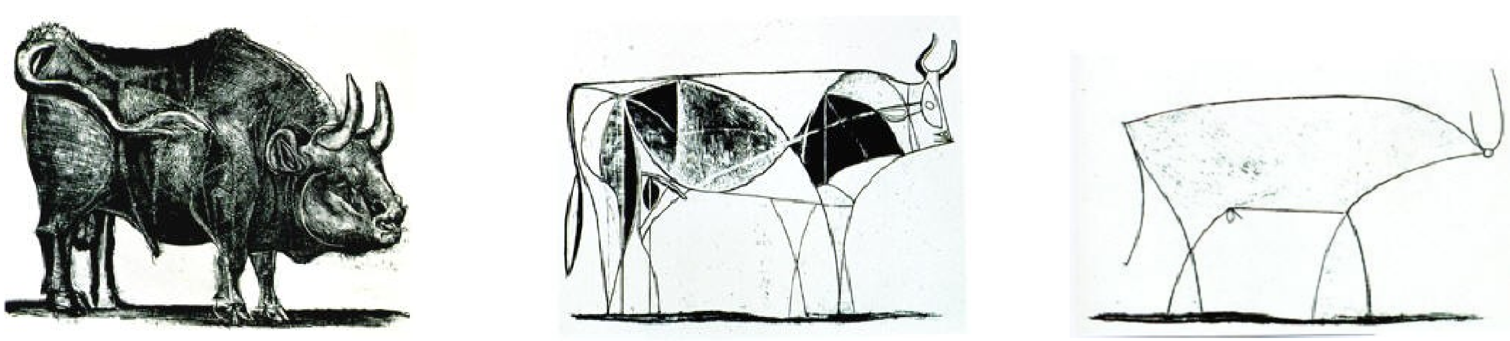
Advanced Statistics for better Cosmological Inference

Amir Aghamousa

Asia Pacific Center for Theoretical Physics, Pohang, South Korea

IBS-KASI Joint Workshop, August 19~20, 2015



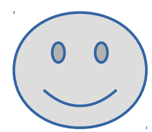
Bull pictures courtesy of “Pablo Picasso”

All models are false, some are useful. (George E. P. Box)

I will talk about:

- Strong Lens Time Delay Challenge II (with Arman Shafieloo)
- Time delay estimation of SDSS J1001+5027 (with Arman Shafieloo)
- Supernova Type Ia classification (with Arman Shafieloo, Alex G. Kim)

If I have time,

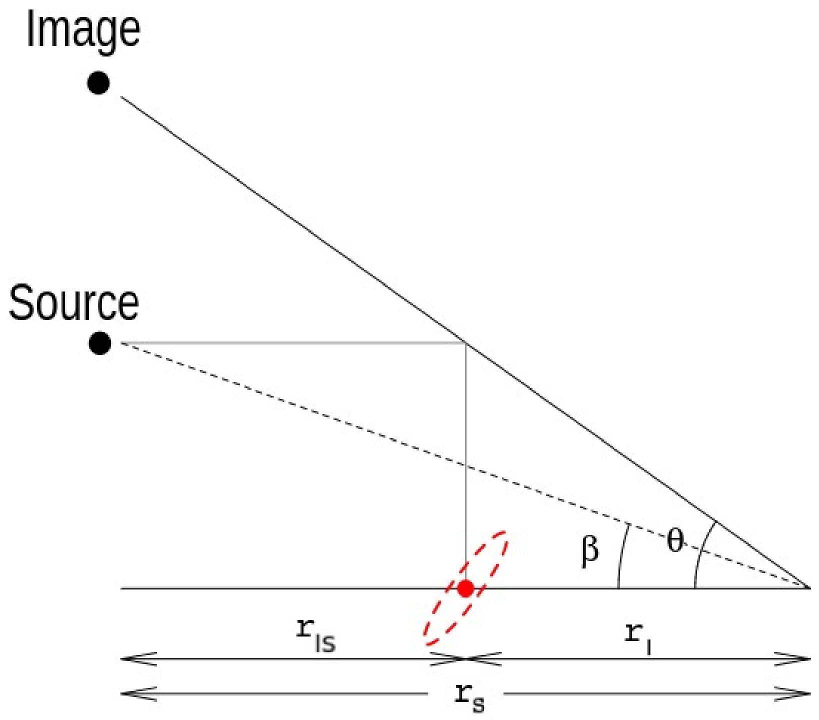


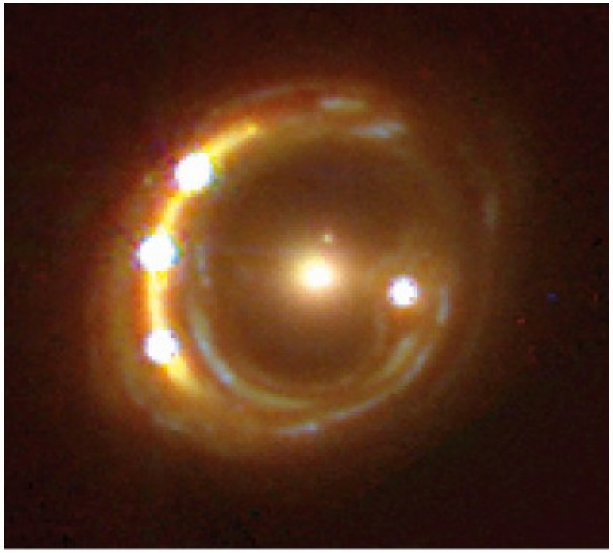
talk about:

- Model-independent estimation of CMB angular power spectrum (with Arman Shafieloo, Mihir Arjunwadkar, Tarun Souradeep)
- Nonparametric test of consistency between cosmological models and CMB data (with Arman Shafieloo)

Strong Lens
Time Delay Challenge II

Strong gravitational lensing





HST ACS image of RXJ1131-1231

Time delay: $$\Delta t(\vec{\theta}, \vec{\beta}) = \frac{r_l r_s}{r_{ls}}(1 + z_l)\phi(\vec{\theta}, \vec{\beta})$$

Fermat potential: $$\phi(\vec{\theta}, \vec{\beta}) = \frac{(\vec{\theta} - \vec{\beta})^2}{2} - \psi(\vec{\theta})$$

geometric delay

lensing potential delay

(Linder 2011, Tewes 2012)

Strong lensing surveys

- Recent survey:
  COSMOGRAIL: the COSmological MOnitoring of GRAvItational Lenses
  (http://www.cosmograil.org)

- Future survey:
  LSST: the Large Synoptic Survey Telescope (LSST) with 10 years observation will be expected to monitor several thousand time delay lens systems.

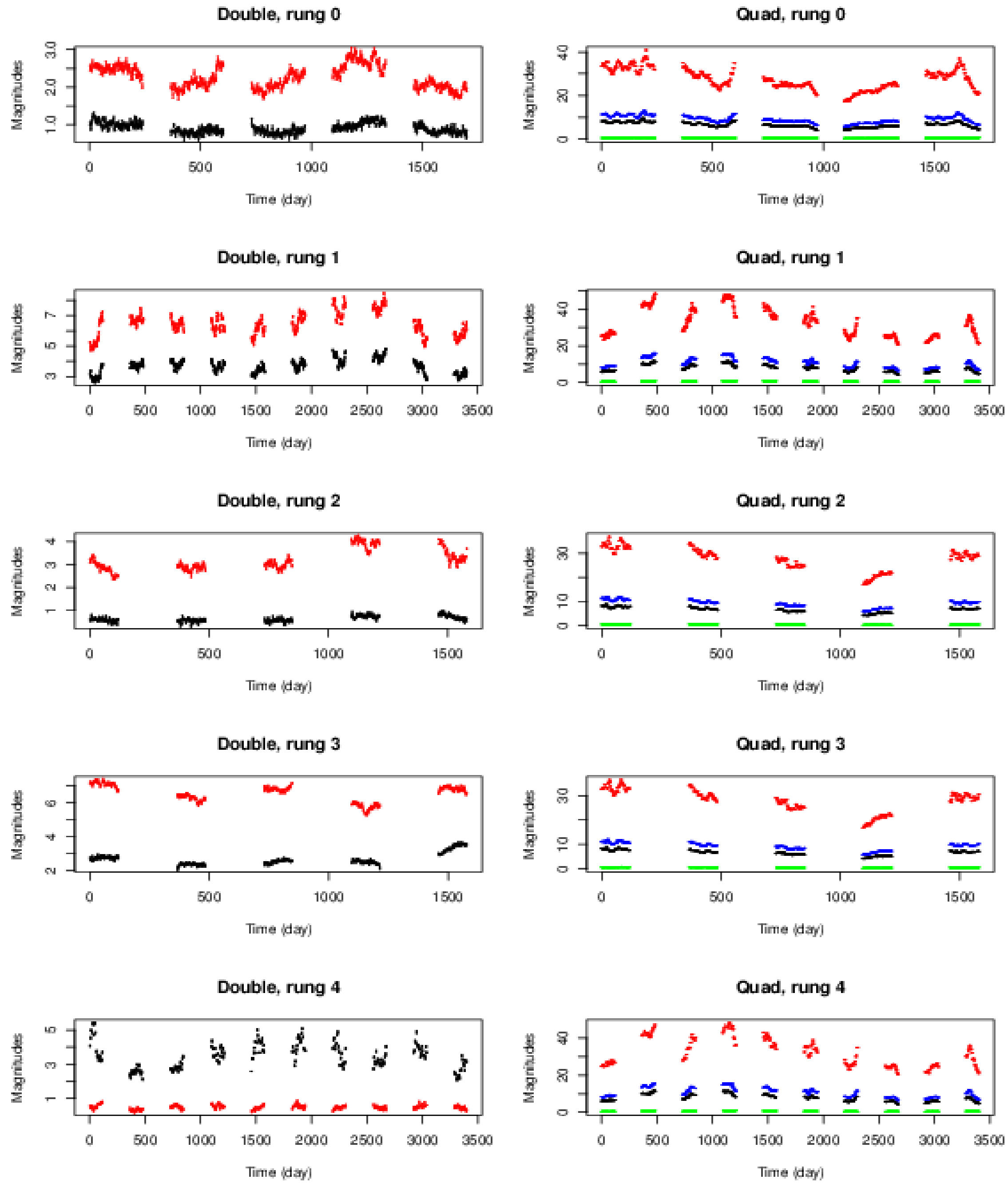
- Need to design the fast and reliable algorithms for the time delay estimation.

- Strong Lens Time Delay Challenge:
  TDC0
  TDC1

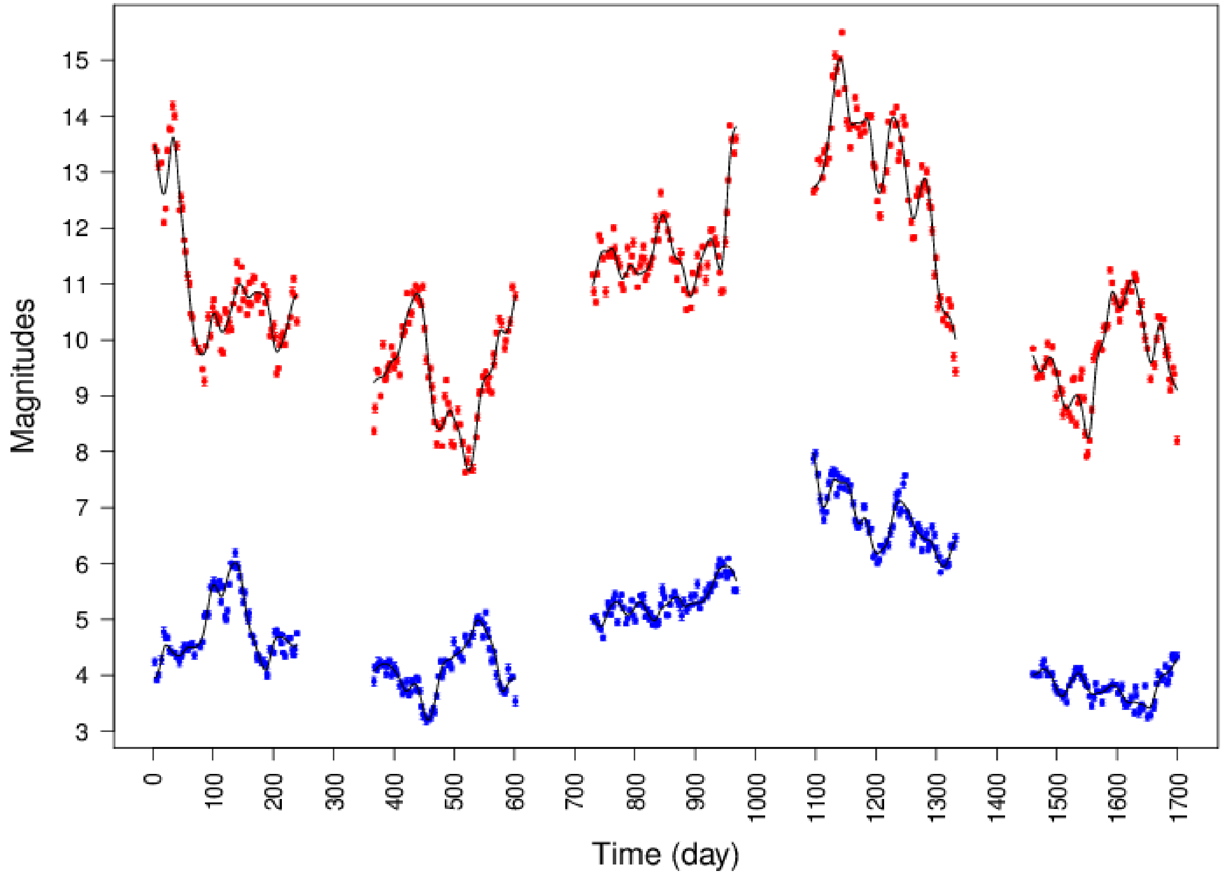
Strong Lens Time Delay Challenge: simulated data

The TDC1 simulated data is provided in five different categories (rungs)

Each rung contains the light curves of 720 Double and 152 Quad image systems.



