{mirror area x detector QE}\]

**Exposure Map** – applies telescope aspect history and coordinate transformations (= area x time). (create with `mkexpmap`).

\[=\text{Instmap Aspect}\]
Problem 3: Exposure map is energy dependent; must assume a spectrum if using a broad band
Event analysis or binned analysis?

• Don't make an image too quickly. If you can get an answer directly from the event list, that's better - binning the data loses information, and collapsing the axes loses information.

• Spatial analysis: make an image (using dmcopy)
  - lose energy and time information

• Spectral analysis: make a 'PHA file' using dmextract (or a grating spectrum using tgeextract)
  - lose spatial and time information

• Temporal analysis: make a light curve using dmextract
The fundamental equation of astronomy

\[ N(E) = A(E)F(E)\Delta T \]

Our instrument makes a spectrophotometric measurement; the sensitivity (“effective area”) \( A(E) \) tells us how to convert from flux to instrumental counts for a given exposure time \( \Delta T \).

But, a real instrument doesn't measure the true energy, it measures instrumental energy \( E' \). The line spread function (“response matrix” in X-rays) \( R(E,E') \) describes how a monochromatic input spectrum is broadened by the instrument (Problem 2).

Let us further assume that the instrumental energy \( E' \) is measured in discrete channels (bins) \( E'_i \). Then

\[ N(E'_i) = \int A(E)R(E, E'_i)F(E)dE\Delta T \]

Of course, you may not be measuring all of the light from the source. Even if it's a point source, there may be an aperture correction. We need the PSF \( P(x-x', y-y') \) and the spatial dependence of the QE, \( q(x,y) \). Then at a given instrument position \( x', y' \)

\[
\begin{align*}
N(E'_i, x'_i, y'_i) &= \int \int A(E)R(E, E'_i)F(E, x, y)P(x-x'_i, y-y'_i)q(E, x'_i, y'_i)dEdxdy\Delta T \\
N(E'_i, x'_i, y'_i) &= \int \int \int A(E)R(E, E'_i)F(E, x, y, t)P(x-x'_i, y-y'_i)q(E, x'_i, y'_i)dEdxdydt
\end{align*}
\]

The source may also be variable in time - we'll ignore this for the purposes of this talk. The detector sensitivity is time-variable on long timescales, but for a single observation you just have to worry about times when the data is filtered - the Good Time Intervals (GTIs).
Pulse height

When you plot an optical spectrum, the wavelength (or energy) axis is really an instrumental quantity. A spectral line is broadened by instrumental effects, so the energies plotted are not the true energies of the photon. However, the instrument is calibrated (i.e. the definition of instrumental energy is rescaled) such that the peak of a line is at the correct energy.

In X-ray astronomy, instead of using the instrumental energy $E'$, we work with the energy bin number. For historical reasons to do with long-forgotten instruments, this bin number is known as the PI channel (for 'pulse invariant' channel) - we'll denote it by $P$. So, for fixed energy bin widths $dE$,

$$E' = P \ dE = [\text{on average}] \ E$$

The instrument actually measures a raw energy bin number $p$, called the PHA channel, or 'pulse height analyser channel'. The scaling of the instrumental energy to real energy depends on position and time:

$$E'(\text{raw}) = p \ dE = g(x,y,t)P \ dE$$

This function $g$ (the gain) is usually assumed to obey

$$g(x,y,t) = g_{\text{spatial}}(x,y) \ g_{\text{t}(t)}$$

and we provide calibrations of both the spatial gain and the temporal gain.
Spectra in Poissonland

We pick a parameterized $F(E)$ such as warm absorber models, lines, thermal plasma codes. Which $F(E)$? You must pick one based on expected physics, but match number of free parameters with quality of data.

With less than 100 counts, we usually just use count ratios (X-ray colors) for spectral analysis.

Does one model fit significantly better than another? Be careful that two physically different models may look quite similar in $F(E)$ space.

Incompletely calibrated instrumental features may show up in residuals, limiting factor in high S/N spectra – these features may include edges. Beware apparent science in regions where $A(E)$ is changing rapidly.

$$N(p) = \int R(E, p) A(E) F(E) dE$$
Modeling and fitting for 1-D and 2-D datasets in any waveband including: spectra, images, surface brightness profiles, light curves, general ASCII data.

Model Poisson and Gaussian data

Calculate confidence levels on the best-fit model parameters

Coded in a Python environment – familiar to the new generation of astronomers and used in other missions.
comes with well-tested, robust optimization methods - e.g. Levenberg-Marquardt, Nelder-Mead Simplex or Monte Carlo/Differential Evolution

comes with statistics for modeling Poisson or Gaussian data

can perform Bayesian analysis with Poisson Likelihood and priors, using Metropolis or Metropolis-Hastings algorithm in the MCMC (Markov-Chain Monte Carlo); allows to include non-linear systematic errors (calibration uncertainties) in the analysis

is extensible (with python and compiled code):
  - is used in CIAO tools and scripts
  - in the Xija Chandra thermal modeling code
  - is used in the TeV HESS data analysis software
  - is used in the IRIS spectral energy distribution program
• The CALDB (Calibration Database) contains everything you need that's not part of your specific observation.

• It's designed as a multimission directory structure. The Chandra files are in $CALDB/data/chandra

• Within that, they are arranged by instrument and kind of calibration. But, with luck, the software will find the CALDB files you need automatically.

• Just make sure that you keep the CALDB up to date! But, be careful - if you start off processing with a given version of the CALDB and CIAO, then upgrade to a new CALDB and CIAO, things are sometimes incompatible. Check the release notes.
Calculating Source Flux

```bash
$ srcflux myevt2.fits "03:29:29.250 +31:18:34.73" myflux
```

Encodes the logic described in six different CIAO threads.

Return count rates and fluxes and errors with all appropriate corrections.

Summary of source fluxes

<table>
<thead>
<tr>
<th>Position</th>
<th>0.5 - 7.0 keV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>3 29 29.25 +31 18 34.7</td>
</tr>
<tr>
<td>90% Conf Interval</td>
<td>Rate 0.0398 c/s (0.0381,0.0415)</td>
</tr>
<tr>
<td>Flux</td>
<td>5.17E-13 erg/cm2/s (4.94E-13,5.39E-13)</td>
</tr>
<tr>
<td>Mod.Flux</td>
<td>4.38E-13 erg/cm2/s (4.2E-13,4.57E-13)</td>
</tr>
</tbody>
</table>

Uses many tools written for the Chandra Source Catalog.

Complementary to it for special cases and fields not covered by the catalog.
srcflux capabilities

- finds auxiliary files automatically, like specextract
- automatically determines PSF-appropriate extraction region size for source and background, or accepts user choice
- uses one of four methods to apply aperture correction
- runs on multiple energy bands including named CSC bands
- accepts one position or a list (catalog of sources)
- calculates count rates using aprates method
- calculates fluxes two different ways (specified spectral model and eff2evt method; however, no spectral fit is performed)
- generates spectral responses for further analysis

Ongoing work: handling of warning flags for hard cases, e.g. chip edge