

Optimal binning and fit statistics for X-ray spectra

Jelle Kaastra

Johan Bleeker

Introduction

- Transition low-resolution → high-resolution
- Needs forward-folding
- Computationally intensive
- Example: RGS properly designed, classical response matrix needs ~100 Mb-sized storage & calculation
- High-quality spectra can have dozens of free parameters; error search
- → Need optimal binning & design

Response matrices

- Distinguish two energies:
 - E = photon energy (entering the instrument)
 - E' = count energy (as reported by instrument)

- $$S(E') = \int_0^{\infty} R(E', E) f(E) dE$$

- Usually replaced by summation:
- $S_i = \sum R_{ij} F_j$

Energy bins

- Bins have different properties:
- Lower limit energy E_{1j}
- Upper limit energy E_{2j}
- Bin centre $E_j = \frac{1}{2} (E_{1j} + E_{2j})$
- Bin width $\Delta E_j = E_{2j} - E_{1j}$
- Photon Flux F_j (photons/s/bin)

How to evaluate the bin flux F_j ?

- 1) Bin center value: $F_j = f(E_j) \Delta E_j$
- Nice for continua
- 2) Integral over bin: $F_j = \int_{E_{1j}}^{E_{2j}} f(E) dE$
- Nice for narrow lines (narrower than bin width)
- Not trivial, which option to choose?

Tacitly assumed....

- Most analysis: *tacitly assume* all photons @ bin centre E_j
- What if line at bin boundary?
- Line centre determined with accuracy σ/\sqrt{N}
- FWHM = 2.35σ for Gaussian, N = number of counts in line
- *Example:* CCD with 50 eV resolution, 10 eV model bins, line 400 counts: \rightarrow line centre determined with $50/2.35/\sqrt{400} = 1$ eV accuracy \rightarrow line @ bin boundary off by 5 eV or 5 sigma significance!
- \rightarrow *Solution: increase number of bins....*
- But what if you study Capella?

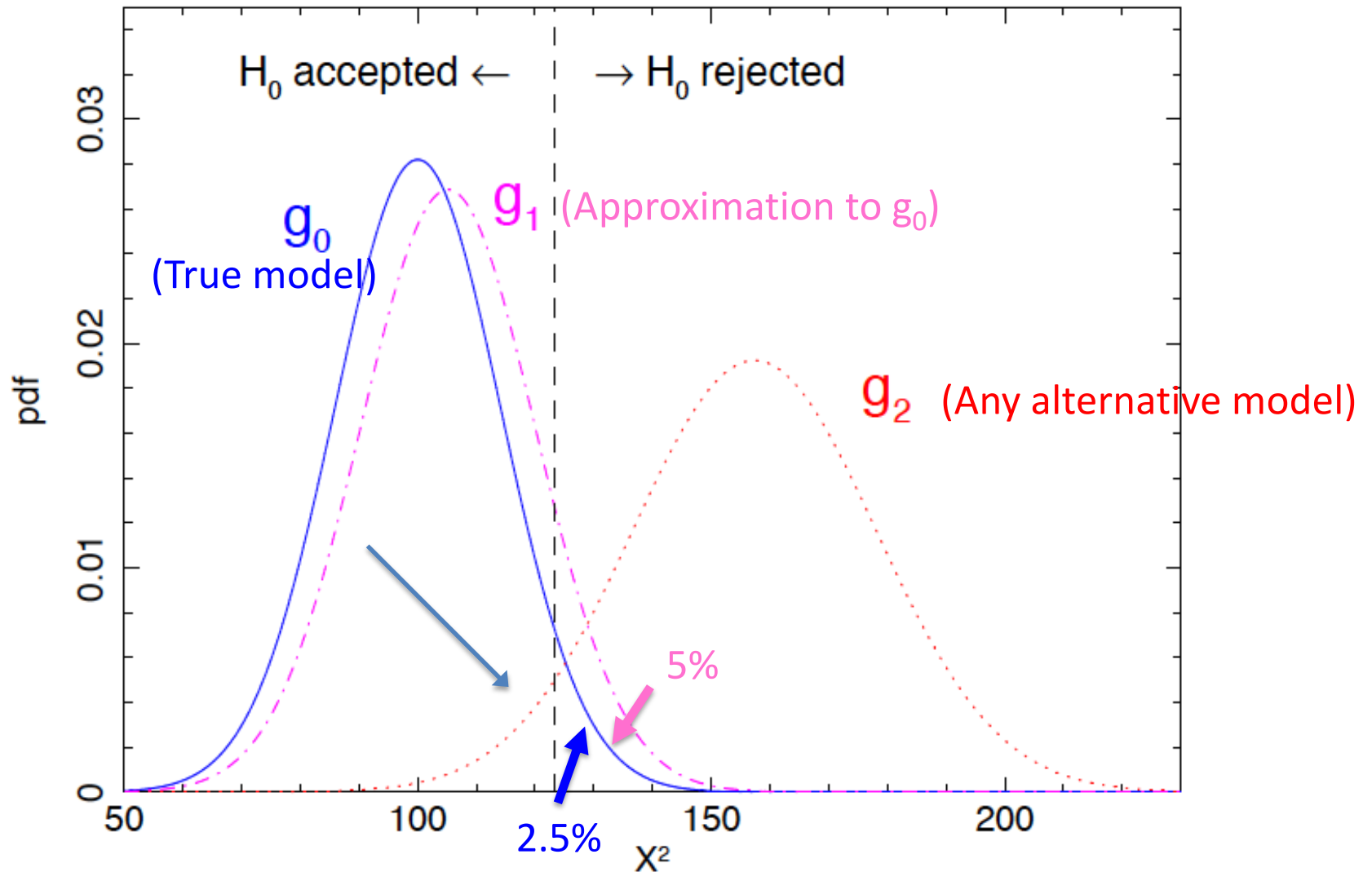
Example

- LETGS Capella spectrum
- Range 1-175 Å
- FWHM 0.040-0.076 Å
- 85 ks obs: 40 000 counts in Fe XVII 15 Å line
- Required model bin width $\sigma/\sqrt{N} = 0.00014$ Å
- = 278 bins per resolution element
- = 1.2×10^6 bins over 1-175 Å range

How to improve on this

1. Resolution usually not constant over range, use *resolution-dependent binning*
2. Highest resolution only needed near high flux / strong lines: *flux-dependent binning*
3. Revisit way spectra are calculated: not only account for number of photons within bin, but also their *average energy*

How to test a model



What test to use?

- Comparing two distributions (spectra):
- χ^2 -test has difficulties for small numbers of counts
- Use here Kolmogorov-Smirnov test (uses maximum deviation between two cumulative distribution functions)

Test criterium:

- G_0 : cumulative pdf of true distribution
- G_1 : cumulative pdf of approximation
- Choose rejection criterion c_α such that:
- $G_0(c_\alpha) = 1 - \alpha$ (e.g. $\alpha = 0.025$)
- $G_1(c_\alpha) = 1 - k\alpha$ (e.g. $k = 2$)

Analytical approximation

- Use that G_1 is close to G_0 in the tail, such that:
- $G_{KS}(c_\alpha) = 1 - \alpha$
- $G_{KS}(c_\alpha - \lambda_k) = 1 - k\alpha$ (shift λ_k)
- For $\alpha=0.025$, $k=2$: $\lambda_k=0.122$ (only weakly dependent on α)
- $\lambda_k = d_{\max} / \sqrt{N}$
- d_{\max} = max difference between both cdfs
- N = number counts or photons

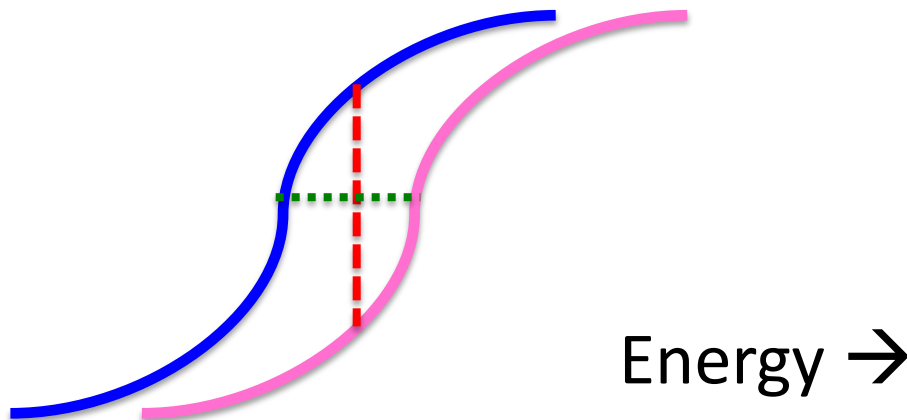
Flux distribution within bin: order 0, 1 or 2 approximations

