School on

X-rays from Star Forming Regions

SPECTRAL AND VARIABILITY ANALYSIS OF X-RAY SOURCES

Antonio Maggio
Istituto Nazionale di Astrofisica
Osservatorio Astronomico di Palermo

Observatories and Instruments

- Several instruments available in Chandra and XMM-Newton

INSTRUMENT		SPACE OBSERVATORY: Chandra			SPACE OBSERVATORY: XMM-Newton		
CLASS	TYPE	HRC	ACIS-I ACIS-S	LETG HETG	EPIC MOS (x2)	EPIC pn	RGS (x2)
Imaging		✓	✓		✓	✓	
Spectrometers	Non-dispersive		✓		✓	✓	
	Dispersive			✓			✓

- Dispersive spectrometers (gratings) employ one of the imaging devices as detector
- All detectors provide timing information

What stellar X-ray spectra tell us

- Optically-thin emission $\Rightarrow$ all photons escape directly
- Collision-dominated hot plasma (no photoionisation) $\Rightarrow$ X-ray luminosity $\propto$ square density $\times$ volume

$$L_X \propto N_e N_H \times V$$

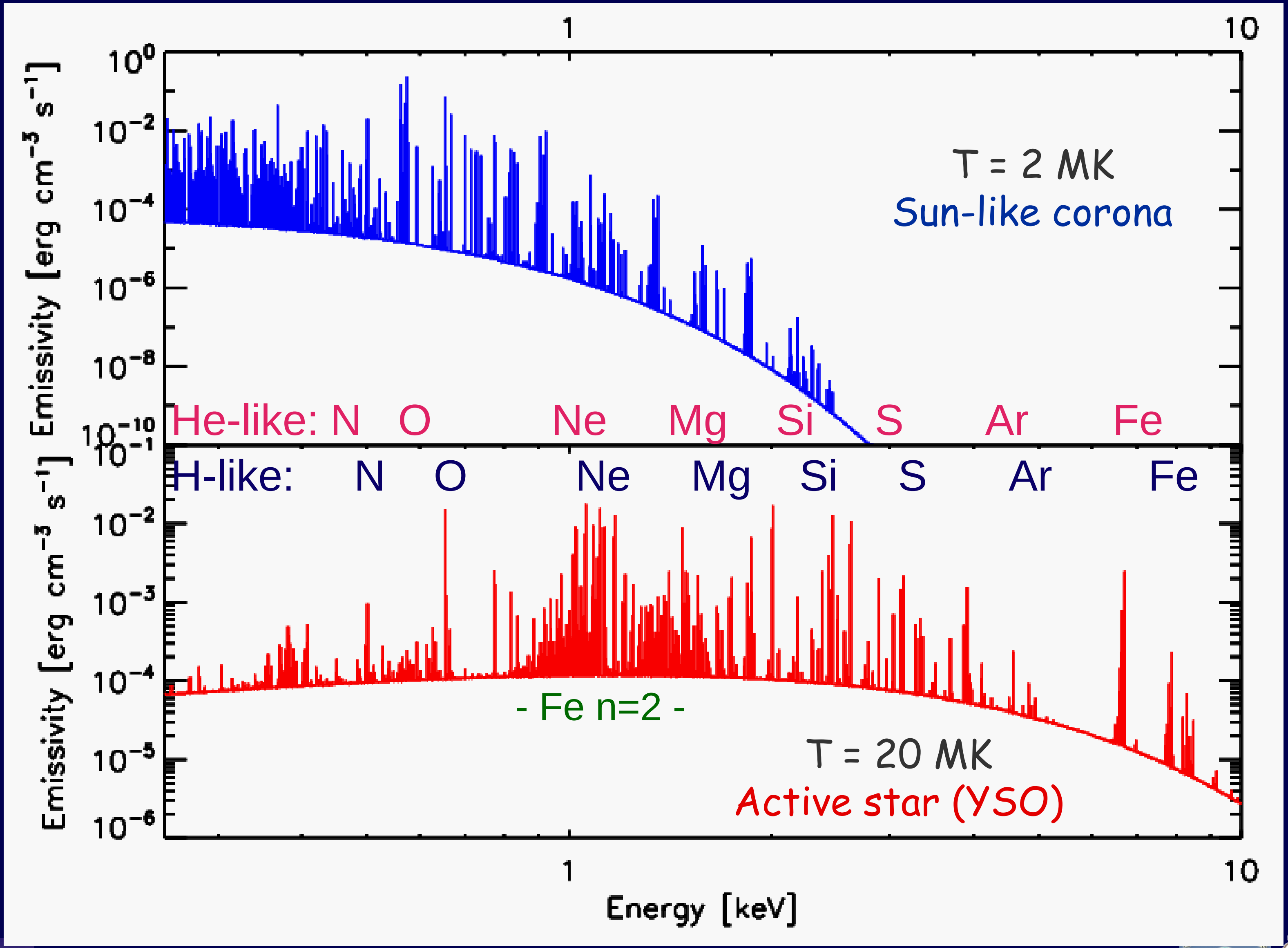
- Free electrons interacting with partially and fully ionized atoms
  - ➢ free-free emission (bremsstrahlung)
  - ➢ free-bound recombination
  - ➢ electron impact excitation and bound-bound transitions

Continua

Lines

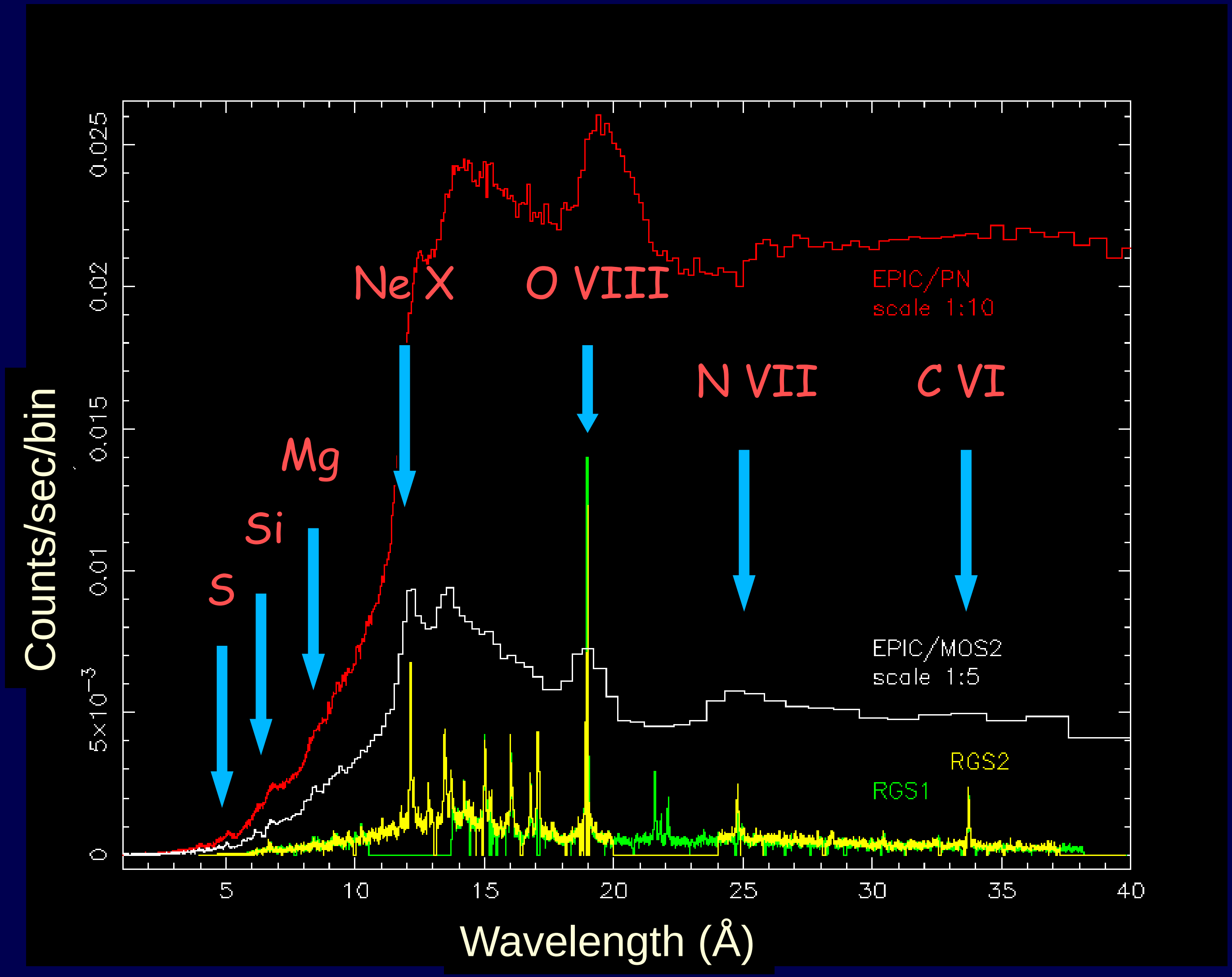
- The actual spectrum depends on the plasma temperature and element abundances

Stellar X-ray Spectra



constellation

Medium- and high-resolution spectroscopy with XMM-Newton



Instruments:

3 CCD detectors (EPIC)

2 reflection grating spectrometers (RGS)

AB Dor, young, active K1 V star. Calibration target

Basics of CCD X-ray spectra

- X-ray data are <u>Poissonian</u>
- Observed source counts are distributed in detector spatial coordinates, <u>s</u>, and energy channels, **PH**
- For a <u>point source</u> with a flux depending on energy, **E**, and time, **t**

Observed photon distribution = $\int dt\, d\underline{s}$ **Instrument Transfer Function** $\otimes$ **Source Flux**

$$C(PH) = \int dt \int d\underline{s} \int dE \quad T(\underline{s}, PH, E) \otimes S(E, t)$$

