Previous lecture:

"real" maximum likelihood method: Poissonian data

Finding periodicities in data

- Lomb-Scarle diagrams
- Phase dispersion minimisation
- Fourier techniques

OAF2 chapter 6.1 & 6.2

Num Res chapter 13.8, 14.3,14.5 & 14.7

Period Finding II

Phase-Dispersion Minimisation: PDM

Fold data given a trial period in M bins

Calculate the variance in each bin

Large variance

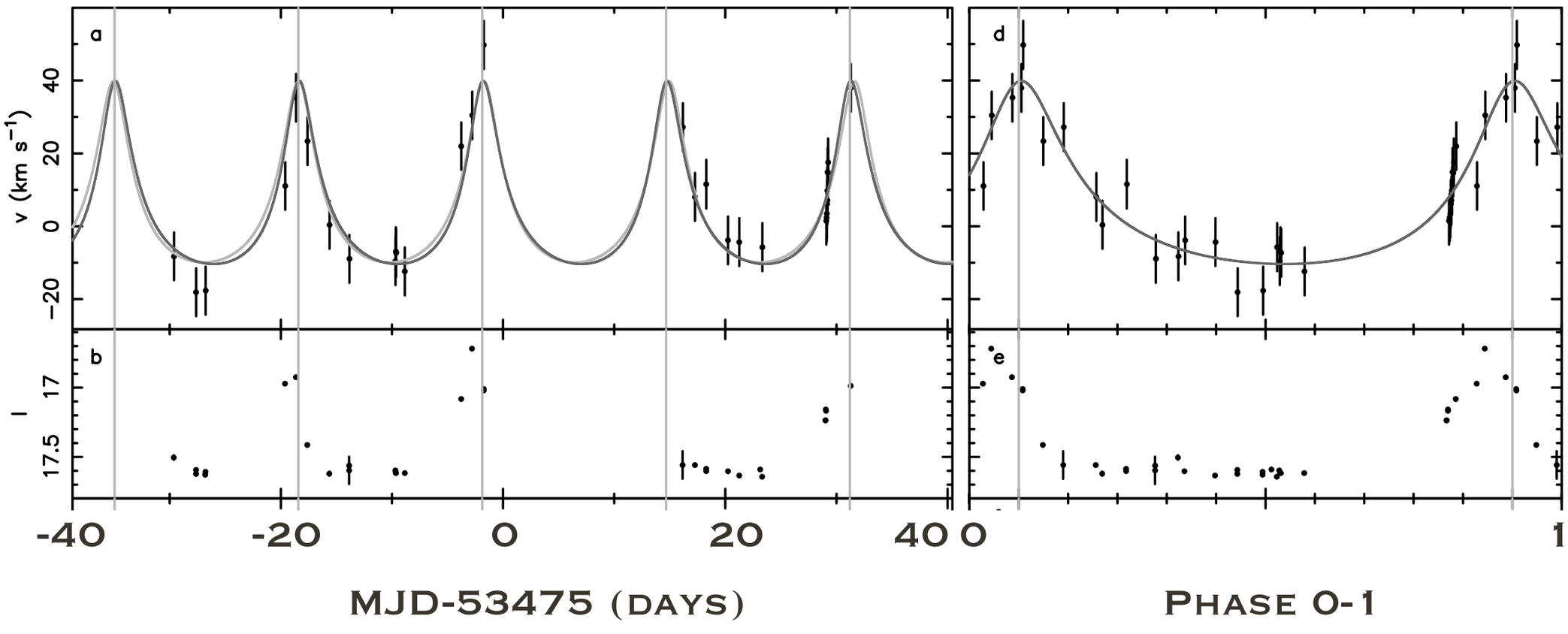
not the right period

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$ variance in the data

$$s_k^2 = \frac{1}{N-1}\sum_{j=1}^{n_k}(x_j - \bar{x})^2$$ variance in one sample

Stellingwerf 1978, ApJ, 224, 953

Light curve and folded light curve



Here independent bins, could also have "sliding window" bins, each data point in more than 1 bin

