# XSPEC Tutorial and Statistics Basic steps for X-ray spectral analysis

# XSPEC Tutorial and Statistics Basic steps for X-ray spectral analysis

Where can I find XSPEC?

- XSPEC is part of the NASA HEASoft software suite (FTOOLS)
- The latest version is HEASoft 6.301 (April 2022) xspec V12.12.1
   <a href="https://heasarc.gsfc.nasa.gov/docs/software/heasoft/">https://heasarc.gsfc.nasa.gov/docs/software/heasoft/</a>
- Supported architectures:
  - macOS/Mac OS X M1/ARM
  - PC Linux Ubuntu (or other Debian-based Linux)
  - PC Linux Fedora (or other RPM-based Linux)

**XSPEC** is a command-driven, interactive, X-ray spectral-fitting program, designed to be completely detector-independent

Alternatively: **Sherpa** https://cxc.cfa.harvard.edu/sherpa/

# Outline

- Grouping' the data. Loading data in xspec & setup commands
- Response matrices
- A quick journey through models in xspec
- Binned vs. unbinned data
- □ Starting model & spectral fit
- Evaluation of the goodness of the fit
- Gaussian vs. Poissonian regime
- Adding spectral components
- F-test
- Contour plots
- Final fitting solution
- Errors on the parameters
- □ Fluxes and luminosities (and errors)
- □ Save/restore working session
- Some additional and useful commands

# Step 1a: 'grouping' spectra (if not already done)

Once X-ray spectra are extracted and response matrices are produced – see Chandra and XMM-Newton Tutorials – four files (fits format) are needed within XSPEC

Chandra ACIS-S data in this example

- source spectrum 3C33\_r3.pi
- background spectrum 3C33\_r3\_bkg.pi
- ARF response matrix 3C33 r3.corr.arf
- RMF response matrix 3C33\_r3.rmf

Before loading these files in XSPEC, it is better to:

- (1) 'associate' a source spectrum with its background file and response matrices (RMF and ARF); this is done automatically for Chandra by running the ciao tool *specextract*;
- (2) group the spectral counts using a binning of e.g. 20 counts/bin (depending on the source photon statistics and the 'sampling' of the instrument spectral resolution) and allow application of the  $\chi^2$  statistics (it is required to be in the Gaussian regime in each spectral bin). Alternatively, in low-photon counting regime, use unbinned (or binned to 1 count/bin) data and **Cash statistics**

input src spectrum output src spectrum list of commands chkey: change key param in the spectrum fits file
 grppha 3C33\_r3.pi 3C33\_r3\_c20.pi comm="group min 20 & chkey BACKFILE 3C33\_r3\_bkg.pi & chkey ANCRFILE 3C33\_r3.corr.arf & chkey RESPFILE 3C33\_r3.rmf & exit"
 → 3C33\_r3\_c20.pi Name all of the files properly!

# Step 1b: loading data in XSPEC and "setup" commands

#### xspec

xspec> data 3C33\_r3\_c20.pi

- xspec> ignore bad
- xspec> ignore \*\*-0.3 7
- xspec> cpd /xw
- xspec> plot Idata
- xspec> setplot command r y

- load the grouped spectrum (produced in step 1a)
- ignore spectral bins flagged as bad (typically, at low/high E)
- ignore spectral bins below 0.3 keV and above 7 keV (to be verified on the data, check src. vs. back level)
- change the plotting device (e.g., on screen here; on PS file)
- plot the data in log scale
- **r**escale the y axis

Energy: with "." Otherwise: interpreted as channels if integer (conversion made via the RMF) o The '-' sign indicates a range o \*\*: means 'everything'

# Loading multiple datasets

In case of **multiple datasets:** adoption of the same model to carry out a simultaneous X-ray spectral analysis (taking advantage of the 'increased' photon statistics)

```
all spectra are fits files (irregardless of the name, as .pi here)
xspec> data 1:1 spectrum1.pi 2:2 spectrum2.pi 3:3 spectrum3.pi
load all datasets at the same time
xspec> ignore 1-3:**-0.3 7.0-** select the proper energy range for all datasets (1-3)
xspec> cpd /xw
```

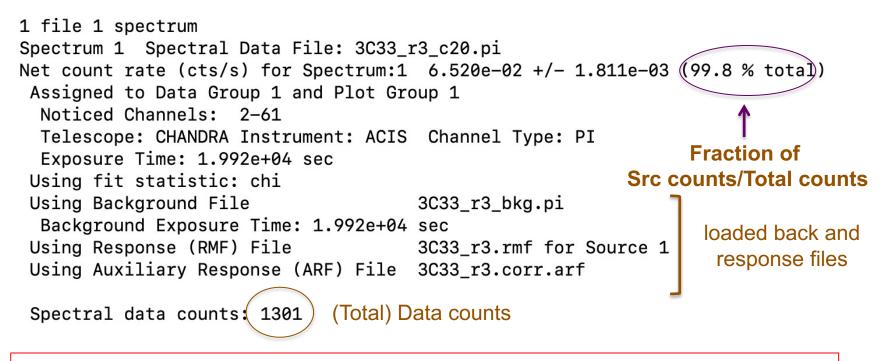
xspec> plot Idata

When **multiple datasets** are used, remember to place in front of all models the **constant model** [e.g., mo **cons**(pha\*po+...)]. This takes into account:

- (a) cross-calibration uncertainties among different instruments of the same telescope (typically, a few percent) and different instruments of different telescopes;
- (b) some possible source flux variability in case of multiple observations not taken simultaneously.

The first constant should be fixed to 1, the others are left free to vary (i.e., are part of the fitting process).

XSPEC version: 12.12.1 Build Date/Time: Thu Mar 31 20:12:13 2022



**Source net** (i.e. background-subtracted) **counts** = data counts × fraction =  $=1301 \times 0.998 \rightarrow$  in this case the source dominates the signal

### Possible binning choices depending on the source photon statistics:

- have enough counts (e.g., 20-25) in each bin and then apply the  $\chi^2$  statistics;
- one count/bin and apply the Cash statistics (C-stat, named W-stat if background is subtracted);
- 'sampling' the spectral resolution of the data

## Response matrices: RMF

RMF: links the instrumental channel scale with the physical energy (wavelength) scale
fv 3C33\_r3.rmf

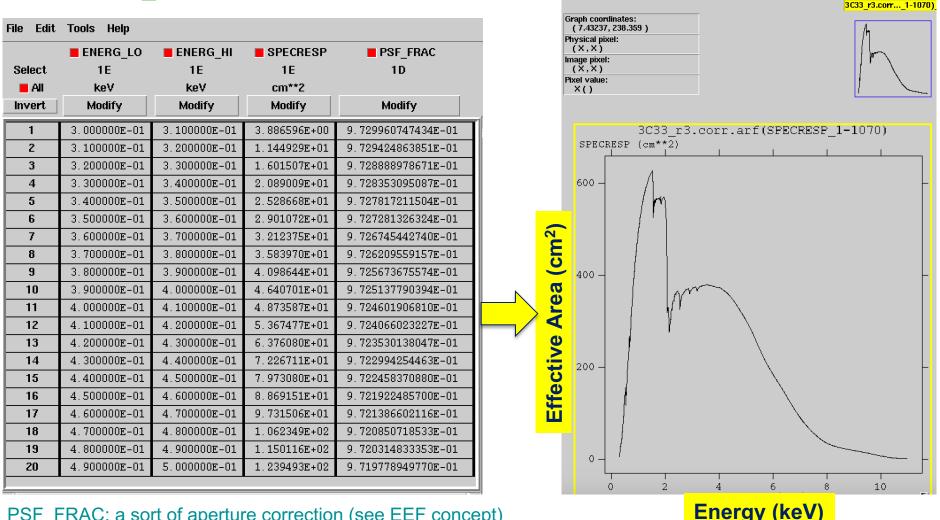
File Edit T	ools Help									File Edit Colors Tools Zoom	Replot Help
Index	Extension	Type D	imension			View					3C33 r3.rmf( 1-1024) 0
0	Primary	Image	0	Header	lm	age (	Т	able		Graph coordinates:	
	MATRIX	Binary 6 col:	: X 1070 rows	Header	Hist	Plot	All	Select		(X,X) Physical pixel:	- / /
2	EBOUNDS	Binary 3 col:	s X 1024 rows	Header	Hist	Plot	All	Select		(X,X) Image pixel:	
File Edit	Tools Help									(X,X) Pixel value:	-
	CHANNEL	E_MIN	E_MAX							×()	
Select	1J	1E	1E								A
E All	channel	keV	keV								_r3.rmf(E_MIN_1-1024)
Invert	Modify	Modify	Modify							E_MIN (keV)	
1	1	7.300000 <b>E</b> -03	1.460000E-02	T						15 -	⊢
2	2	1.460000E-02	2.920000 <b>E</b> -02								
3	3	2.920000E-02	4.380000E-02								
4	4	4.380000E-02	5.840000E-02								
5	5	5.840000E-02	7.300000E-02								
6	6	7.300000E-02	8.760000 <b>E</b> -02						S		
7	7	8.760000 <b>E</b> -02	1.022000E-01					N	e	10 -	
8	8	1.022000E-01	1.168000E-01						Ľ		
9	9	1.168000E-01	1.314000E-01						>		
10	10	1.314000E-01	1.460000E-01						<b>D</b>	<mark>מ</mark>	
11	11	1.460000E-01	1.606000E-01					V	<b>e</b>		
12	12	1.606000E-01	1.752000E-01						Energy (keV)	5	
13	13	1.752000E-01	1.898000E-01								
14	14	1.898000E-01	2.044000E-01								
15	15	2.044000E-01	2.190000E-01	_							
16	16	2.190000E-01	2.336000 <b>E-</b> 01	_							
17	17	2.336000E-01	2.482000E-01	-							
18	18	2.482000E-01	2.628000E-01	_						0-	Channel _
19	19	2.628000E-01	2.774000E-01	_							
20	20	2.774000E-01	2.920000E-01							0 200	400 600 800 1000