Integrated in time | **Integrated in detector space** | **Convolved with**

Forward-modeling approach

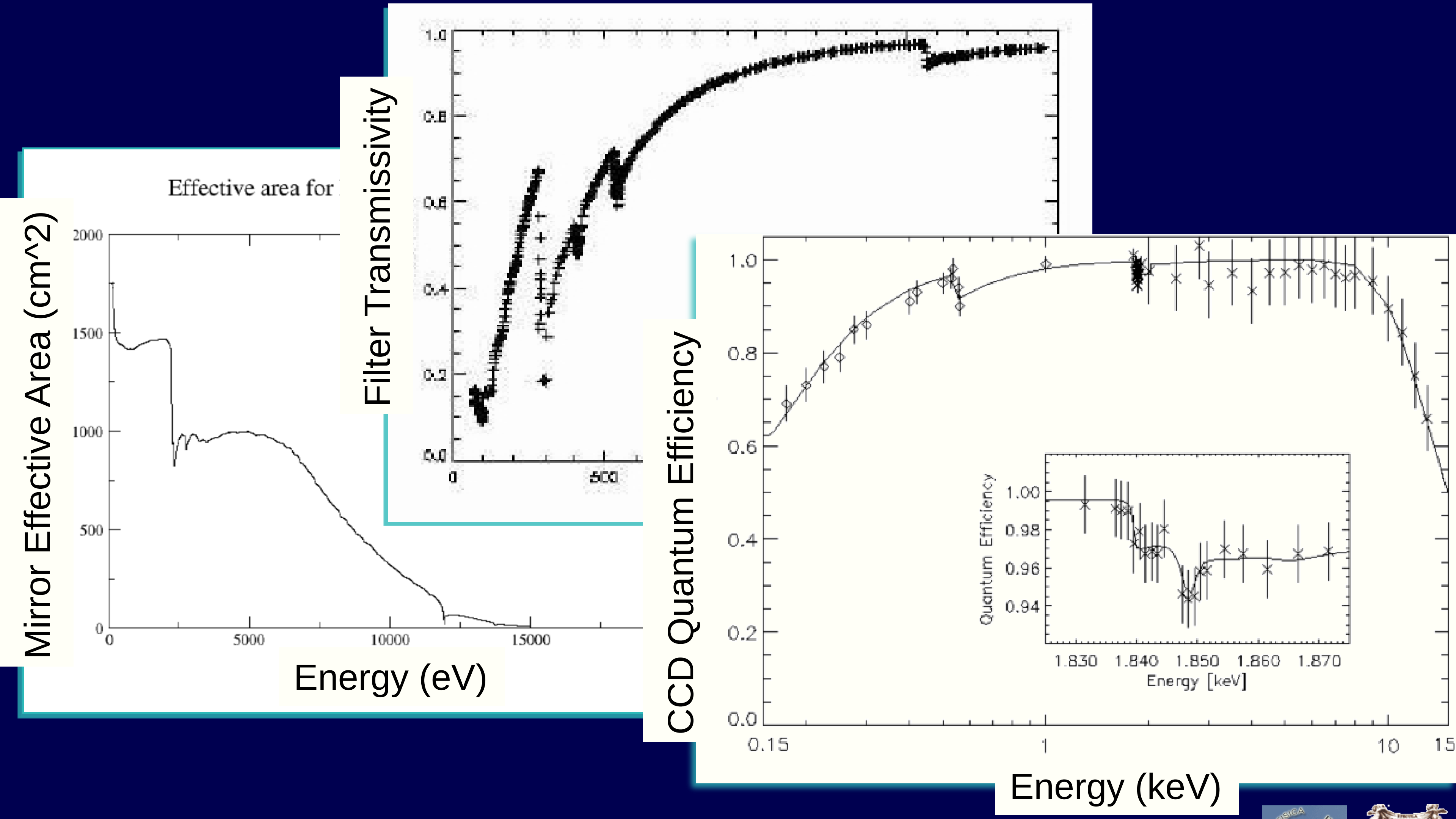
- The aim of the spectral analysis is to derive properties of the X-ray emitting plasma from the science data
- If the Transfer function is not diagonal (as usually happens) the above equation cannot be inverted
- Alternative: assume a physical model, convolve it with the instrument response, and compare the predicted spectrum with the observation, using an appropriate statistical indicator of goodness-of-fit quality

Instrument Response

- The instrument transfer function, T(<u>s</u>, PH, E), describes
  - How many photons are collected by optics+detector (Effective Area)
  - The spatial distribution of the events in detector coordinates, <u>s</u> (Point Spread Function)
  - The probability that a photon of energy E triggers an event in energy channel PH (spectral Response Matrix)

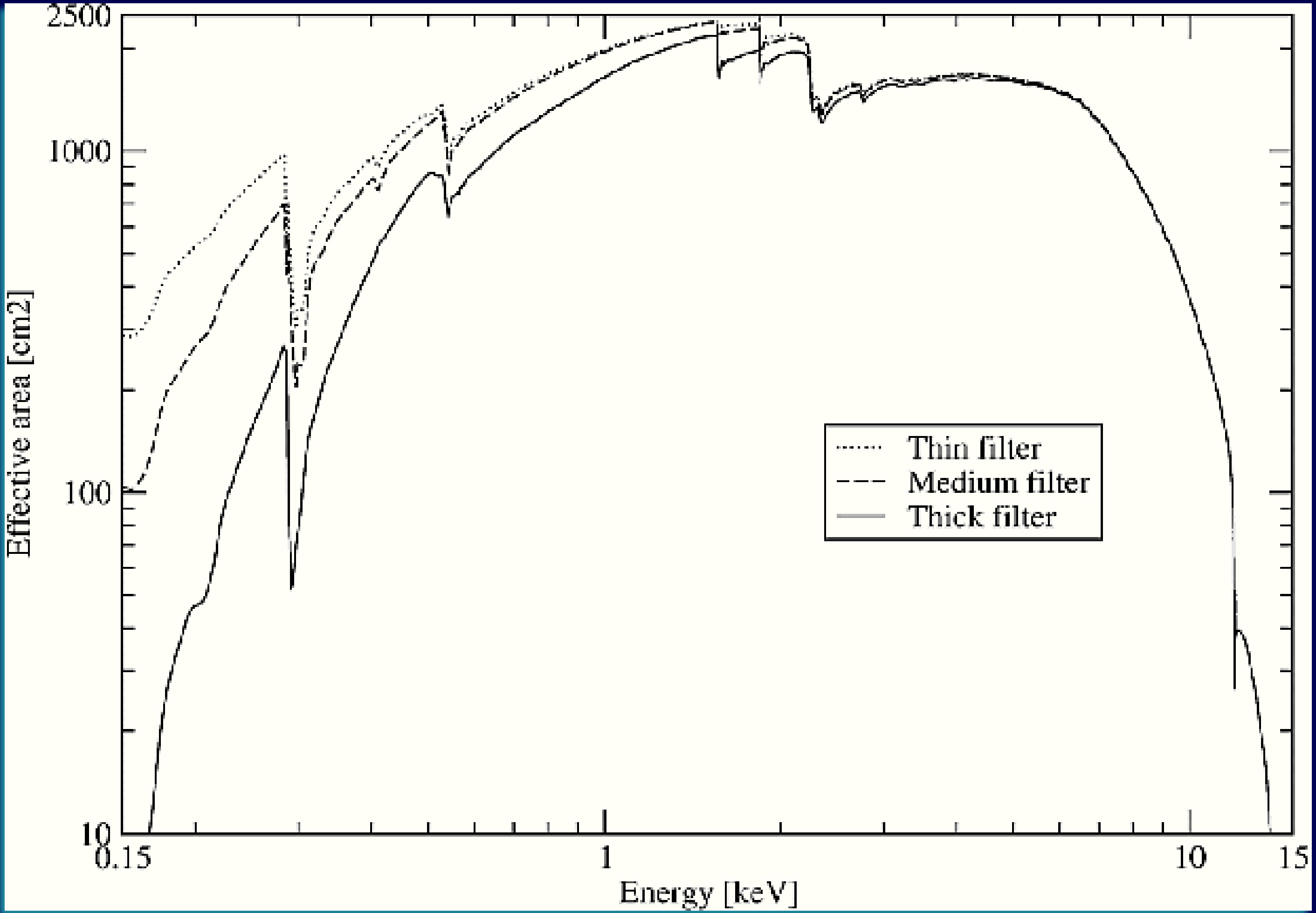
Effective Area: energy dependence

- Three contributions: mirror, filter, CCD



Effective Area: energy dependence

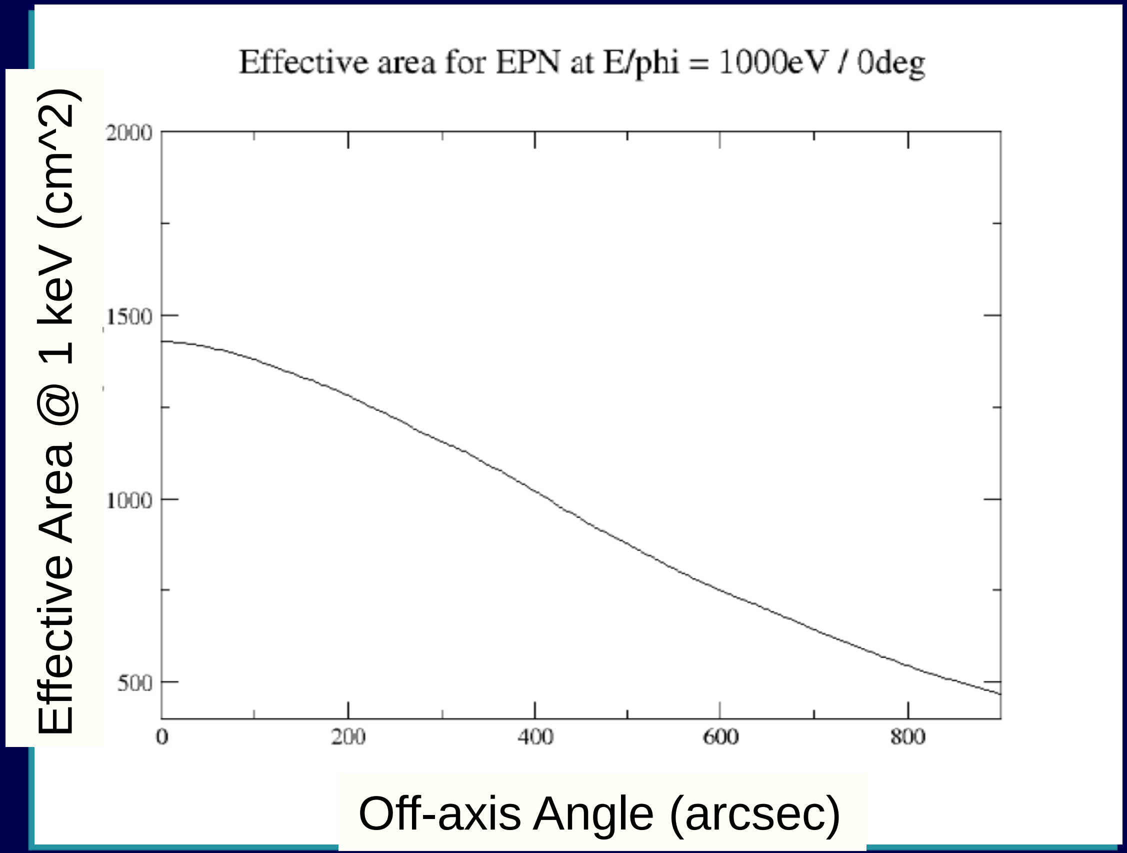
- For the XMM-Newton EPIC-pn, mirror × filter(s) × QE(CCD) =



Then, we need to correct also for CCD gaps, bad pixels and offset columns

Effective Area: spatial dependence

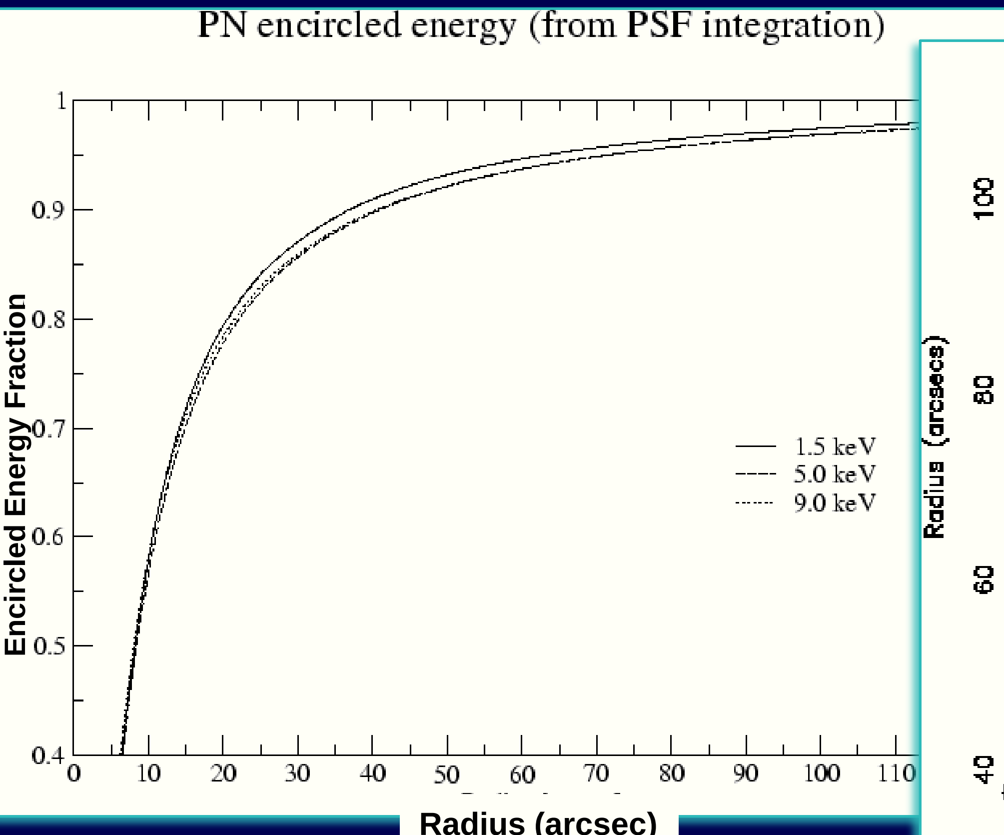
- Vignetting effect of the optics: the Effective Area decreases for increasing off-axis angle.