PDM (CONTINUED)

$$s^2 = \frac{\sum_{k=1}^{M}(n_k - 1)s_k^2}{\sum_{k=1}^{M} n_k - M}$$ VARIANCE IN THE SAMPLES

→ $$\theta = \frac{s^2}{\sigma^2}$$ WRONG PERIOD $\Theta \approx 1$

VARIANCE IN THE SAMPLES=VARIANCE IN THE DATA

"RIGHT" PERIOD $\Theta \ll 1$

SCRAMBLE DATA IN A MONTE CARLO SIMULATION TO CALCULATE SIGNIFICANCES

Today

Correlations

Student's t

Pearson's r

Spearman's rank

Science projects at SRON

Student's T

How different are two means?

$$s_d = \left(\frac{\sum_{i \epsilon A} (x_i - \overline{x_A})^2 + \sum_{i \epsilon B} (x_i - \overline{x_B})^2}{N_A + N_B - 2} \left(\frac{1}{N_A} + \frac{1}{N_B} \right) \right)^{0.5}$$

standard error on the difference in the means

$$t = \frac{\overline{x_A} \quad \overline{x_B}}{s_d}$$

significance in t: is distributed as t's distribution with

$N_A + N_B - 2$ degrees of freedom

What is student's t distribution?

numerical recipes Chapter 6.4

$$A(t|\nu) = \frac{1}{\nu^{1/2} B(\frac{1}{2}, \frac{\nu}{2})} \int_{-t}^{t} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu}{}} dx$$

$A(t|\nu)$ is often numerically calculated with an incomplete Beta function (see Num Res equation 6.4.9)

this can be expanded to cases where the two measurement sets A and B have different variances.