- $f_0(E)$ = true distribution
- $f_{1,0} = N \delta(E-E_j)$ with E_j bin centre
- $f_{1,1} = N \delta(E-E_a)$ with $E_a = \int_{E_{1j}}^{E_{2j}} f(E)E dE$
- $f_{1,2} = N \exp[-(E-E_a)^2/2\tau^2]$ with

$$\tau^2 = \int_{E_{1j}}^{E_{2j}} f(E)(E - E_a)^2 dE$$

Comparing distributions in practice

- Take “true” line @bin boundary E_{1j} , distribution f_0
- Take “normal” line @ bin centre approximation for $f_{1,0}$ ($=f_0$ shifted by $-\Delta E_j/2$)
- Convolve both with instrument (Gaussian, width σ) and determine cumulative cdf; calculate d_{\max}



Extending to full spectra

- Above formalism for narrow lines (with Gaussian instrumental broadening)
- Full spectrum in resolution element r written as $S_r = \sum w_i F_i$
- In general $\sqrt{\sum w_i^2} \leq \sum w_i$
- \rightarrow error in S_r in broadband-spectrum smaller than error if all flux in a line
- \rightarrow use $d_{\max} = \max_r \sqrt{N_r D_r}$

Extending to full spectra II

- For total of R resolution elements:
- $G_0(c_\alpha)^R = 1 - \alpha$
- $G_1(c_\alpha)^R = 1 - k\alpha$
- (this makes use of distribution of maximum of a set of R variables)

Back to approximations order 0, 1, 2

- Can be shown that
- $\Delta_0 < 2.1289 \lambda_k N^{-1/2}$
- $\Delta_1 < 2.4418 \lambda_k^{1/2} N^{-1/4}$
- $\Delta_2 < 2.2875 \lambda_k^{1/3} N^{-1/6}$

Model bin size (result of Monte Carlo calculations)

$$\frac{\Delta}{\text{FWHM}} = \min(1, y), \quad (17)$$

with

$$\text{order 0} : y = \frac{0.5707}{x^{1/2}} \left(1 + \frac{1.0}{x}\right), \quad (18)$$

$$\text{order 1} : y = \frac{1.404}{x^{1/4}} \left(1 + \frac{18}{x}\right), \quad (19)$$

$$\text{order 2} : y = \frac{1.569}{x^{1/6}} \left(1 + \frac{1.14}{x^{1/3}}\right), \quad (20)$$

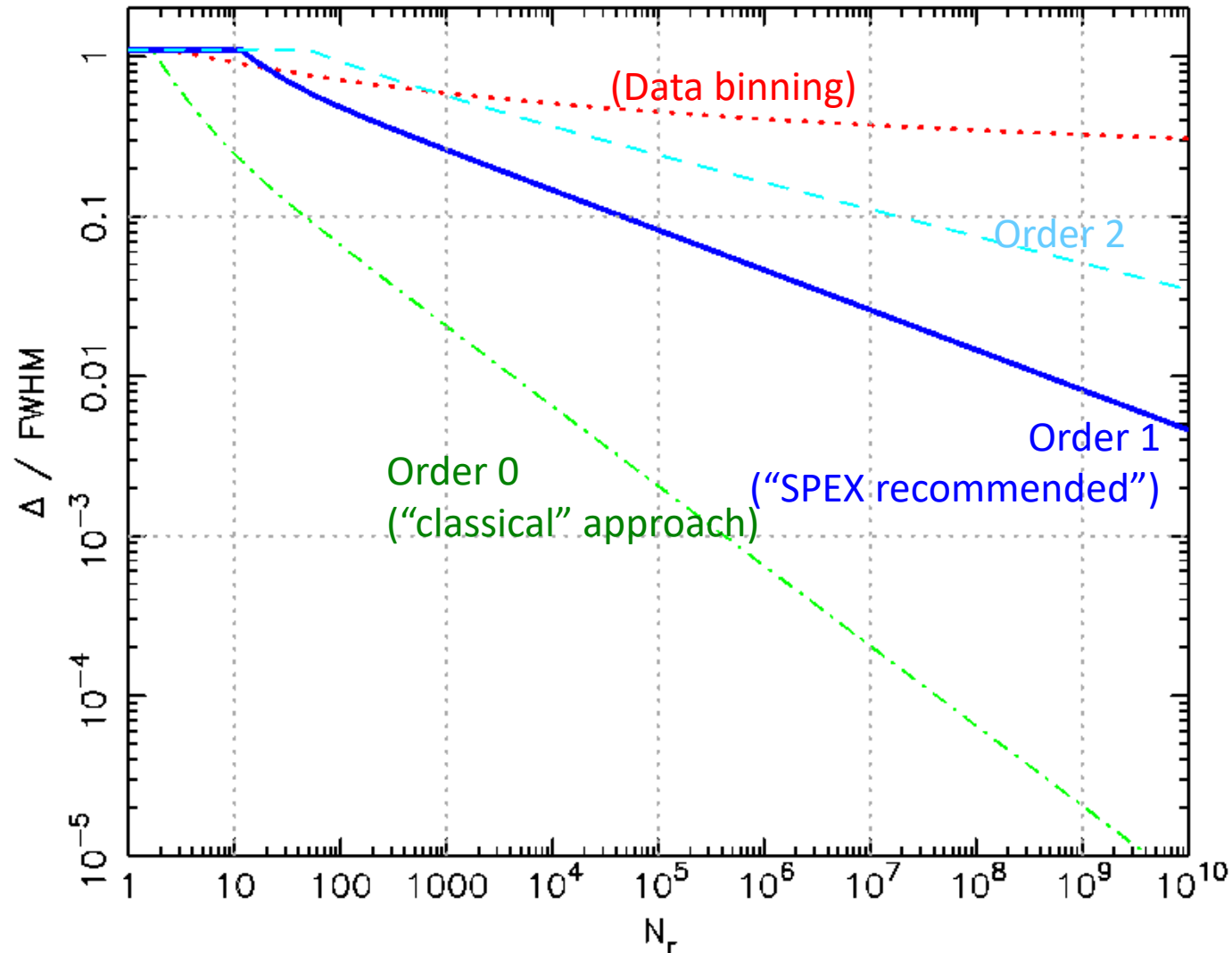
where

$$\text{order 0} : x = N_r(1 + 0.3 \ln R), \quad (21)$$

$$\text{order 1} : x = N_r(1 + 0.1 \ln R), \quad (22)$$

$$\text{order 2} : x = N_r(1 + 0.6 \ln R). \quad (23)$$

Monte Carlo results



Effective area curvature

$$A_j(E) = A(E_j) + A'(E_j)(E - E_j) + \dots \quad (24)$$

$$\epsilon_{\max} = \frac{1}{8}(\Delta E)^2 A''(E_j)/A(E_j),$$

Taking several steps in between (see paper):

$$\frac{\Delta}{\text{FWHM}} = 1.5877 \left(\frac{d \ln E}{d \ln A} \right) \left(\frac{E_j}{\text{FWHM}} \right) N_r^{-0.25}. \quad (30)$$

$$\frac{\Delta}{\text{FWHM}} = \frac{1}{1/w_1 + 1/w_a}, \quad (31)$$

Optimal binning

- From order 0 to order 1 improvement order of magnitude
- From order 1 to order 2 only slight improvement at lot of computational expense
- → *choose order 1 as best option*

Consequence for response design

If response matrix is R_{ij} with i observed count energy bin & j photon energy bin:

Classical
Response
matrix

New term:
response
derivative

$$S_i = R_{ij} F_j + R'_{ij} (E_{a,j} - E_j) F_j$$

with the prime indicating differentiation with respect to photon energy (second index)

S_i is count spectrum, F_j incoming photon spectrum

Data binning: Shannon theorem

The Shannon (1949) sampling theorem states the following: Let $f(x)$ be a continuous signal. Let $g(\omega)$ be its Fourier transform, given by

$$g(\omega) = \int_{-\infty}^{\infty} f(x)e^{i\omega x} dx. \quad (32)$$

If $g(\omega) = 0$ for all $|\omega| > W$ for a given frequency W , then $f(x)$ is band-limited, and in that case Shannon has shown that