Methodology: smoothing

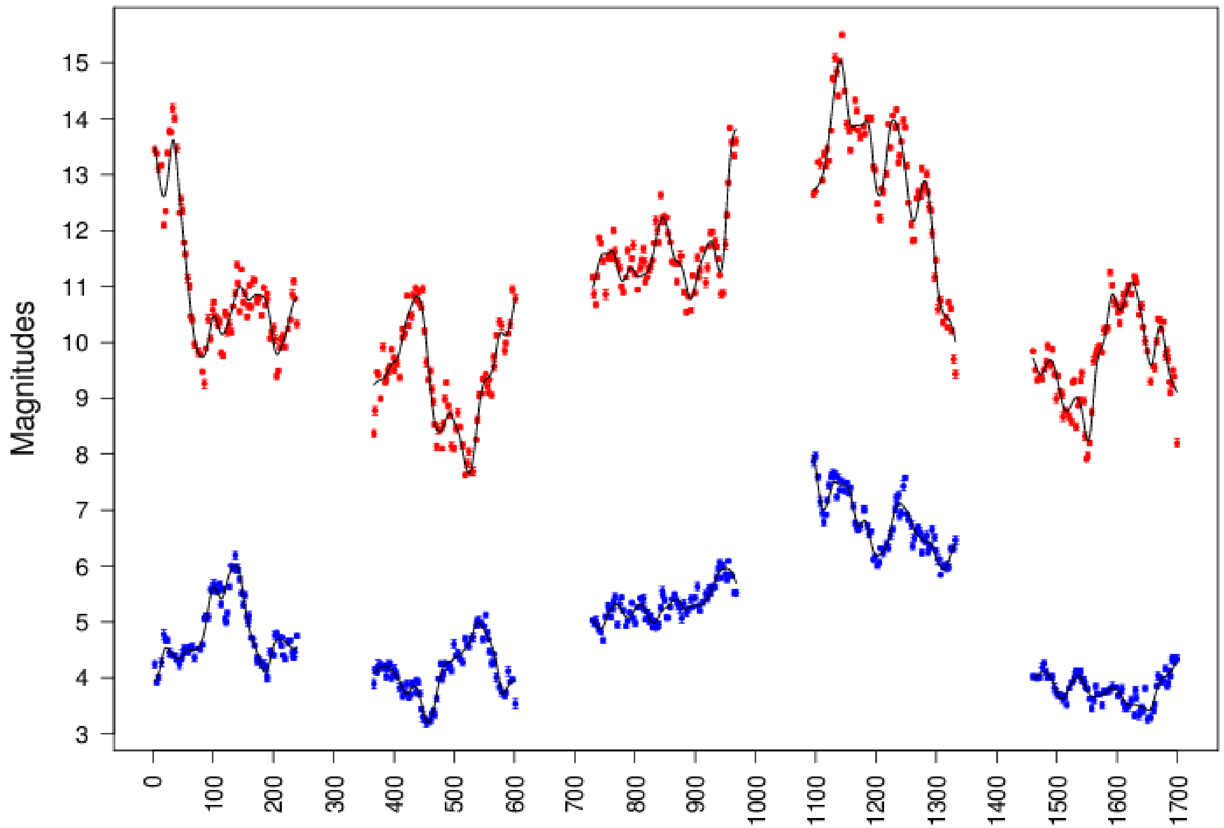


$$A^s(t) = A^g(t) + N(t)\sum_i \frac{A^d(t_i) - A^g(t_i)}{\sigma_d^2(t_i)} \times exp\left[-\frac{(t_i - t)^2}{2\Delta^2}\right]$$

where $$N(t)^{-1} = \sum_i exp\left[-\frac{(t_i - t)^2}{2\Delta^2}\right]\frac{1}{\sigma_d^2(t_i)}$$

(Shafieloo et al. 2006, Shafieloo 2007, Shafieloo & Clarkson 2010)

Methodology: cross-correlation



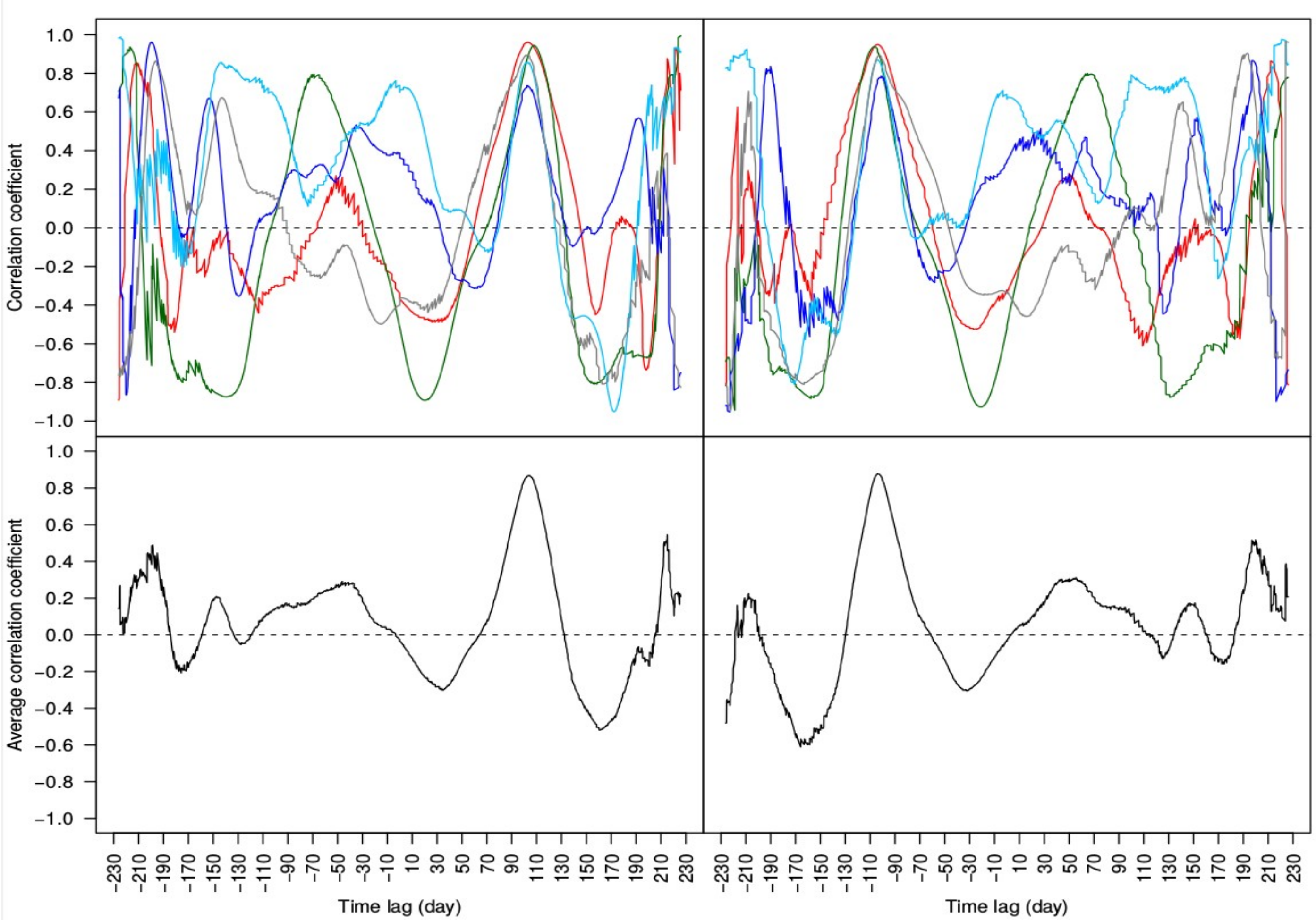
$$\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \Rightarrow r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}},$$

Microlensing

$$\left.\begin{aligned} X' &= a_x X + b_x \\ Y' &= a_y Y + b_y \end{aligned}\right\} \Rightarrow \rho_{XY} = \rho_{X'Y'}$$

(Peterson 2001, Wasserman 2004)

Methodology: time delay estimation



Correlation > 0.6

$$|\tilde{\Delta} t_{A_1 A_2}| = \frac{|\tilde{\Delta} t_{A_1;A_2^s}| + |\tilde{\Delta} t_{A_2;A_1^s}|}{2} \qquad \sigma^{ini}_{A_1 A_2} = \sqrt{2} \times \frac{\left| |\tilde{\Delta} t_{A_1;A_2^s}| - |\tilde{\Delta} t_{A_2;A_1^s}| \right|}{2}$$

(Aghamousa, Shafieloo, APJ, 2015)

Methodology: error estimation, using Quad systems

The light curves of a Quad image are labeled A1 , A2 , B1 and B2. For every Quad system we should have:

$$\underbrace{\tilde{\Delta} t_{A_1A_2} - (\tilde{\Delta} t_{A_1B_1} + \tilde{\Delta} t_{B_1A_2})}_{T_{dif}} \pm \underbrace{\sqrt{(\sigma^{ini}_{\tilde{\Delta} t_{A_1A_2}})^2 + (\sigma^{ini}_{\tilde{\Delta} t_{A_1B_1}})^2 + (\sigma^{ini}_{\tilde{\Delta} t_{B_1A_2}})^2}}_{\sigma_{T_{dif}}} \equiv 0$$

$$|T_{dif}| \leqslant \sigma_{T_{dif}}$$

Yes → We can assume that all time delays and their corresponding errors are estimated consistently.

No → $\sigma^2_{ec} = |T_{dif}|^2 - \sigma^2_{T_{dif}}$ → $\sigma^{new}_{\tilde{\Delta} t_{A_1A_2}} = \sqrt{(\sigma^{ini}_{\tilde{\Delta} t_{A_1A_2}})^2 + \frac{\alpha}{3}\sigma^2_{ec}}$

(Aghamousa, Shafieloo, APJ, 2015)

Methodology: error estimation, using Quad systems

Error profile for Double systems

For **Rung 0**,

if $|\tilde{\Delta t}| \leqslant 20 \Rightarrow \sigma_R = 0.06 \times |\tilde{\Delta t}|$, if $|\tilde{\Delta t}| > 20 \Rightarrow \sigma_R = 1.2$

For **Rung 1**,

if $|\tilde{\Delta t}| \leqslant 20 \Rightarrow \sigma_R = 0.06 \times |\tilde{\Delta t}|$, if $|\tilde{\Delta t}| > 20 \Rightarrow \sigma_R = 1.3$

For **Rung 2**,

if $|\tilde{\Delta t}| \leqslant 30 \Rightarrow \sigma_R = 0.07 \times |\tilde{\Delta t}|$, if $|\tilde{\Delta t}| > 30 \Rightarrow \sigma_R = 1.3$

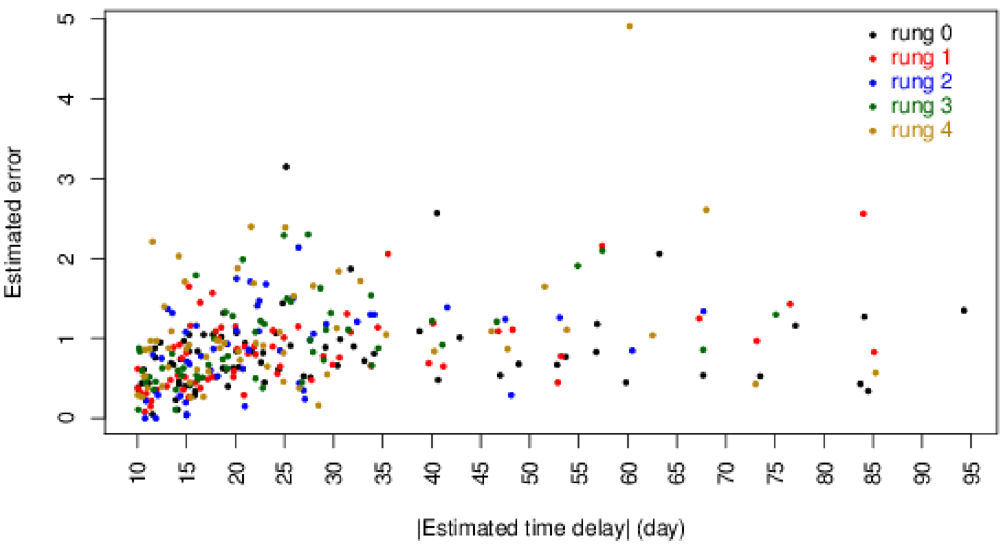
For **Rung 3**,

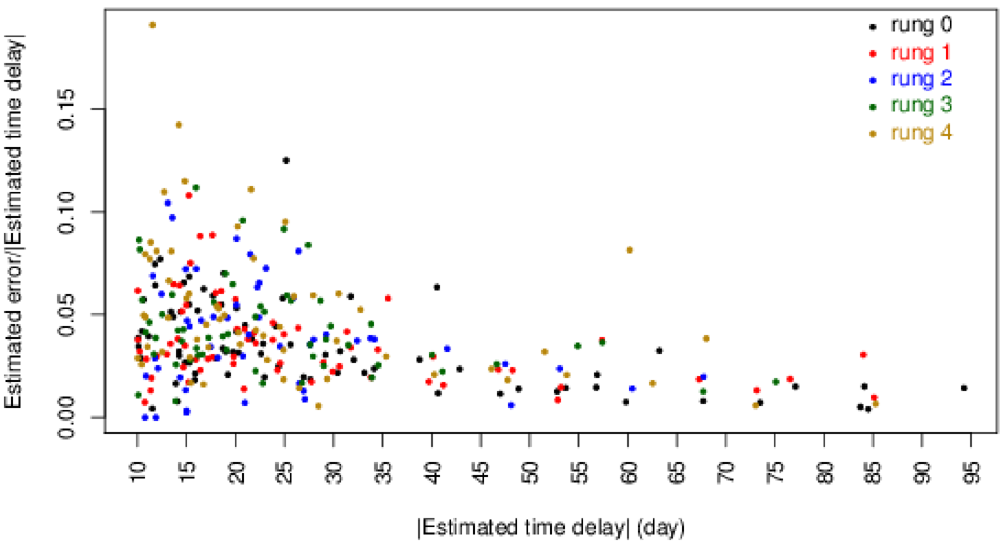
if $|\tilde{\Delta t}| \leqslant 30 \Rightarrow \sigma_R = 0.08 \times |\tilde{\Delta t}|$, if $|\tilde{\Delta t}| > 30 \Rightarrow \sigma_R = 1.5$

For **Rung 4**,

if $|\tilde{\Delta t}| \leqslant 25 \Rightarrow \sigma_R = 0.08 \times |\tilde{\Delta t}|$, if $|\tilde{\Delta t}| > 25 \Rightarrow \sigma_R = 1.5$

$$\sigma_{\tilde{\Delta t}_{A_1A_2}} = \sqrt{(\sigma^{ini}_{\tilde{\Delta t}_{A_1A_2}})^2 + \sigma_R^2}$$





(Aghamousa, Shafieloo, APJ, 2015)

Results: TDC1 paper

STRONG LENS TIME DELAY CHALLENGE: II. RESULTS OF TDC1

Kai Liao[1,2*], Tommaso Treu[2*], Phil Marshall[3], Christopher D. Fassnacht[4], Nick Rumbaugh[4], Gregory Dobler[5,20], Amir Aghamousa[9], Vivien Bonvin[13], Frederic Courbin[13], Alireza Hojjati[6,7], Neal Jackson[12], Vinay Kashyap[17], S. Rathna Kumar[14], Eric Linder[8,18], Kaisey Mandel[17], Xiao-Li Meng[15], Georges Meylan[13], Leonidas A.Moustakas[11], Tushar P. Prabhu[14], Andrew Romero-Wolf[11], Arman Shafieloo[9,10], Aneta Siemiginowska[17], Chelliah S. Stalin[14], Hyungsuk Tak[15], Malte Tewes[19], and David van Dyk[16]

Draft version September 5, 2014

ABSTRACT

We present the results of the first strong lens time delay challenge. The motivation, experimental design, and entry level challenge are described in a companion paper. This paper presents the main challenge, TDC1, which consisted in analyzing thousands of simulated light curves blindly. The observational properties of the light curves cover the range in quality obtained for current targeted efforts (e.g. COSMOGRAIL) and expected from future synoptic surveys (e.g. LSST), and include "evilness" in the form of simulated systematic errors. 7 teams participated in TDC1, submitting results from 78 different method variants. After a describing each method, we compute and analyze basic statistics measuring accuracy (or bias) A, goodness of fit χ^2, precision P, and success rate f. For some methods we identify outliers as an important issue. Other methods show that outliers can be controlled via visual inspection or conservative quality control. Several methods are competitive, i.e. give $|A| < 0.03$, $P < 0.03$, and $\chi^2 < 1.5$, with some of the methods already reaching sub-percent accuracy. The fraction of light curves yielding a time delay measurement is typically in the range f =20–40%. It depends strongly on the quality of the data: COSMOGRAIL-quality cadence and light curve lengths yield significantly higher f than does sparser sampling. We estimate that LSST should provide around 400 robust time-delay measurements, each with $P < 0.03$ and $|A| < 0.01$, comparable to current lens modeling uncertainties. In terms of observing strategies, we find that A and f depend mostly on season length, while P depends mostly on cadence and campaign duration.