### **Response matrices: ARF**

**ARF**: indicates the effective response (hence sensitivity) of the mirrors+instrument at a given source position on the detector

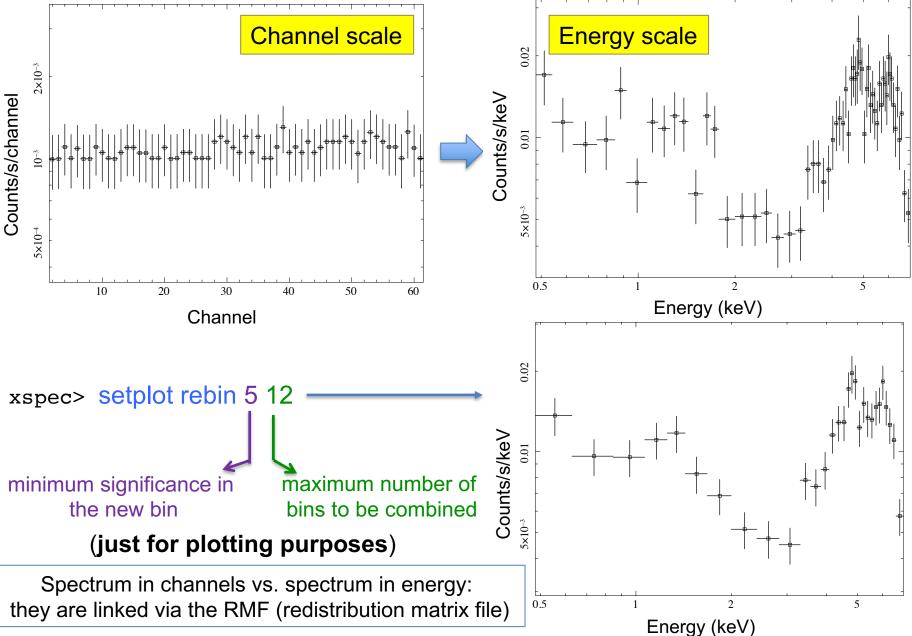
Edit Colors Tools Zoom Replot Help

• fv 3C33 r3.corr.arf



PSF FRAC: a sort of aperture correction (see EEF concept)

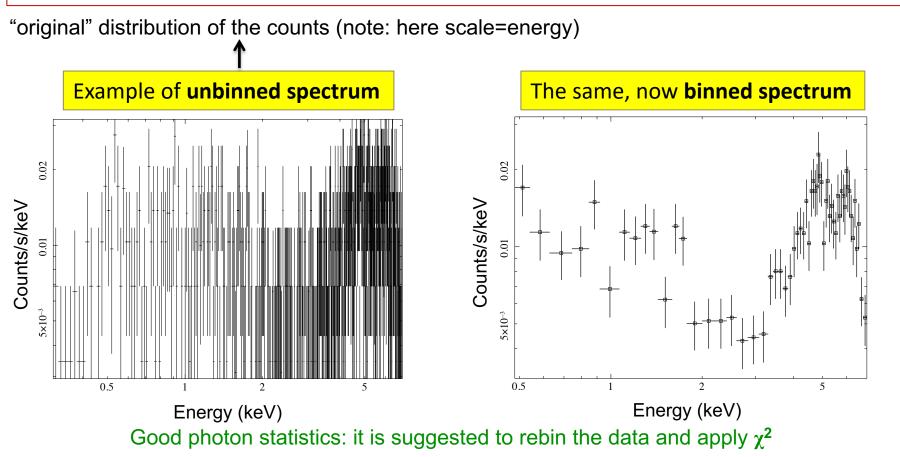
### xspec> setplot energy Channel scale = instrument scale Energy scale = "physical" scale



# What does 'binning' (grouping) mean?

To apply the  $\chi^2$  test, we need that in every bin the statistics is nearly in the Gaussian regime, i.e., there is a sufficiently high number of counts in each spectral bin (datapoint)

Using the ftool grppha (or similar), we can require that each spectral bin contains at least a given number of counts (see step 1a)



# Step 2: `families' of xspec models

XSPEC models used like in math operations

### Additive models

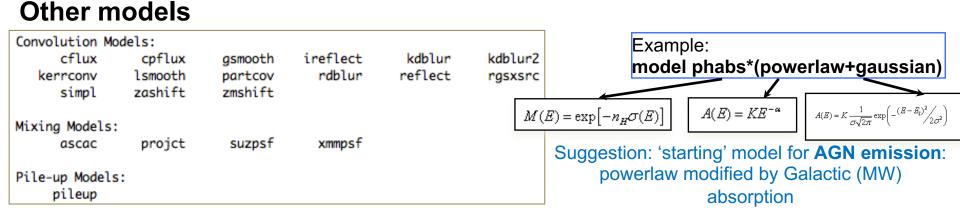
agauss	apec	bapec	bbody	bbodyrad	bexrav	
bexriv	bkn2pow	bknpower	bmc	bremss	bvapec	
bvvapec	c6mekl	c6pmekl	c6pvmkl	c6vmekl	cemekl	
cevmkl	cflow	compLS	compPS	compST	compTT	
compbb	compmag	comptb	compth	cplinear	cutoffpl	
disk	diskbb	diskir	diskline	diskm	disko	
diskpbb	diskpn	eplogpar	eqpair	eqtherm	equil	
expdec	ezdiskbb	gadem	gaussian	gnei	grad	
grbm	kerrbb	kerrd	kerrdisk	laor	laor2	
logpar	lorentz	meka	mekal	mkcflow	nei	
npshock	nsa	nsagrav	nsatmos	nsmax	nsmaxg	
nsx	nteea	nthComp	optxagn	optxagnf	pegpwrlw	
pexmon	pexrav	pexriv	plcabs	posm	powerlaw	
pshock	raymond	redge	refsch	rnei	sedov	
sirf	smaug	srcut	sresc	step	vapec	
vbremss	vequil	vgadem	vgnei	vmcflow	vmeka	
vmekal	vnei	vnpshock	vpshock	vraymond	vrnei	
vsedov	vvapec	vvgnei	vvnei	vvnpshock	vvpshock	
vvrnei	vvsedov	zagauss	zbbody	zbremss	zgauss	
zpowerlw		_	-		-	

### **Multiplicative models**

SSS_ice	TBabs	TBgrain	TBvarabs	absori	acisabs
cabs	constant	cyclabs	dust	edge	expabs
expfac	gabs	heilin	highecut	hrefl	lyman
notch	pcfabs	phabs	plabs	pwab	recorn
redden	smedge	spexpcut	spline	swind1	uvred
varabs	vphabs	wabs	wndabs	xion	zTBabs
zbabs	zdust	zedge	zhighect	zigm	zpcfabs
zphabs	zredden	zsmdust	zvarabs	zvfeabs	zvphabs
zwabs	zwndabs	zxipcf			

#### Syntax: M1\*M2\*(A1+A2+M3\*A3)

# **M=multiplicative model:** modifies incident flux **A=additive model**: source of emission



# Step 3a: starting model + spectral fit

### Absorption due to our Galaxy (MW): you need to **include it in all spectral models**. All photons pass through our own Galaxy

#### xspec> nh

>>>>>> NH version 3

Equinox (d/f 2000)[2000]

RA in hh mm ss.s or degrees[159.386] 01 08 52.86 DEC in dd mm ss.s or degrees[56.171] 13 20 14.2

Your source's coordinates (3C33 here)

```
>> Using map h1_nh_HI4PI.fits
```

```
LII , BII 129.448839 -49.313559
Requested position at X and Y pixel 932.06 321.88
Search nH in 12 X 12 pixels box
Each pixel is 0.083 deg 0.083 deg
```