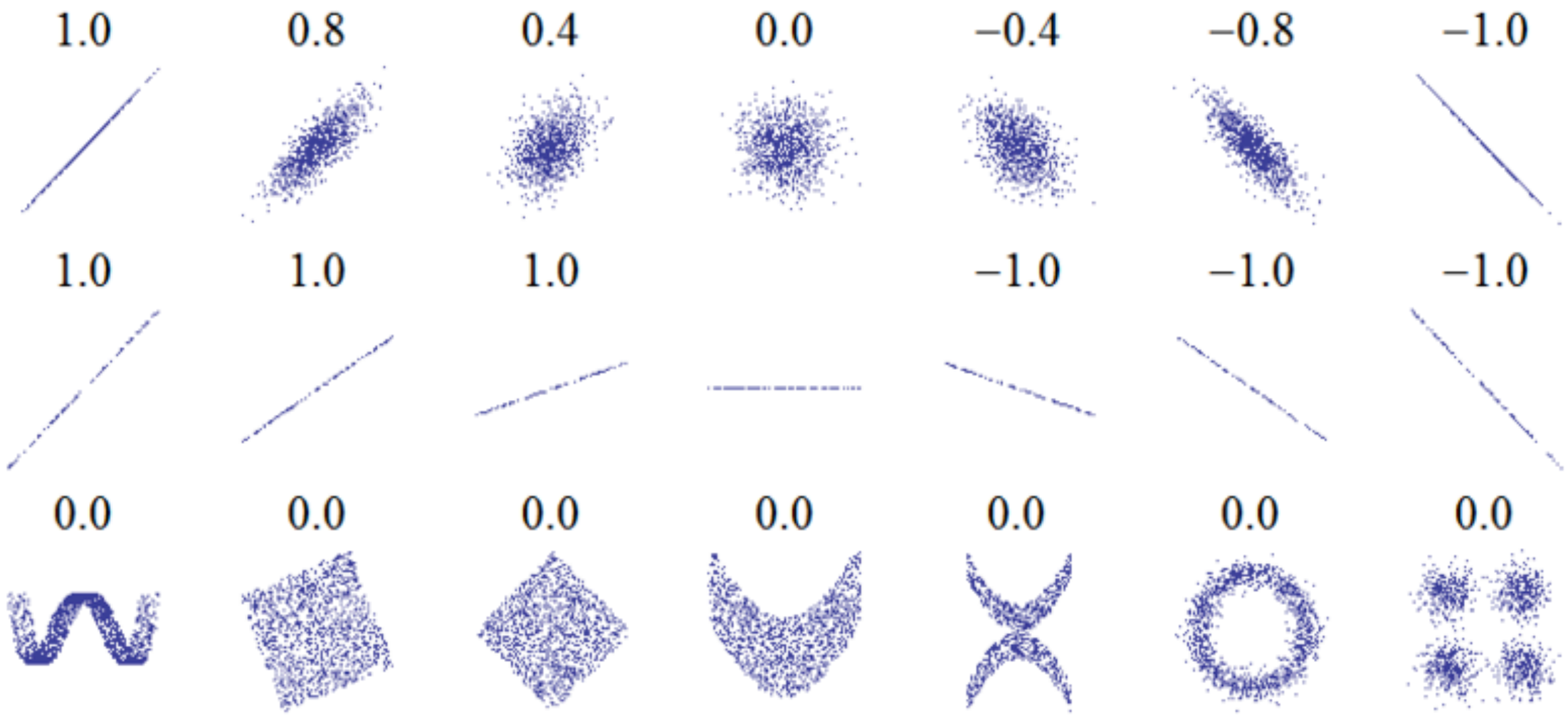
Linear correlation

Pearson's r Formal definition of Pearson's r

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{E((X-\mu_x)(Y-\mu_y))}{\sigma_x \sigma_y}$$

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

from Wikipedia



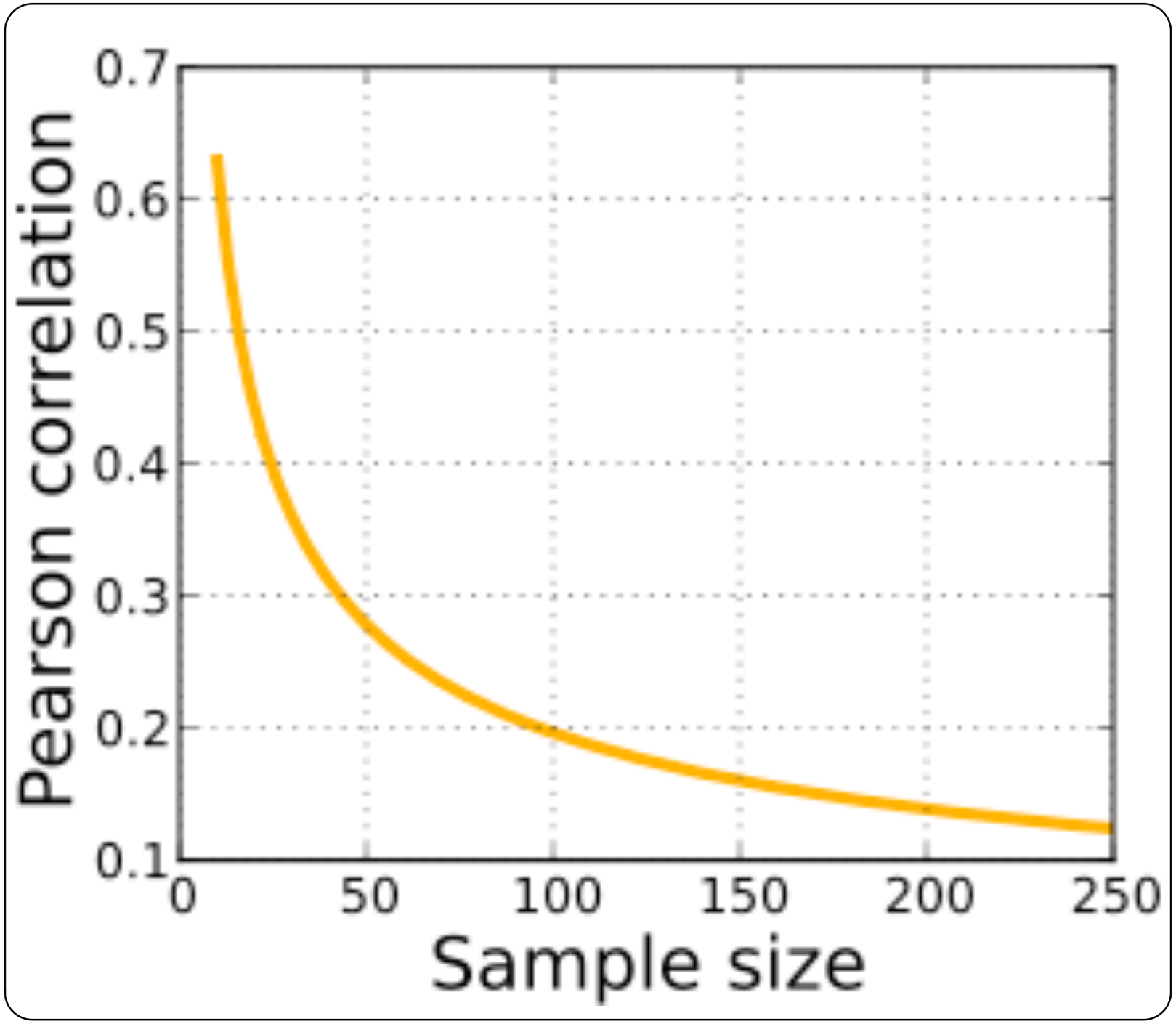
Significance & confidence region? Bootstrap!