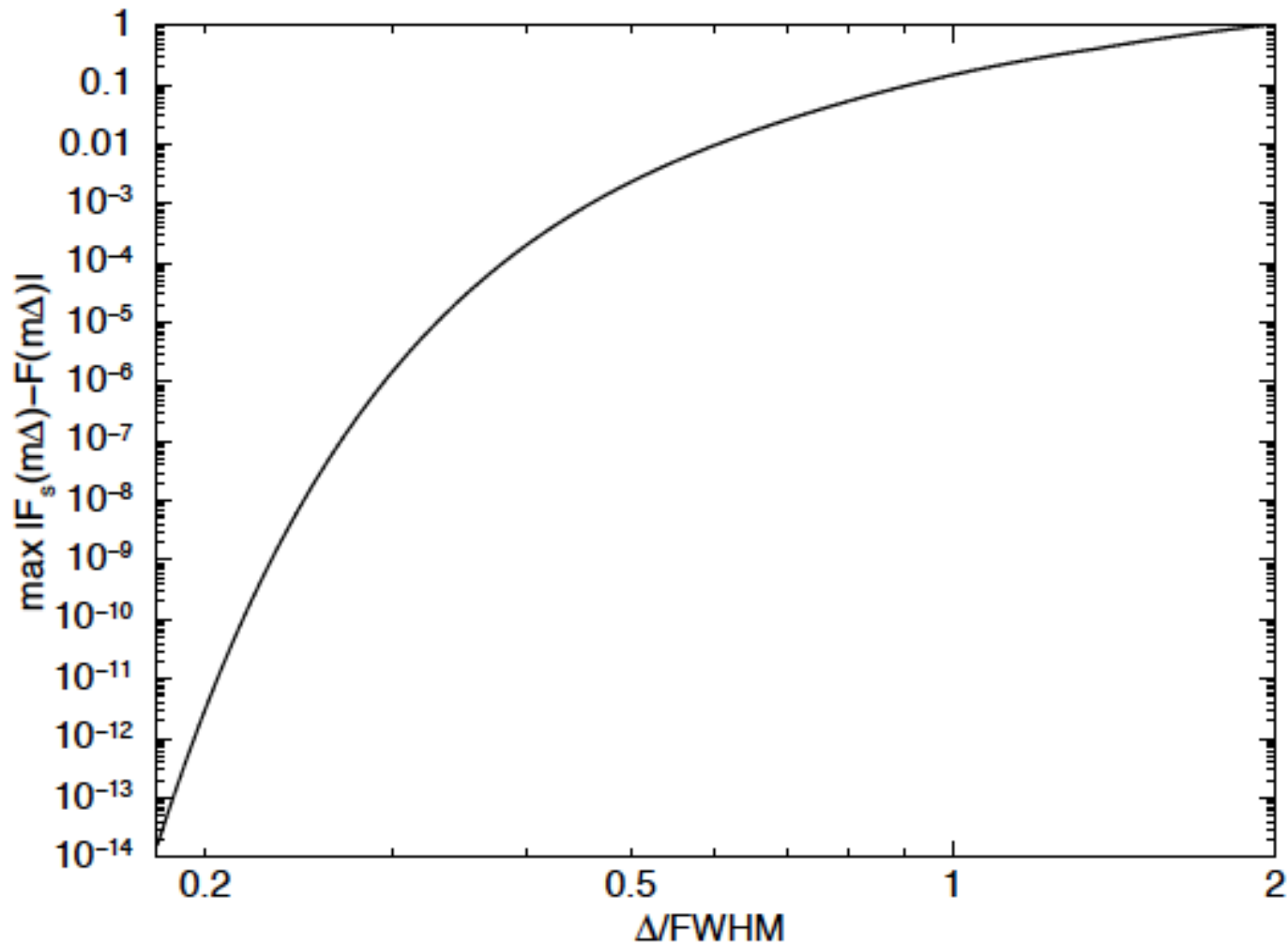
$$f(x) = f_s(x) \equiv \sum_{n=-\infty}^{\infty} f(n\Delta) \frac{\sin \pi(x/\Delta - n)}{\pi(x/\Delta - n)}. \quad (33)$$

Integrated Shannon theorem

$$F_s(m\Delta) = \frac{\Delta}{\pi} \sum_{n=-\infty}^{\infty} f(n\Delta) \left\{ \frac{\pi}{2} + \text{Si}[\pi(m-n)] \right\}.$$

X-ray spectra usually given as histograms, so need to use integrated form of Shannon theorem: use cumulative distribution at bin boundaries as only info available

Max difference Shannon approximation from true distribution for a Gaussian redistribution



Analytical approximation data bin size Δ

$$\frac{\Delta}{\text{FWHM}} = \begin{cases} 1 & \text{if } x \leq 2.119; \\ \frac{0.08 + 7.0/x + 1.8/x^2}{1 + 5.9/x} & \text{if } x > 2.119, \end{cases} \quad (36)$$

with

$$x \equiv \ln[N_r(1 + 0.20 \ln R)]. \quad (37)$$

N_r : number of counts in resolution element considered

R : total number of resolution elements in the spectrum

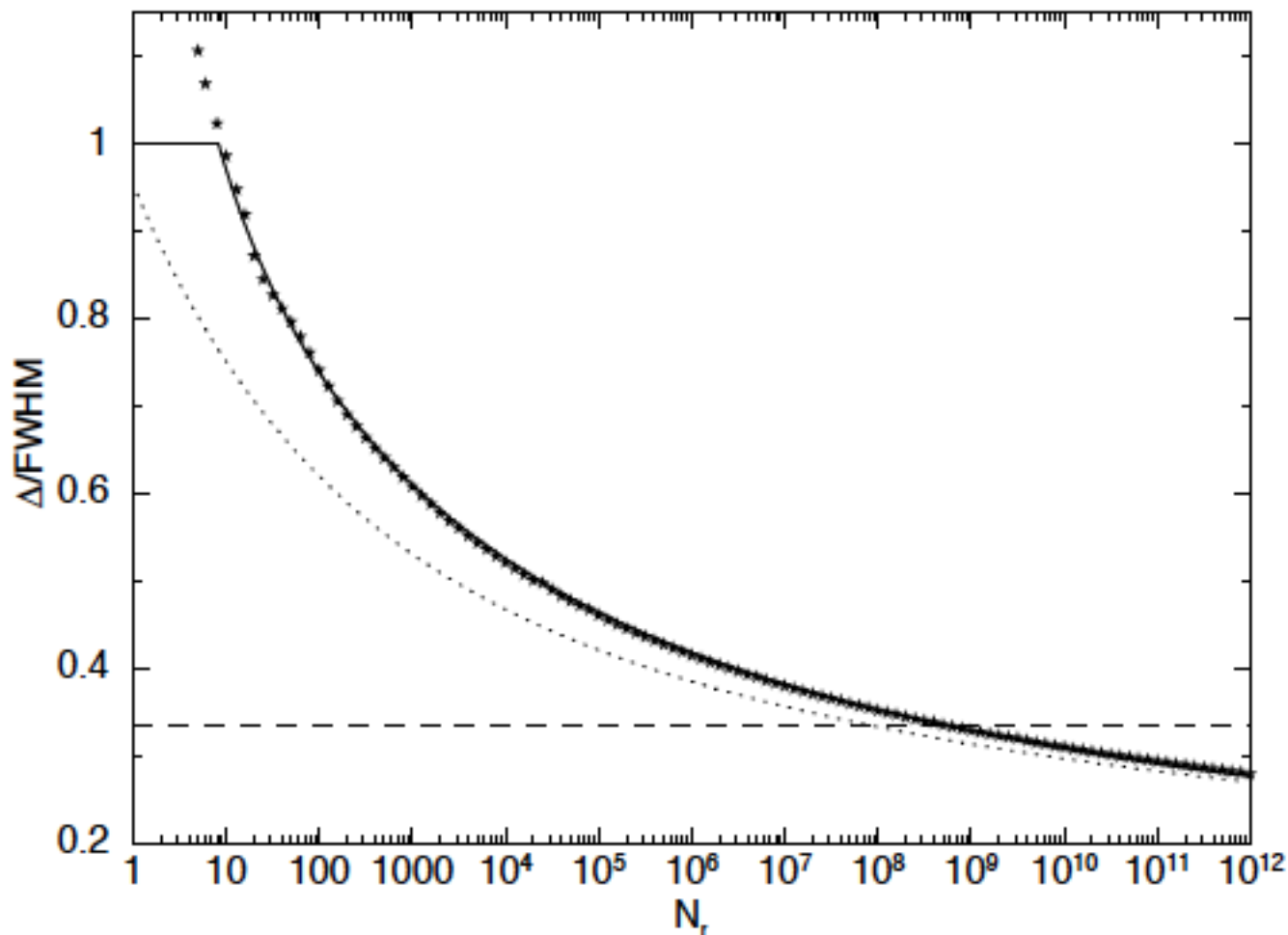


Fig. 3. Optimal bin size Δ for data binning with a Gaussian *lsf*, as a function of the number of counts per resolution element N_r . Dotted curve: analytical approximation using Sect. A.3; stars: results of Monte Carlo calculation; solid line: our finally adopted bin size Eq. (36), based on a fit to our Monte Carlo results for large values of N_r ; dashed line: commonly adopted bin size 1/3 FWHM.