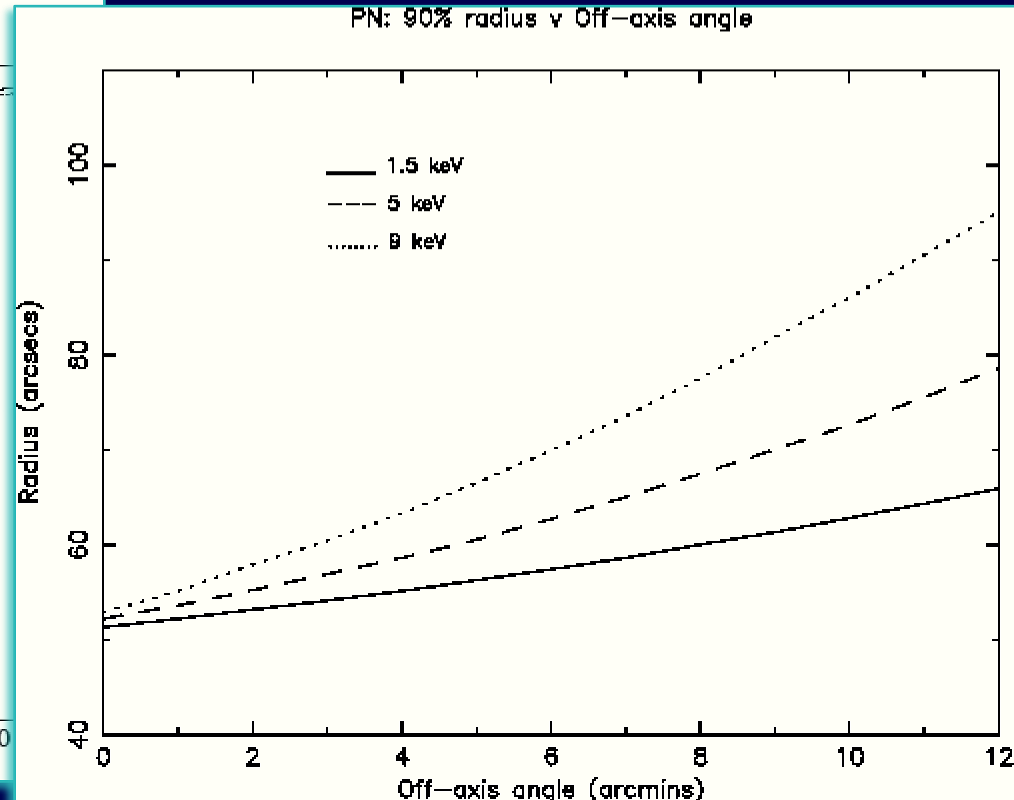


- Note that EPIC/MOS also has an azimutal variation due to the gratings on the light path

Encircled Energy Fraction

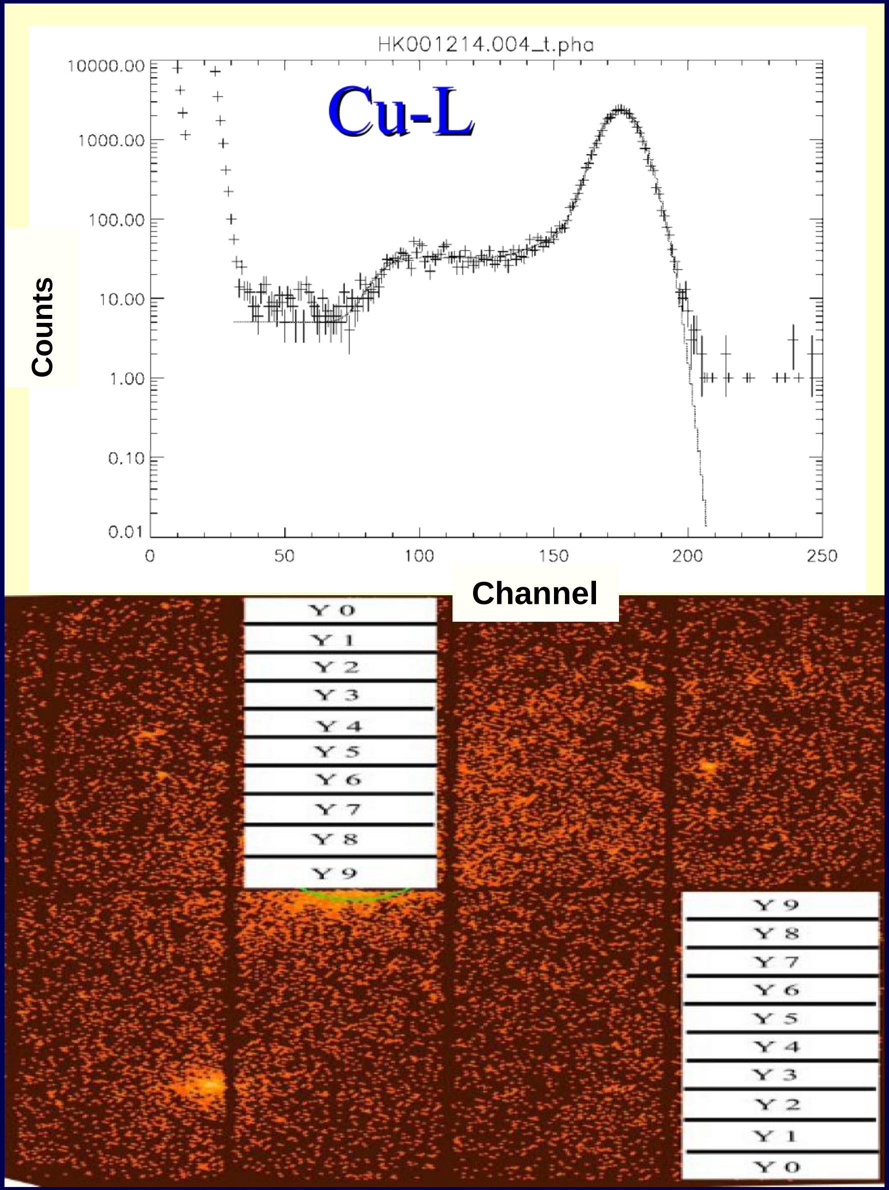
- Fraction of source counts collected in a finite source region (circle), as determined by integration of the PSF





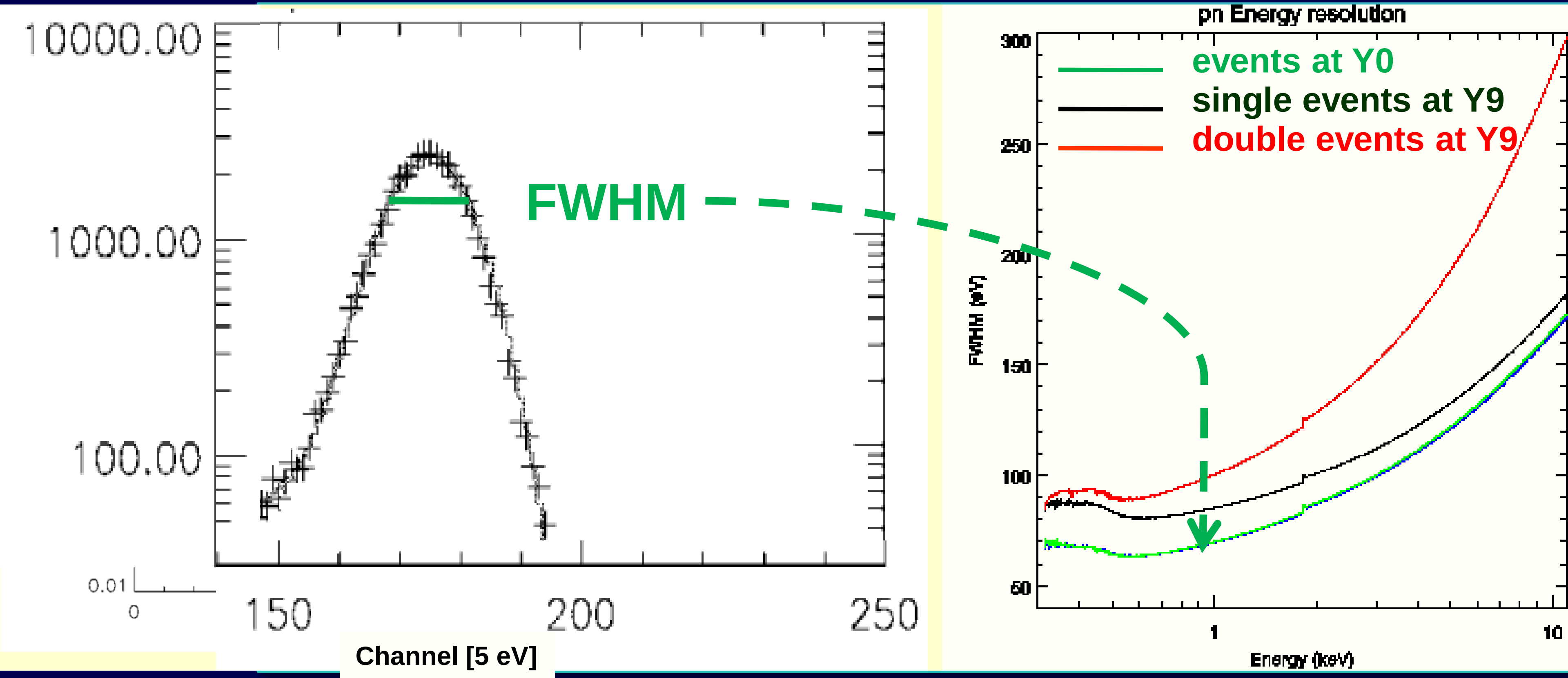
- In general, this correction depends on energy and off–axis angle

Spectral Response Matrix



- Statistical description of the distribution of events in different instrument energy channels, PH, corresponding to source photons with any fixed energy value, E
- It depends on the frame rate and on the position of the source in the FoV (distance from the read-out node) because of the Charge Transfer Efficiency

Spectral Resolution



- In CCD detectors, it is slightly dependent on energy (for EPIC-pn, less than a factor ~3 between 0.3-10 keV), but it depends also on the event type (pattern) and source position due to Charge Transfer Efficiency

Data products for spectral analysis

- **Effective area**: transfer function of optics+detector as a function of energy (includes correction for the Encircled Energy Fraction, bad pixels, etc.)
- **Redistribution matrix**: probability that a photon of a given energy is registered in a given channel
- **Source spectrum**: number of photons collected in a suitable source region, binned in energy
- **Background spectrum**: number of photons collected in a suitable background region, binned in energy
- **Exposure time**: effective integration time related to the actual source and background regions in the FoV