make new random pairs X_I and Y_I and repeat calc r

Analytical approx

Probability that $|r| > |\mathrm{r_{obs}}|$ **by chance** $= \mathrm{erfc}\left(\frac{|r|\sqrt{N}}{\sqrt{2}}\right)$

in the case of the null hypothesis of uncorrelated x and y



from Wikipedia

Non-parametric correlation

Spearman's rank

$$\rho = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$$

$$d_i = x_i - y_i$$

difference between ranks x_i and y_i

Significance & confidence region? Bootstrap!

make new random rank pairs X_I and Y_I and repeat

calc rho

IF SAME DATA VALUE APPEARS MORE THEN ONCE:

$$\rho = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\left(\sum_i (x_i - \overline{x})^2 \sum_i (y_i - \overline{y})^2\right)^{1/2}}$$

ASSIGN SAME RANK TO THE DATA WHERE VALUE IS THE SAME: AVERAGE RANK!

	Rank	Rank
x_1 = 1.2	1 or 2	1.5
x_2 = 3.2	4	4
x_3 = 2.8	3	3
x_4 = 1.2	1 or 2	1.5

Example Spearman's rank (from Wikipedia)

IQ Hours of TV per week

X_i	Y_i	$rankx_i$	$ranky_i$	d_i	d_i^2
86	0	1	1	0	0
97	20	2	6	-4	16
99	28	3	8	-5	25
100	27	4	7	-3	9
101	50	5	10	-5	25
103	29	6	9	-3	9
106	7	7	3	4	16
110	17	8	5	3	9
112	6	9	2	7	49
113	12	10	4	6	36

Sum over d_i^2 = 194, n=10, rho =-0.176

Science at SRON staff members

- Wim Hermsen: pulsars, INTEGRAL data
- Jean in 't Zand: type I X-ray bursts (mostly broadband X-ray spectra
- Jelle Kaastra: Clusters of galaxies, ISM studies (high-resolution X-ray spectroscopy)
- Elisa Costantini: ISM studies, AGN (high-resolution X-ray spectroscopy)
- Lucien Kuiper: pulsars, INTEGRAL sources (X-ray timing, broadband X-ray spectra)