Subject headings: gravitational lensing — methods: data analysis

$$f \equiv \frac{N_{submitted}}{N}$$

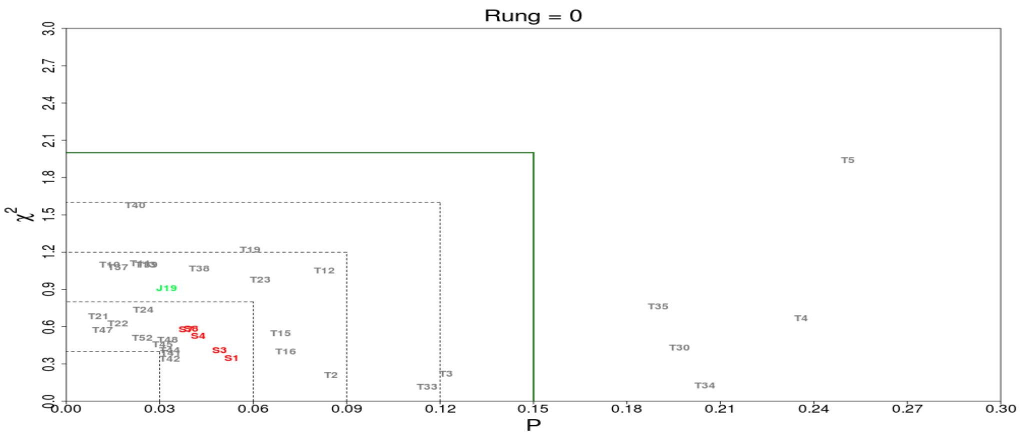
$$P = \frac{1}{fN}\sum_i \left(\frac{\sigma_i}{|\Delta t_i|}\right)$$

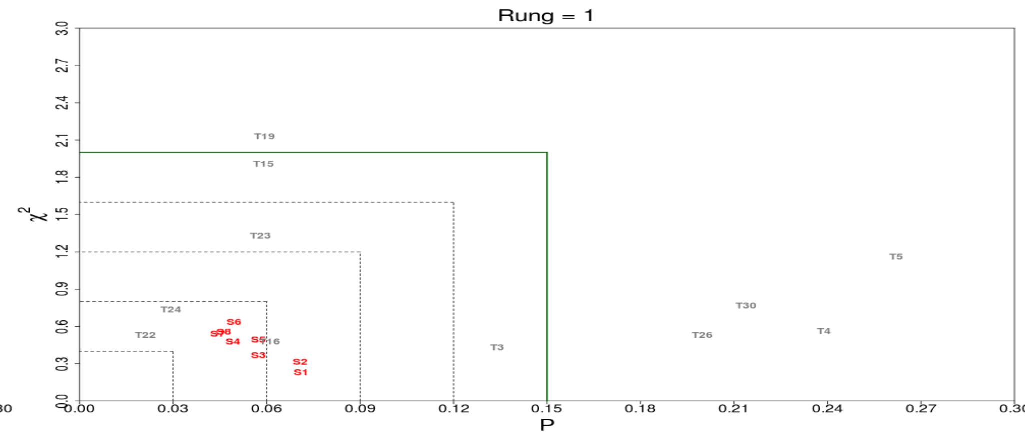
$$A = \frac{1}{fN}\sum_i \left(\frac{|\tilde{\Delta t_i}| - |\Delta t_i|}{|\Delta t_i|}\right)$$

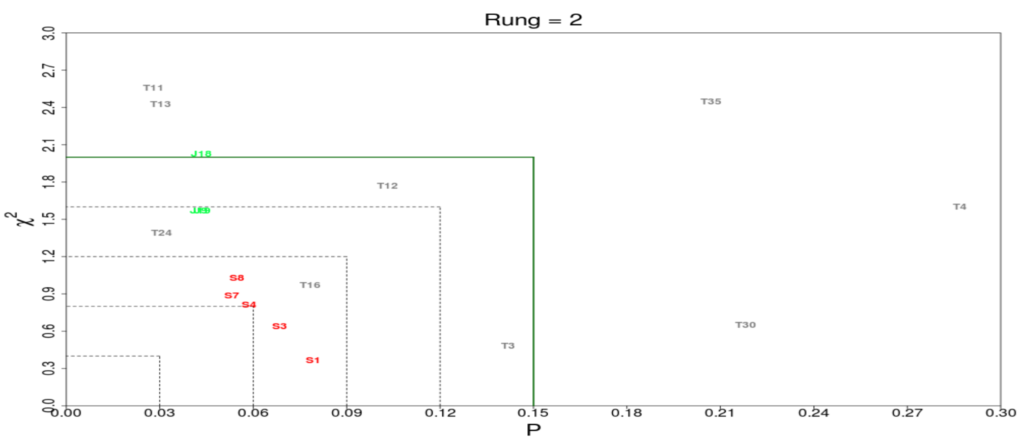
$$\chi^2 = \frac{1}{fN}\sum_i \left(\frac{\tilde{\Delta t_i} - \Delta t_i}{\sigma_i}\right)^2$$

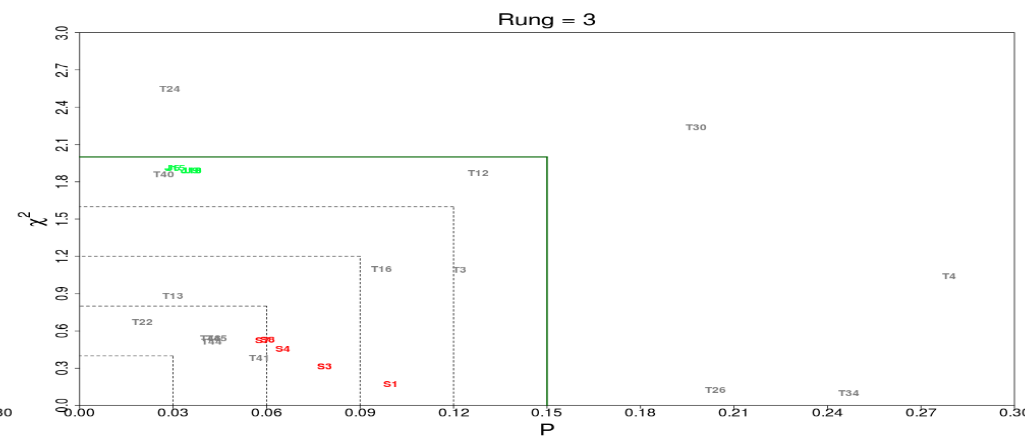
Rung	f	χ^2	P	A
0	0.529	0.579	0.038	-0.018
1	0.366	0.543	0.044	-0.022
2	0.350	0.885	0.053	-0.025
3	0.337	0.524	0.059	-0.021
4	0.346	0.608	0.056	-0.024

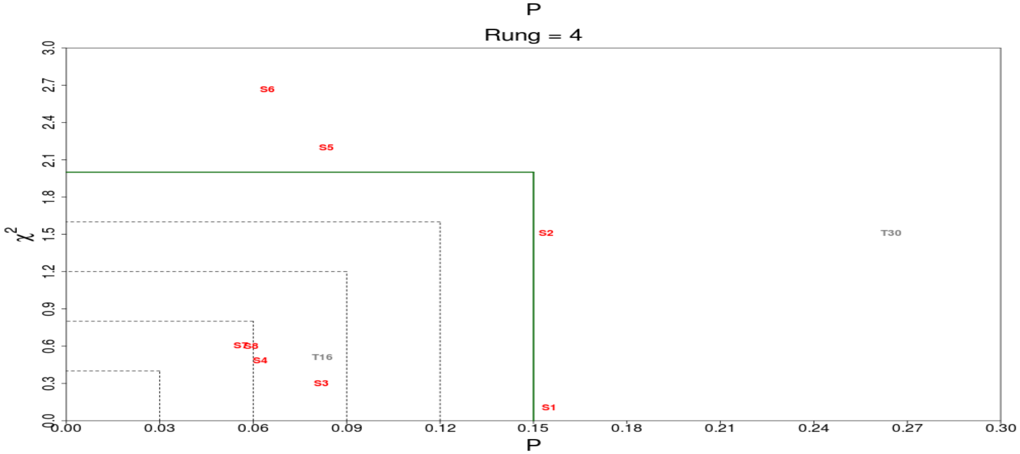
Results: estimated vs true time delay











f>0.3

(H6, K2, J19 are after feedbacks.)

Rumbaugh(R)

Hojjati(H)

Kumar(K)

Jackson(J)

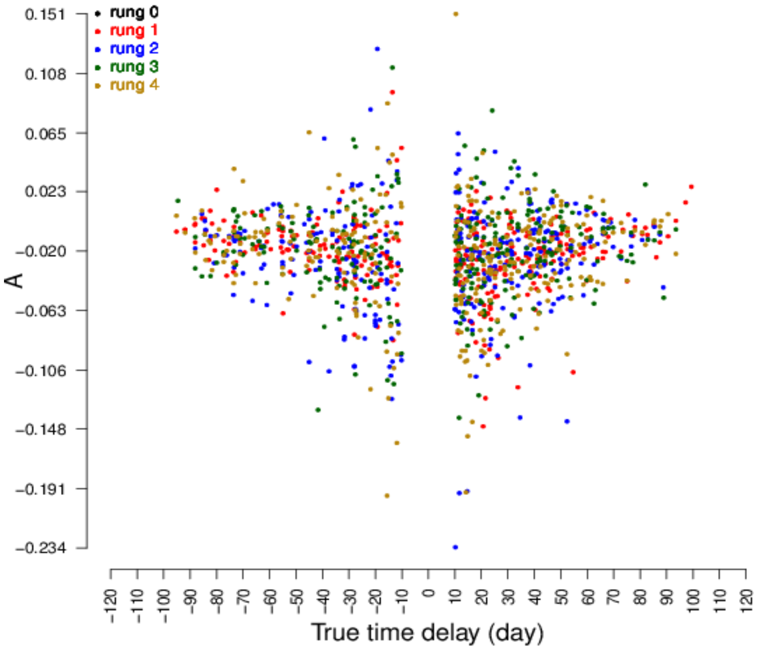
Shafieloo(S)

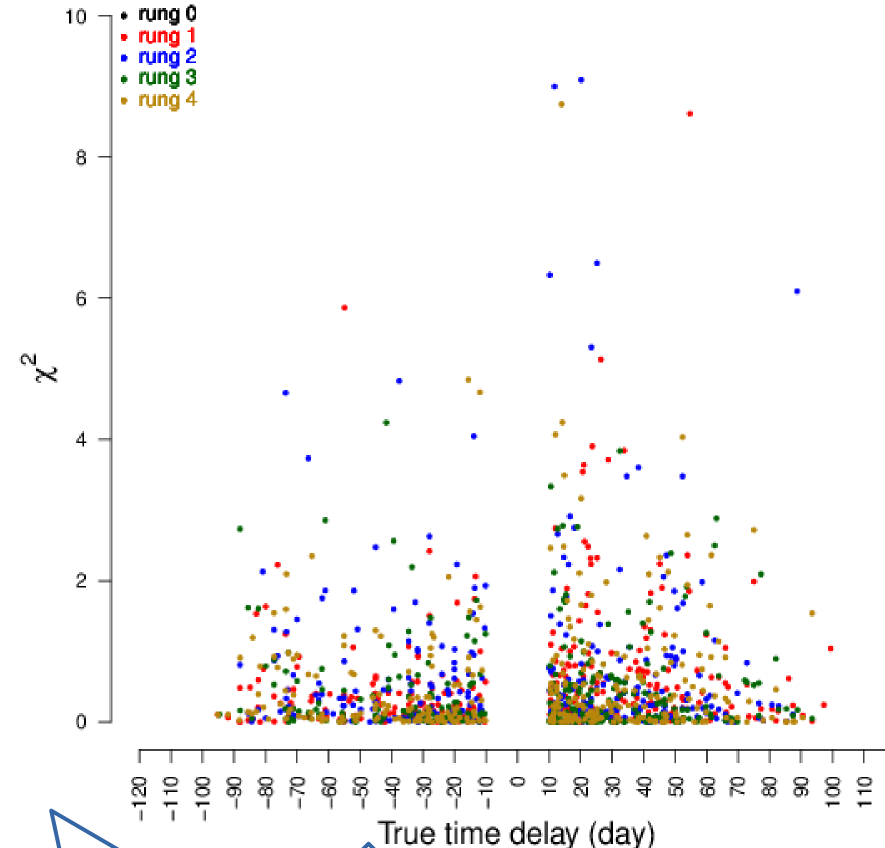
Tewes(T)

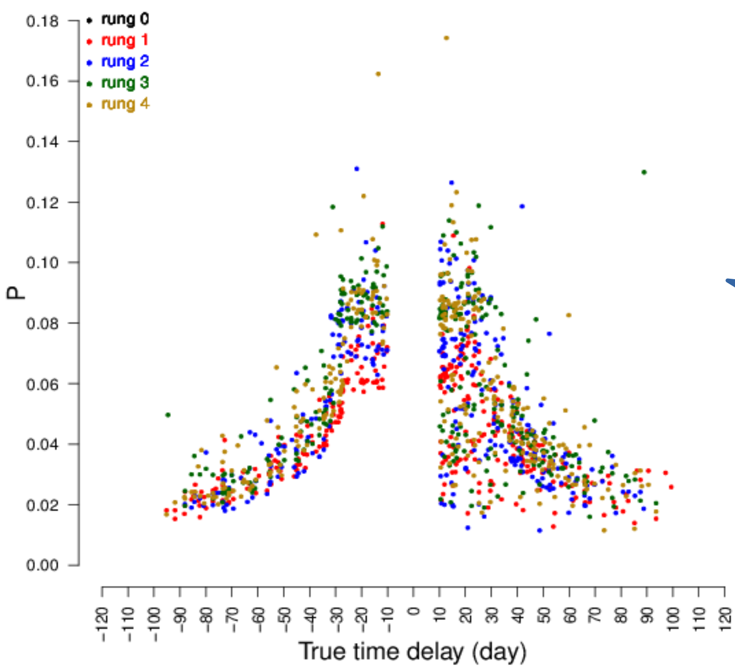
DeltaTBayes(D)

