Highly Structured Models in High Energy Astrophysics

David A. van Dyk
Department of Statistics, University of California at Irvine

Astro-Statistics

In recent years, there has been an avalanche of new data in observational high-energy astrophysics. Recently launched or soon-to-be launched space-based telescopes that are designed to detect and map ultra-violet, X-ray, and γ-ray electromagnetic emission are opening a whole new window to study the cosmos. Because the production of high-energy electromagnetic emission requires temperatures of millions of degrees and is an indication of the release of vast quantities of stored energy, these instruments give a completely new perspective on the hot and turbulent regions of the universe. The new instrumentation allows for very high resolution imaging, spectral analysis, and time series analysis. The *Chandra X-ray Observatory*, for example, produces images at least thirty times sharper than any previous X-ray telescope. The complexity of the instruments, the complexity of the astronomical sources, and the complexity of the scientific questions leads to subtle inference problems that requires sophisticated statistical tools. For example, data are subject to non-uniform stochastic censoring, heteroscedastic errors in measurement, and background contamination. Astronomical sources exhibit complex and irregular spatial structure. Scientists wish to draw conclusions as to the physical environment and structure of the source, the processes and laws which govern the birth and death of planets, stars, and galaxies, and ultimately the structure and evolution of the universe.

Figure 1: The Launch of the *Chandra X-ray Observatory*.

X-ray Astronomy: The sky in X-rays looks very different from that in visible light. X-rays are the signature of accelerating, energetic charged particles, such as those accelerated in very strong magnetic fields, extreme gravity, explosive nuclear forces, or strong shocks. Thus, X-ray telescopes can be used to study nearby stars (like our Sun) with active coronae, the remnants of exploding stars, areas of star formation, regions near the event horizon of a black hole, very distant but very turbulent galaxies, or even the glowing gas embedding a cosmic cluster of galaxies.

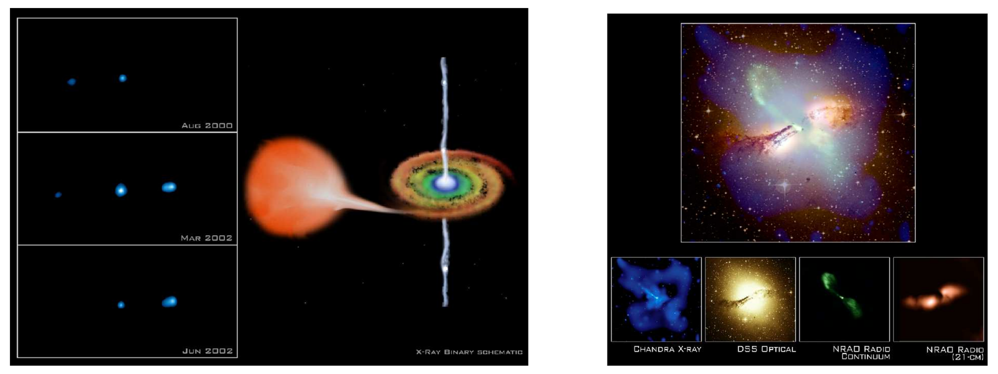

Figure 2: An X-ray image of an accreting black hole and a multi-wavelength image of two colliding galaxies.

Spectral Analysis

X-ray Spectra: An electromagnetic spectrum is the distribution of the energies of the photons that a source radiates. In X-ray astronomy photons are counted in each of a large number of energy bins. We explicitly model photon arrivals as a Poisson process, and thus have no difficulty with high resolution/low count high-energy spectral data.

The Basic Spectral Model: The spectral model has two components:

1. The *continuum*, a Generalized Linear Model for the baseline spectrum.
2. Several emission line profiles (i.e., a finite mixture of Gaussian or Lorentzian distributions added to the continuum).

An example of a simple spectral model of this form is given in Figure 3.