Background subtraction

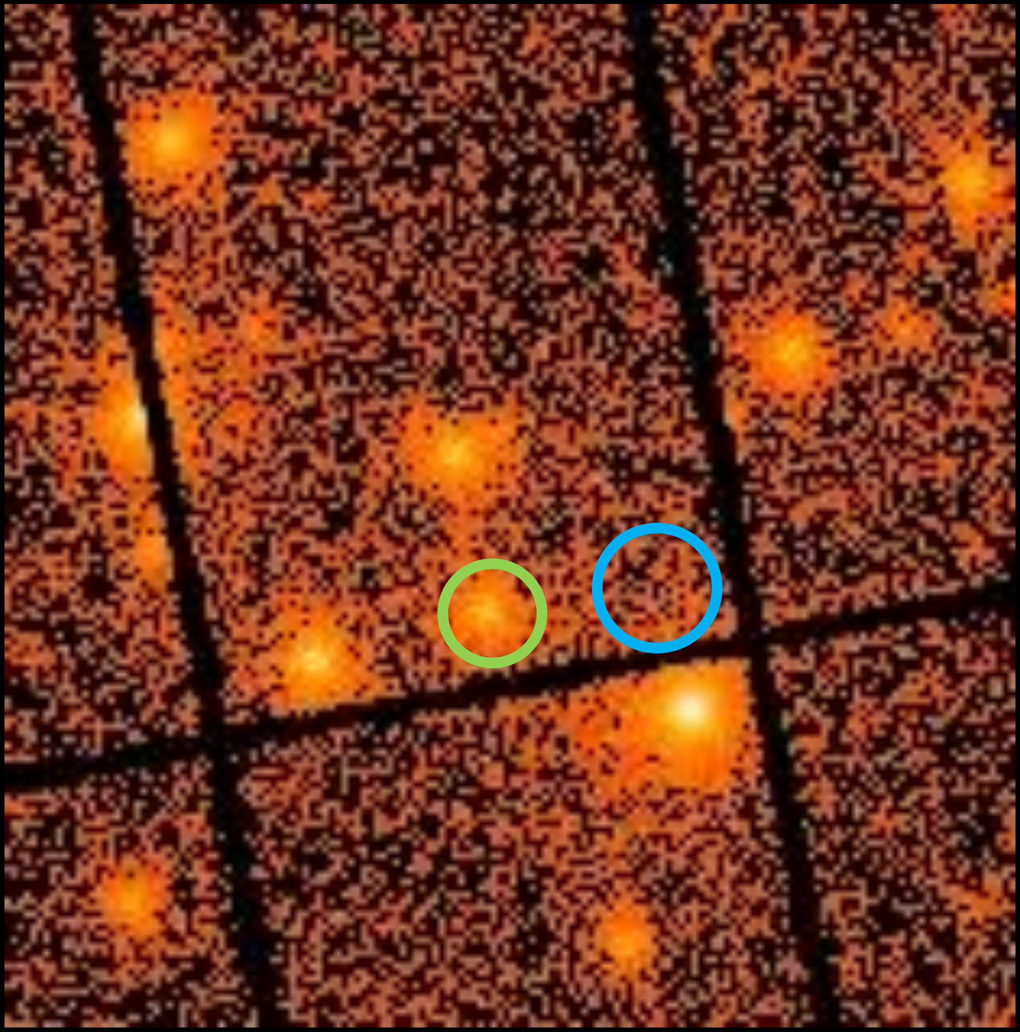
- Counts collected in the source region, **S(PH)**, include also background (bkg) emission
- A separate bkg spectrum, **B(PH)**, is required
- Two approaches are possible:
  a) subtract the background spectrum from the source+bkg spectrum, to get a “net” bkg-subtracted spectrum on which spectral analysis will be performed

  $$C(PH) = S(PH)/T_{SRC} - A_{SRC}/A_{BKG} \cdot B(PH)/T_{BKG}$$

  where T_{SRC} and T_{BKG} are the source and bkg exposure times, A_{SRC} and A_{BKG} are the source bkg region areas

  b) perform a simultaneous spectral analysis of S(PH) and B(PH) with independent models describing the source and the bkg

Extracting spectra from event files



- Choose a source region (usually a circle) avoiding nearby sources
- Extract the list of events in that region with their energy information: they are source+background events
- Extract background counts in a nearby region
- Create appropriate instrument response file
- Fit your spectrum

Basics of spectral fitting

Load spectra and instrument response files → **Choose statistics and decide confidence threshold** → **Define a model** → **Fit the model (parametric hypothesis testing)** → **Good fit?**

NO
Try again

YES
Go on

Compute statistical confidence regions on model parameters

Optionally, try an alternative model → **Compare fit quality of competing models**

Statistical hypothesis testing

- **Goodness-of-fit test:** How well does the model describe my data ?
- Two tests are commonly used

$$\chi^2 = \Sigma [C(PH_i) - M(PH_i)]^2 / \sigma_i^2(PH_i)$$

  - Simple and well-known distribution: if the fit is good, $\chi^2/DoF \approx 1$, **where DoF is the number of Degrees of Freedom (= number of data points – number of free parameters)**
  - Requires that distribution of $C(PH_i)$ in each spectral channel is Gaussian ⇒ Data rebinning required ⇒ possible loss of spectral resolution
  - Requires that the estimate of the variance, $\sigma_i^2(PH_i)$ is uncorrelated with $C(PH_i)$ ⇒ not true for Poisson variates ⇒ Possible remedy: bins with equal S/N ratio

- **C-statistics** (Cash 1979):

$$C = 2\Sigma[m(PH_i) - S(PH_i) \times \log(m(PH_i)) + \log(S(PH_i)!)]$$

where $S(PH_i)$ are source+bkg counts and $m(PH_i)$ is the corresponding model

  - It can be used with low-count spectra without rebinning
  - It cannot be applied to background-subtracted spectra
  - It does not provide us with an absolute estimate of the quality of the fit

Spectral binning vs. resolution

- In general, instrument energy channels oversample the spectral resolution
  - Example: for EPIC-pn the spectral resolution at 1 keV is FWHM≃70 eV , while energy channels are 5 eV wide over the whole instrument bandpass

  ⇒ measurements in adjacent channels are correlated
- A sampling with a bin size ~FWHM/3 is sufficient to reconstruct the spectral characteristics of the source (see Nyquist-Shannon information theory)

  ⇒ You can safely rebin your spectrum up to FWHM/3
- To employ χ^2 statistics, a S/N ≥ 5 per bin is usually recommended, and data points should be independent
  - the higher the number of counts per bin, the less important are bias effects in parameter estimation

  ⇒ Rebinning is a must!

Accepting/rejecting a best-fit model

- Any quantity used for the statistical analysis is itself of statistical nature, hence it has an error
  - Example: $E(\chi^2) = \sqrt{(2DoF)} = \sqrt{(2(n - \nu))}$
    (r.m.s. in the asymptotic limit $n \to \infty$)
  - A correct model yields $\chi^2 \approx n - \nu$ with a r.m.s. $E(\chi^2)$
  - At any best-fit χ^2 value corresponds a probability, P, that the model is acceptable (confidence level)
  - Due to systematic errors in the data calibration or in the model, you might not get very large values of P
- To reject a wrong model we ask that

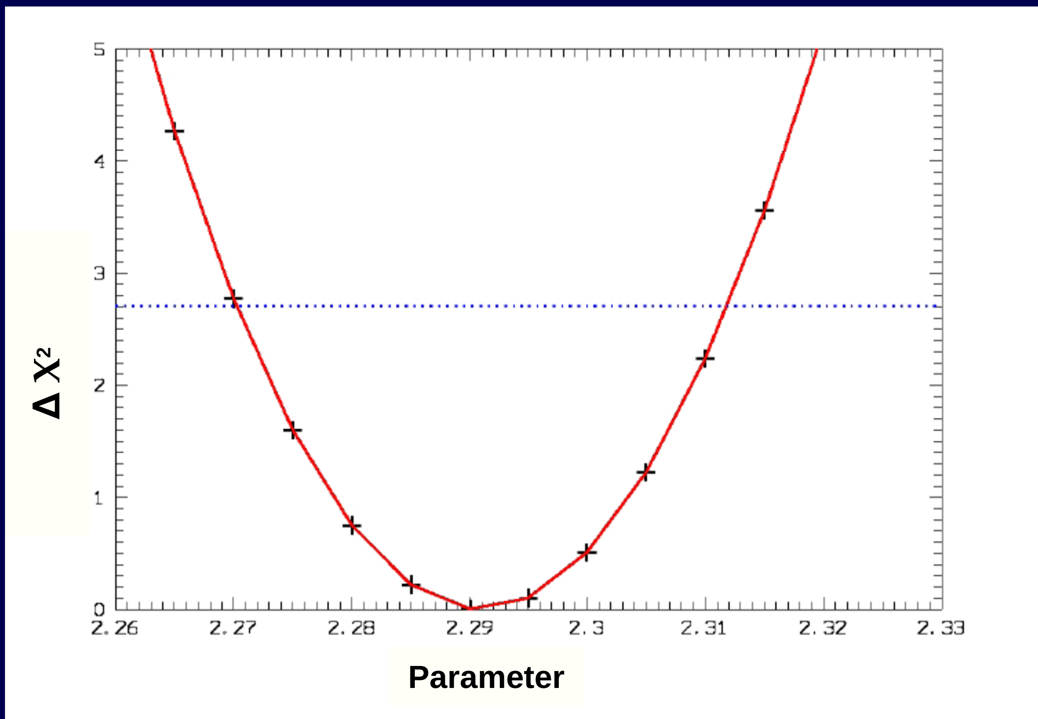
  $$\chi^2 > n - \nu + \sigma \times \sqrt{(2(n - \nu))}$$

  where σ is related to the confidence level we choose
  - If the model is wrong
    $\chi^2 \approx n - \nu + N$ where N is $\approx$ total number of events

  $\Rightarrow$ in order to reject a wrong model $N > \sigma \times \sqrt{(2(n - \nu))}$ is required

Statistical uncertainties

- Method: search for the values of the parameter which yield an increase of the χ^2 , corresponding to a certain confidence level, with respect to the best-fit model (Lampton et al. 1976)
  - Example: the 90% confidence level of one interesting parameter corresponds to $\Delta\chi^2 = 2.71$, hence search for which values of the parameter p $\chi^2 = \chi^2_{min} + 2.71$



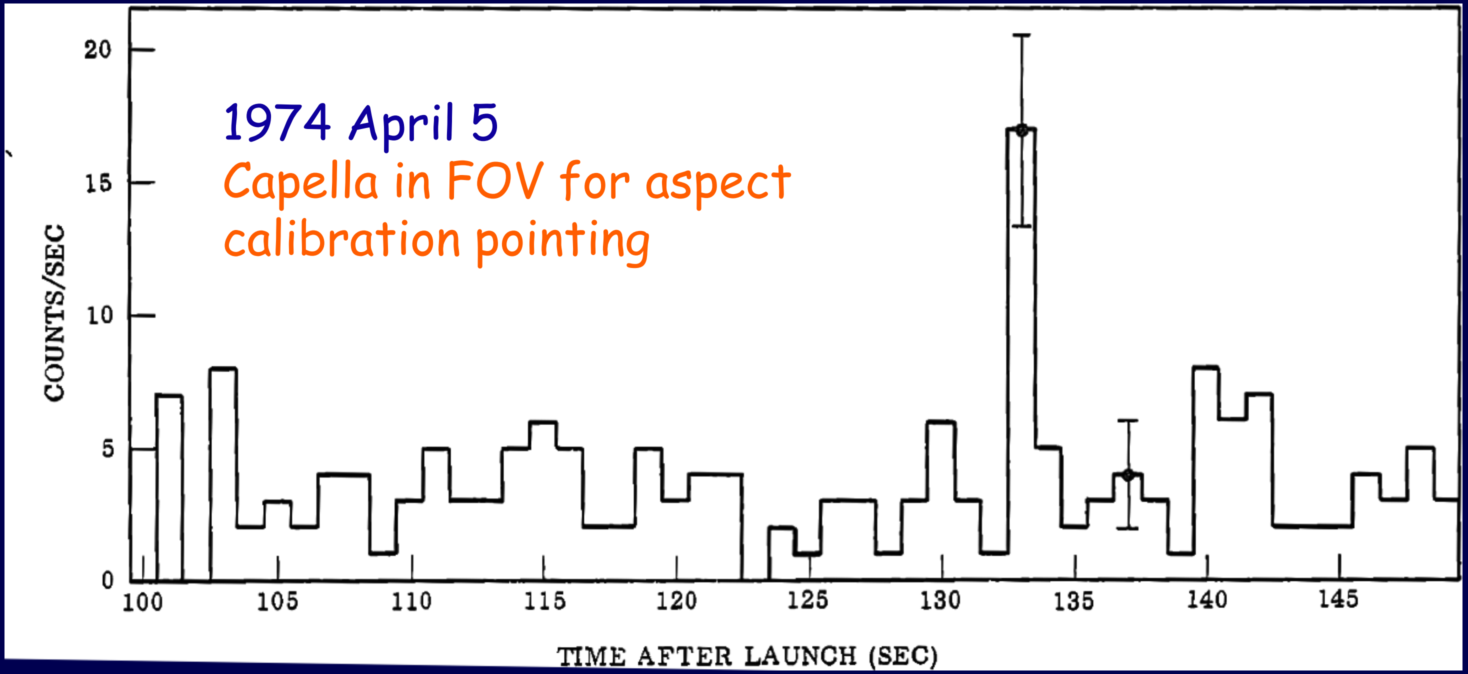
Confidence Level	Interesting parameters 1	2	3
68%	1.00	2.30	3.50
90%	2.71	4.61	6.25
99%	6.63	9.21	11.30