JPL(P)

Results: histogram of metrics







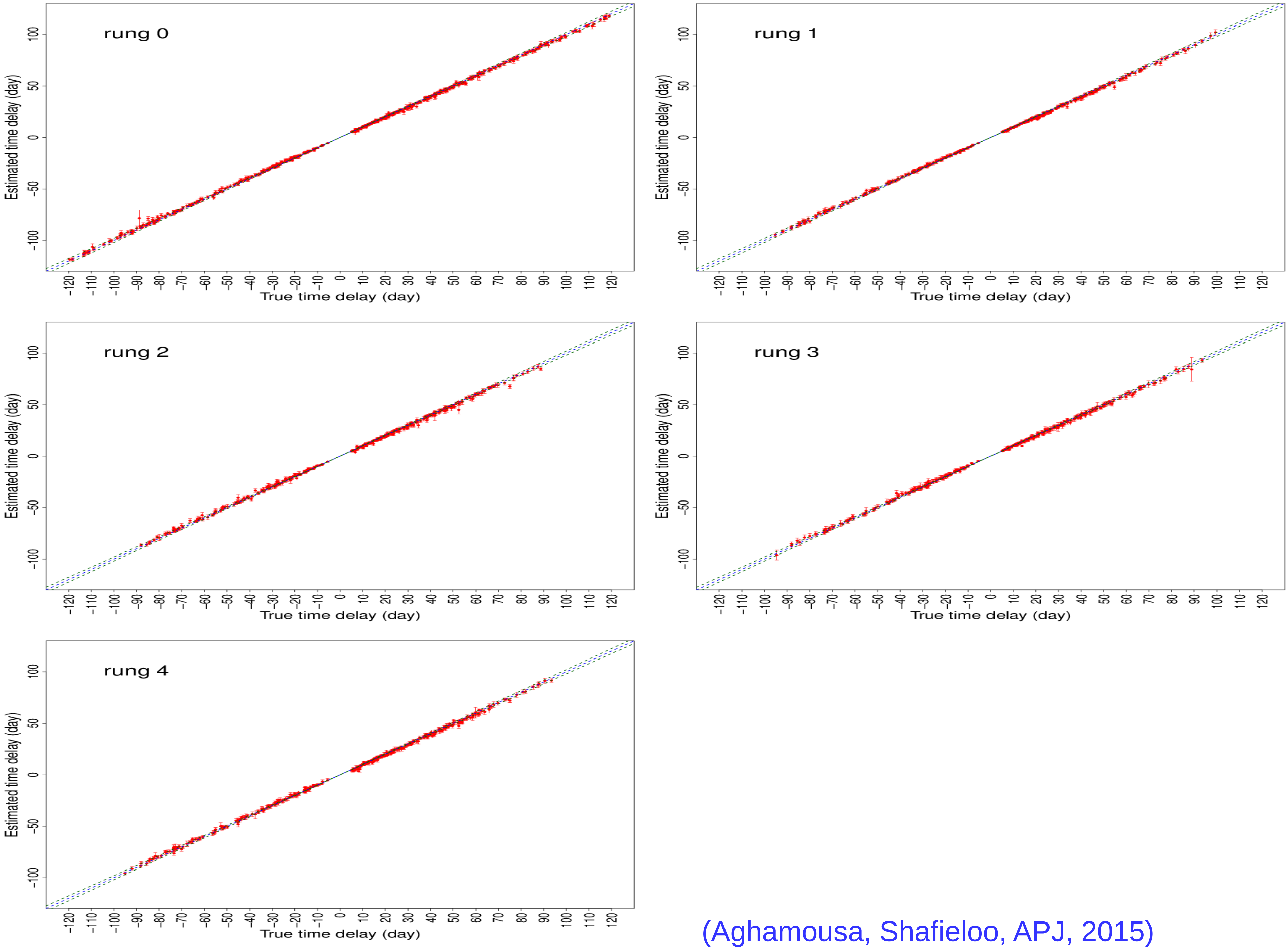
there are only 5 items with

$\chi^2 \geq 10$

out of few thousand entries

(Aghamousa, Shafieloo, APJ, 2015)

Results: estimated vs true time delay



(Aghamousa, Shafieloo, APJ, 2015)

Results: calibration

Rung	f	χ^2	P	A
0	0.529	0.579	0.038	-0.018
1	0.366	0.543	0.044	-0.022
2	0.350	0.885	0.053	-0.025
3	0.337	0.524	0.059	-0.021
4	0.346	0.608	0.056	-0.024

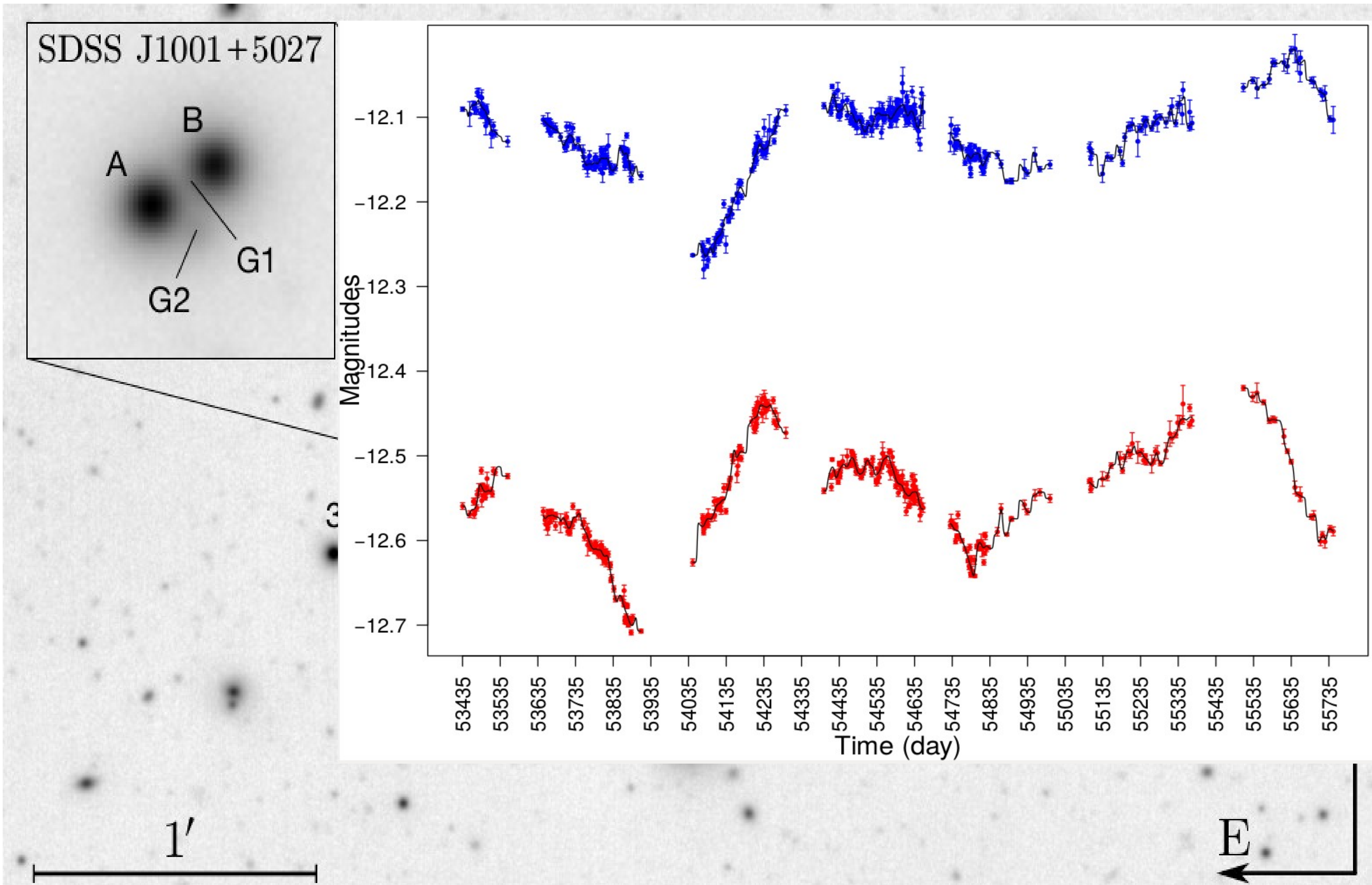
$$\tilde{\Delta t_i} + 0.5 \quad \text{and} \quad \sigma_i/\sqrt{2}$$

Rung	f	χ^2	P	A
0	0.529	0.792	0.027	-0.0014
1	0.366	0.660	0.031	-0.0036
2	0.350	1.439	0.038	-0.0058
3	0.337	0.766	0.041	-0.0010
4	0.346	0.868	0.040	-0.0048

(Aghamousa, Shafieloo, APJ, 2015)

Time delay estimation of SDSS J1001+5027

Light curves



Methodology

algorithm

$$A^s(t) = A^g(t) + N(t)\sum_i \frac{A^d(t_i) - A^g(t_i)}{\sigma_d^2(t_i)} \times exp\left[-\frac{(t_i - t)^2}{2\Delta^2}\right]$$

Weighted Correlation:

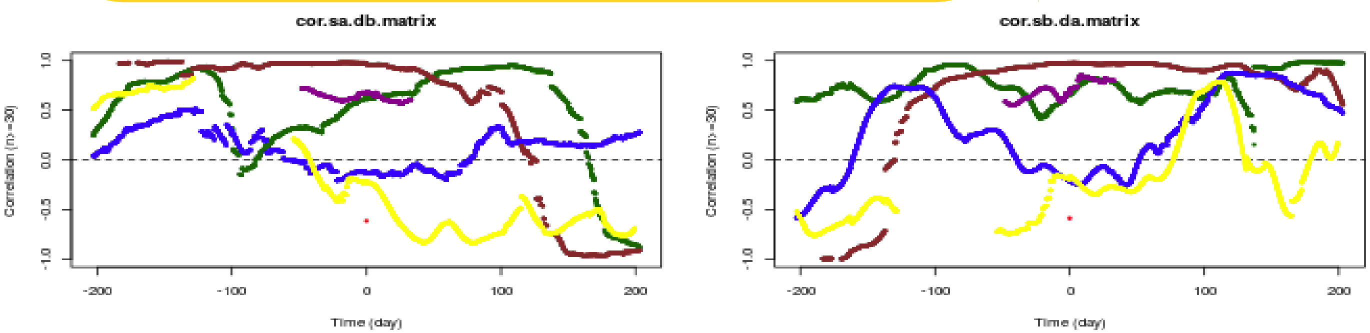
$$r_w = \frac{\sum_{i=1}^{n} w_i(x_i - \bar{x}_w)(y_i - \bar{y}_w)}{\sqrt{\sum_{i=1}^{n} w_i(x_i - \bar{x}_w)^2 \sum_{i=1}^{n} w_i(y_i - \bar{y}_w)^2}}$$

$$\bar{x}_w = \sum_{i=1}^{n} w_i x_i, \quad \bar{y}_w = \sum_{i=1}^{n} w_i y_i \quad w_i = \frac{1}{\sigma_i^2}$$

$\Delta = ?$

$n = ?$

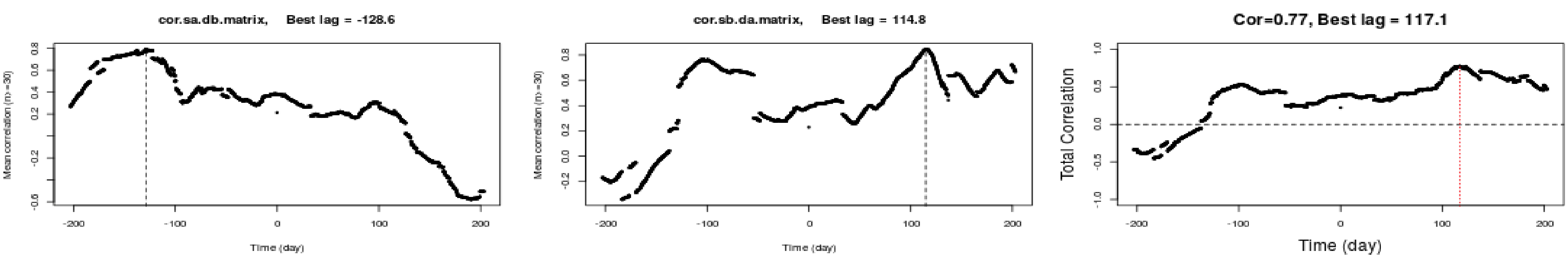
Mean or Mirror estimator?