DEC	Dist(deg	g) nH
13.2623	0.0812	2.93E+20
13.2564	0.0940	2.89E+20
13.3530	0.0709	3.03E+20
13.3472	0.0140	2.95E+20
13.3413	0.0890	2.91E+20
	13.2623 13.2564 13.3530 13.3472	13.26230.081213.25640.094013.35300.070913.34720.0140

nH calculated using all points within 0.1000 deg from input position
h1\_nh\_HI4PI.fits >> Average nH (cm\*\*-2) 2.94E+20
h1\_nh\_HI4PI.fits >> Weighted average nH (cm\*\*-2) 2.96E+20

NH\_HI4PI.fits >> Weighted average nH (cm\*\*-2) 2.96E+20

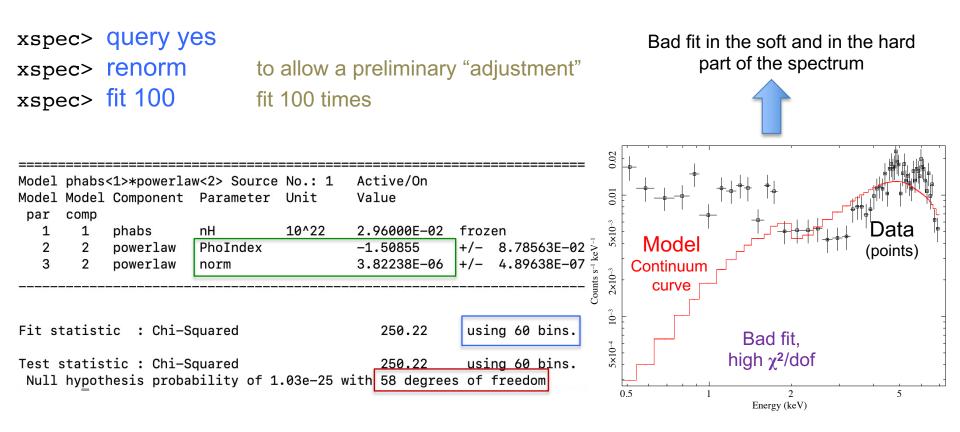
Alternatively (web tool): <u>https://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3nh/w3nh.pl</u> based on the HI4PI Survey (N.B. Bekhti et al. 2016, A&A, 594, A116)

#### Suggestion: start with a simple modeling (as a powerlaw modified by Galactic absorption)

### xspec> mo pha\*po

**pha**: accounts for the Galactic N<sub>H</sub> (multiplicative model)**po**: powerlaw model (additive model) for the primary AGN comp.

[XSPEC12>mo pha*po	It is possibl	e to provide va step of th			parameters at every cess
[1:phabs:nH>2.96e-2-1 1 [2:powerlaw:PhoIndex>1.9	.001( 0.01 0.01( 0.01 9 0.01( 0.01 -1 n	) 0 ) -3	0 –2 0 eter (the sa	10000 1e+2 ame as	9 10 20 1e+24 5 using the
	Parameter Uni		+/- 0.0	83	Alternatively, you may assign later a value to a parameter using xspec> newpar 2 1.9
Fit statistic : Chi-So Test statistic : Chi-So Null hypothesis probab Current_data and mode	quared bility of 3.89e	1084.55 1084.55 -189 with 58 degre	using 6	50 bins	
rameter number of umber the compone	nt				Model: pha(po)



dof=degrees of freedom=(number of datapoints – number of free parameters)=60-2=58

- χ<sup>2</sup>/dof close to unity means that it is a good fit (here: 250/58: not in this case!) see lesson on statistics and the following slides
- <u>Null hypothesis probability</u>=probability that the model is a good representation of the datapoints (i.e., good if close to 1)

All the adopted models should be physically motivated according to the known source (multi-wavelength) properties & classification Model: pha(po)

# To evaluate the goodness of the fit: the $\chi^2$ statistical test

Test to compare the observed distribution of the results with that expected

$$\chi^2 = \sum_{k=1}^n \frac{(O_k - E_k)}{\sigma_k^2}^2$$

O<sub>k</sub>=observed values (spectral datapoints)

E<sub>k</sub>=expected values (model)

 $\sigma_k$ =error on the measured values (error on each spectral bin) k=number of datapoints (bins after rebinning)

$$\chi^2 / dof \approx 1$$



the observed and expected distributions are similar

### Applicability of $\chi^2$ statistics

### $\chi^2$ statistic

$$S = \sum_{i} \left( \frac{S_{i} - B_{i}t_{s}}{O_{\mathsf{K}}} / t_{b} - \frac{m_{i}t_{s}}{\mathsf{E}_{\mathsf{K}}} \right)^{2} / \left( \frac{(\sigma_{s})_{i}^{2} + (\sigma_{B})_{i}^{2}}{\mathsf{\sigma}_{\mathsf{K}}} \right)^{2}$$

- $S_i = \text{src counts in the } I=\{1,...,N\} \text{ data bins with exposure } t_S;$
- $\circ$  B<sub>i</sub> = background counts with exposure t<sub>B</sub>;
- $\circ$  m<sub>i</sub> = model predicted count rate;
- $\circ$  (σ<sub>S</sub>)<sup>2</sup> and (σ<sub>B</sub>)<sup>2</sup> = variance on the src and background counts, typically estimated by S<sub>i</sub> and B<sub>i</sub>

### BUT

the  $\chi^2$  statistic fails in low-counting regime (few counts in each data bin)

### Alternative solutions in case of low photon statistics

- i. To rebin the data so that each bin contains a large enough number of counts
- BUT: loss of information and dependence on the rebinning method adopted
- ii. To modify S so the it performs better in low-count regime (e.g., by estimating the variance for a given data bin using the average counts from the surrounding bins; Churazov+96)
- BUT: it would need Montecarlo simulations to properly support the result
- iii. To construct a **maximum-likelihood estimator** based on the Poisson distribution of the detected counts (Cash79; Wachter+79). ML means finding the best fit of parameters that maximizes the Poisson likelihood

xspec> statistic chi (default) xspec> statistic cstat

Binned data,  $\chi^2$  statistics  $\Leftrightarrow$  Gaussian statistics Unbinned data, C-statistics  $\Leftrightarrow$  Poisson statistics

# $\chi^2$ in a nutshell

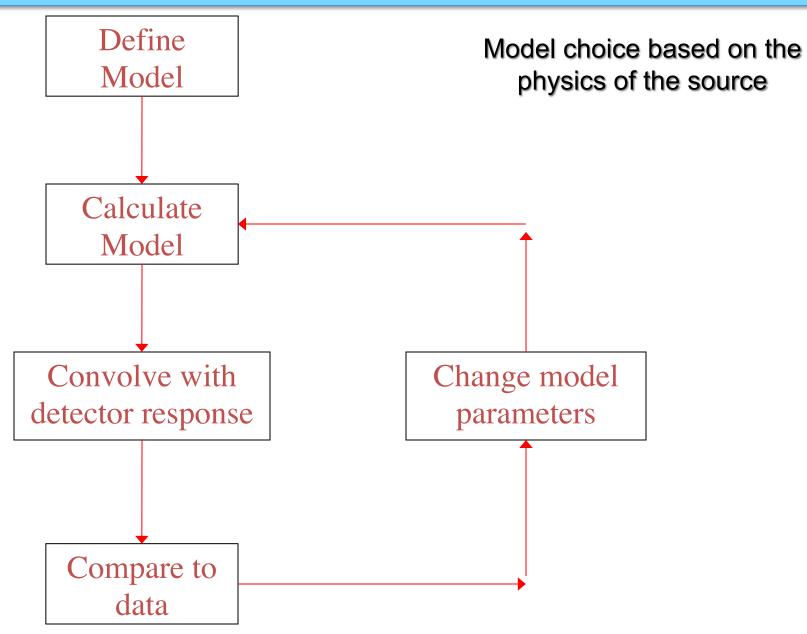
# Reduced $\chi^2$ large $\leftarrow \rightarrow P(\chi^2)$ small

- a. Errors are under-estimated
- b. The model does not describe the data correctly

# Reduced $\chi^2$ small $\leftarrow \rightarrow P(\chi^2)$ large

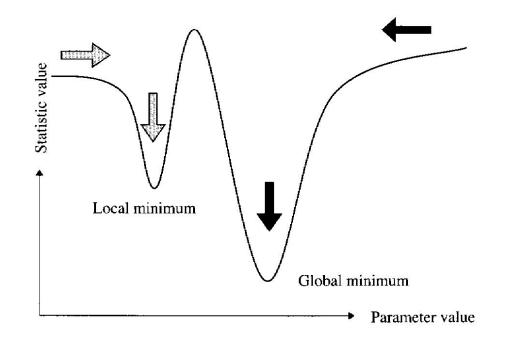
- c. Errors are over-estimated
- d. Were data "selected" in a particular way?