Example: Capella LETGS spectrum

- 85 ks Chandra LETGS observation
- Range 0.86 – 175.55 Å
- Resolution 0.040 – 0.076 Å
- Strongest line (Fe XVII) has 15000 counts

Example: Capella (continued)

Binning	Data bins	Model bins	Resp. elements	comments
A	8.52×10^3	1.26×10^6	1.01×10^7	Constant bin size (20 mÅ data, 0,14 mÅ model)
B	6.20×10^3	9.14×10^5	7.31×10^6	Bin size follows resolution ΔE
C	5.12×10^3	1.23×10^5	9.84×10^5	Bin size follows resolution ΔE & N; photons @ bin centre E_j
D	5.12×10^3	8.21×10^3	6.57×10^4	Bin size follows resolution ΔE & N; photons @ average energy E_a

Appendix B.1: Creating the data bins

Given an observed spectrum obtained by some instrument, the following steps should be performed to generate an optimally binned spectrum.

1. Determine for each original data channel i the nominal energy E_{j0} , defined as the energy for which the response at channel i reaches its maximum value. In most cases, this is the nominal channel energy.
2. Determine for each data channel i the limiting points $(i1, i2)$ for the FWHM in such a way that $R_{k,j0} \geq 0.5 R_{i,j0}$ for all $i1 \leq k \leq i2$, while the range of $(i1, i2)$ is as broad as possible.
3. By (linear) interpolation, determine for each data channel the points (fractional channel numbers) $c1$ and $c2$ near $i1$ and $i2$, where the response is actually half its maximum value. By virtue of the previous step, the absolute difference $|c1 - i1|$ and $|c2 - i2|$ never can exceed 1.
4. Determine for each data channel i the FWHM c_i in units of channels, by calculating $c2 - c1$. Assure that c_i is at least 1.
5. Determine for each original data channel i the FWHM in energy units (e.g. in keV). Call this W_i . This and the previous steps may of course also be performed directly using instrument calibration data.
6. Determine the number of resolution elements R by the following approximation:

$$R = \sum_i \frac{1}{c_i}. \quad (\text{B.1})$$

7. Determine for each bin the effective number of events N_r from the following expressions:

$$C_r = \sum_{k=i1-1}^{i2+1} C_k, \quad (\text{B.2})$$

$$h_r = \sum_{k=1}^{N_c} R_{k,j0} / \sum_{k=i1-1}^{i2+1} R_{k,j0}, \quad (\text{B.3})$$

$$N_r = C_r h_r. \quad (\text{B.4})$$

In the above, C_k is the number of observed counts in channel k , and N_c is the total number of channels. Take care that in the summations $i1 - 1$ and $i2 + 1$ are not out of their valid range $(1, N_c)$. If for some reason there is not a first-order approximation available for the response matrix $R_{k,j}$ then one might simply approximate h_r from e.g. the Gaussian approximation, namely $h_r = 1.314$; cf. Sect. 5.4. This is justified since the optimal bin size is not a strong function of N_r ; cf. Fig. 3. Even a factor of two error in N_r in most cases does not affect the optimal binning too much.

8. Using (36), determine for each data channel the optimal data bin size in terms of the FWHM. The true bin size b_i in terms of number of data channels is obtained by multiplying this by c_i calculated above during step 4. Make b_i an integer number by ignoring all decimals (rounding it to below), but take care that b_i is at least 1.

9. It is now time to merge the data channels into bins. In a loop over all data channels, start with the first data channel. Name the current channel i . Take in principle all channels k from channel i to $i + b_i - 1$ together. However, check that the bin size does not decrease significantly over the rebinning range. In order to do that check, determine for all k between i and $i + b_i - 1$ the minimum a_i of $k + b_k$. Extend the summation only from channel i to $a_i - 1$. In the next step of the merging, a_i becomes the new starting value i . The process is finished when a_i exceeds N_C .

Appendix B.2: Creating the model bins

After having created the data bins, it is possible to generate the model energy bins. Some of the information obtained from the previous steps that created the data bins is needed.

The following steps need to be taken:

1. Sort the FWHM of the data bins in energy units (W_i) as a function of the corresponding energies E_{j0} . Use this array to interpolate any true FWHM later. Also use the corresponding values of N_r derived during that same stage. Alternatively, one may use directly the FWHM as obtained from calibration files.
2. Choose an appropriate start and end energy, e.g. the nominal lower and upper energy of the first and last data bin with an offset of a few FWHMs (for a Gaussian, about 3 FWHM is sufficient). In the case of a lsf with broad wings (like the scattering due to the RGS gratings), it may be necessary to take an even broader energy range.
3. In a loop over all energies, as determined in the previous steps, calculate the bin size in units of the FWHM using (19).
4. Also, determine the effective area factor $\frac{d \ln E}{d \ln A}$ for each energy; one may do that using a linear approximation.
5. For the same energies, determine the necessary bin width in units of the FWHM using eqn. (31). Combining this with the FWHMs determined above gives for these energies the optimal model bin size ΔE in keV.

6. Now the final energy grid can be created. Start at the lowest energy $E_{1,1}$, and interpolate in the ΔE table the appropriate $\Delta E(E_{1,1})$ value for the current energy. The upper bin boundary $E_{2,1}$ of the first bin is then simply $E_{1,1} + \Delta E(E_{1,1})$.
7. Using the recursive scheme $E_{1,j} = E_{2,j-1}$, $E_{2,j} = E_{1,j} + \Delta E(E_{1,j})$ determine all bin boundaries until the maximum energy has been reached. The bin centres are simply defined as $E_j = 0.5(E_{1,j} + E_{2,j})$.
8. Finally, if there are any sharp edges in the effective area of the instrument, it is necessary to add these edges to the list of bin boundaries. All edges should coincide with bin boundaries.

Final words

- Can use **different binning** for different components of the matrix:
- Example: CCD spectrum
 - **Diagonal** needs fine binning but is narrow
 - **Partial events tail** is broad but coarse binning sufficient
- This can save a lot of memory (& computing time!)
- All implemented in SPEX, now the others ...
- Spread the news!

Part 2: Cstat

Why use C-stat?

- Classical χ^2 fitting drawbacks, in particular for low # counts
- Bias in flux, even for high counts ~ 1 count (Mighell 1999)
- Rebinning to >25 counts/bin may wash out spectral details

What is Cstat?