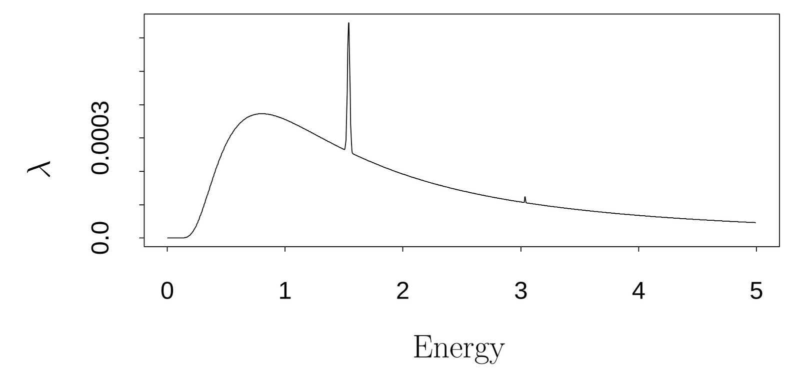

Figure 3: A Typical Spectral Model. The model includes a smooth continuum and two narrow Gaussian emission lines.

Distortion of the Data: Unfortunately, the photon counts reflect several layers of distortion in the data:

1. Photon absorption, stochastic partial non-homogeneous censoring;
2. Blurring of photons—energy and coordinates are recorded with error;
3. Photon pile up—if more than one photon arrives during the same time bin with the same sky coordinates, they are recorded as one photon with energy equal to the sum of their actual energies; and
4. Background contamination of the data.

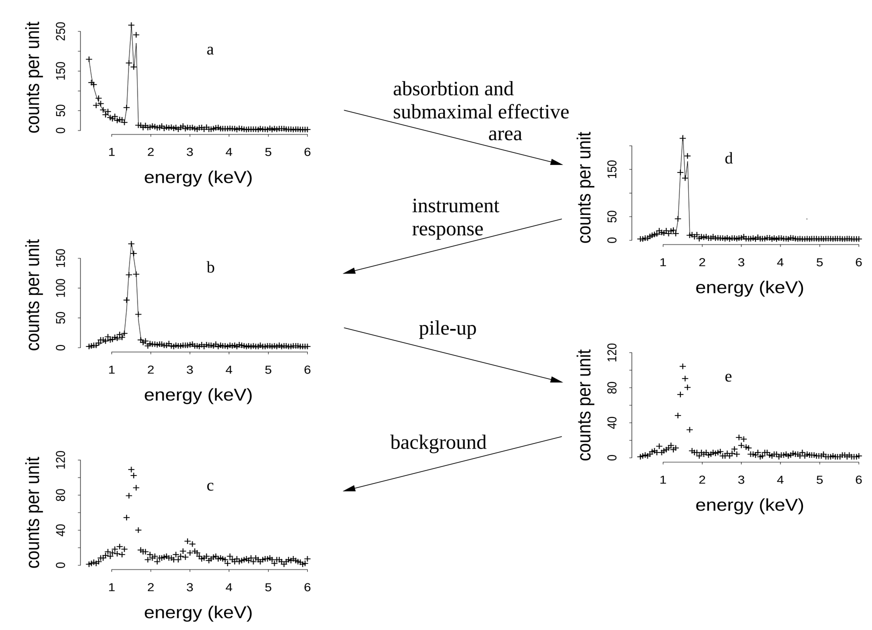

Figure 4: The Deterioration of Spectral Data.

BHiRXS: BHiRXS (Bayesian fitting of High Resolution X-ray Spectra) is a free statistical software that uses efficient EM-type algorithms, Markov chain Monte Carlo, and posterior predictive checks to:

- fit a variety of spectral models including several continuum models and allows for various added model components such as spectral lines,
- account for the data distortion processes described above,
- compute ppp-values to compare models of varying complexity, and
- preform model diagnostics such as residual plots based on the posterior predictive distribution.

Image Analysis

The Basic Image Model: Image reconstruction is similar to spectral analysis in that the basic model involves Poisson counts (but in a 2D array of pixels) and must account for similar data distortion processes. The image model, however, is far less structured than the spectral model. Instead of a highly parameterized model we use a multi-scale wavelet-like smoothing prior distribution on the image.