# Forward-fitting algorithm



### Global vs. local minimum

Data analysis



If the fit process is started at the "right place", then it will converge to the true minimum

The more complicated the model and the more highly correlated the parameters, then the more likely that the algorithm will hardly find the true minimum

To 'move' the fit from a local miminum, you can change some of the parameters using the *newpar* command and then fit again

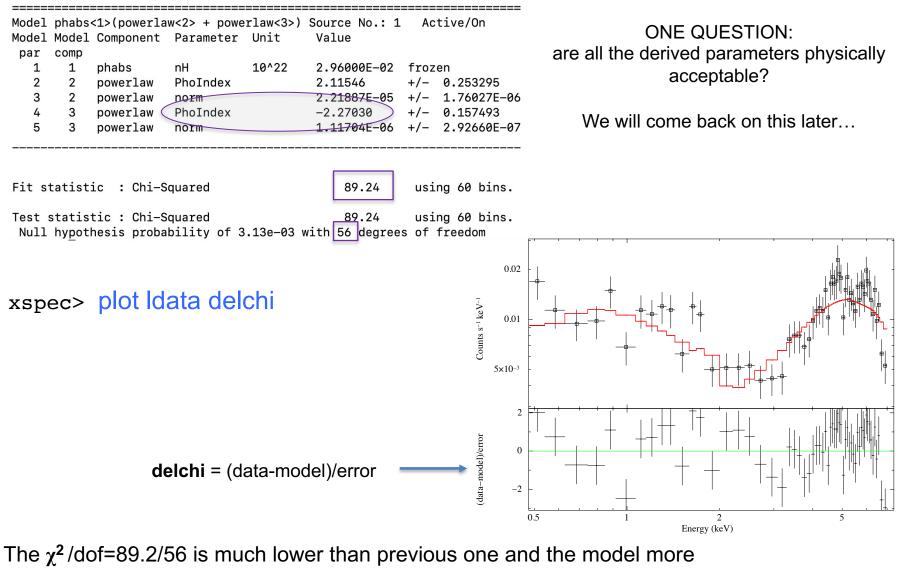
### Step 3b: adding components and fit. I

xspec> addcomp 2 powerlaw adding a powerlaw as # component (#=order in the model)

Input parameter	c value,	delta, min	, bot, to	op, and max va	lues for	• • •	
1	L	0.01(	0.01)	-3	-2	9	10
2:powerlaw:Phol	Index>1.	8					
1	L (	0.01(	0.01)	0	0	1e+20	1e+24
3:powerlaw:norm	n>1e-5						
Fit statistic	: Chi-S	quared		164.18	using 6	0 bins.	
	a				• •		
Test statistic			(7- 10		using 6		(
Null hypothesi Current data a				vith 56 degree	s of free	aom	
Guilent uata a	ind mode.	I NOT TIT y	el.				
							SC
Model phabs<1>(	powerla	w<2> + powe	rlaw<3>)	Source No.: 1	Active	/0n	(s
Model Model Com	ponent	Parameter	Unit	Value			(0
par comp							
1 1 pha	abs	nH	10^22	2.96000E-02	frozen	_	
	verlaw	PhoIndex		1.80000	+/- 0.0		
3 2 pov	verlaw	norm		1.00000E-05	+/- 0.0		
	verlaw	PhoIndex		-1.50636	+/- 0.0		
5 3 pov	verlaw	norm		3.83424E-06	+/- 0.0		

Inclusion of an additional powerlaw component to account for the residuals observed at low energies: scattering component? phenomenological (simplistic) parameterization of something different?

### xspec> fit 100



properly reproduces the observed spectral datapoint distribution. There are yet *some residuals* (bottom panel: data-model, i.e. deviations in units of  $\sigma$ =stat. error)

Model: pha(po+po)

# Step 3b: adding components and fit. II

#### xspec> addcomp 3 zpha adding zpha=absorption intrinsic to the source as third component

Input parameter value, delta, min, bot, top, and max values for 1 0.001( 0.01) 0 0 100000	1e+06
4:zphabs:nH>1	
0 -0.01( 0.01) -0.999 -0.999 10	10
5:zphabs:Redshift>0.06 z=0.06	
Fit statistic : Chi-Squared 93.47 using 60 bins.	J
	CO
Test statistic : Chi-Squared 93.47 using 60 bins.	
Null hypothesis probability of 9.34e-04 with 55 degrees of freedom	
Current data and model not fit yet.	
Model phabs<1>(powerlaw<2> + zphabs<3>*powerlaw<4>) Source No.: 1 Active	′0n –
Model Model Component Parameter Unit Value	
par comp	
1 1 phabs nH 10^22 2.96000E-02 frozen	
2 2 powerlaw PhoIndex 2.11546 +/- 0.253295	
3 2 powerlaw norm 2.21887E-05 +/- 1.76027E-06	
4 3 zphabs nH 10^22 1.00000 +/- 0.0	
5 3 zphabs Redshift 6.00000E-02 frozen	
6 4 powerlaw PhoIndex -2.27030 +/- 0.157493	
7 4 powerlaw norm 1.11704E-06 +/- 2.92660E-07	

Justification: the inverted slope of the continuum (negative photon index) may be ascribed to the presence of obscuration

 $\rightarrow$  Is the newly derived photon index (after the inclusion of  $N_{H_2}$  parameter 6, component 4) more consistent with what is expected in case of an AGN (Γ=1.8-2)?

Model: pha(po+zpha(po))

### xspec> fit 100

=====	======			========			=====
Model	phabs	<1>(powerla	w<2> + zpha	bs<3>*pow	erlaw<4>) Sou	rce No.: 1	Active/On
Model	Model	Component	Parameter	Unit	Value		
par	comp						
1	1	phabs	nH	10^22	2.96000E-02	frozen	
2	2	powerlaw	PhoIndex		1.14238	+/- 0.1618	801
3	2	powerlaw	norm		2.45188E-05	+/- 1.6921	7E-06
4	3	zphabs	nH	10^22	22.0419	+/- 5.5921	.3
5	3	zphabs	Redshift		6.00000E-02	frozen	
6	4	powerlaw	PhoIndex		-0.539341	+/- 0.5286	88
7	4 🖵	powerlaw	norm		3.81227E-05	+/- 3.4321	<del>.7E-0</del> 5

Fit statistic : Chi-Squared

76.04 using 60 bins.

Counts s-1 keV-1

data-model)/error

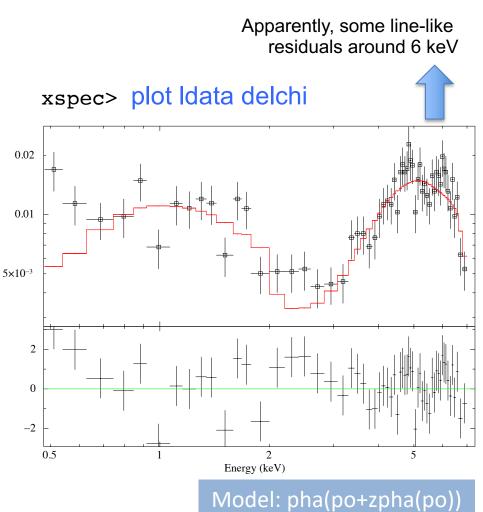
Test statistic : Chi-Squared 76.04 using 60 bins. Null hypothesis probability of 3.16e-02 with 55 degrees of freedom

### $\chi^2$ /dof=76.0/55 now $\rightarrow$ improvement in the fit

However, the photon index (parameter 6, component 4) is still negative

→ likely, the hard-band powerlaw and the column density are *degenerate parameters*, hence constraining both over the limited hard band of Chandra is challenging

VIABLE SOLUTIONS: Link the photon indices of the two powerlaw as in the case of Thomson scattering in the soft band) – see the following slides



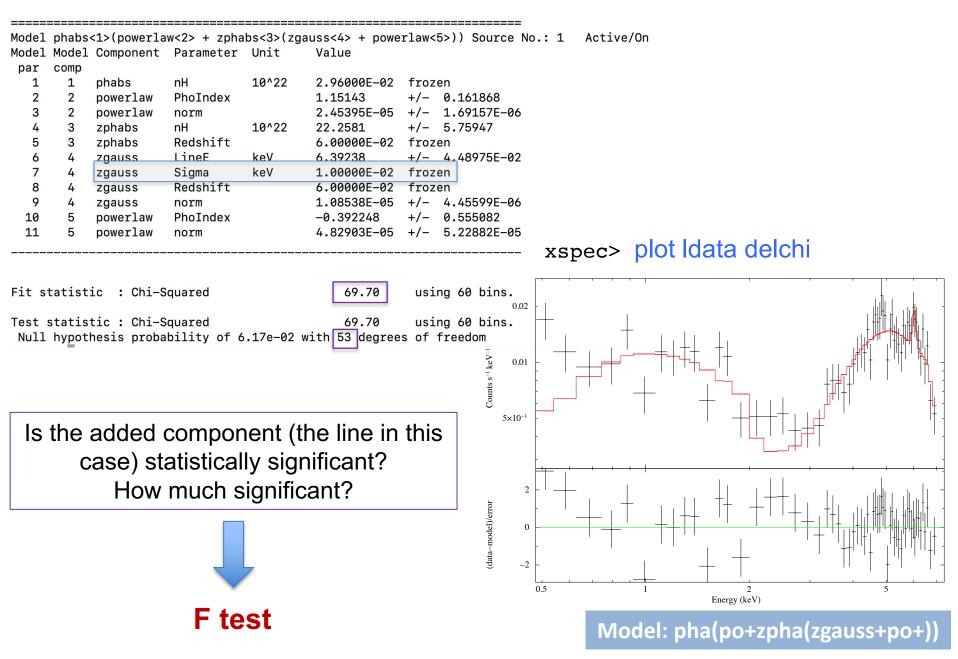
## Step 3b: adding components and fit. III