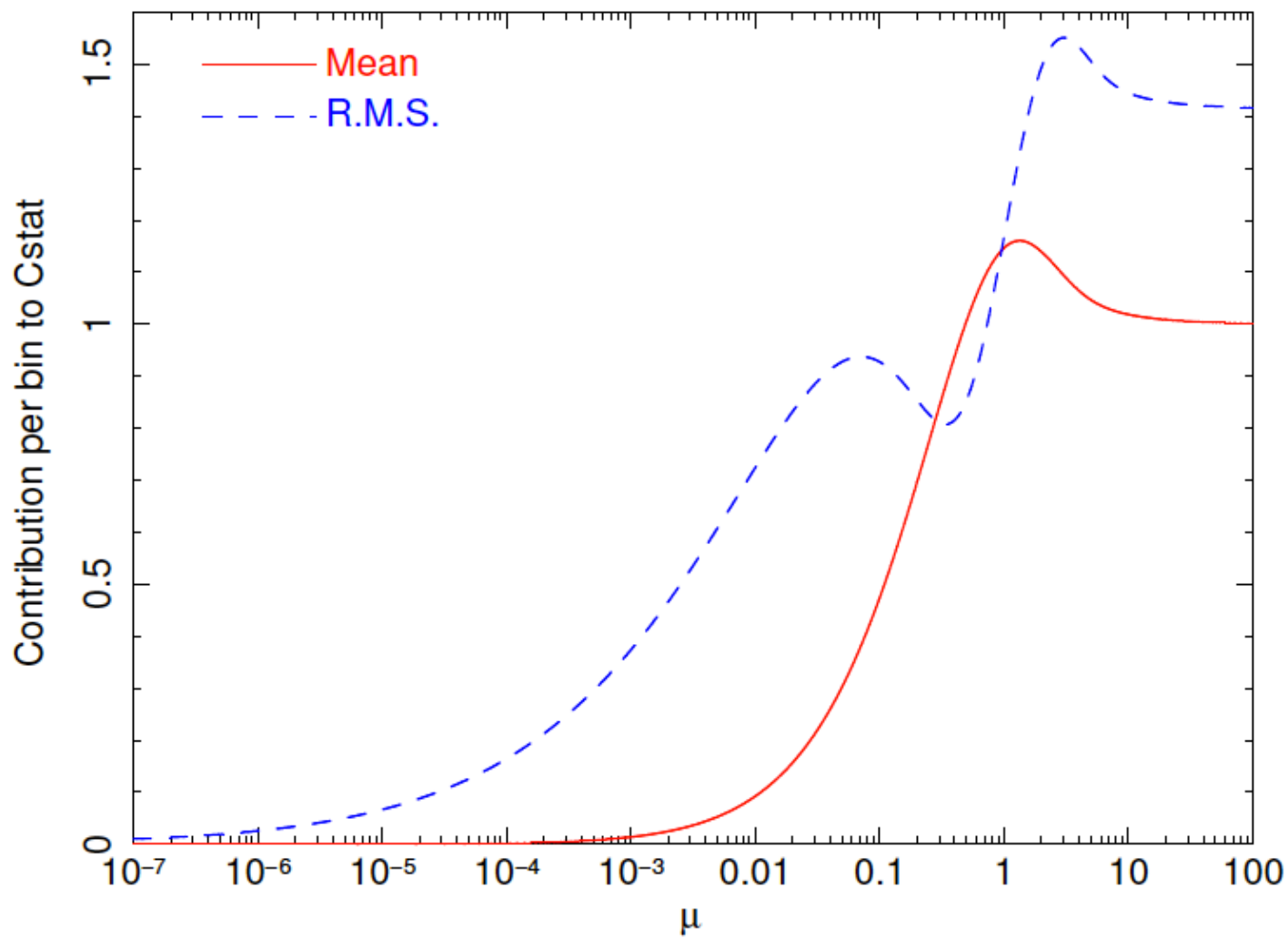
- s_i model in bin i (model for source+background)
- N_i counts (source + background) in bin i
- $C = 2 \sum_i [s_i - N_i + N_i \ln(N_i/s_i)]$

Criticism on Cstat

- Nousek & Shue (1989): *The principal disadvantage of the C statistic is that there is no value corresponding to the reduced χ^2 value with which we can measure the goodness of the fit.*
- Humphrey et al. (2009): *Since the absolute value of the C-statistic cannot be directly interpreted as a goodness-of-fit indicator observers typically prefer instead to minimize the better-known χ^2 fit statistic.*

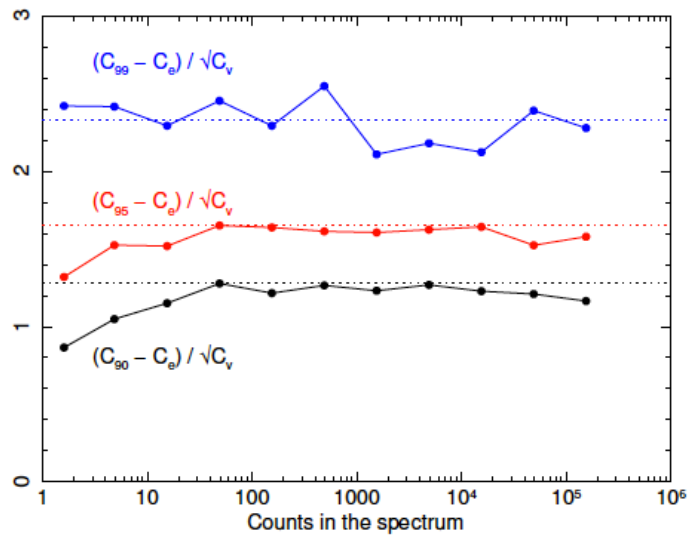
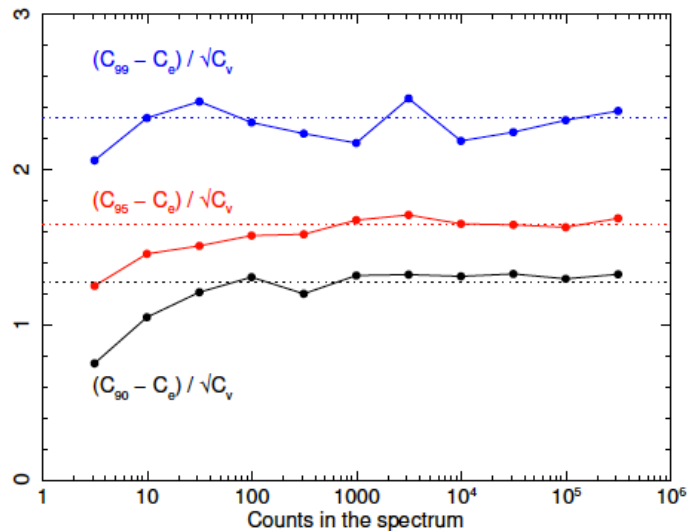
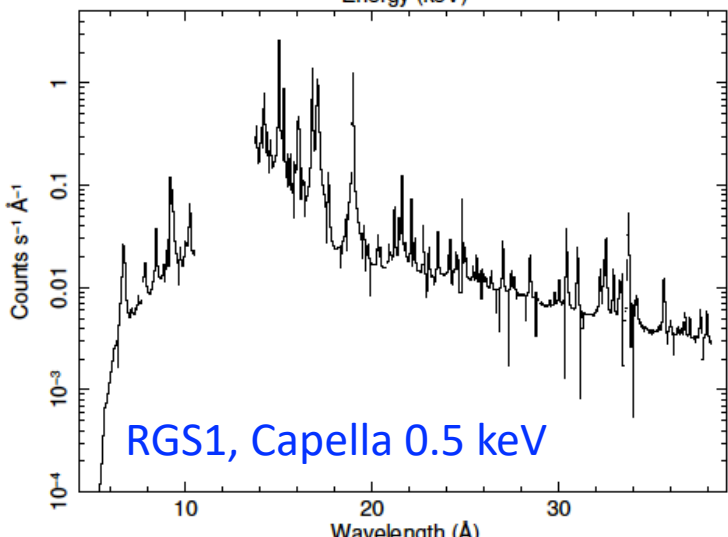
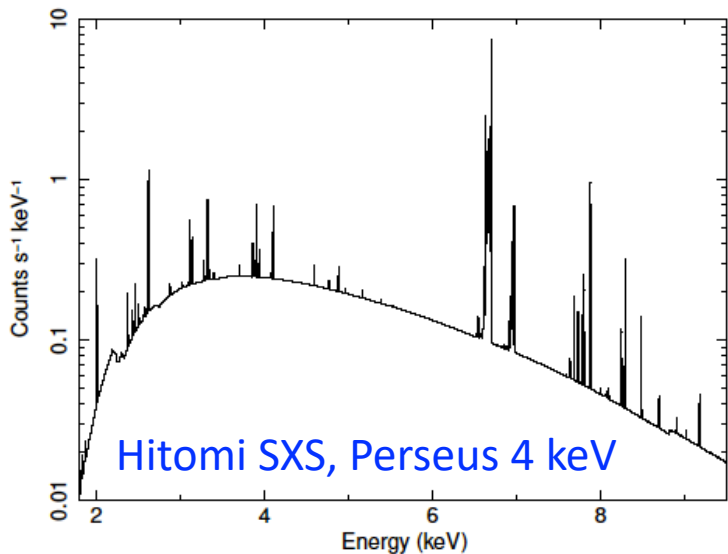
Cstat is a statistic

- $C = 2 \sum_i [s_i - N_i + N_i \ln(N_i/s_i)]$
- C is a statistic: given model s_i , N_i is Poissonian distributed with mean s_i , so the expected value, variance etc. of C can be determined!
- Done in Kaastra (2017), gives numerical parameterization of expected value and variance of C



Simulations

(1000 runs/point)



On proper use of Cstat (and no use of Wstat)