An Example: Figure 5 shows optical and X-ray images of NGC 6240, a galaxy that is the product of the collision of two smaller galaxies. The X-ray image is produced using Gaussian smoothing which blurs details but highlights the two bright black holes near the galaxy's center.

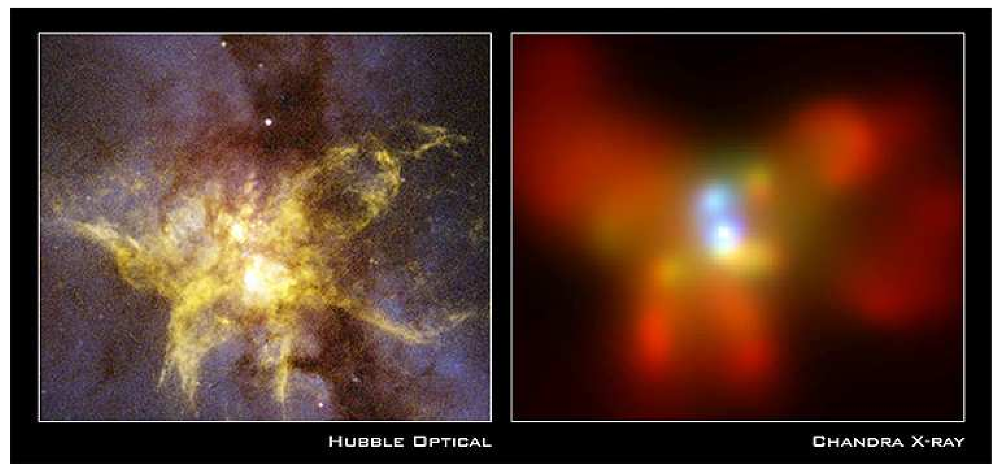

Figure 5: Optical and X-ray images of NGC 6240.

The raw data ("original"), the posterior mean under our model-based method ("EMC2 image"), and two Maximum Likelihood ("R-L") images appear in Figure 6. The ML reconstructions become more grainy with more EM iterations. Thus, EM is stopped before convergence in practice.