Peter Jonker: low-mass X-ray binaries, neutron stars & (Intermediate mass) black holes

optical: spectroscopy/photometry
X-ray: astrometry, photometry, spectroscopy

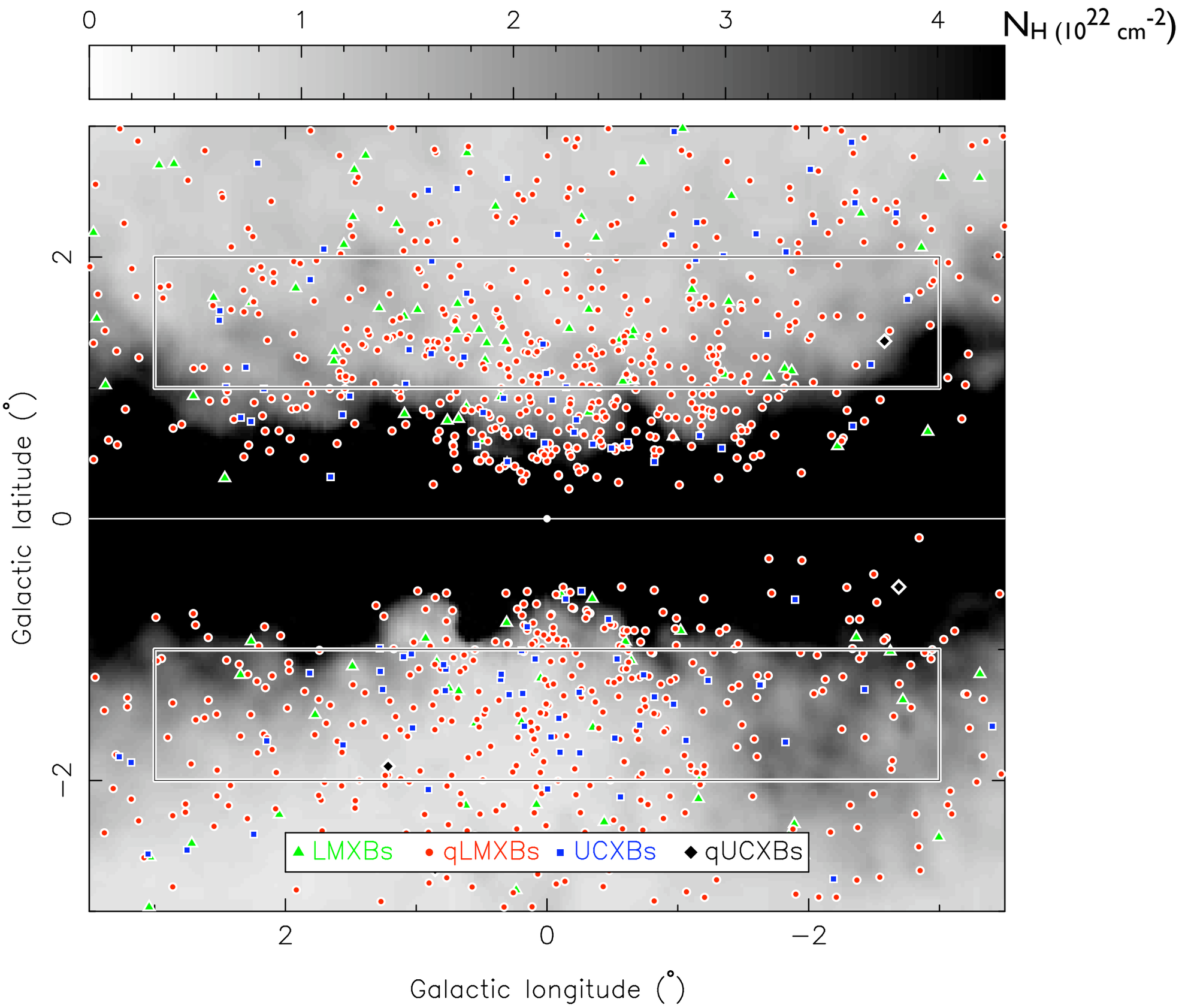
Predictions: ~400 LMXBs

CVs ~450

qLMXBs ~350