Standard model

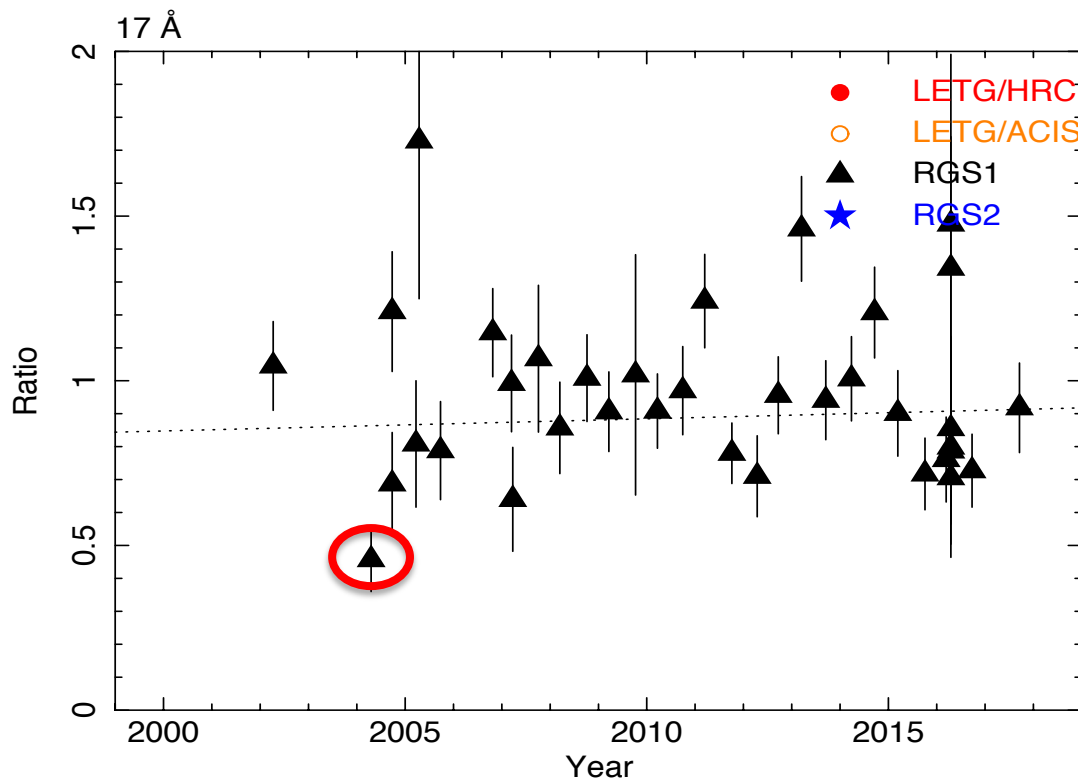
- Blackbody spectrum (SPEX: *bb*)
- Galactic absorption (SPEX: *hot*, with $T=0.5$ eV fixed (neutral) and Lodders et al. abundances)
- Fudge factor for individual spectra (SPEX: *knak*, i.e. piecewise multiplication with local power law, prescribed at grid of λ_i with correction factors f_i , interpolated with powerlaws)
- Choose spacing of 2 \AA for λ_i , starting at 15 \AA

LETGS fit parameters

- Fit to 15-80 Å range
- C-stat = 19172 (expected 18839 ± 193)
- BB norm = 0.228 ± 0.005 [*adopt 0.228*]
- BB kT = 62.3 ± 0.2 [*adopt 62.3*]
- NH = $0.984 \pm 0.014 \times 10^{24} \text{ m}^{-2}$ [*adopt 0.985*]

Binning issues

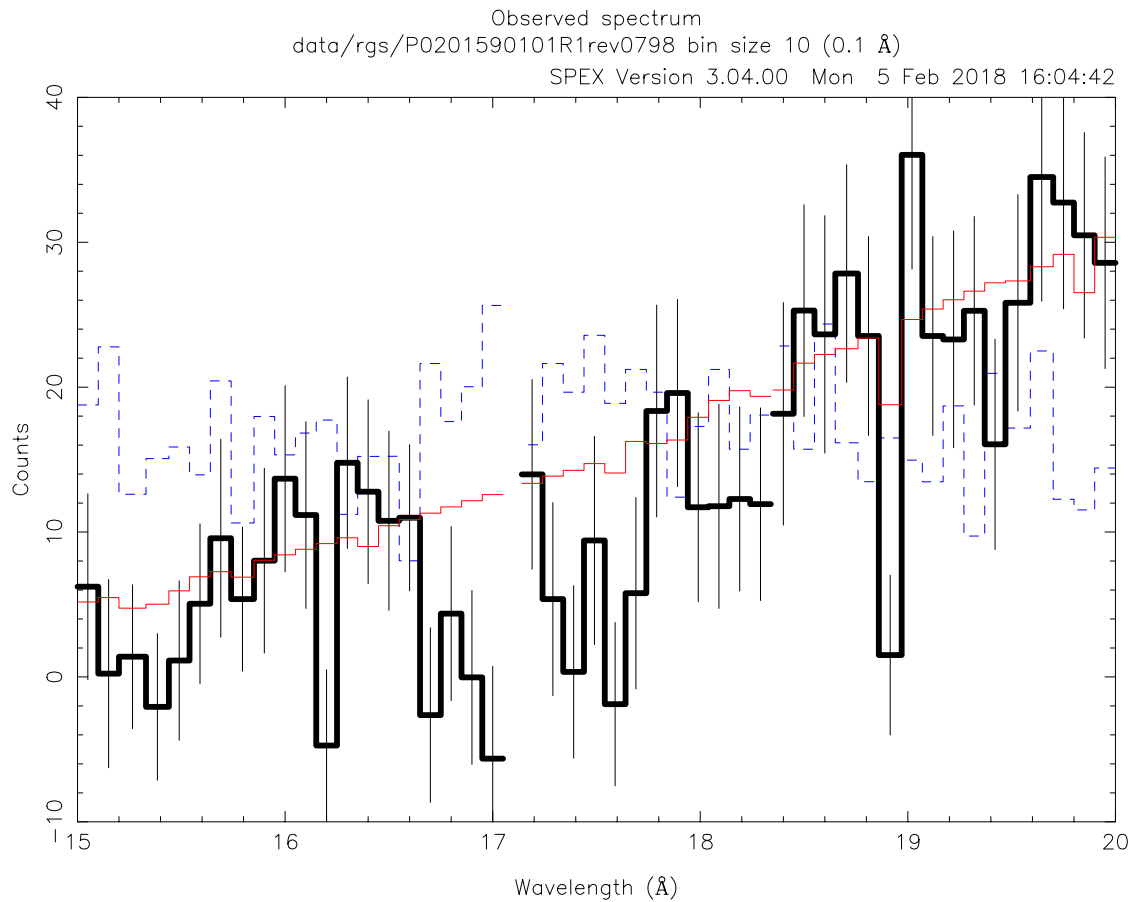
- Some data showed unexpected behavior:



Example is 17 Å fudge factor for RGS1 of obsid 0201590101, revolution 798
Has correction factor of 0.5 ± 0.1 , hence -5σ deviation