### xspec> addcomp 4 zgauss adding zgauss=emission line as fourth component

		6.5		, bot, to .065)	p, and max va 0	alues for 0	 1e+06	1e+06	
[6:zgau	uss:Li	neE>6.4 0.1	0.05( 0	.001)	0	0	10	20	I have
7:zgau	uss:Sig	gma>0.01,-1							line at
0				0.01)	-0.999 -	-0.999	10	10	the n
8:zgal	uss:Re	dshift>0.06 1		0.01)	0	0	1e+20	1e+24	
[9:zgau	uss:no:	rm>1e-6	0.01(	0.01)	U	U	TETZO	16+24	
Fit s	tatist	ic : Chi-S	quared		74.92	using 6	0 bins.		
Null	hypotl				74.92 ith 53 degree	using 6 es of free			
====== Model	nhahs	======================================	======================================	======================================	======================================	======================================	Source No	• 1 Ac	tive/On
			Parameter		Value	/110W(0///	oddiec no.	• • • •	, CIVC/ 011
par	comp								
1	1	phabs	nH	10^22	2.96000E-02	frozen			
2	2	powerlaw	PhoIndex		1.14238	+/- 0.1	61801		
3	2	powerlaw	norm		2.45188E-05	+/- 1.6			
4	3	zphabs	nH	10^22	22.0419	+/- 5.5	9213		
5	3	zphabs	Redshift		6.00000E-02	frozen			
6	4	zgauss	LineE	keV	6.40000	+/- 0.0			
7	4	zgauss	Sigma	keV	1.00000E-02	frozen			
8	4	zgauss	Redshift		6.00000E-02	frozen			
9	4	zgauss	norm		1.00000E-06	+/- 0.0			
10	5	powerlaw	PhoIndex		-0.539341	+/- 0.5	28688		
11	5	powerlaw	norm		3.81227E-05	+/- 3.4	3217E-05		

I have included a narrow ( $\sigma$ =10 eV, fixed) line at the expected rest-frame energy of the neutral iron K $\alpha$  transition (6.4 keV)

### xspec> fit 100

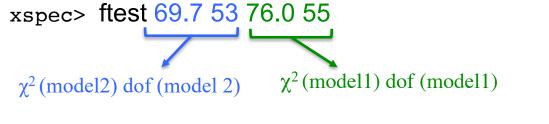


## Step 4: the F-test

(here applied to estimate the statistical significance of the inclusion of an emission line)

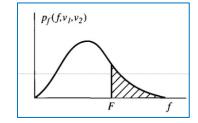
- > Model 1: double powerlaw + obscuration:  $\chi^2$  /dof=76.0/55
- > Model 2: double powerlaw + obscuration + iron emission line:  $\chi^2$  /dof=69.7/53

 $\Delta \chi^2 / \Delta dof = 6.3/2$ 



F statistic value = 2.39527 and probability 0.10095

XSPEC12>ftest 69.7 53 76.0 55



Large F value → low probability (of exceeding that value, see tables) → highly significant improvement due to the additional component

→ The iron line has low significance: P(real line)=1–0.10095~0.90 → ~1.6 $\sigma$ 

Use the F-test to evaluate the improvement to a spectral fit due to the assumption of a different model, with additional terms

Conditions:

(a) the simpler model is nested within the more complex model;

(b) the extra parameters have Gaussian distribution (not truncated by the parameter space boundaries) – BUT see also Protassov+02 on caveats

 $P_f(f;v_1,v_2) = \frac{\chi_1/v_1}{\chi_2/v_2} \propto \Delta \chi^2/k$ 

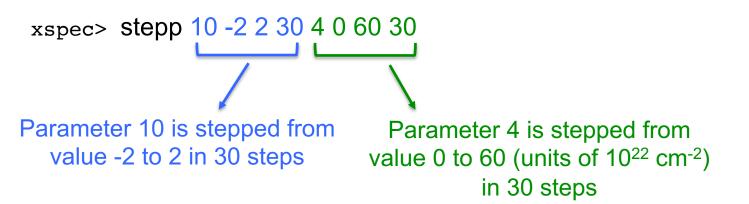
The larger this ratio, the larger the improvement in the spectral fitting

k=number of additional parameters

## Step 5a: contour plots and error computation

Use the *steppar* command to compute errors simultaneously for two parameters and visualize them using contour plots it performs a fit while stepping the values of two parameters through a given range

Here: photon index (param. 10) vs. column density (param. 4)

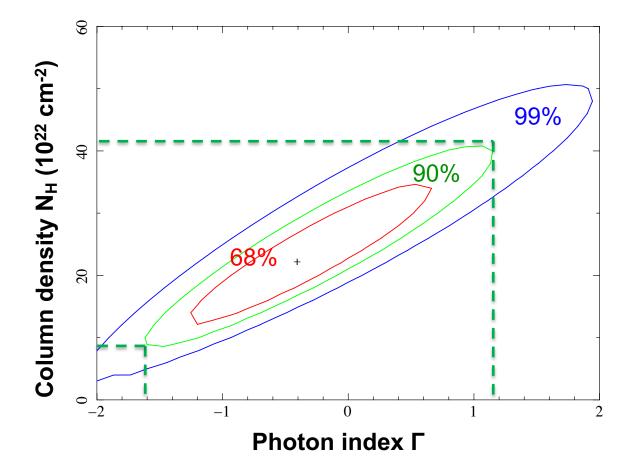


It provides how one parameter varies wrt. another parameter (i.e., the error range for sets of parameters)

Useful also to understand whether two spectral parameters are correlated (not necessarily from a physical point of view)

È th	at each step he process of 'mov rough the selected ranges of the two parameters) Chi-Squared	<sup>d</sup> wrt.		in $\chi^2$ (hence, $\Delta$ st-fitting solution		nH		
	chi Squarcu	Chi-Squared		10		4		Parameters involved in the fit
1		chir oquarea	, i					
	82.257	12.553	0	-2	0		0	
	97.112	27.408	1	-1.6	0		0	
	130.85	61.143	2	-1.2	0		0	
	186.48	116.78	3	-0.8	0		0	
	265.11	195.4	4	-0.4	0		0	
	363.95	294.25	5	0	0		0	
	474.76	405.06	6	0.4	0		0	
	585.89	516.18	7	0.8	0		0	
	687.28	617.57	8	1.2	0		0	
	772.51	702.81	9	1.6	0		0	
	840.9	771.2	10	2	0		0	
	462.99	393.28	10	2	1		6	
	406.58	336.87	9	1.6	1		6	
	347.9	278.19	8	1.2	1		6	
	288.93	219.23	7	0.8	1		6	
	232.32	162.61	6	0.4	1		6	
	181.08	111.38	5	0	1		6	
	138.18	68.48	4	-0.4	1		6	
	105.93	36.229	3	-0.8	1		6	
	85.501	15.797	2	-1.2	1		6	
	76.667	6.963	1	-1.6	1		6	
	77.912	8.2079	0	-2	1		6	
	83.7	13.996	0	-2	2		12	
	74.987	5.2834	1	-1.6	2		12	

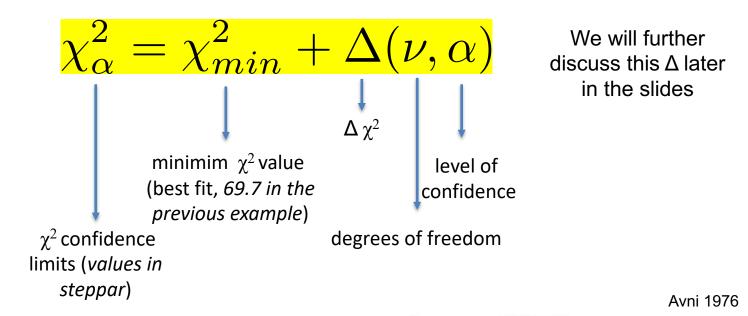
### xspec> plot contour



- 90% confidence level: the photon index varies in the range ~ [-1.6, 1.2], while the column density varies in the interval ~ [9-41] × $10^{22}$  cm<sup>-2</sup>
- The photon index and the column density are degenerate parameters
- We can decide to link the photon indices of the two powerlaws (as in case of scattering)