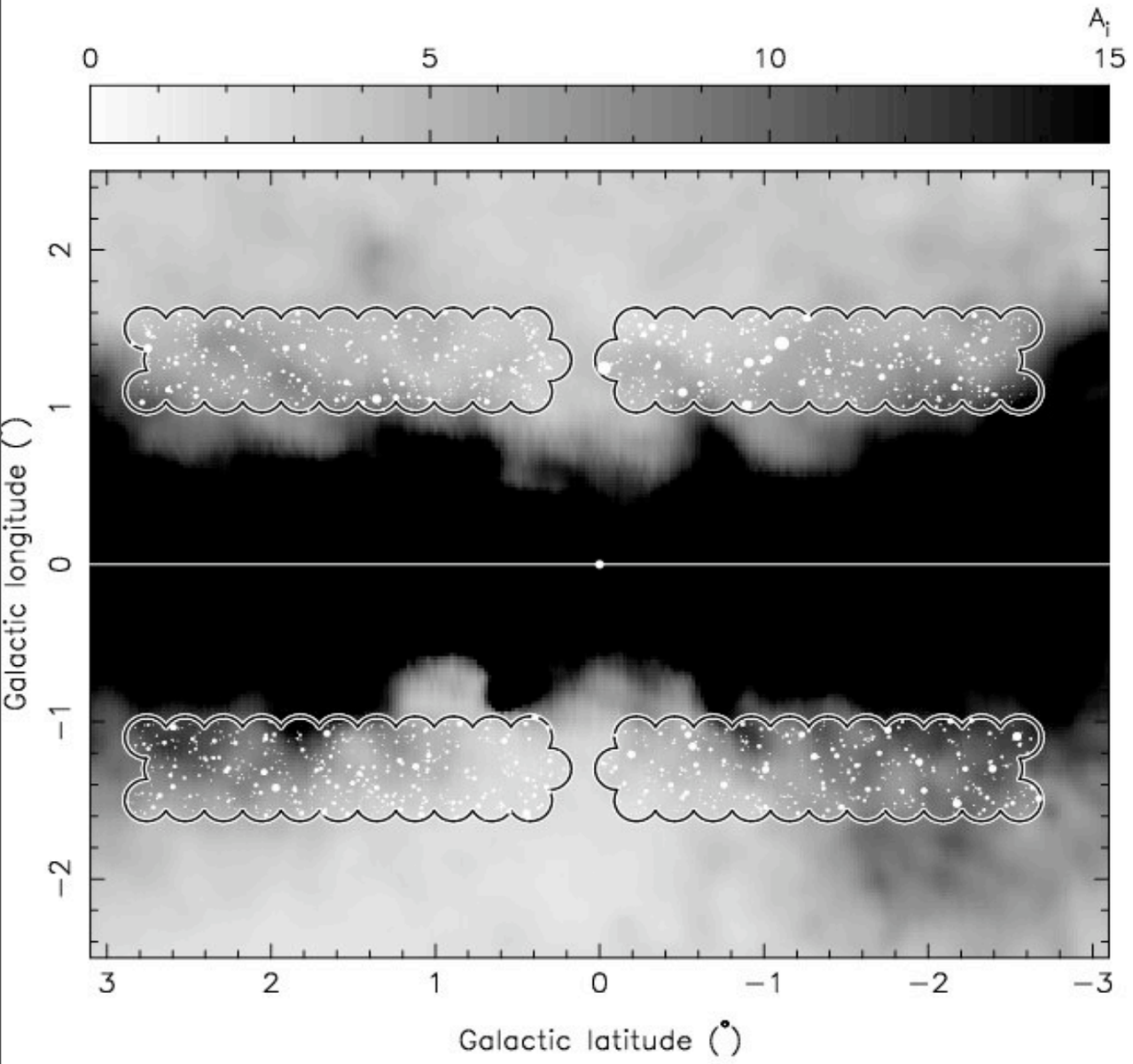
UCXBs ~50

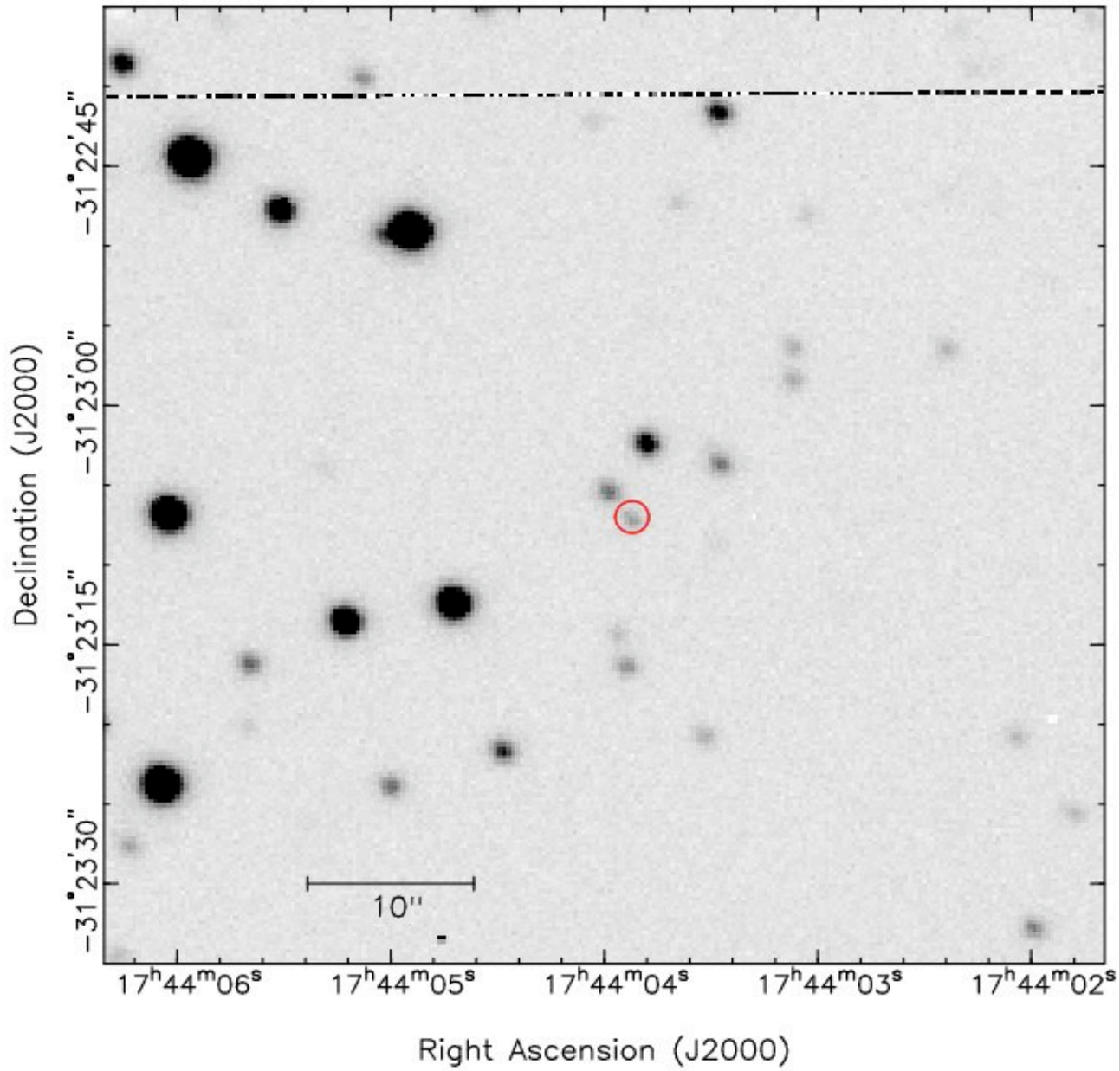
RS CVn ~550



Details of Galactic model: Nelemans et al. 2004

Galactic Bulge Survey





Core group of people working on the GBS:

me (PI)

Eva Ratti (PhD student)

Gijs Nelemans (Nijmegen Univ)
Lennart van Haaften (PhD student, Nijm.)

Danny Steeghs (Warwick Univ, UK)

Manuel Torres (Harvard-Smithsonian CfA becomes my post-doc, May 1)

Several student projects, including master student projects