Statistical uncertainties

- In general χ^2 (and hence $\Delta\chi^2$) is not a smooth function of the parameter p with a single minimum
  - Local minima of χ^2 in the domain of the model parameter are possibly present
  - ⇒ Repeat your fit with different initial values of p
  - ⇒ plot $\Delta\chi^2$ as a function of the parameter
  - the interesting parameter is often not independent from other parameters
  - ⇒ plot two-dimensional $\Delta\chi^2$ contour maps

Discovery of stellar coronal emission

beside the solar one (Catura, Acton, Johnson 1975, ApJ 196, L47)



- Capella (α Aurigae) detected for 1.2 sec (22 photons) with a X-ray detector (0.2-1.6 keV band) during a rocket flight

The first X-ray spectrum of a stellar corona and its interpretation

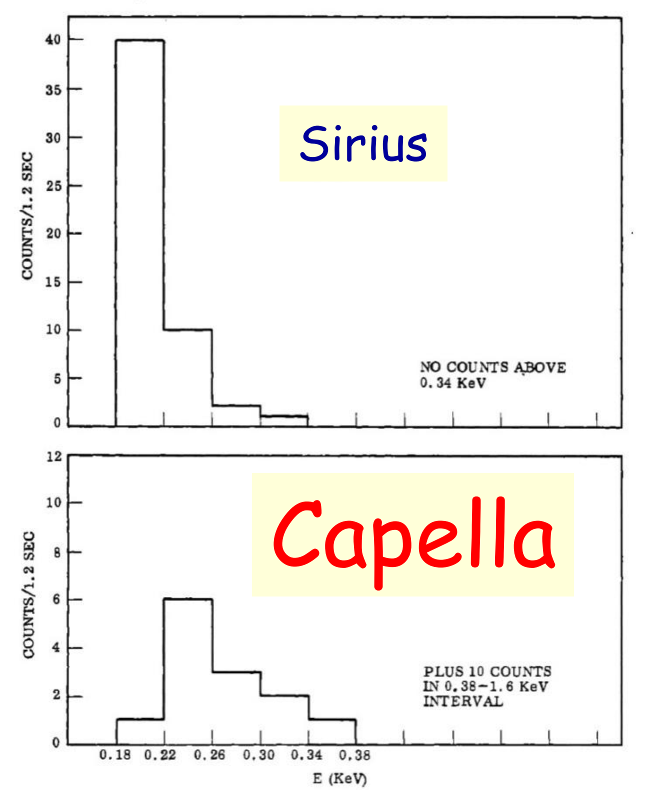


FIG. 2.—Energy spectra of pulses observed when the rocket was pointed at Sirius and Capella.

- Detected signal not due to photospheric UV radiation
- Thermal bremsstrahlung model yields $T = 8^{+7}_{-3} \times 10^6$ K
- No indication of interstellar absorption ⇒ nearby source
- $L_X \sim 10^{31}$ erg s^{-1}
- Point-like source
- Not detected in previous observations ⇒ variable or transient source

New class of Galactic soft (E < 2 keV) X-ray sources!

Discovery of intense line emission

(Cash et al. 1978, ApJ 223, L21)

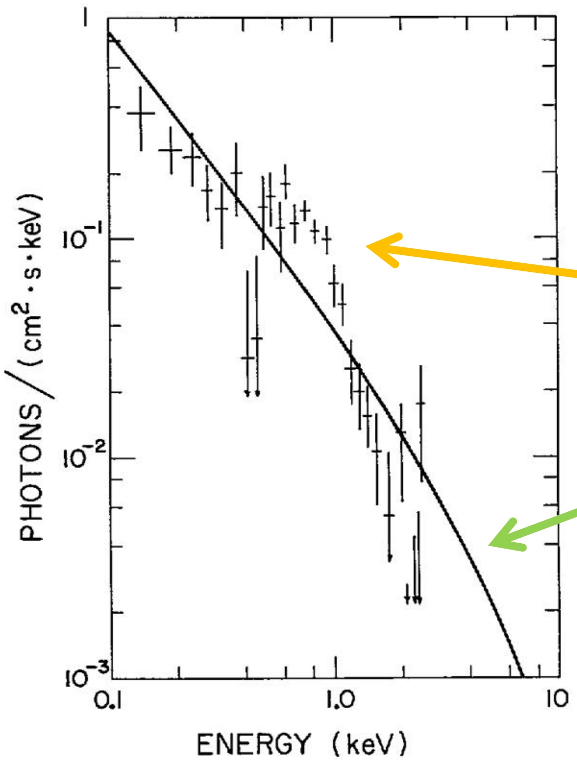


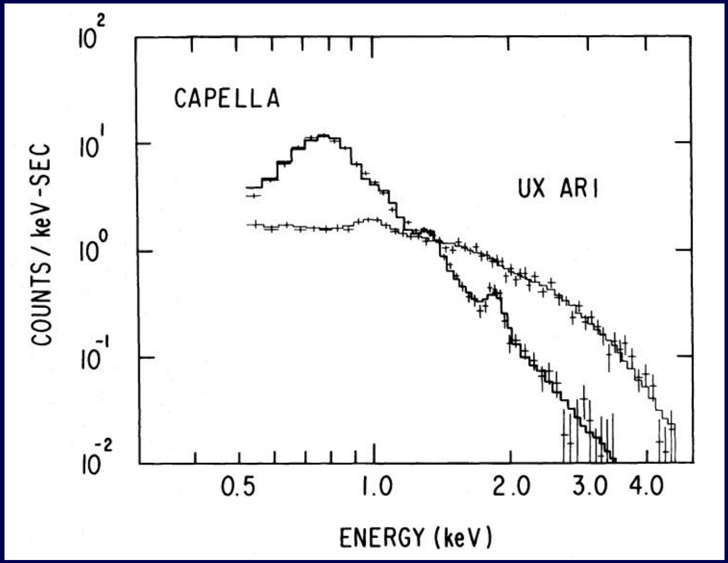
FIG. 1.—Soft X-ray spectrum of Capella. The solid line represents the best-fit simple bremsstrahlung spectrum ($T = 5.6 \times 10^7$). Line emission is clearly visible at 0.85 keV.

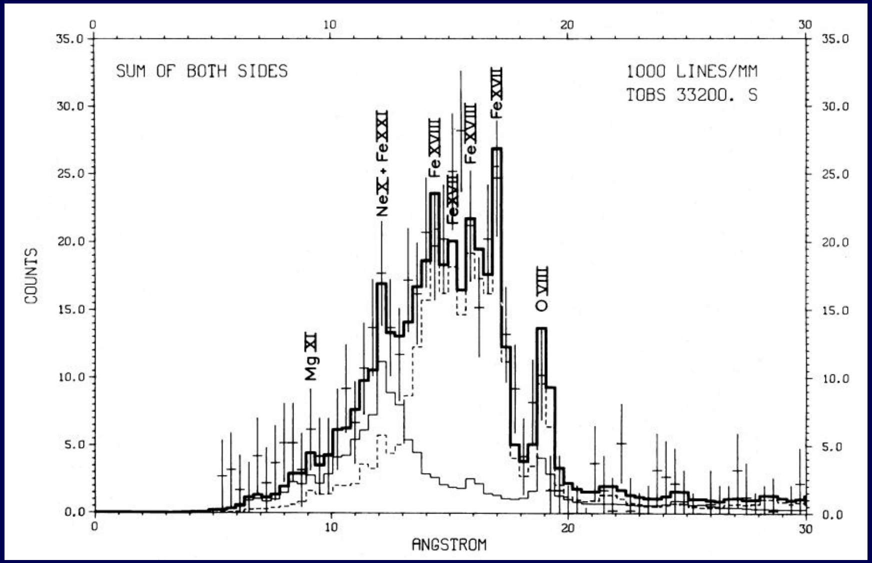
- Capella observed with HEAO-1 gas scintillation proportional counter
- **Emission excess between 0.65-1 keV** with respect to thermal bremsstrahlung model spectrum

$\Rightarrow$ Evidence of **thermal emission from optically-thin plasma in collisional equilibrium** at T $\approx 10^7$ K

Capella is 5 times hotter and 10^3 times more intense than the solar corona.

Discovery of intense line emission





- *Einstein* Observatory (HEAO-2)
  Solid State Spectrometer (SSS)
  ($\Delta E \sim 160$ eV, $E/\Delta E \sim 6$ at 1 keV)

⇒ Evidence of unresolved emission line complexes from Fe, Mg, Si, and S
(Swank et al. 1981, ApJ 246, 208)