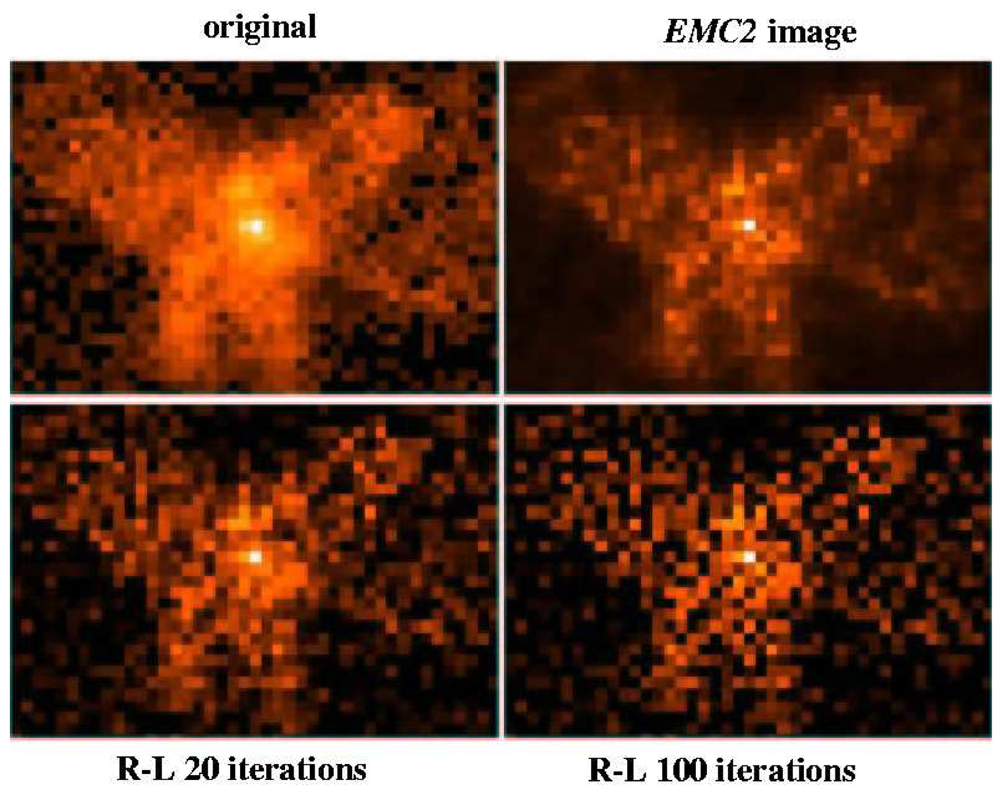

Figure 6: Model based reconstructions of the NGC 6240 X-ray image.

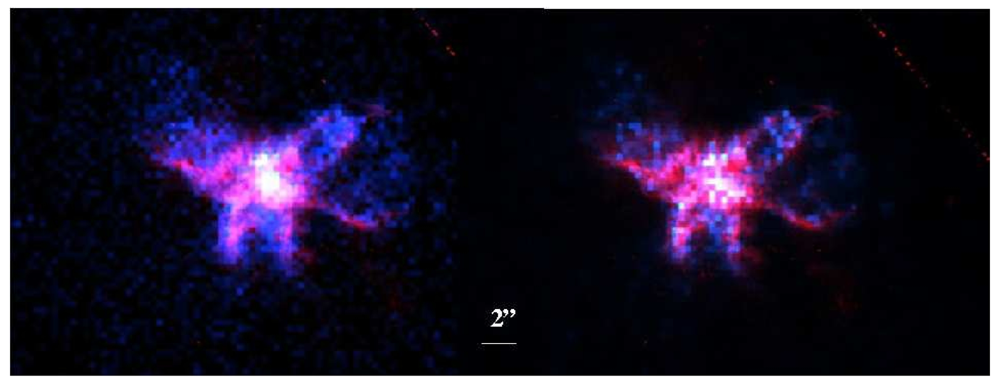

Figure 7: A comparison of the optical image (red) with the raw X-ray data (blue on left) and our reconstruction of the NGC 6240 X-ray image (blue or right).

We can use our MCMC simulation of the posterior distribution of the image to explore the variability of structures in the image. These are illustrated in the movies on the compute screen.

Exploring the Environment of a Stellar Corona

The Solar Corona: Figure 8 shows the a large area of sunspot activity in three wavelengths of light. The complex structure of the X-ray emission is a tracer of the temperature and density in the coronal plasma.