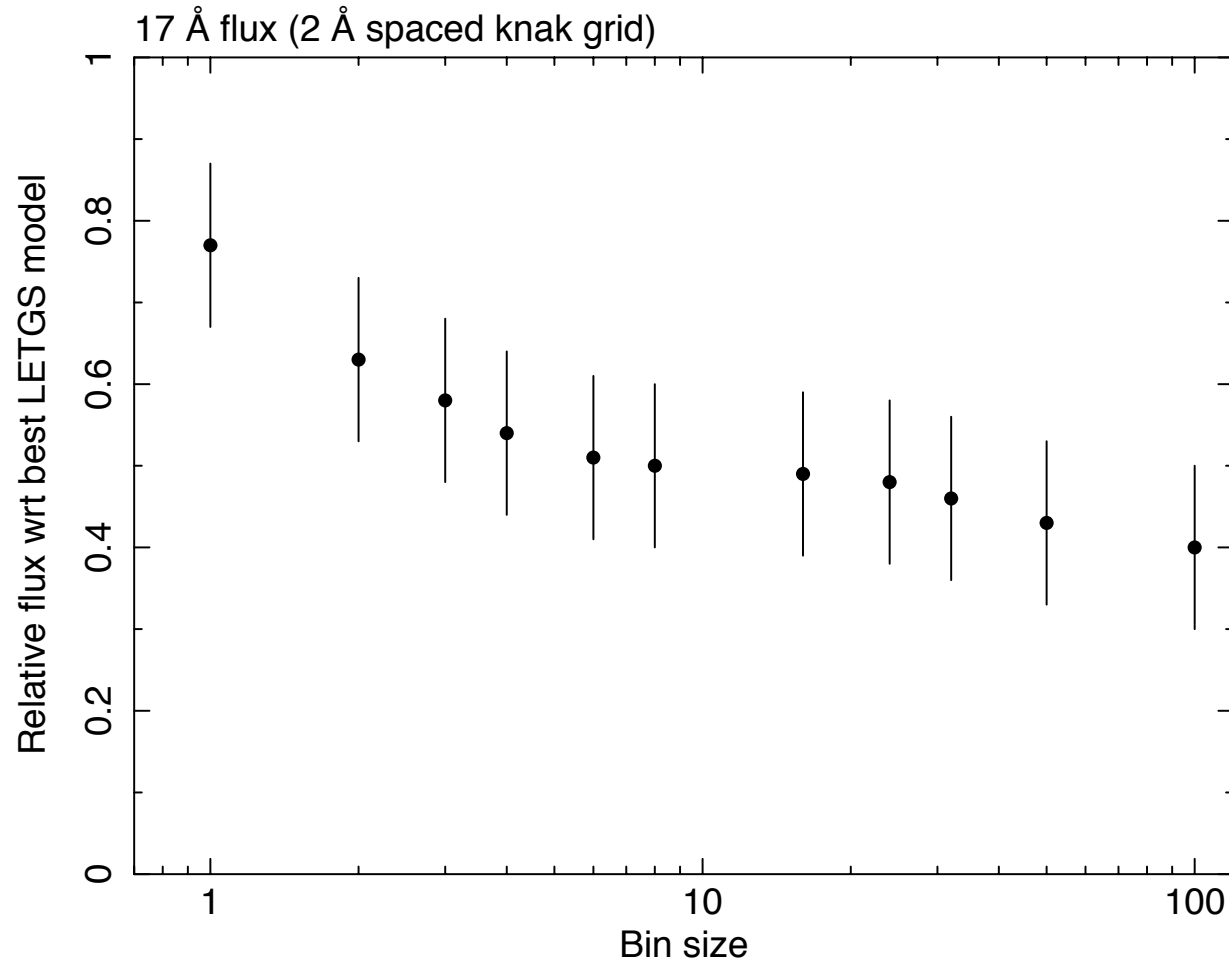
Deviant spectrum

Low spectrum at 17 Å, with relatively high subtracted background



Fudge factor depends on bin size

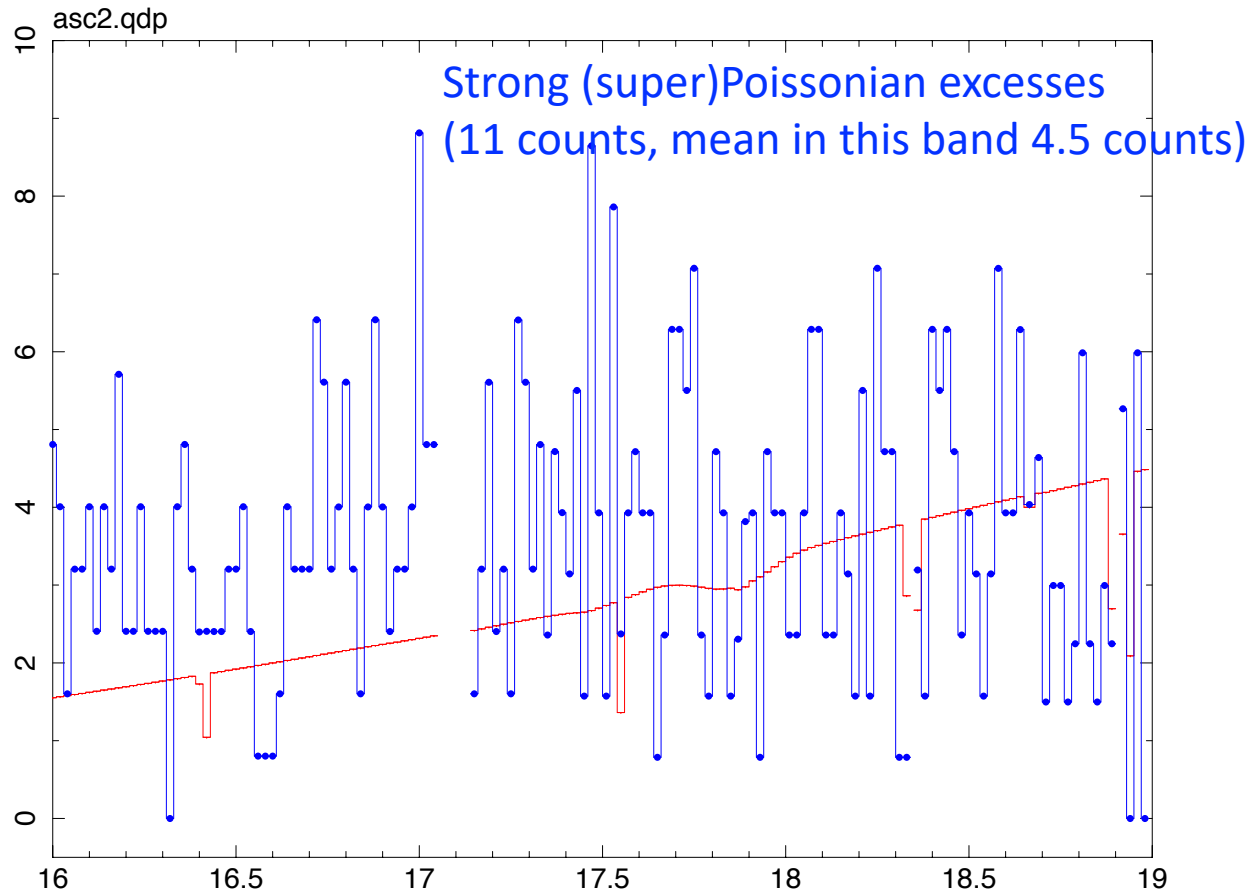
(multiples of 0.01 Å)



What is going on?

- Issues with fitting method
- χ^2 fitting known to cause bias, so used already cstat
- However, for cstat method $\sum(\text{source}+\text{background model})$ needed
- For RGS spectra (& many other instruments) measured background (scaled counts) used instead
- For source < background, this gives bias in weights

Binning: factor 2 (0.02 \AA) – Model for RXJ 1856 – subtracted background



$$f_i = \frac{S_i + B_i - (t_s + t_b)m_i + d_i}{2(t_s + t_b)}$$

Wstat

$$W = 2 \sum_{i=1}^N t_s m_i + (t_s + t_b) f_i - S_i \ln(t_s m_i + t_s f_i) - B_i \ln(t_b f_i) - S_i(1 - \ln S_i) - B_i(1 - \ln B_i)$$

$$f_i = \frac{S_i + B_i - (t_s + t_b)m_i + d_i}{2(t_s + t_b)}$$

$$d_i = \sqrt{[(t_s + t_b)m_i - S_i - B_i]^2 + 4(t_s + t_b)B_i m_i}$$

Is Wstat the solution?

- In XSPEC, when CSTAT is used and data & background is Poissonian, they use WSTAT instead (but call it cstat...)
- For full definition, see WSTAT
- W determined as maximum likelihood estimate from background & source (that order)
- Background estimate depends on source model, so during fit (adjusting source model), the accounted background changes

Is Wstat the solution? II

- For low count rates, this gives problems:
- XSPEC manual: *In practice, it works well for many cases but for weak sources can generate an obviously wrong best fit. It is not clear why this happens although binning to ensure that every bin contains at least one count often seems to fix the problem.*
- This binning is undesired!