# The meaning of contour plots/confidence regions

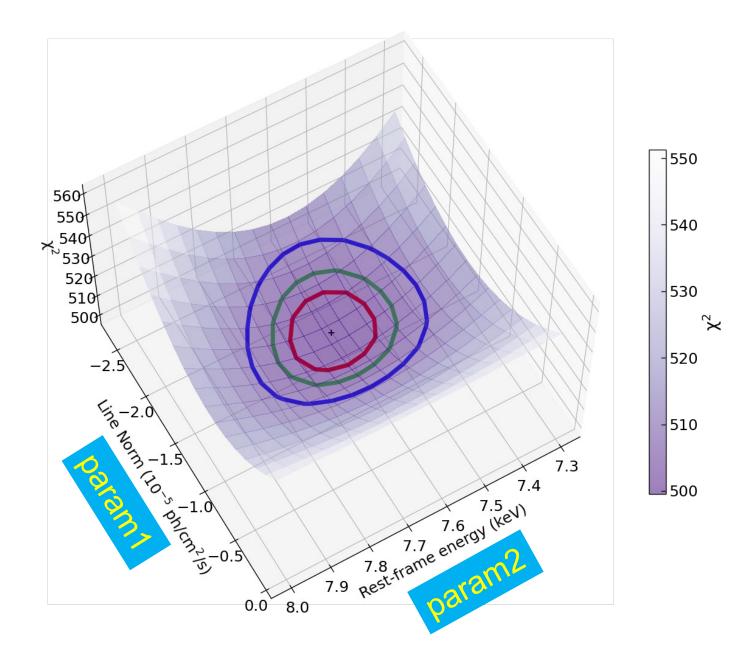
The contour plots define a confidence region in the parameter space (i.e., the "statistical surface") within which the true parameters lie with a certain confidence (hence, 68, 90, 99% in XSPEC by default). They represent regions of constant probability



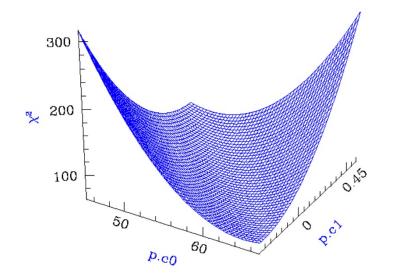
 $\Delta$  depends only on the number of parameters involved in the fit (see previous slides)

CONSTANTS FOR CALCULATING CONFIDENCE REGIONS

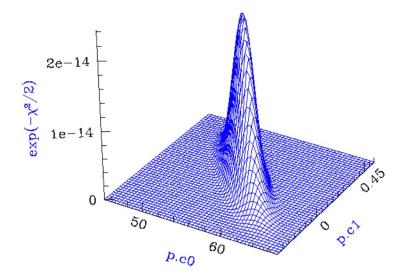
	q (No. of Interesting Parameters)							
(%)	1	2	3					
68	1.00	2.30	3.50					
90	2.71	4.61	6.25					
99	6.63	9.21	11.30					



Courtesy of E. Bertola

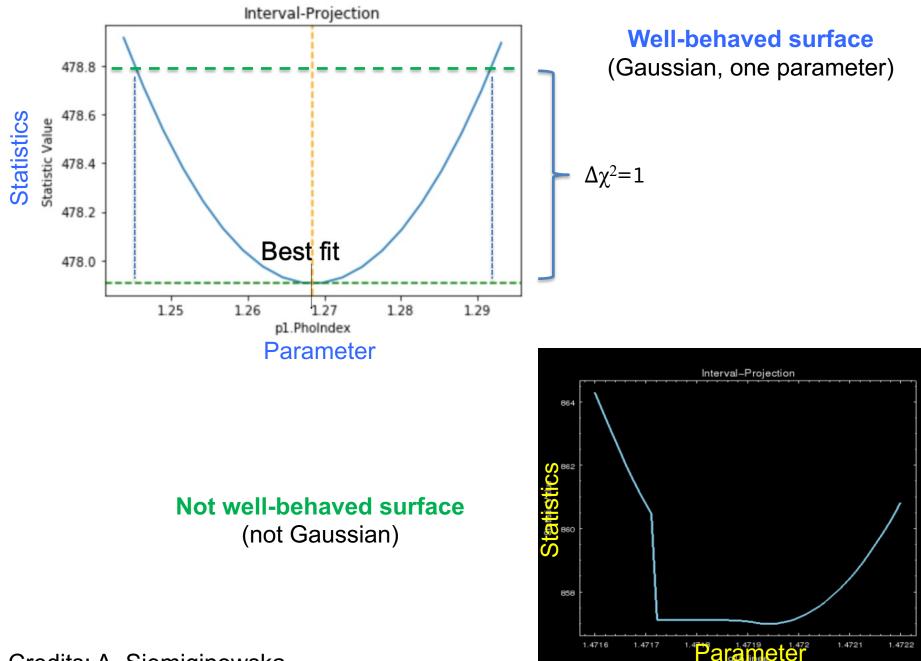


Calculating Confidence Limits means Exploring the Parameter Space - Statistical Surface



Example of a "well-behaved" statistical surface in parameter space, viewed as a multi-dimensional paraboloid ( $\chi^2$ , top), and as a multi-dimensional Gaussian (exp(- $\chi^2$  /2) ≈ L, bottom).

Credits: A. Siemiginowska



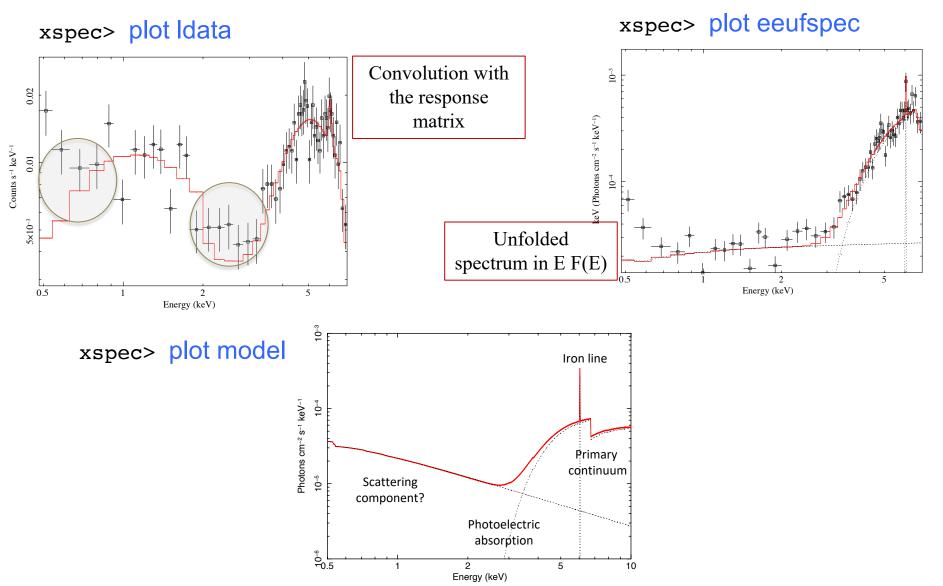
Credits: A. Siemiginowska

## Step 5b: towards the final fitting solution

xspec> newpar 2 = 10 We link the photon index of the secondary (soft) component to that of the primary one (as expected in case of scattering)

Model	phabs	<1>(powerl	aw<2>	+ zpha	bs<3>(zgau	ss<4> + powe	rlaw<5>))	Source No.: 1	1 Act	ive/On
		Component				Value				
par	comp									
1	1	phabs	nH		10^22	2.96000E-02	frozen			
2	2	powerlaw	PhoI	ndex		-0.392248	= p10			
3	2	powerlaw	norm	1		2.45395E-05				
4	3	zphabs	nH			22.2581	+/- 0.0			
5 6 7	3	zphabs		hift		6.00000E-02				
6	4	zgauss	Line			6.39238	+/- 0.0			
7	4	zgauss	Sigm			1.00000E-02				
8	4	zgauss		hift		6.00000E-02				
9	4	zgauss	norm			1.08538E-05				
10	5	powerlaw		ndex		-0.392248	+/- 0.0			
11	5	powerlaw	norm			4.82903E-05	+/- 0.0			
xspe	9C>					aw<2> + zpha Parameter nH PhoIndex norm nH Redshift LineE		uss<4> + powe Value 2.96000E-02 0.935292 2.35020E-05 37.6379 6.00000E-02 6.39929	froze = p10 +/- +/- froze	n 1.69263E-06 2.73295
			7	4	-	Sigma	keV	1.00000E-02		
			8	4	zgauss	Redshift	NEV	6.00000E-02	froze	
			o 9	4	zgauss zgauss	norm		1.43843E-05		<sup>n</sup> <sub>5.</sub> Model: pha(po+zpha
			9 10	4 5	powerlaw	PhoIndex		0.935292	+/-	
			11	5	powerlaw	norm		6.78472E-04		0.1 (zgauss+po+))
									+/-	$\Gamma_1 = \Gamma_2$