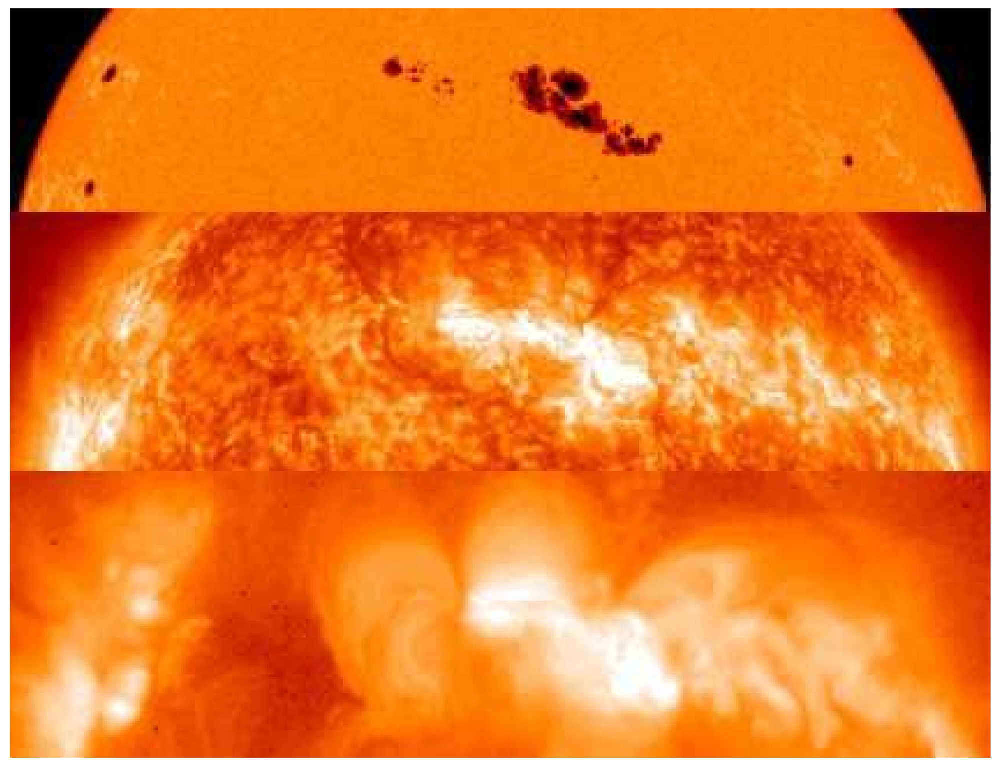

Figure 8: The Sun in visible, extreme ultra-violet, and X-ray wave lengths.

Stellar Coronae: Although no star except the sun can be imaged, ultra high-resolution X-ray spectral data can reveal much about the environment of a stellar corona. Figure 9 illustrates the forest of spectral lines in the X-ray spectrum of the star *Capella*. Photons are emitted when electrons fall to lower energy shells of an ion in the coronal plasma. Because the energy differences are particular to ions and the shell transitions, the energy of the emitted photons can identify the ions and the temperature of the plasma.

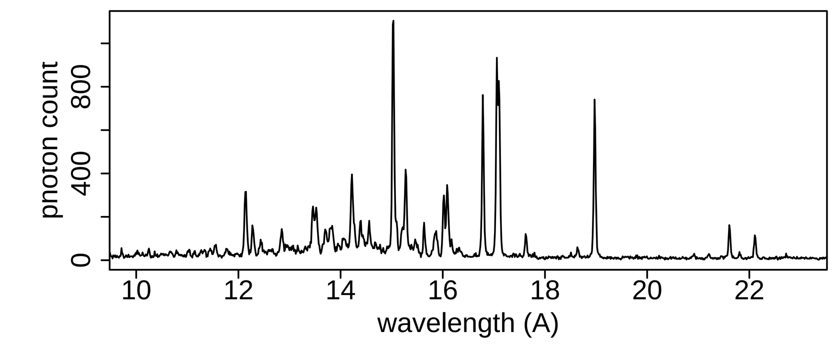

Figure 9: The forest of spectral lines of the star *Capella* in a narrow range of X-ray wavelengths.

Decoding the stellar spectrum requires careful analysis that combines ultra high-resolution spectral data with detailed quantum physical calculations. A Bayesian analysis of the resulting highly structured model can reveal the the temperature distribution of the stellar corona and the relative abundance of ions in the corona, see Figure 10.

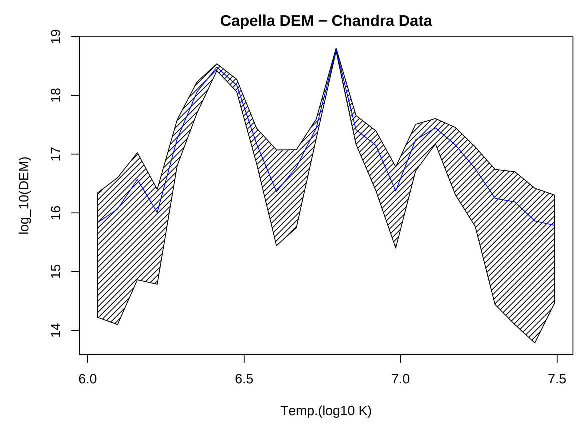

Figure 10: The log p.d.f. of the temperature of the plasma in the corona of Capella. The shaded area represents point wise 95% posterior intervals.