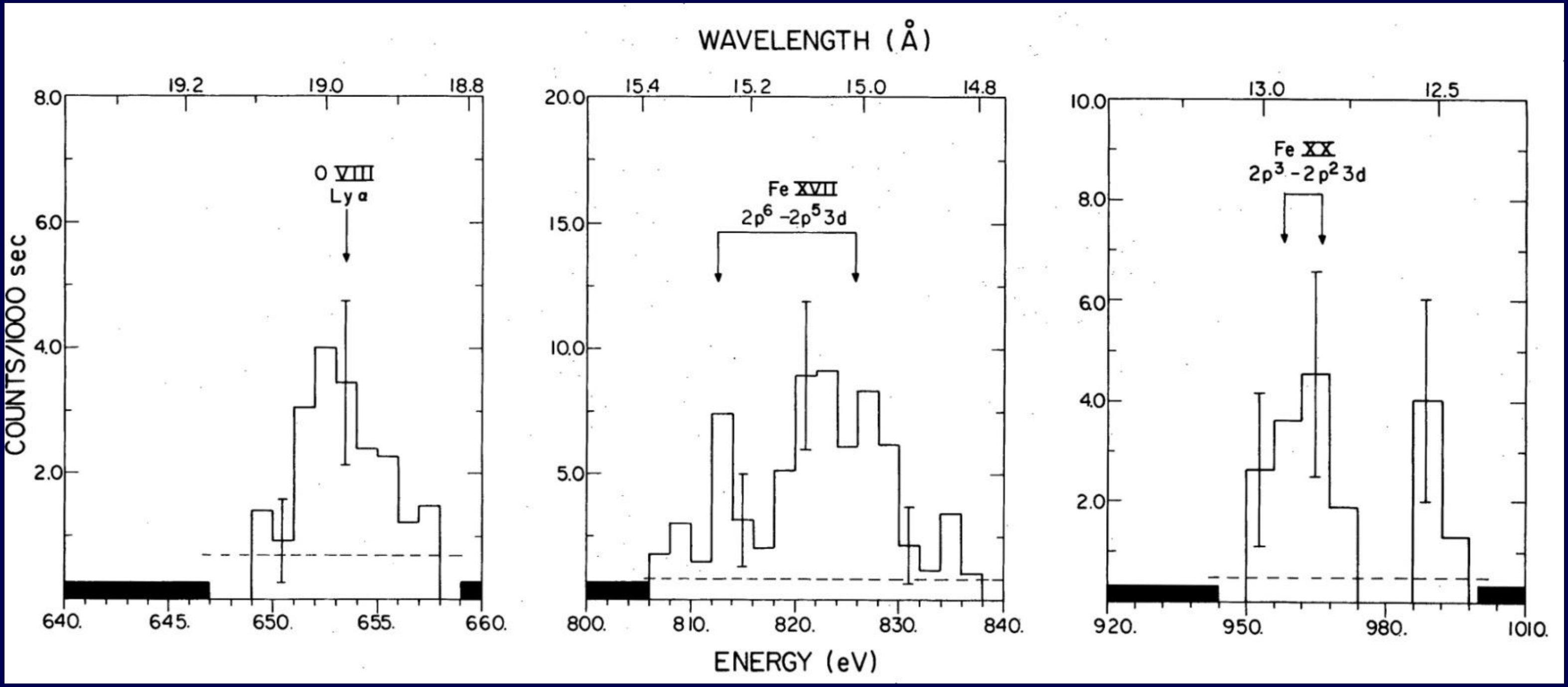
- *Einstein* Observatory
  Objective Grating Spectrometer
  ($\Delta\lambda \sim 1$ Å, 5-30 Å, $E/\Delta E \sim 12$ at 1 keV)

⇒ Line identifications

⇒ Thermal model with two discrete components
(Mewe et al. 1982, ApJ 260, 233)

First measurements of line intensities

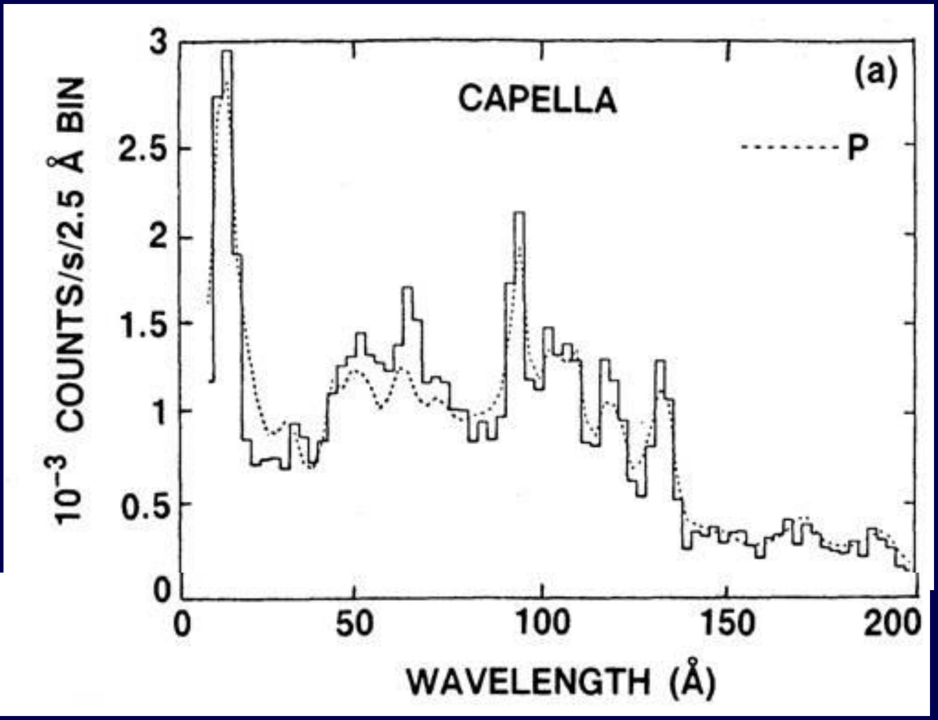
(Vedder & Canizares 1983, ApJ)



- *Einstein* Focal Crystal Spectrometer ($R = E/\Delta E \sim 30$)
- First attempts of emission measure analysis $\Rightarrow$ line intensities consistent with isothermal plasma at $T \sim 6 \times 10^6$ K or with emission measure distribution with peak at $T \sim 3 \times 10^6$ K

Detailed multi-temperature analysis

(Lemen et al. 1989)

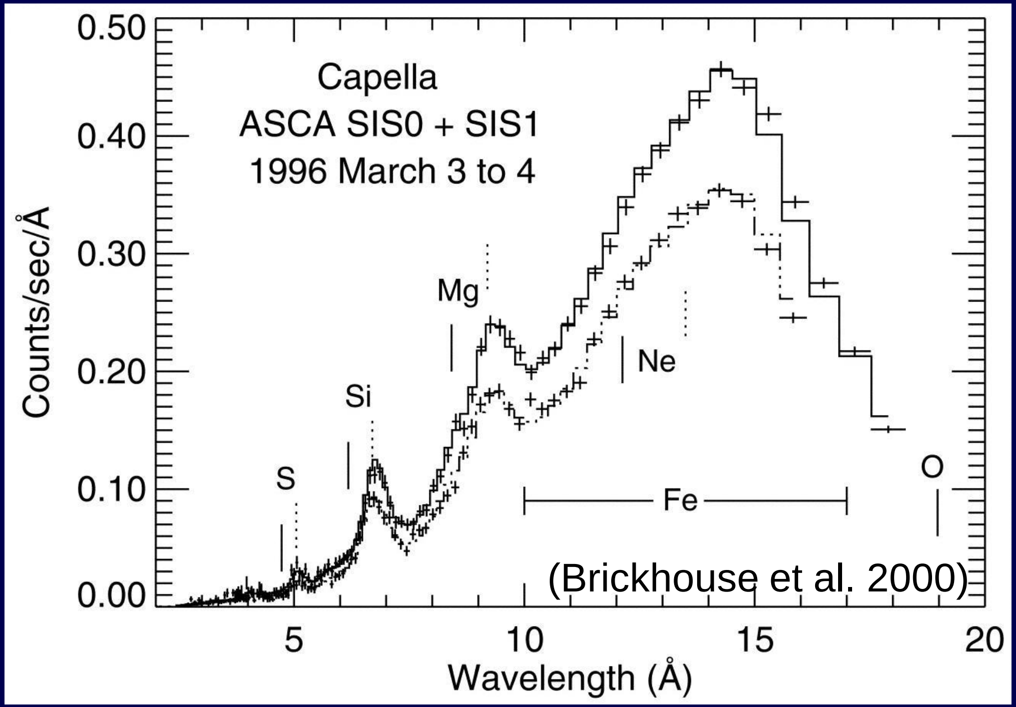


- Capella observed with EXOSAT Transmission Grating Spectrometer ($\Delta\lambda \sim$ 3Å, 10–200 Å range, R = 3–60)
- Results consistent with those of 2-T models

⇒ Coronae apparently dominated by plasma in two relatively narrow temperature intervals

⇒ Interpretation in terms of two classes of coronal magnetic structures

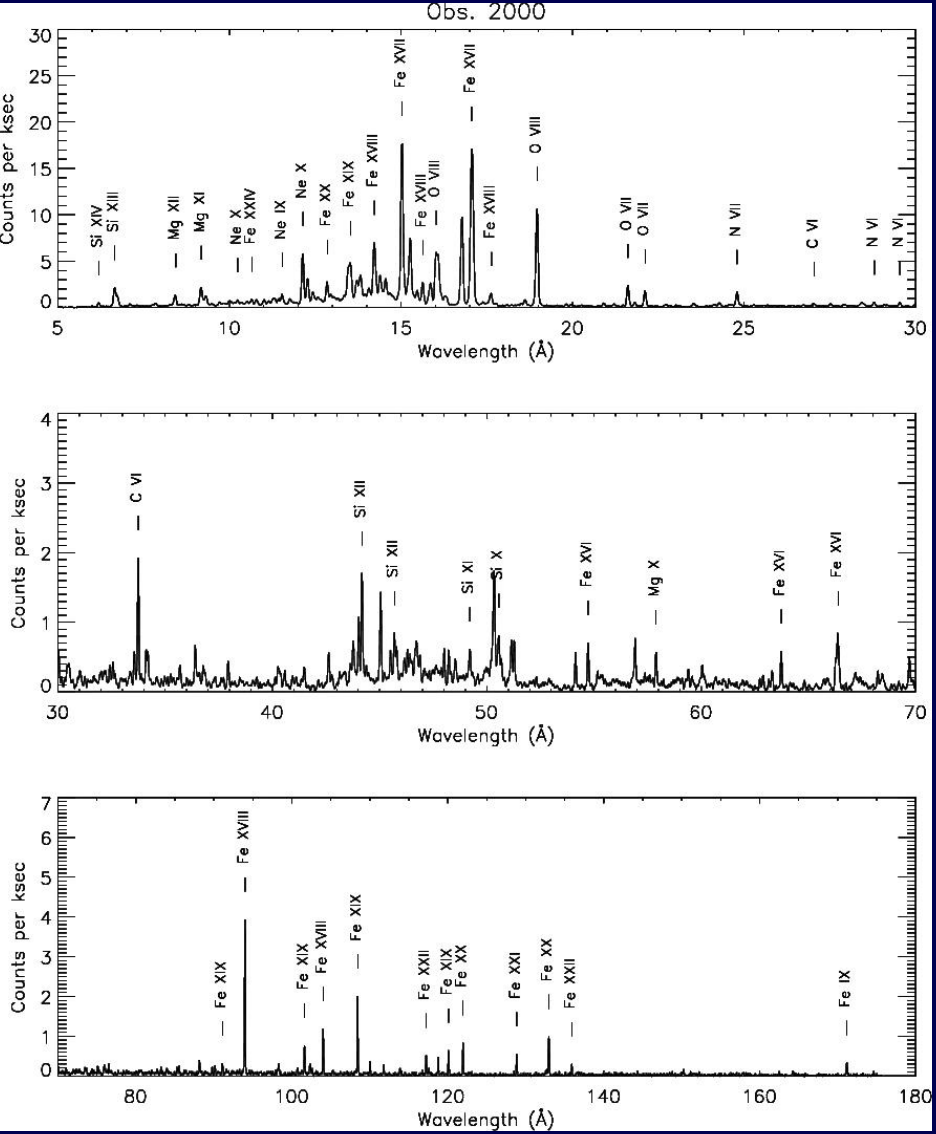
First abundance measurements



- **ASCA** Solid-state Imaging Spectrometer (CCD-based detector $E/\Delta E \sim 15$ at 1 keV, $E/\Delta E \sim 50$ at 6 keV)

- Line complexes due to O, N, Ne, Mg, Si, S, and Fe
- 1-, 2-, 3-component thermal models adopted + individual element abundances as free parameters
- $\Rightarrow$ Anomalous (non-solar) abundances found for most magnetically (i.e. X-ray luminous) active stars

High-resolution X-ray spectroscopy



(Argiroffi et al. 2003)

- Capella observed with <u>**Chandra**</u> Low-Energy Transmission Grating ($\Delta\lambda \sim$ 0.0125 Å, 5–170 Å)
- Several tens of emission lines identified and measured $\Rightarrow$ plasma emission measure vs. temperature
- $\Rightarrow$ element abundances
  $\Rightarrow$ plasma densities
  $\Rightarrow$ plasma dynamics