- Some residuals in the soft band and around 2 keV, where the soft and hard components 'connects' each other
- The photon index is still (nominally) lower than expected → reflection component needed (but <u>not</u> accounted for in this tutorial)



# Step 5c: estimate of parameters uncertainties

To compute errors: error and uncertainty commands in xspec

1odel par	Model comp	Component	Parameter	Unit	Value	[		_					
1 2 3 4 5 6 7 8 9	1 2 3 3 4 4 4 4 4	phabs powerlaw powerlaw zphabs zphabs zgauss zgauss zgauss zgauss zgauss	nH PhoIndex norm nH Redshift LineE Sigma Redshift norm	10^22 10^22 keV keV	2.96000E-02 0.935354 2.35026E-05 37.6728 6.00000E-02 6.39928 1.00000E-02 6.00000E-02 1.43840E-05	<pre>frozen = p10 +/- 1.69262E +/- 2.73288 frozen +/- 4.51013E frozen frozen +/- 5.64652E</pre>	-02	Т		1σ	e the for t met	that	
10 11	5 5	powerlaw powerlaw	PhoIndex norm		0.935354 6.78558E-04	+/- 0.137424 +/- 1.83102E							
							$\Delta \chi^2$ as	a Function	n of Confid	ence Leve	l and Deg	rees of Fre	edom
Test statistic : Chi-Squared       73.33       using 60 bir       90%       2.71       4.61       6.25       7.78       9.24       10.6         Null hypothesis probability of 4.11e-02 with 54 degrees of freedom       99%       6.63       9.21       11.3       13.3       15.1       16.8										6 7.04 10.6 12.8 16.8 20.1 27.8			

Parameter Confidence Range (2.706) 4 33.101 42.8442 (-4.57177,5.17145) 90% confidence level (2.706) for one parameter of interest (Avni 1976; Lampton et al. 1976) – it is the  $\Delta$  parameter seen before

- Ex.1: Error at 90% confidence level for one parameter of interest: xspec> error 2.71 #param
- Ex. 2: Error at 90% confidence level for two parameters of interest: xspec> error 4.61 #param
- Ex. 3: Error at 99% confidence level for one parameter of interest: xspec> error 6.63 #param

Confidence	sigma	delta_chi-square	1 parameter of interest
68.3%	1.0	1.00	
90.0%	1.6	2.71	
95.5%	2.0	4.00	
99.0%	2.6	6.63	
99.7%	3.0	9.00	

1 parameter of interest: as only one parameter at each time would vary

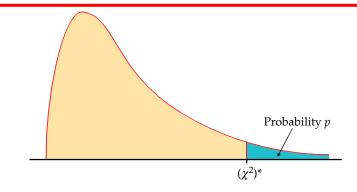
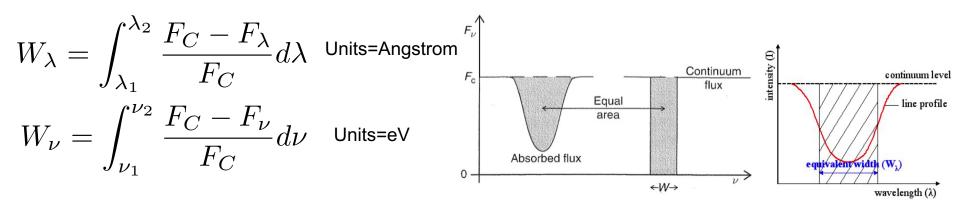


Table entry for *p* is the critical value  $(\chi^2)^*$  with probability *p* lying to its right.

	TABLE F $\chi^2$ distribution critical values												
							Tail prob	ability p					
	df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
Parameters of interest	1   2   3   4   5   1	1.32 2.77 4.11 5.39 6.63	1.64 3.22 4.64 5.99 7.29	2.07 3.79 5.32 6.74 8.12	2.71 4.61 6.25 7.78 9.24	3.84 5.99 7.81 9.49 11.07	5.02 7.38 9.35 11.14 12.83	5.41 7.82 9.84 11.67 13.39	6.63 9.21 11.34 13.28 15.09	7.88 10.60 12.84 14.86 16.75	9.14 11.98 14.32 16.42 18.39	10.83 13.82 16.27 18.47 20.51	12.12 15.20 17.73 20.00 22.11

#### **Uncertainties on the line EW measurement. I**

Wavelength/frequency space definition [Angstrom/keV units] EW is a measure of how prominent a line is  $(F_{\lambda}, F_{v})$  wrt. the continuum  $(F_{c})$ 



xspec> eqw 4

#=model component associated with the Gaussian line

```
Data group number: 1
Additive group equiv width for Component 4: 0.100872 keV
```

xspec> eqw 4 err 100 90 EW including errors at 90% confidence level doing 100 trials

Data group number: 1 Additive group equiv width for Component 4: 0.100872 keV Parameter distribution is derived from fit covariance matrix. Equiv width error range: 0.0414912 - 0.184698 keV

EW=101 [41-185] eV

#### Uncertainties on the line EW measurement. II

Alternatively: assuming that the dominant contribution to the EW error comes from the line intensity (so, limited contribution from the uncertainty on the continuum emission), one can (1) compute the 90% error on the line normalization, (2) place the upper 90% value as line normalization and (3) type *eqw* again (without fitting), then (4) place the lower 90% value as line normalization and (5) type *eqw* again (without fitting)

	<b>U</b> 17	、 <i>,</i> , ,							5	0,
===== Model	phabs	======================================	======================================	======================================	======================================	===== rlaw<	======================================			
		Component		-	Value					
par	comp									
1	1	phabs	nH	10^22	2.96000E-02					
2	2	powerlaw	PhoIndex		0.935354	= p1	10			
3	2	powerlaw	norm		2.35026E-05	+/-	1.69262E-06			
4	3	zphabs	nH	10^22	37.6728	+/-	2.73288			
5	3	zphabs	Redshift		6.00000E-02	froz	zen			
6	4	zgauss	LineE	keV	6.39928	+/-	4.51013E-02			
7	4	zgauss	Sigma	keV	1.00000E-02	froz	zen			
8		zgauss	Redshift		6.00000E-02					
9	4	zgauss	norm		1.43840E-05		5.64652E-06			
10	5	powerlaw	PhoIndex		0.935354		0.137424			
11	5	powerlaw	norm		6.78558E-04	+/-	1.83102E-04			
	-	Totally		66] eV nt with th	e		XSPEC12>ne 9 2.36522e-05 Fit statistic : Chi-Squa Test statistic : Chi-Squa Null hypothesis probabil Current data and model r XSPEC12>eqw 4 Data group number: 1 Additive group equiv widt XSPEC12>ne 9 4.95654e-06 Fit statistic : Chi-Squa	ared ared Lity of 2.48e-02 with not fit yet. Th for Component 4	76.24 76.24 54 degrees 0.165868 ke	
							Test statistic : Chi-Squa Null hypothesis probabil Current data and model r XSPEC12>eqw 4 Data group number: 1 Additive_group equiv widt	Lity of 2.43e-02 with not fit yet.	76.34 54 degrees 0.0347591 k	

# Step 6: source flux and luminosity

xspec> flux 2 10 xspec> newpar 4 0 xspec> lum 2 10 0.06 z=0.06 flux in the observed-frame 2–10 keV band absorption set to 0 luminosity in the rest-frame 2-–10 keV band

command cosmo to change the cosmology

XSPEC12>flux 2 10 Model Flux 0.00035807 photons (3.8799e-12 ergs/cm^2/s) range (2.0000 - 10.000 keV) XSPEC12>ne 4 0 Fit statistic : Chi-Squared 327095.1 using 60 bins. Test statistic : Chi-Squared 327095.1 using 60 bins. Null hypothesis probability of 0.0e+00 with 54 degrees of freedom Current data and model not fit yet. XSPEC12>lum 2 10 .06 Model Luminosity 8.2876e+43 ergs/s (2.0000 - 10.000 keV rest frame) (z = 0.0600 H0 = 70.0 q0 = 0.00 Lambda0 = 0.730)

□ Flux is *observed* (typically, no correction for absorption) and in the *observed-frame* band (units: erg/cm<sup>2</sup>/s)

□ Luminosity needs to be *intrinsic/de-absorbed* (so, put  $N_H = 0$  and do **not** fit again) and is reported in the *source rest frame* (units: erg/s)

### **Uncertainties on fluxes and luminosities. I**

xspec> flux 2 10 error 100 90 (100 trials to compute the error at 90% c.l., 2–10 keV band)