Alternative: make model for background

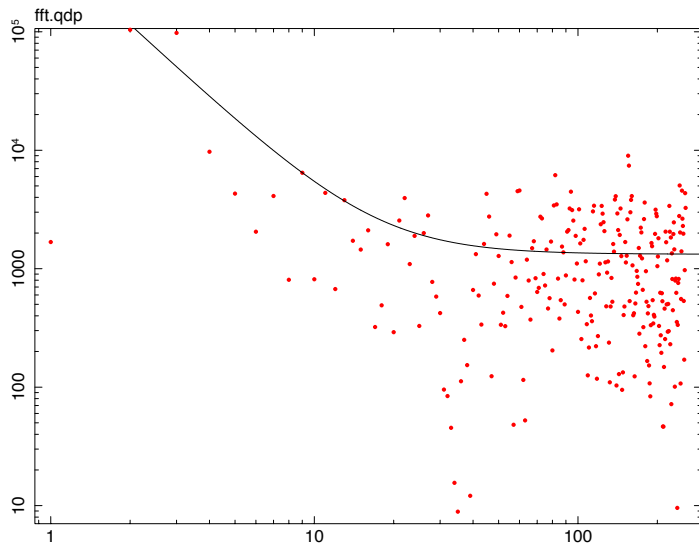
Method chosen: Wiener filtering (Numerical Recipes)

1. Make FFT of background spectrum
2. Determine noise level N^2 at high frequency
3. Make crude model for “true” signal S^2
4. *Filter = $S^2 / (S^2 + N^2)$*
5. Multiply FFT by filter
6. Transform back

Pros and contras

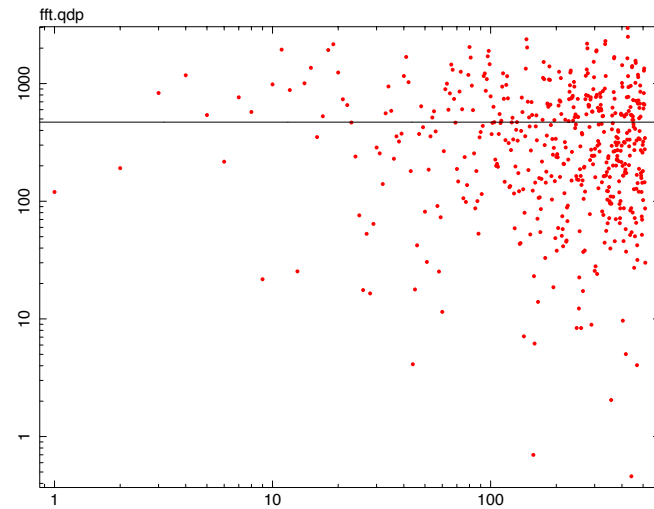
- Pro:
 - Average background level preserved
 - Smooth background, reduces noise significantly
- Contra:
 - Needs to extend spectrum with “mean” value to power of 2 number of bins
 - May cause small deviations at edges spectrum
 - FFT model not always “automatically” to obtain

Example ffts



kaastra 5-Apr-2018 09:20

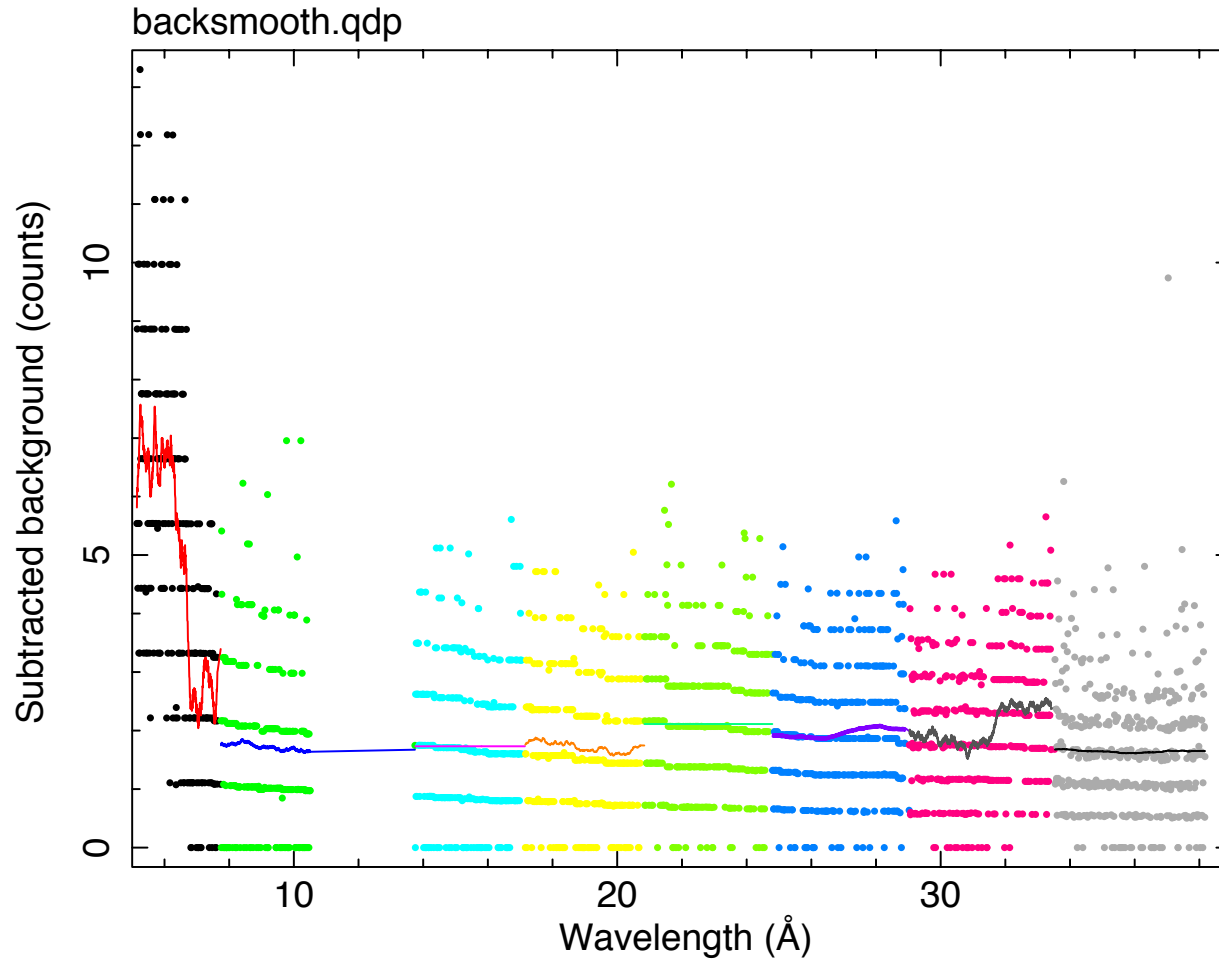
RGS1 CCD9



kaastra 5-Apr-2018 09:20

RGS1 CCD6

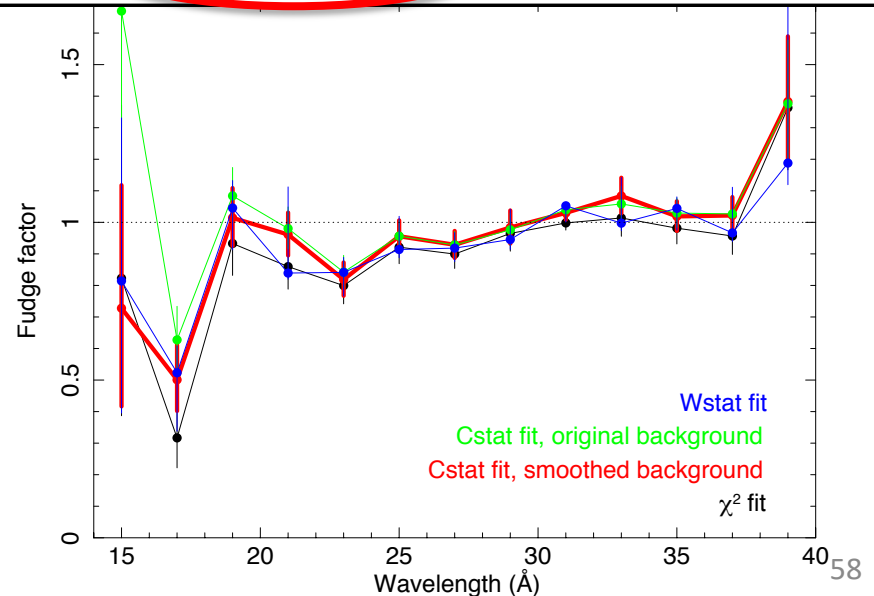
Old & new filtered background



Comparison methods

Method	χ^2	Cstat (data bg)	Wstat	Cstat (Model bg)
Statistic	1188	1405	1154	1161
Dof/range	1102	1136±48	1102	1135±48
# evaluations	3084	2888	12873	3527
F (17 Å)	0.32±0.10	0.63±0.10	0.52(-0.18,+0.09)	0.50±0.10
F (27 Å)	0.90±0.05	0.93±0.04	0.92±0.04	0.93±0.04
F (37 Å)	0.96±0.06	1.03±0.06	0.97(-0.03,+0.14)	1.02±0.06

- χ^2 & Cstat with data bg:
→ bias
- Wstat: unstable, asymmetric errors, 4 x more model evaluations; bias?



Comparison methods