Weighted Correlation

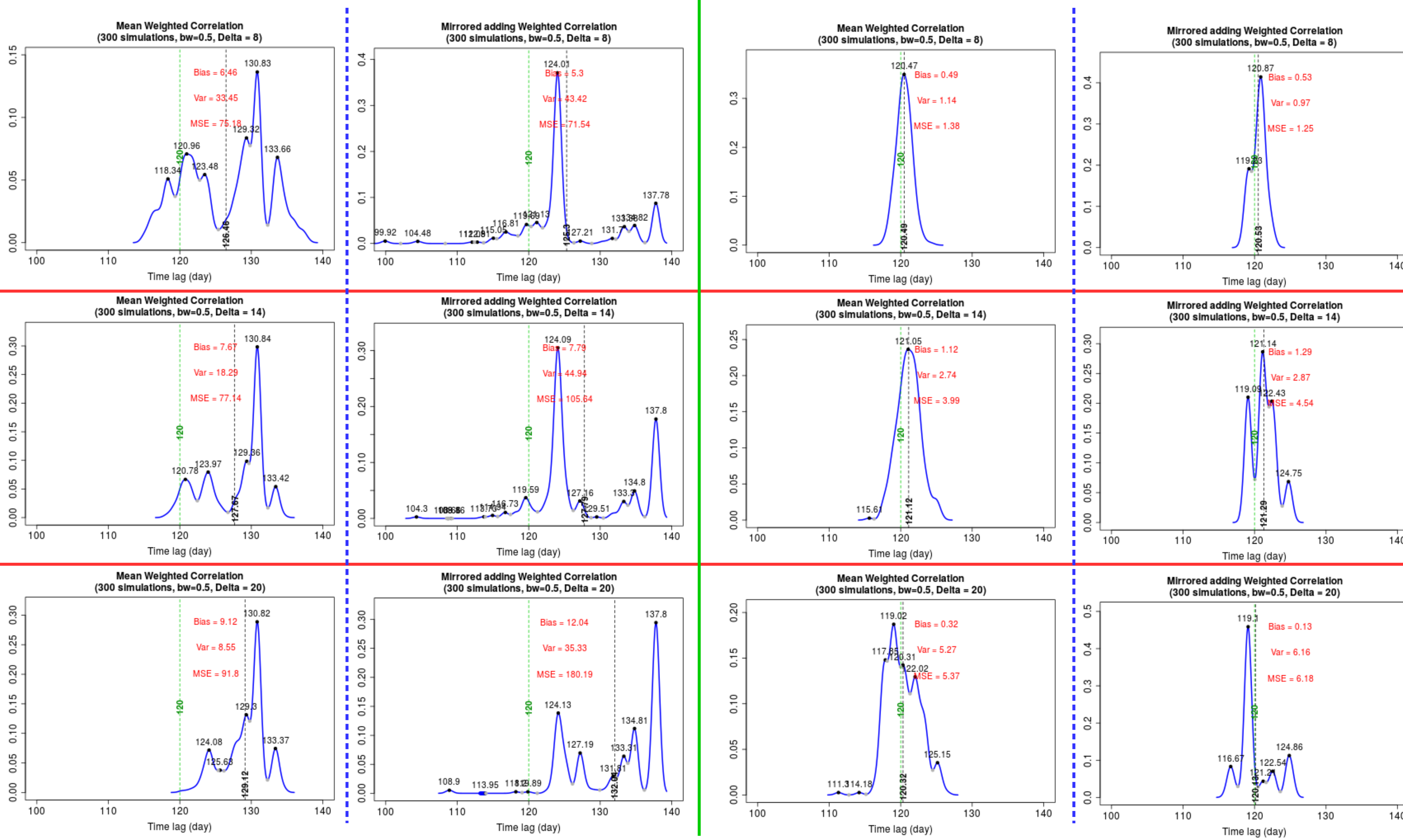
Delta = 8, n>30

Mean = 121.7 Mirror = 117.1



Simulation

	$n > 3$		$n > 30$	
	Mean	Mirror	Mean	Mirror
$\Delta = 8$				
$\Delta = 14$				
$\Delta = 20$				



Mean Squared Error (MSE)

$\hat{\theta}_n$ is the **estimator** of θ **unknown parameter/value**.

Bias:

$$\mathrm{Bias}(\hat{\theta}_n) = \mathbb{E}(\hat{\theta}_n) - \theta$$

Variance:

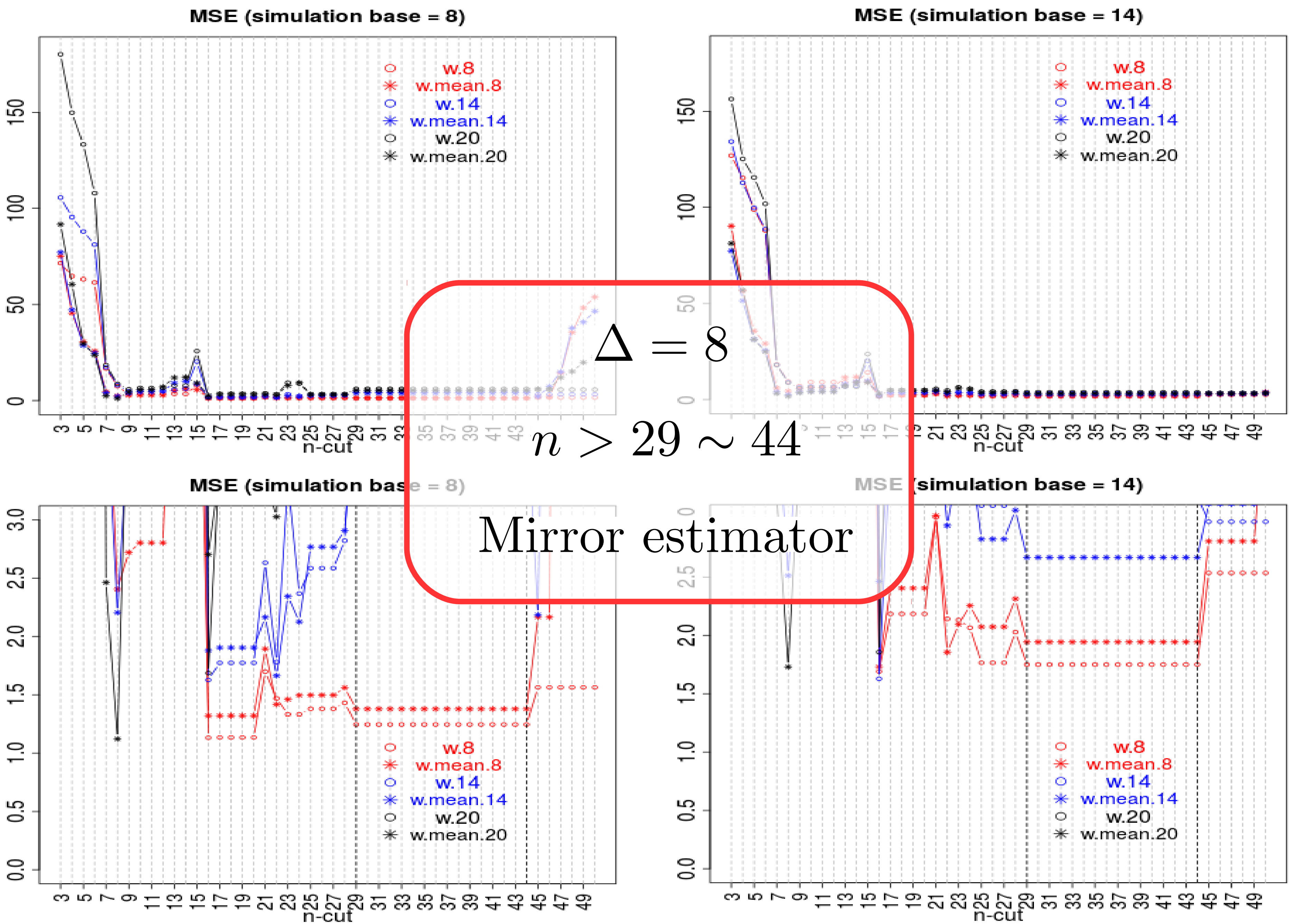
$$\mathrm{Var}(\hat{\theta}_n) = \mathbb{E}\left[(\hat{\theta}_n - \bar{\theta}_n)^2\right] \qquad \text{where} \quad \bar{\theta}_n = \mathbb{E}(\hat{\theta}_n)$$

Mean Squared Error (MSE)

$$\mathrm{MSE}(\hat{\theta}_n) = \mathbb{E}\left[(\hat{\theta}_n - \theta)^2\right] = \cdots = \left(\mathbb{E}(\hat{\theta}_n) - \theta\right)^2 + \mathbb{E}\left[(\hat{\theta}_n - \bar{\theta}_n)^2\right]$$

$$\mathrm{MSE} = \mathrm{Bias}^2 + \mathrm{Var}$$

Optimum estimator



$$\Delta = 8$$

$$n > 29 \sim 44$$

Mirror estimator

Error estimation

- **Confidence set.**

$$\mathbb{P}(a \leq \theta \leq b) = (1-\alpha) \equiv \psi\%$$

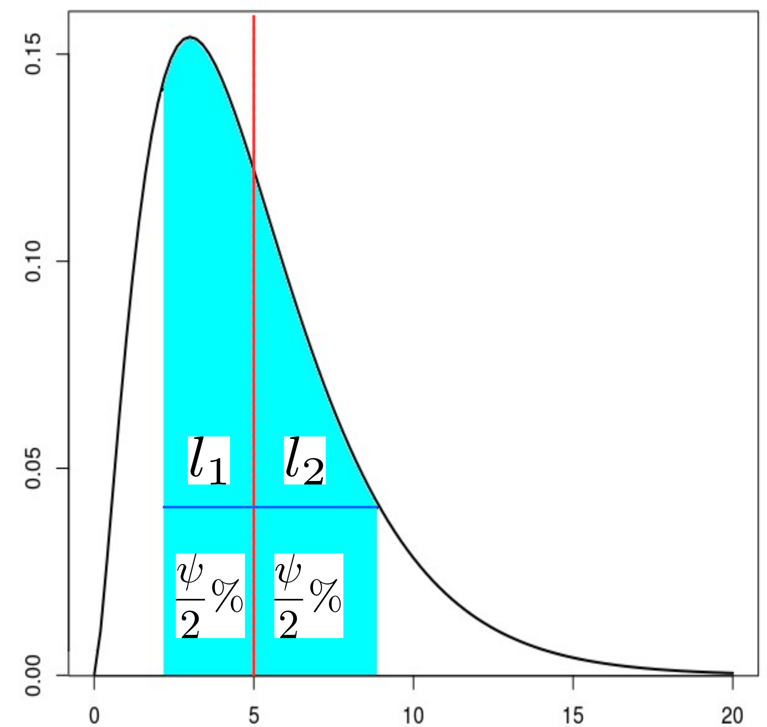
  - To find a confidence interval we need the **probability distribution of estimator** which can be known from **statistical properties of estimator** or can be achieved by **simulation**.
  - In some cases we know the probability distribution of a quantity which has a relation to estimator.

$$\mathbb{P}\left(\mathbb{E}(\hat{\theta}_n) - l_1 \leq \quad \hat{\theta}_n \quad \leq \mathbb{E}(\hat{\theta}_n) + l_2\right) = (1-\alpha) \equiv \psi\%$$

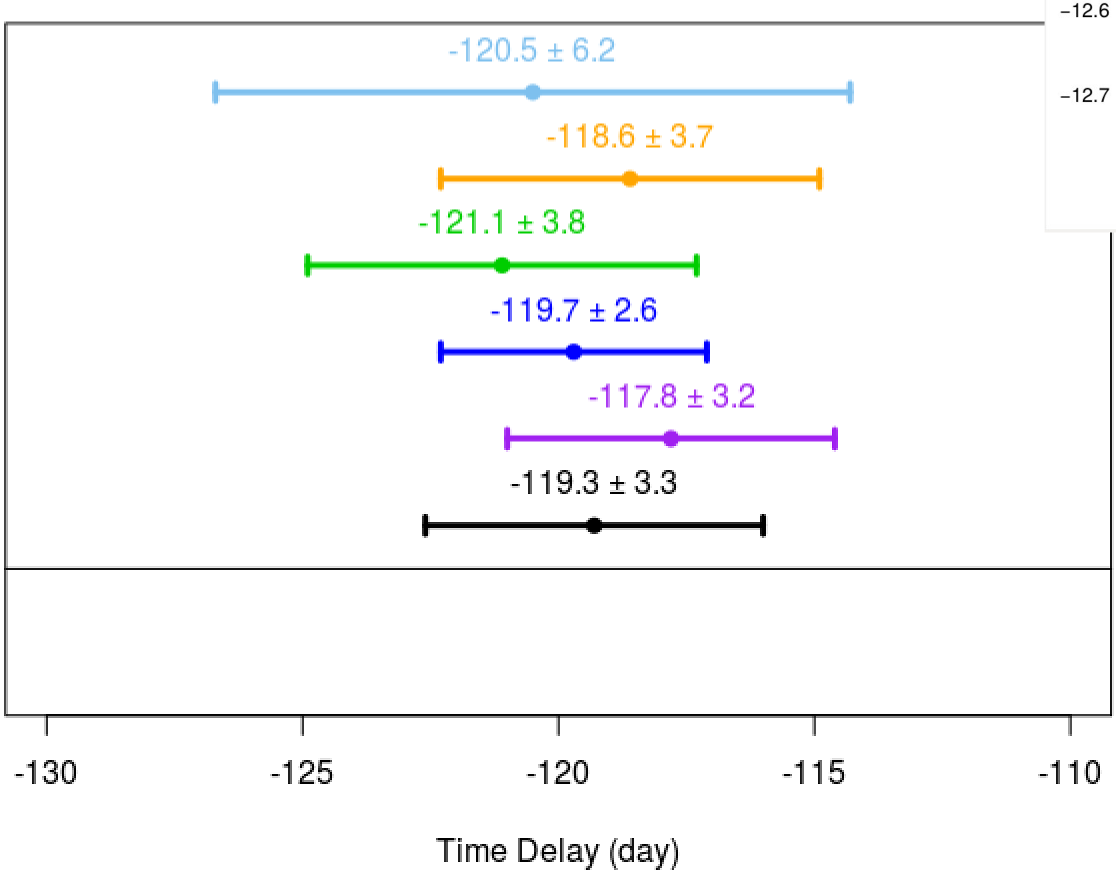
$$\mathbb{P}\left(\mathbb{E}(\hat{\theta}_n) - \theta) - l_1 \leq \hat{\theta}_n - \theta \leq (\mathbb{E}(\hat{\theta}_n) - \theta) + l_2\right)$$