XSPEC12>flux 2 10 error 100 90 Parameter distribution is derived from fit covariance matrix. Model Flux 0.00035807 photons (3.8799e-12 ergs/cm^2/s) range (2.0000 - 10.000 keV) Error range 0.0003102 - 0.0003806 (3.310e-12 - 4.166e-12) (90.00% confidence

F(2-10 keV)=3.9 [3.3-4.2]×10<sup>-12</sup> erg/cm<sup>2</sup>/s

For what concerns the luminosity, we cannot apply the same method: if we place  $N_H=0$  to have intrinsic values, xspec requires the data to be fit again

cflux and clumin commands

#### **Uncertainties on fluxes and luminosities. II**

- o cflux and clumin are multiplicative model components.
- cflux (clumin) are placed in front of model component(s). At least one of the additive models should have the normalization fixed (frozen) to a non-zero value.
- o cflux/clumin must be treated as the other model components (as part of the fit)
- Example: model pha\*cflux\*zpha\*pow

xspec> addcomp 3 cflux

#### Input parameter value, delta, min, bot, top, and max values for ... -0.1( 0.5 0.005)0 0 1e+06 1e+06 4:cflux:Emin>2 -0.1( 0.1) 0 0 1e+06 1e+06 10 5:cflux:Emax>10 0.01( 0.12)-100-12-100100 100 6:cflux:lg10Flux>-12 Setting the range (2-10 keV) where the flux is computed xspec> freeze 14 fix the powerlaw normalization (as required by the cflux tool) xspec> fit 100

### **Uncertainties on fluxes and luminosities. Ill**

Model phabs<1>*powerlaw<2> + cflux<3>*zphabs<4>(zgauss<5> + powerlaw<6>)Model Model ComponentParameter UnitValuepar comp11 phabsnH10^222.96000E-02 frozen22powerlawPhoIndex0.935615= p1332powerlawnorm2.35038E-05 +/-1.69261E-0643cfluxEminkeV2.00000frozen	=====:	=====;		:======================================	:======;			
par comp 1 1 phabs nH 10^22 2.96000E-02 frozen 2 2 powerlaw PhoIndex 0.935615 = p13 3 2 powerlaw norm 2.35038E-05 +/- 1.69261E-06 Boundaries for the computati		•				abs<4>(zgauss<	5> + powerlaw<6>)	
1       1       phabs       nH       10^22       2.96000E-02       frozen         2       2       powerlaw       PhoIndex       0.935615       = p13         3       2       powerlaw       norm       2.35038E-05       +/-       1.69261E-06       Boundaries for the computation	Model	Model	Component	Parameter	Unit	Value		
22powerlawPhoIndex0.935615= p1332powerlawnorm2.35038E-05+/-1.69261E-06Boundaries for the computation of the computatio	par	comp						
3 2 powerlaw norm 2.35038E-05 +/- 1.69261E-06 Boundaries for the computation	1	1	phabs	nH	10^22	2.96000E-02	frozen	
A 2 of lux [min] key 2 00000 frozon	2	2	powerlaw	PhoIndex		0.935615	= p13	
	3	2	powerlaw	norm		2.35038E-05	+/- 1.69261E-06	Boundaries for the computation
	4	3	cflux	Emin	keV	2.00000	frozen	-
5 3 cflux Emax keV 10.0000 frozen Emax=10 keV	5	3	cflux		keV		frozen	Emax=10 keV
6 3 cflux lg10Flux cgs -11.4505 +/- 2.32119E-02 lg10Flux is the log of the flux in	6	3	cflux	lg10Flux	cgs	-11.4505	+/- 2.32119E-02	Ig10Flux is the log of the flux in the
7 4 zphabs nH 10^22 37.6633 +/- 2.73227 observed energy range	7	4	zphabs	nH	10^22	37.6633	+/- 2.73227	observed energy range
8 4 zphabs Redshift 6.0000E-02 frozen Emin-Emax	8	4	zphabs	Redshift		6.00000E-02	frozen	Emin-Emax
9 5 zgauss LineE keV 6.39925 +/- 4.50821E-02	9	5	zgauss	LineE	keV	6.39925	+/- 4.50821E-02	
	10	5	zgauss	Sigma	keV	1.00000E-02	frozen	The powerlaw normalization has
11 5 zgauss Redshift 6.0000E-02 frozen been frozen	11	5	zgauss	Redshift		6.00000E-02	frozen	been frozen
12 5 zgauss norm 1.43960E-05 +/- 6.89801E-06	12	5	zgauss	norm		1.43960E-05	+/- 6.89801E-06	
13 6 powerlaw PhoIndex 0.935615 +/- 0.137435	13	6	powerlaw	PhoIndex		0.935615	+/- 0.137435	
14 6 powerlaw norm 6.79101E-04 frozen	14	6	powerlaw	norm		6.79101E-04	frozen	

Fit statistic : Chi-Squared73.33using 60 bins.Test statistic : Chi-Squared73.33using 60 bins.Null hypothesis probability of 4.11e-02 with 54 degrees of freedomXSPEC12>error 6ParameterConfidence Range (2.706)6-11.4923-11.41(-0.0418036,0.0405441)

 $LogF_{2-10 \text{ keV}} = -11.45 [-11.49, -11.41] \rightarrow F_{2-10 \text{ keV}} = 3.5 [3.2-3.9] \times 10^{-12} \text{ erg/cm}^{2}/\text{s}$ consistent with previous value within errors

#### **Uncertainties on fluxes and luminosities. IV**

For what concerns the X-ray luminosity, it must be computed as *intrinsic* (*de-absorbed*), i.e. placing  $N_H=0$  without subsequent fitting the spectrum. However, errors can be computed only using **clumin**. It works similarly to *cflux*. Example: model pha\*zpha\*clumin\*pow

xspec> addcomp 5 clumin

xspec> [...]; fit 100

Model phabs<1>\*powerlaw<2> + zphabs<3>(zgauss<4> + clumin<5>\*powerlaw<6>) Model Model Component Parameter Unit Value par comp 1 1 phabs nH 10^22 2.96000E-02 frozen 2 2 powerlaw PhoIndex 0.935227 = p143 2 2.35016E-05 +/- 1.69269E-06 powerlaw norm 4 3 zphabs nH 10^22 37.6737 +/- 2.73077 5 3 zphabs Redshift 6.00000E-02 frozen 6 4 zgauss LineE +/- 4.51252E-02 keV 6.39933 7 4 zgauss Sigma keV 1.00000E-02 frozen 8 4 zgauss Redshift 6.00000E-02 frozen 9 4 1 43828E-05 +/- 5.64716E-06 zgauss norm 10 5 clumin Emin keV 2.00000 frozen 5 11 clumin keV 10.0000 frozen Emax 5 6.00000E-02 12 clumin Redshift frozen 13 5 clumin la10Lum 43.8978 +/- 3.64824E-02 cqs powerlaw PhoIndex 14 6 0.935227 +/- 0.137412 15 6 powerlaw 6.79101E-04 frozen norm

Boundaries for the computation Emin=2 keV Emax=10 keV Ig10Lum is the log of the luminosity in the *rest-frame energy range* Emin-Emax The powerlaw normalization has been frozen

```
Fit statistic : Chi-Squared73.33using 60 bins.Test statistic : Chi-Squared73.33using 60 bins.Null hypothesis probability of 4.11e-02 with 54 degrees of freedomImage: Confidence Range (2.706)Image: Confidence Range (2.706)1343.822743.9725(-0.0749676.0.0748433)Image: Confidence Range (2.706)
```

LogL<sub>2-10 keV</sub>= 43.90 [43.82 − 43.97] →  $L_{2-10 keV}$ =7.9 [6.6−9.3]×10<sup>43</sup> erg/s

# Step 7: save data+model and 'recover' all in XSPEC later

To save the current data + model you may use the command

xspec> save all po\_zpha\_po\_zgauss

A file po\_zpha\_po\_zgauss.**xcm** is saved with model and data (name these files properly!)

To recover the settings + data + model later, you can use the command

xspec> @po\_zpha\_po\_zgauss.xcm
xspec> fit

XSPEC will ask you to fit the data again

# Other possibly useful commands – some already discussed

#### in XSPEC

- > setplot rebin #1 #2 (to rebin the data; #1 indicates the number of  $\sigma$ )
- show all
- show files
- show notice
- script filename [save all the commands in a file (filename here; default: xspec.xcm]
- save model bestmodel.xcm (save only the best fit model, without the data)
- setplot command redshift # (set the energy axis to redshift # of the source)
- setplot background; plot (plot the background; to remove it: setplot noback; plot)

#### In IPLOT (plotting environment for XSPEC)

#### xspec> iplot

- time off (to remove the date in the bottom-right part of the plot)
- csize 2 (character size)
- msize (marker size)
- label top (title of the plot)
- label filename (title of the file)
- hardcopy nomefile.ps/cps (save a figure)
- plot
- wen namefile [writes two files (.qdp and .pco), one with data and the other with plot settings]