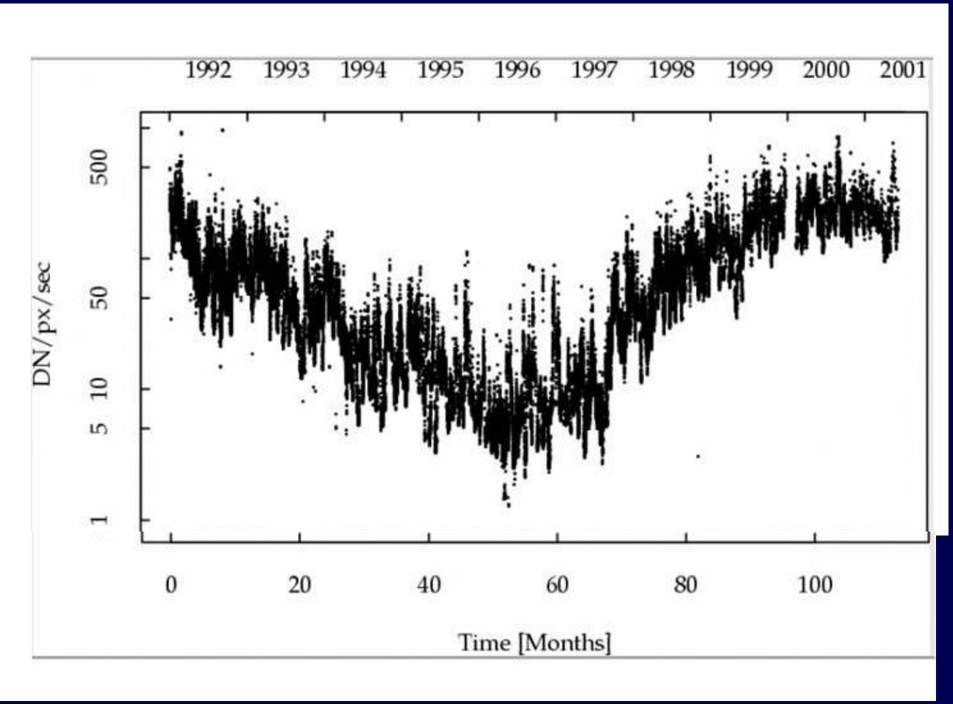
Variability analysis

- Time series obtained from X-ray observations have some peculiarities
  - Low-count statistics (data are Poissonian)
  - Time series are non-uniform:
    - periodic discontinuities due to spacecraft orbit
    - observation gaps between Good Time Intervals
- Variability analysis must cope with these peculiarities
- Stellar X-ray variability is non-periodic on typical observation lengths (10-100 ksec)

Stellar X-ray source variability

- Stellar coronal sources are known to vary on several time scales
  - Short-term (from minutes to a few days) variability due to flares
  - Medium-term variability (from a few hours to tens of days): rotational modulation
  - Long-term variability (years) due to magnetic cycles
- X-ray emission from YSOs may vary, at least in principle, also due to
  - Variable accretion rate
  - Absorption by the circumstellar disk

X-ray variability of the solar corona

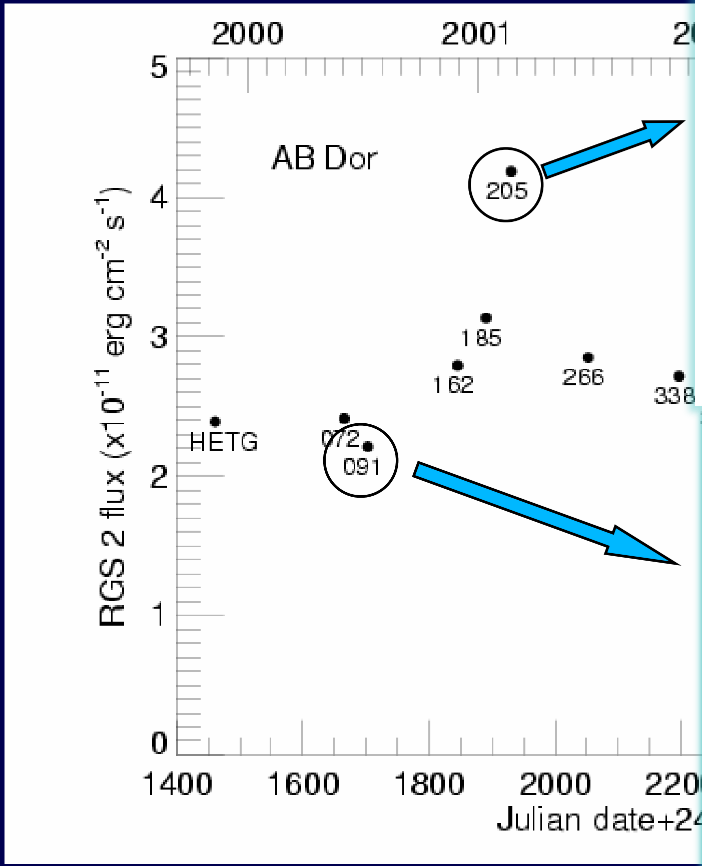


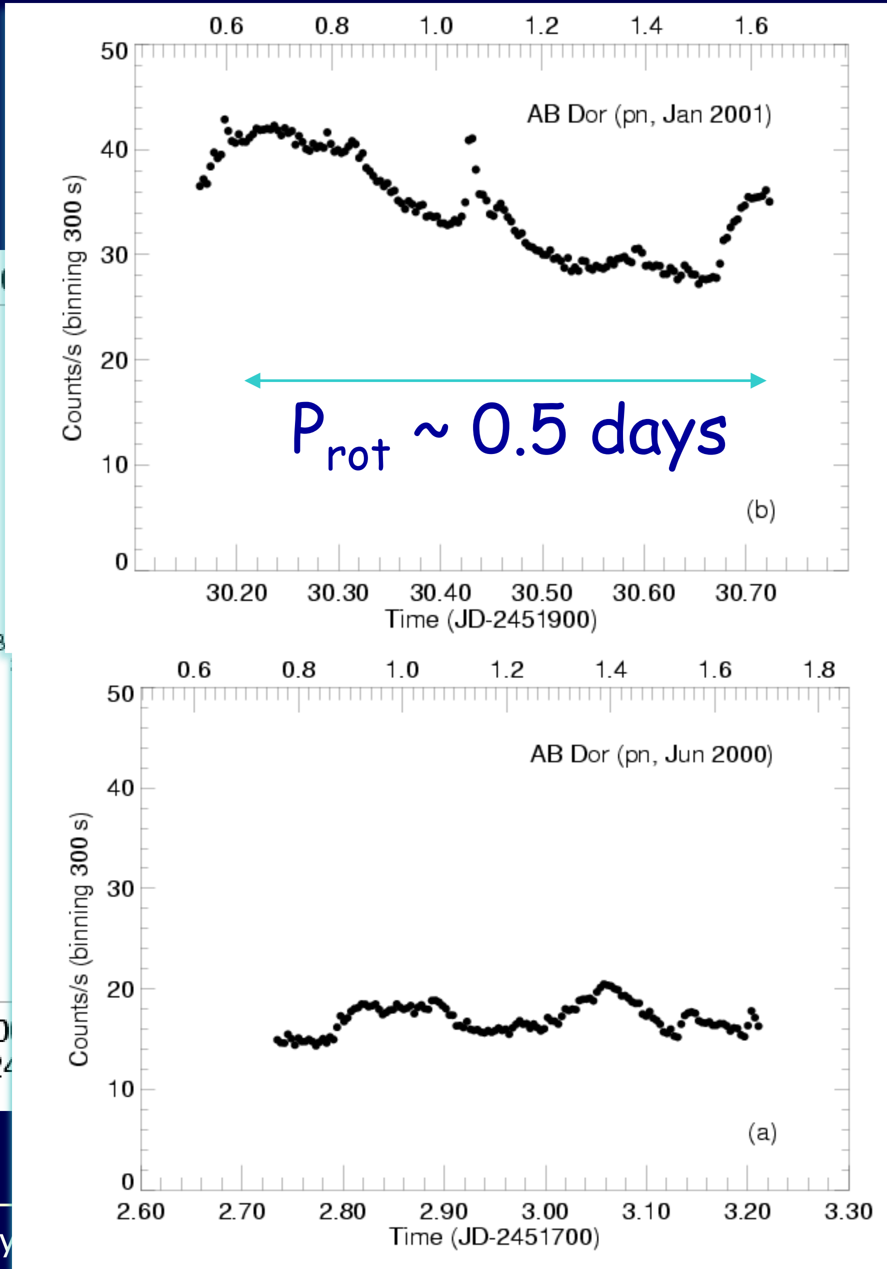
(Micela & Marino 2003)

- Solar X-ray emission observed with Yohkoh/SXT (0.7 – 2.5 keV band)
- Variability observed at several time scales:
  - ✓ flares (τ < 1 day)
  - ✓ rotational modulation ($\tau \sim$ 28 days)
  - ✓ magnetic cycle (11 years)

Stellar X-ray Variability

Do active stars exibit magnetic cycles, rotational modulation, and flares like the Sun ?