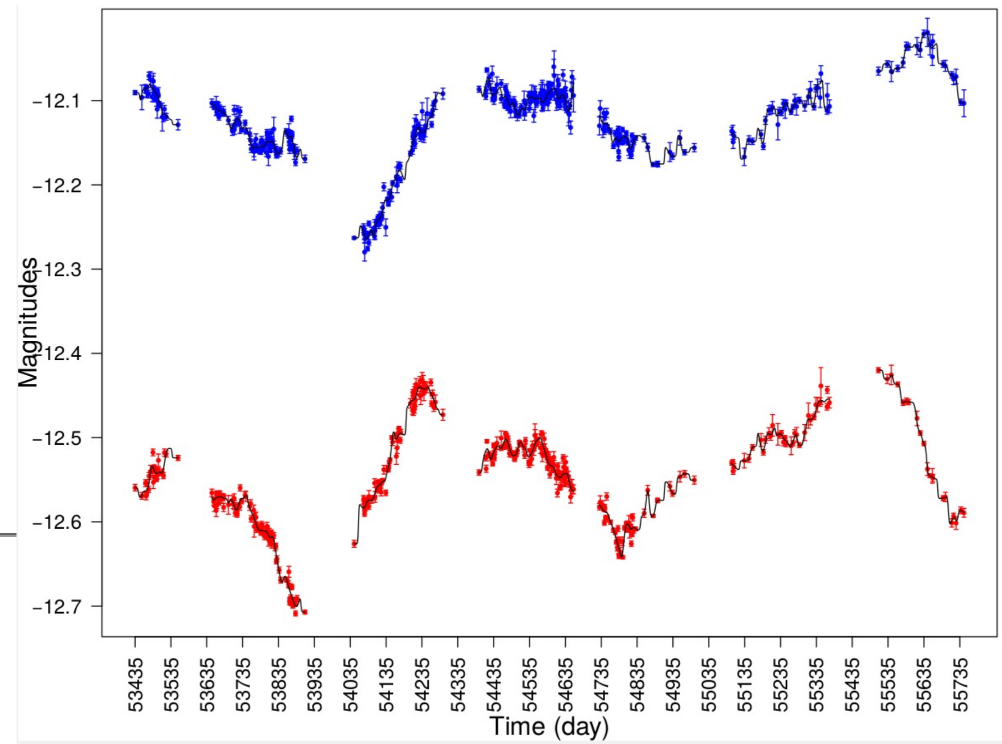
$$\mathbb{P}\left(\underbrace{\hat{\theta} - l_2 - \overbrace{(\mathbb{E}(\hat{\theta}) - \theta)}^{\text{Bias}}}_{a} \leq \theta \leq \underbrace{\hat{\theta} + l_1 - \overbrace{(\mathbb{E}(\hat{\theta}) - \theta)}^{\text{Bias}}}_{b}\right)$$

$$\mathbb{P}(a \leq \theta \leq b) = (1-\alpha) \equiv \psi\%$$



Results





$$\Delta = 8$$

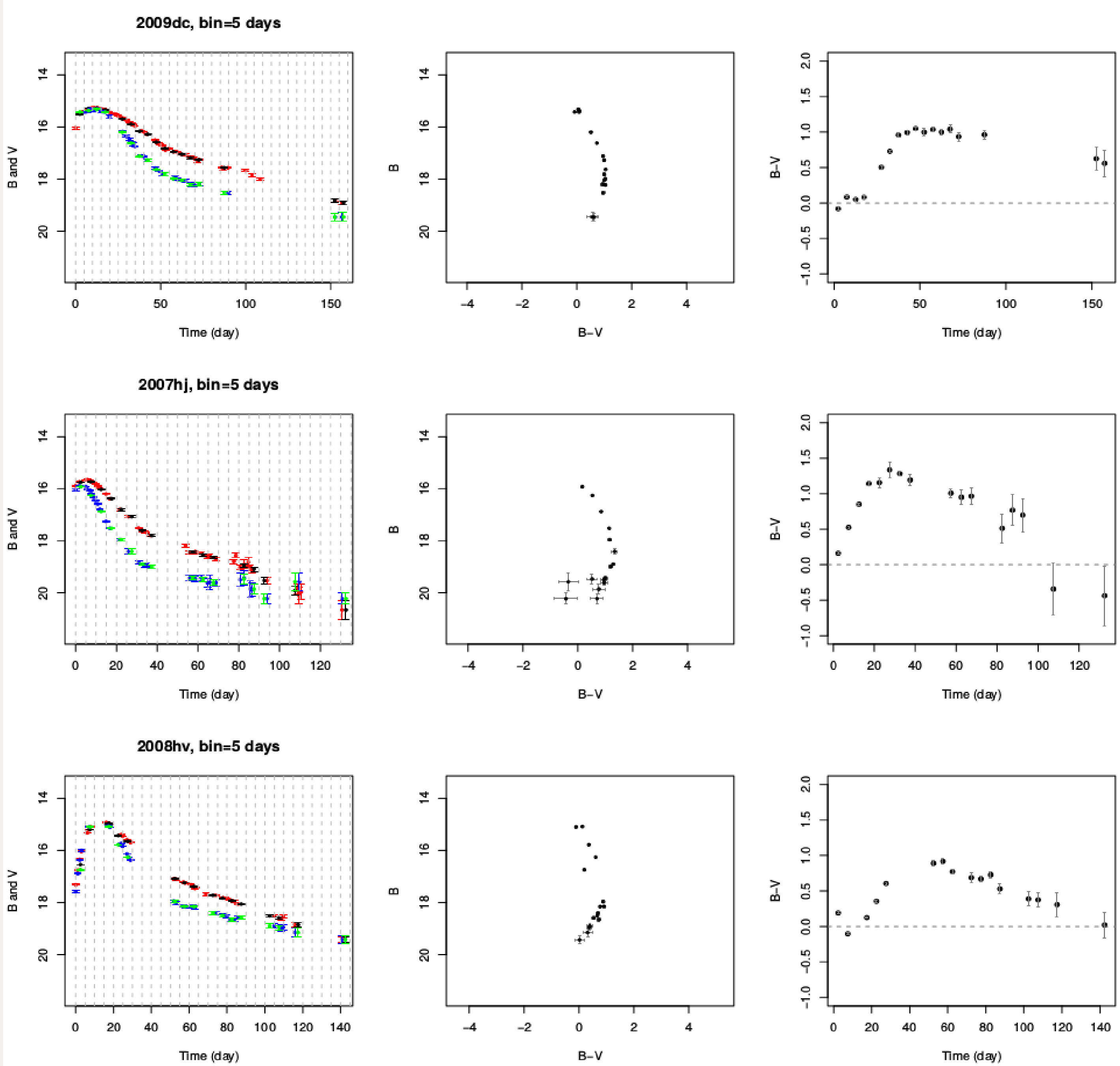
$$n > 29 \sim 44$$

Mirror estimator

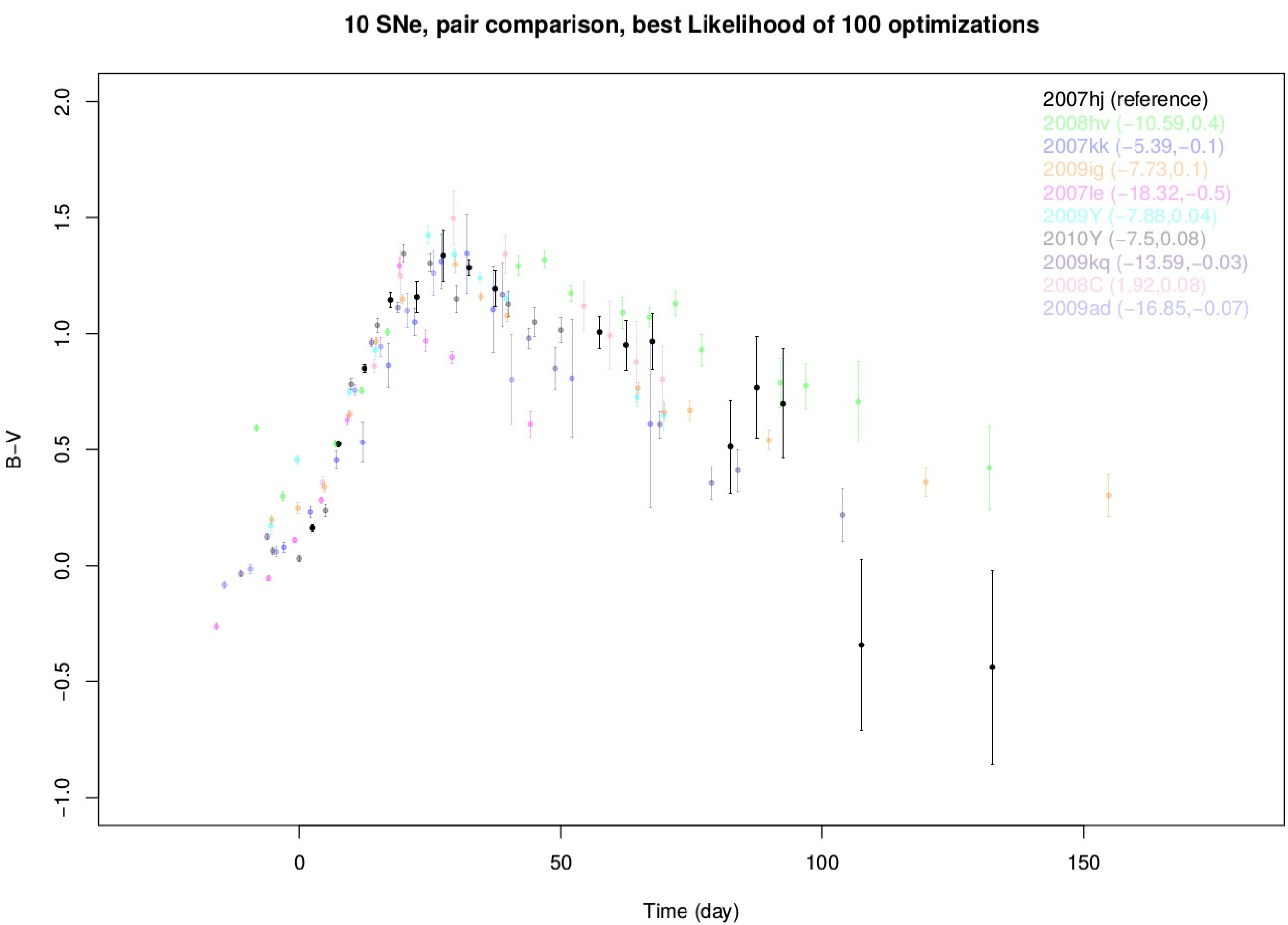
$$\mathbb{P}(a \leq \theta \leq b) = (1-\alpha) \equiv \psi\%$$

Supernova Type Ia classification

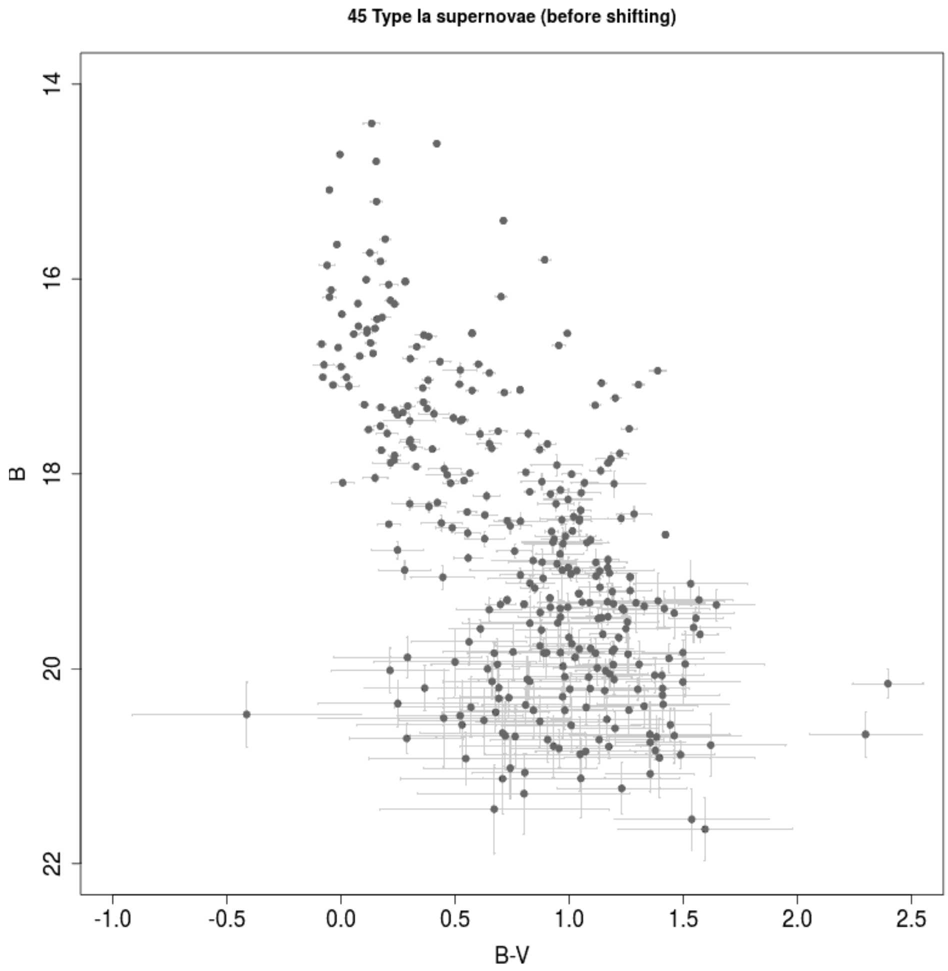
Binned light curves



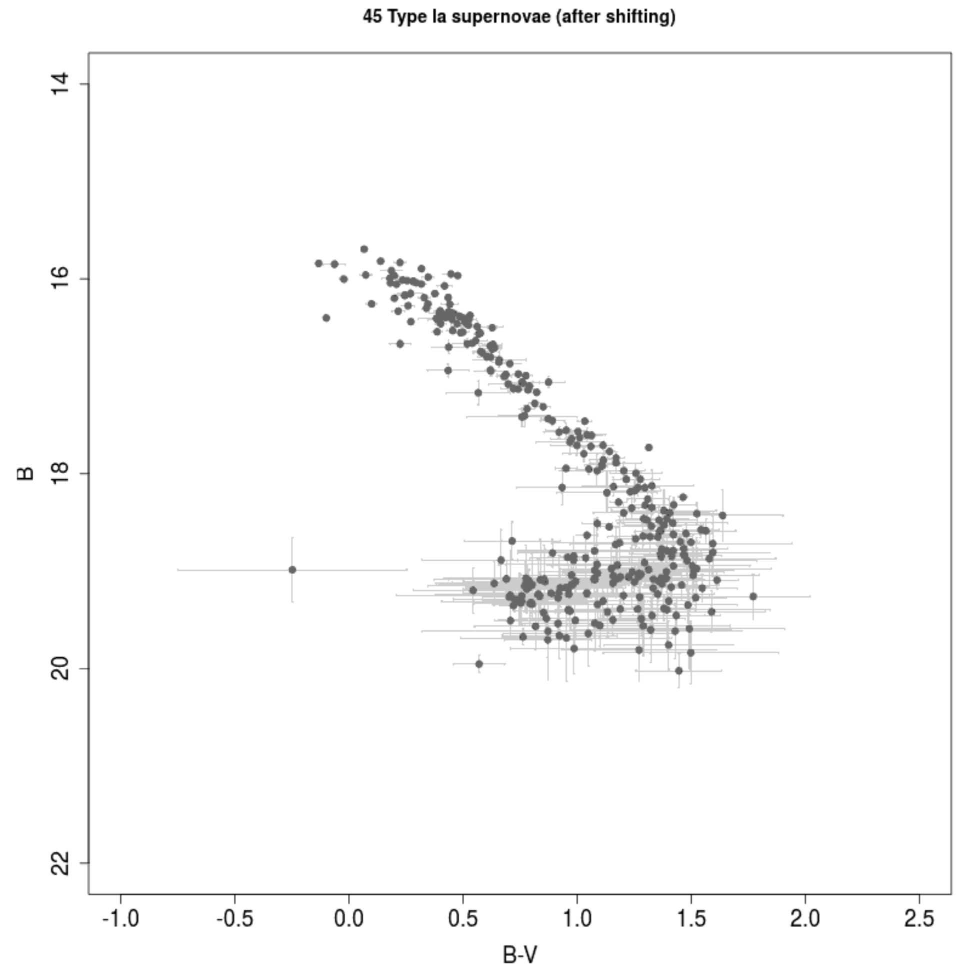
Results


10 SNe, pair comparison, best Likelihood of 100 optimizations
2007hj (reference)
2008hv (−10.59,0.4)
2007kk (−5.39,−0.1)
2009ig (−7.73,0.1)
2007le (−18.32,−0.5)
2009Y (−7.88,0.04)
2010Y (−7.5,0.08)
2009kq (−13.59,−0.03)
2008C (1.92,0.08)
2009ad (−16.85,−0.07)
B−V
Time (day)
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
0
50
100
150

Results



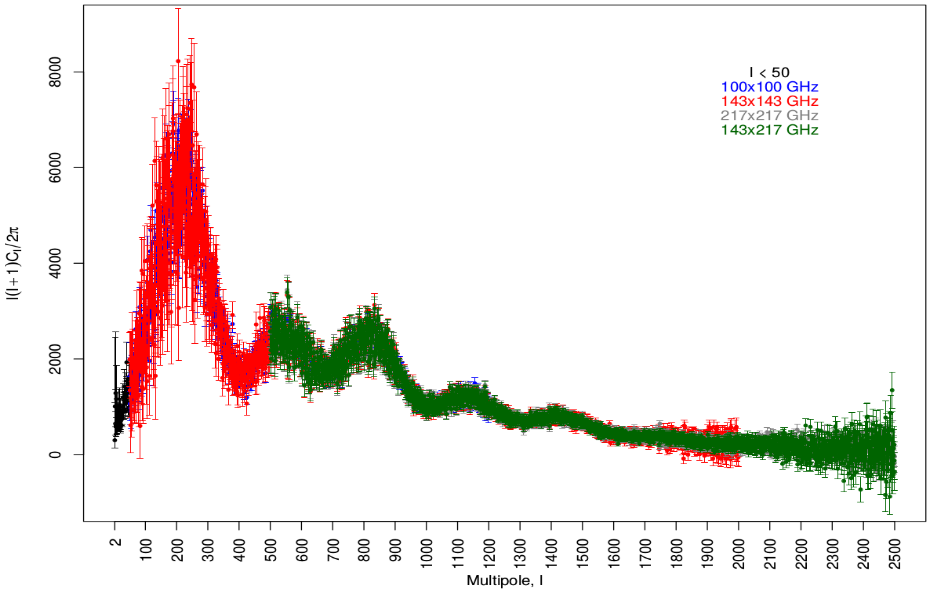
45 Type Ia supernovae (before shifting)
B
B-V
14
16
18
20
22
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5

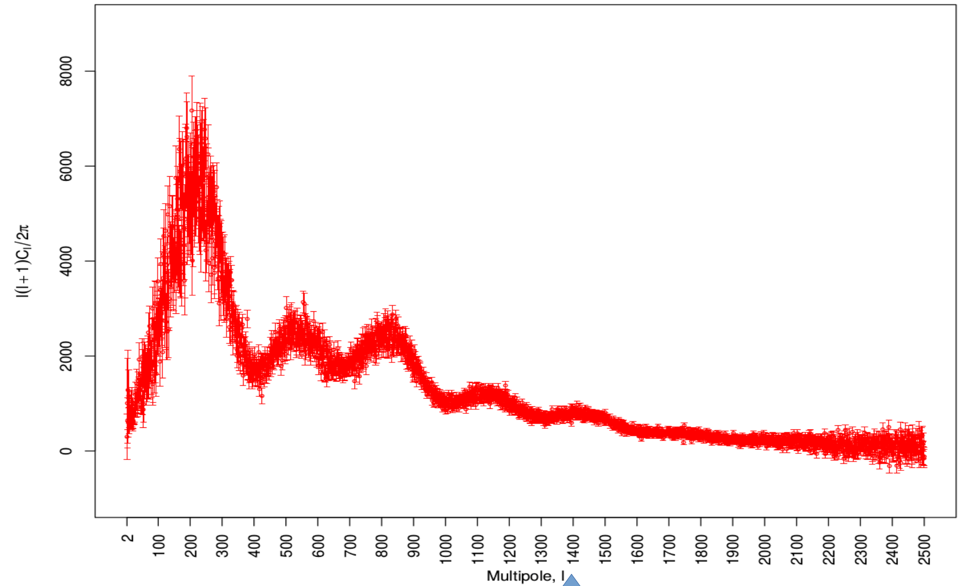


45 Type Ia supernovae (after shifting)
B
B-V
14
16
18
20
22
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5

Model-independent estimation of CMB angular power spectrum

CMB Angular power spectrum data





$$\bar{\mathcal{D}}_\ell = \frac{\sum_{ch} w_\ell^{ch} \mathcal{D}_\ell^{ch}}{\sum_{ch} w_\ell^{ch}}, \quad w_\ell^{ch} = (\sigma_\ell^{ch})^{-2}$$

$$\mathrm{Cov}(\bar{\mathcal{D}}_\ell, \bar{\mathcal{D}}_{\ell'}) = \frac{1}{\sum_{ch} w^{ch} \sum_{ch'} w^{ch'}} \sum_{ch,ch'} w_\ell^{ch} w_{\ell'}^{ch'} \mathrm{Cov}(\mathcal{D}_\ell^{ch}, \mathcal{D}_{\ell'}^{ch'})$$

(Aghamousa, Shafieloo, Arjunwadkar and Souradeep, JCAP, 2015)

Parametric regression

$Y_i = f(x_i) + \epsilon_i$

- Assume $f(x) = ax + b$.
- Assume noise $\epsilon_i \sim N(0, \sigma^2)$ IID.
- Likelihood function

$$L(a, b|\text{data}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(Y_i - (ax_i + b))^2}{2\sigma^2}\right)$$

- To estimate a, b: Maximize $L(a, b|\text{data})$ w.r.t. a, b.
- This is same as linear least-squares regression, under the assumptions made.

REACT: nonparametric regression

- $Y_i = f(x_i) + \epsilon_i$, with $\epsilon_i \sim N(0, \sigma^2)$ IID, σ^2 known.
- Assume $f \in L_2(a, b)$ and a complete orthonormal basis $\{\phi_j(x)\}$.

$$f(x) = \sum_{j=0}^{\infty} \beta_j \phi_j(x), \; \beta_j = \int_a^b f(x)\phi_j(x)dx$$

- Regression estimator $\widehat{f}(x)$:

$$f(x) = \sum_{j=0}^{n-1} \widehat{\beta}_j \phi_j(x) + \text{(some truncation bias)}$$

$\widehat{\beta}_j := \lambda_j Z_j$ with $1 \geq \lambda_0 \geq \ldots \geq \lambda_{n-1} \geq 0.$ and $Z_j = \sum_{i=1}^{n} Y_i \phi_j(x_i)$

- Inverse-noise-weighted squared loss function

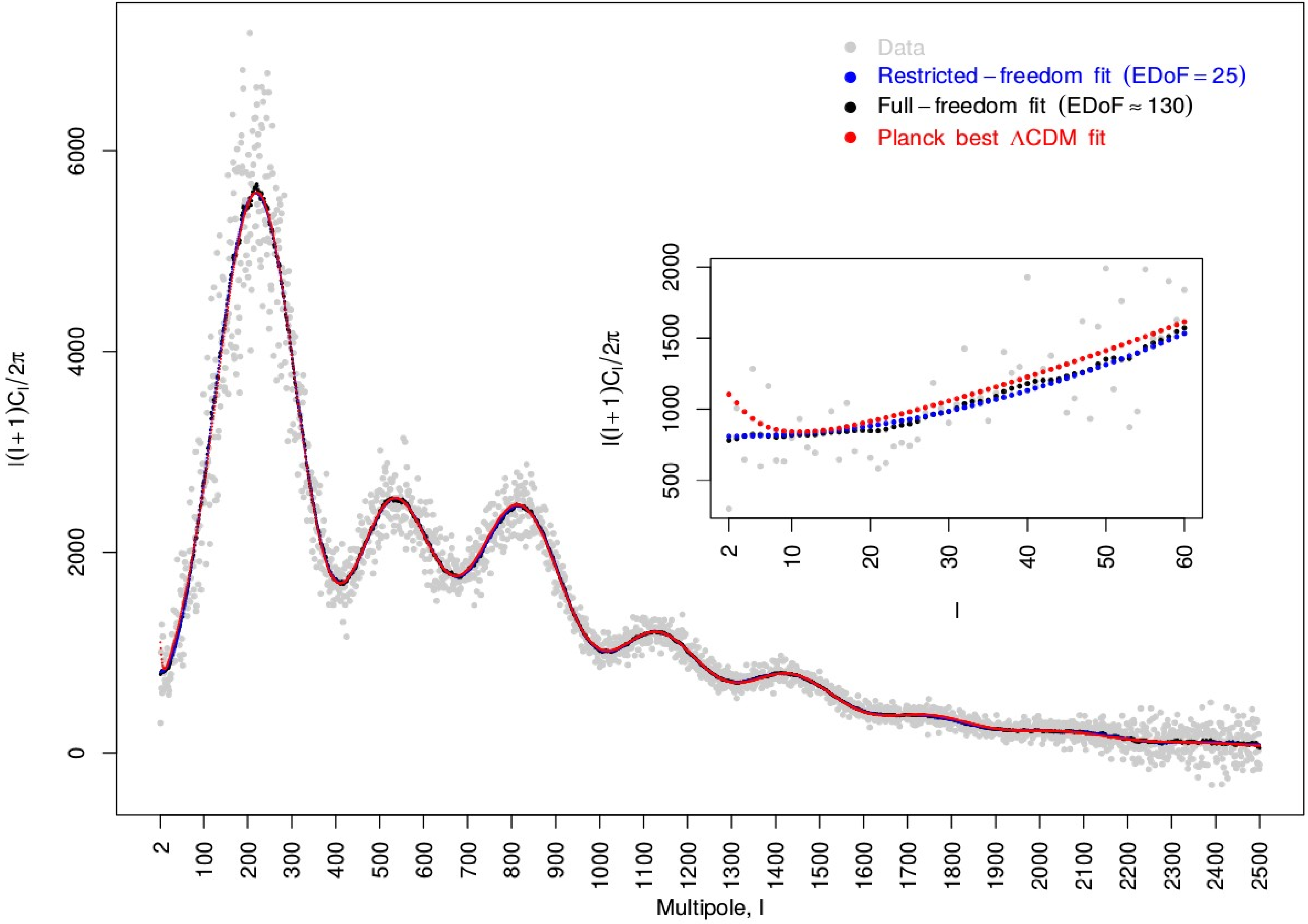
$$L(\widehat{f}, f) = \int \left(\frac{\widehat{f}(x) - f(x)}{\sigma(x)} \right)^2 dx.$$

- Risk estimator

$$\widehat{R}(\lambda) = Z^T \bar{D} W \bar{D} Z + \mathrm{tr}(DWDB) - \mathrm{tr}(\bar{D} W \bar{D} B),$$

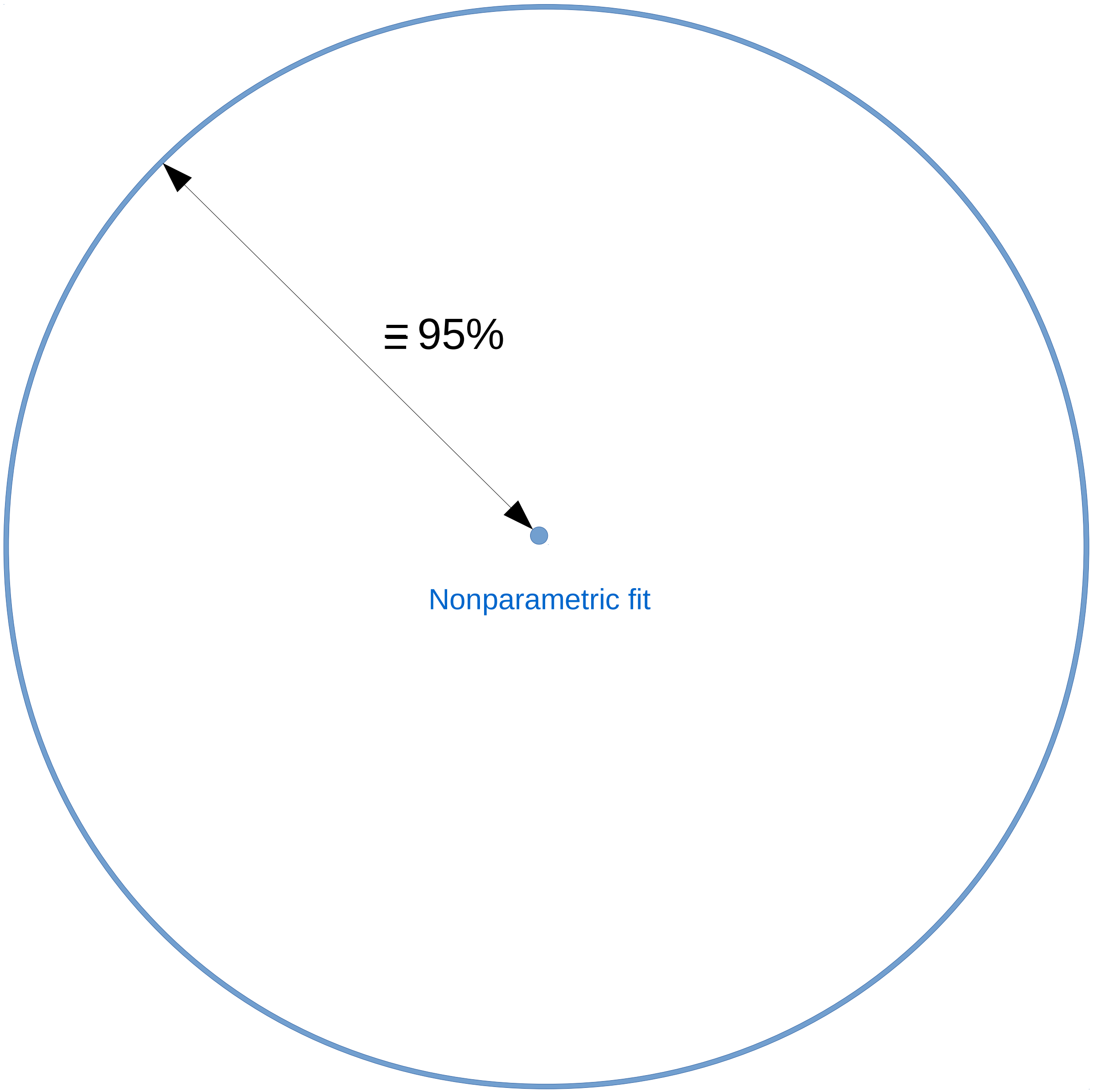
subject to the constraint $1 \geq \lambda_0 \geq \ldots \geq \lambda_{n-1} \geq 0.$