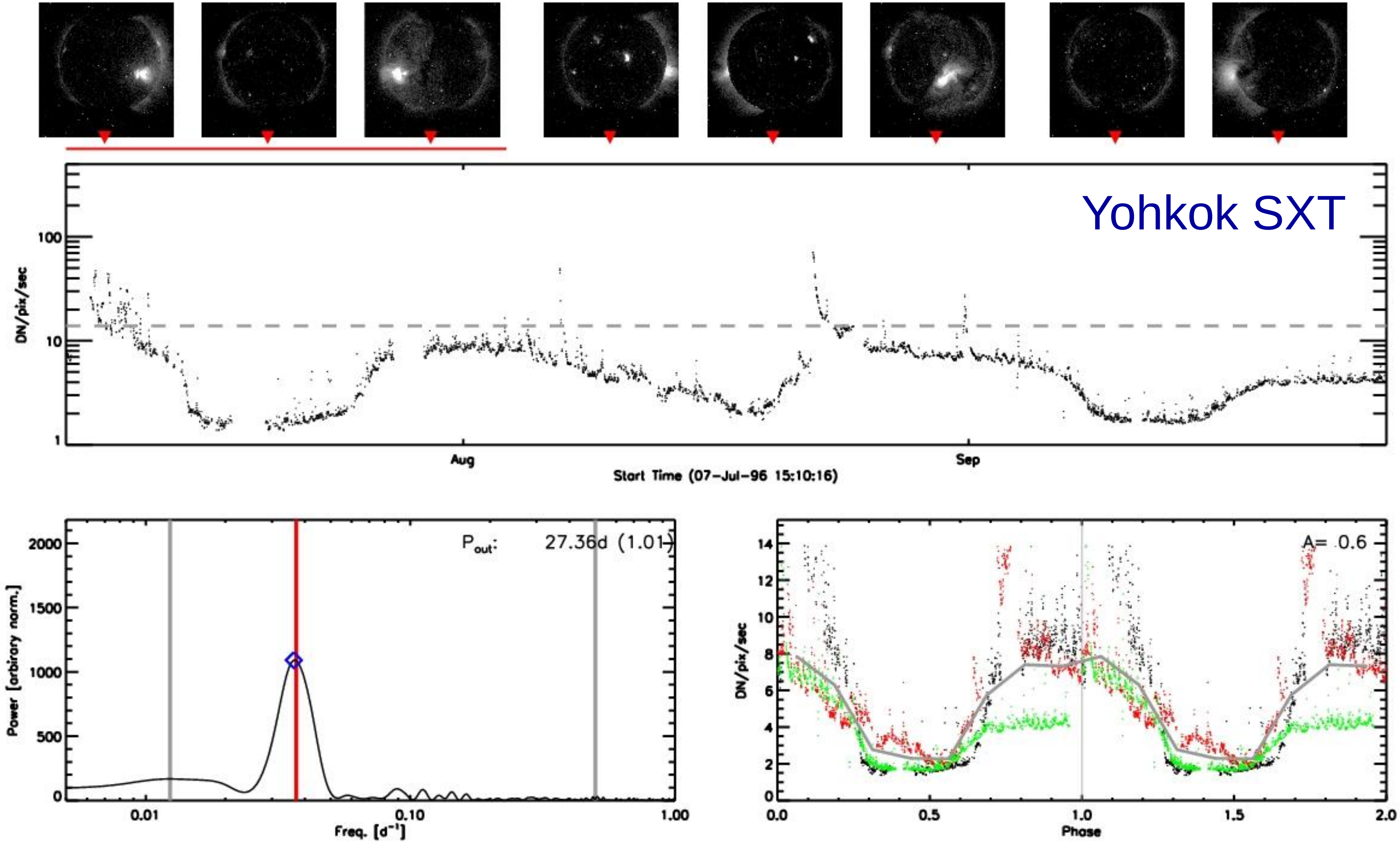
Sanz-Forcada, Maggio, Micela 2003

Variability analysis: methods

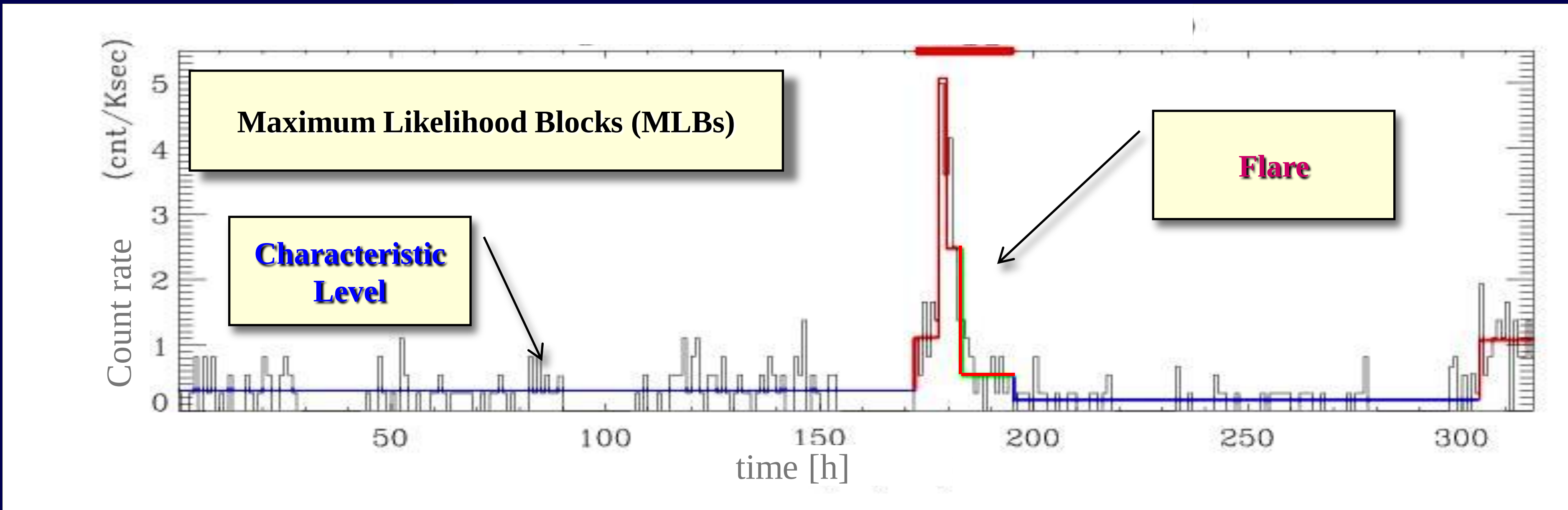
(just a few examples among many)

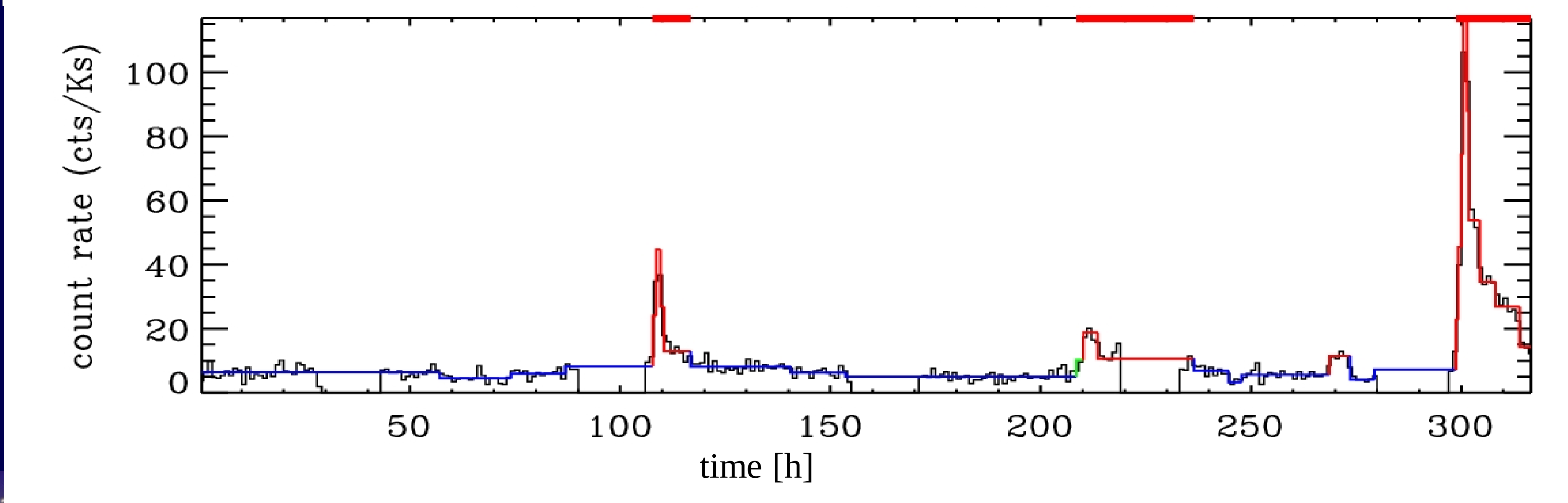
- Non-periodic, non-flaring variability on time scales of typical X-ray observations
  - Kolmogorov-Smirnov test (statistical textbooks)
  - χ^2 methods (e.g. Collura et al. 1987)
- Periodic variability on medium-long time scales
  - Lomb-Scargle periodograms: frequency analysis of unequally spaced data (Lomb 1976; Scargle 1982)
- Flare identification
  - Maximum-Likelyhood Bayesian Blocks (Scargle 1998)

Periodogram of Solar X-ray emission

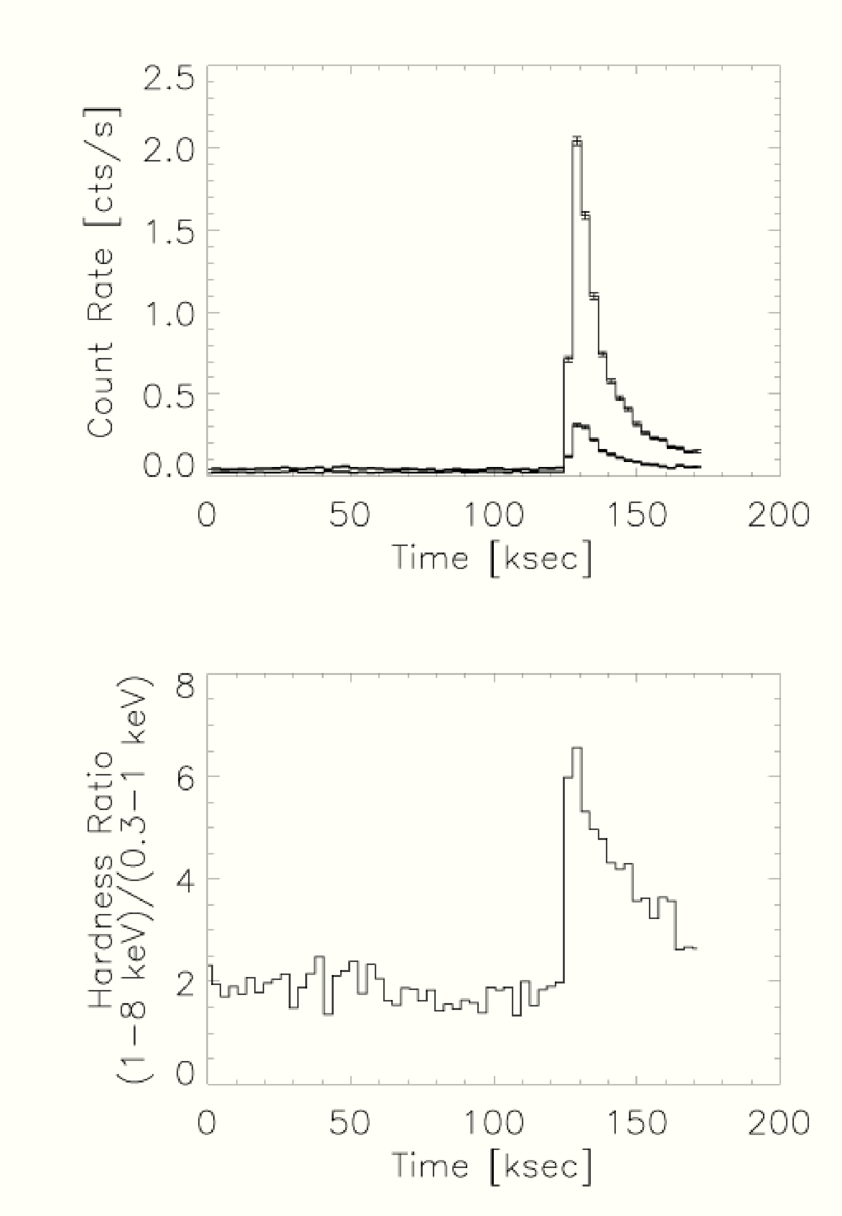


Byesian Block analysis





X-ray light curves and Hardness Ratios



- Flares from compact coronal structures are characterized by a steep rise in X-ray flux and in temperature
- In order to test the latter, time-resolved spectroscopy could be employed (if you have enough photons) or plots of **hardness ratios**
- The rise in temperature is expected and usually observed to precede the rise in emission measure (i.e. X-ray flux)

The End

(Enjoy your analysis session)

Data Analysis Sessions

- 8 data sets (3 from Chandra, 5 from XMM-Newton)
- 16 exercises
- 19 workstations available (3 reserved)
- 32 students (2 per exercise and per workstation)
- 10 tutors
  - Costanza Argiroffi
  - Paola Ballerini
  - Fabrizio Bocchino
  - Marilena Caramazza
  - Francesco Damiani
  - Ettore Flaccomio
  - Elena Franciosini
  - Mario Guarcello
  - Antonio Maggio
  - Beate Stelzer

Data Analysis Sessions

- Relevant subdirectories
  - **bin** (environment set-up)
  - **Documents** (tutorials, manuals)
  - **DATA/<ExerciseName>**
    - Different data sets linked in each directory
  - **Results**
- Linux operating system
  - **bash (default) or csh available**
- Software
  - Data analysis: **CIAO** for Chandra, **SAS** for XMM-Newton, **pwdetect** for source detection
  - Image handling: SAOimage (**ds9**)
  - Database queries: **SIMBAD** (via Mozilla Web Browser)
  - Spectral analysis: **XSPEC**
  - Text editors, PS/PDF viewers, etc. (see README)

Data Analysis Sessions

- First day Session
  - **Data visualization, filtering/screening**
  - **Source Detection**
  - **Source identification**
- Second day Session
  - **Individual source and background extraction**
  - **Spectral analysis**