Planck 2013 power spectrum estimation

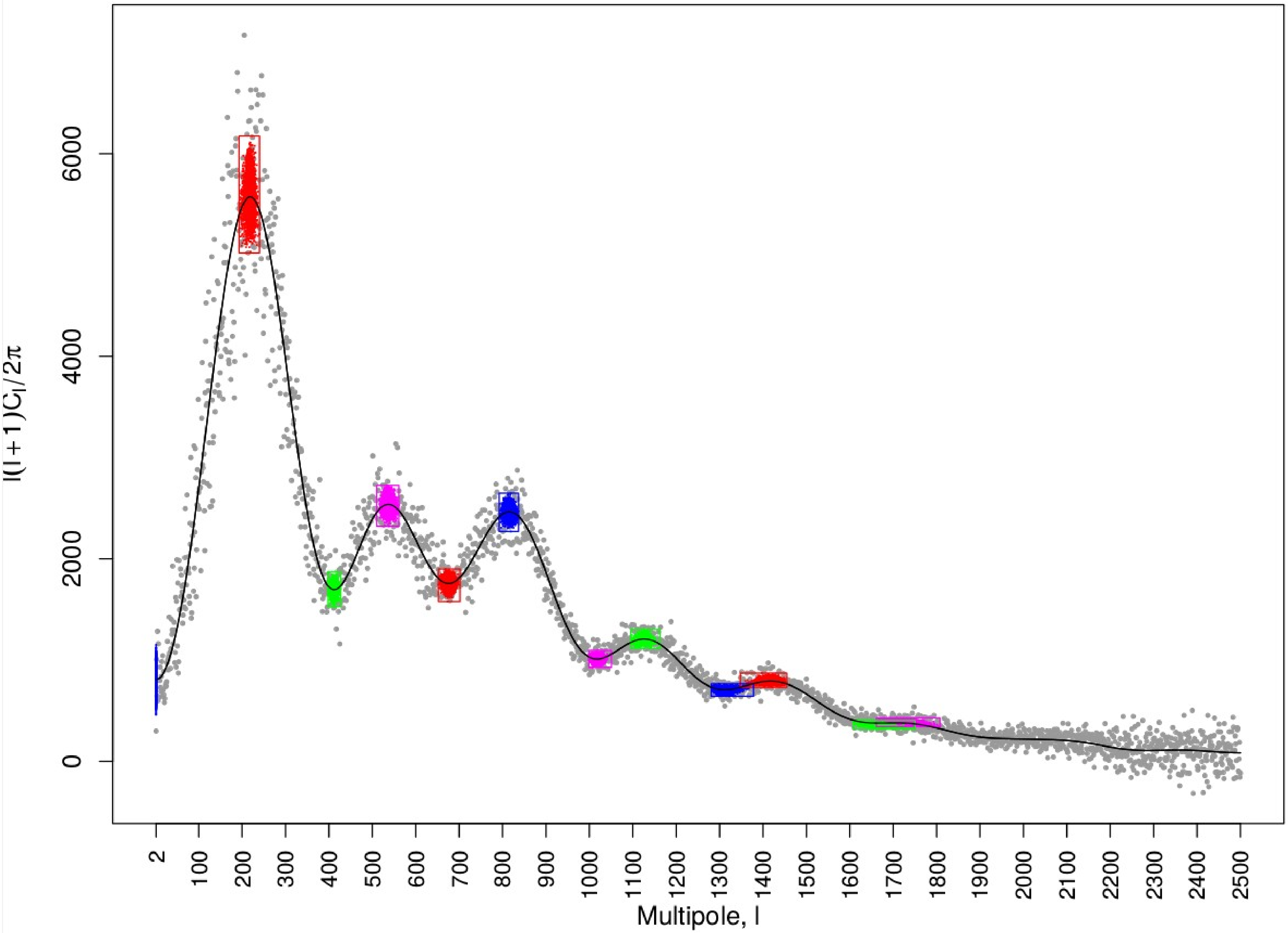


(Aghamousa, Shafieloo, Arjunwadkar and Souradeep, JCAP, 2015)

Confidence set in Function space

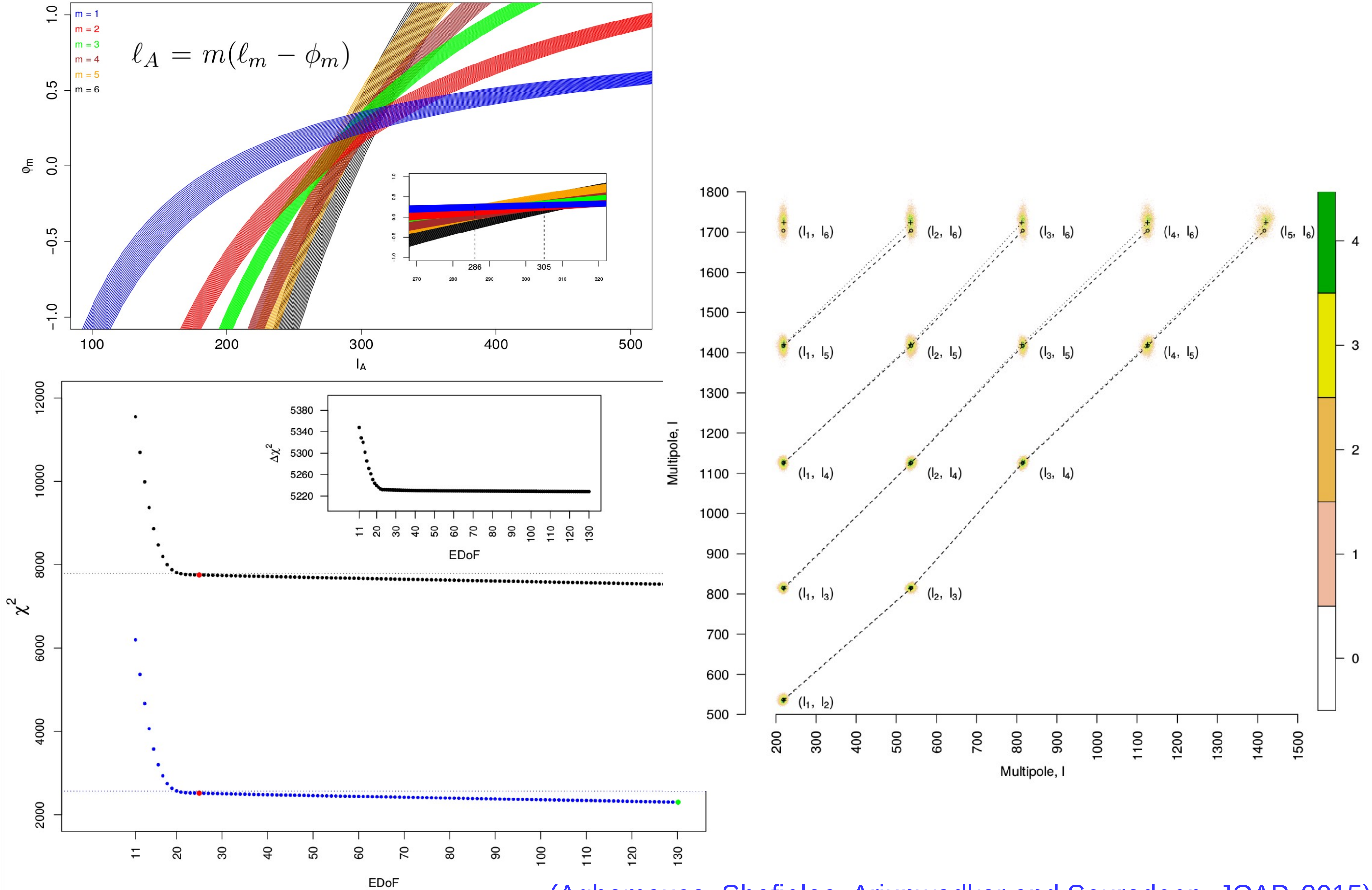


Picks and dips



(Aghamousa, Shafieloo, Arjunwadkar and Souradeep, JCAP, 2015)

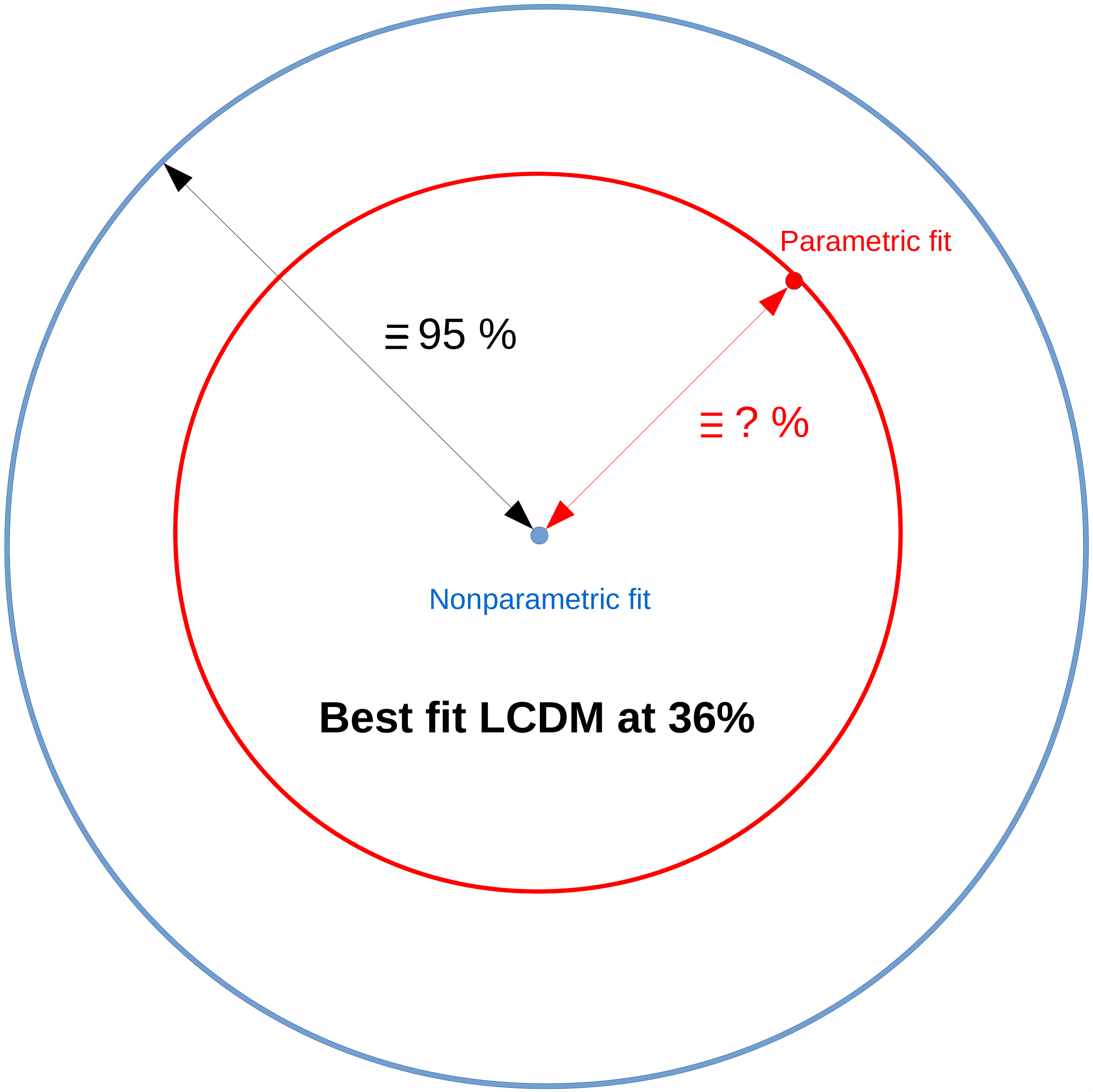
Other results



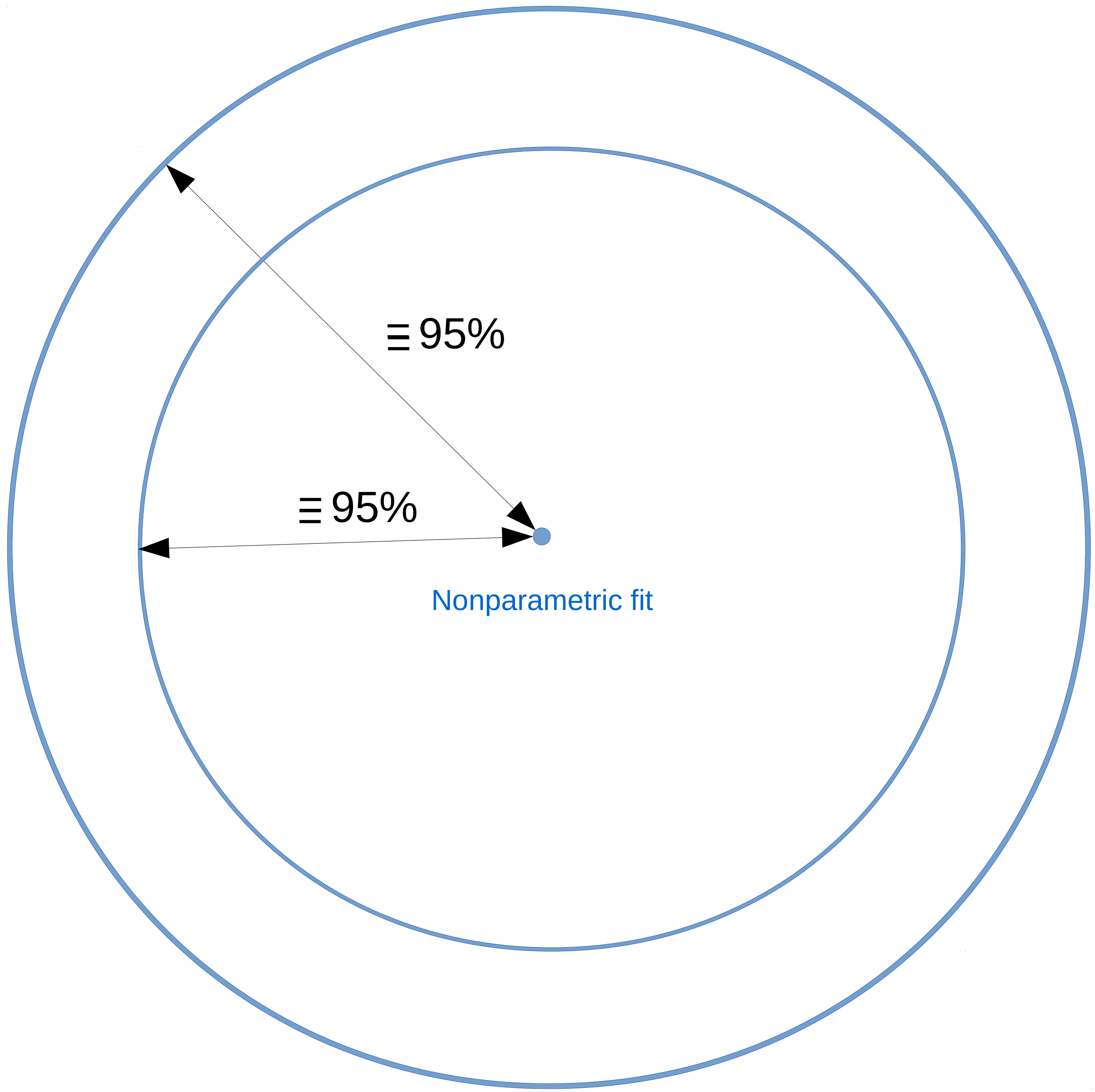
(Aghamousa, Shafieloo, Arjunwadkar and Souradeep, JCAP, 2015)

Nonparametric test of consistency between cosmological models and CMB data

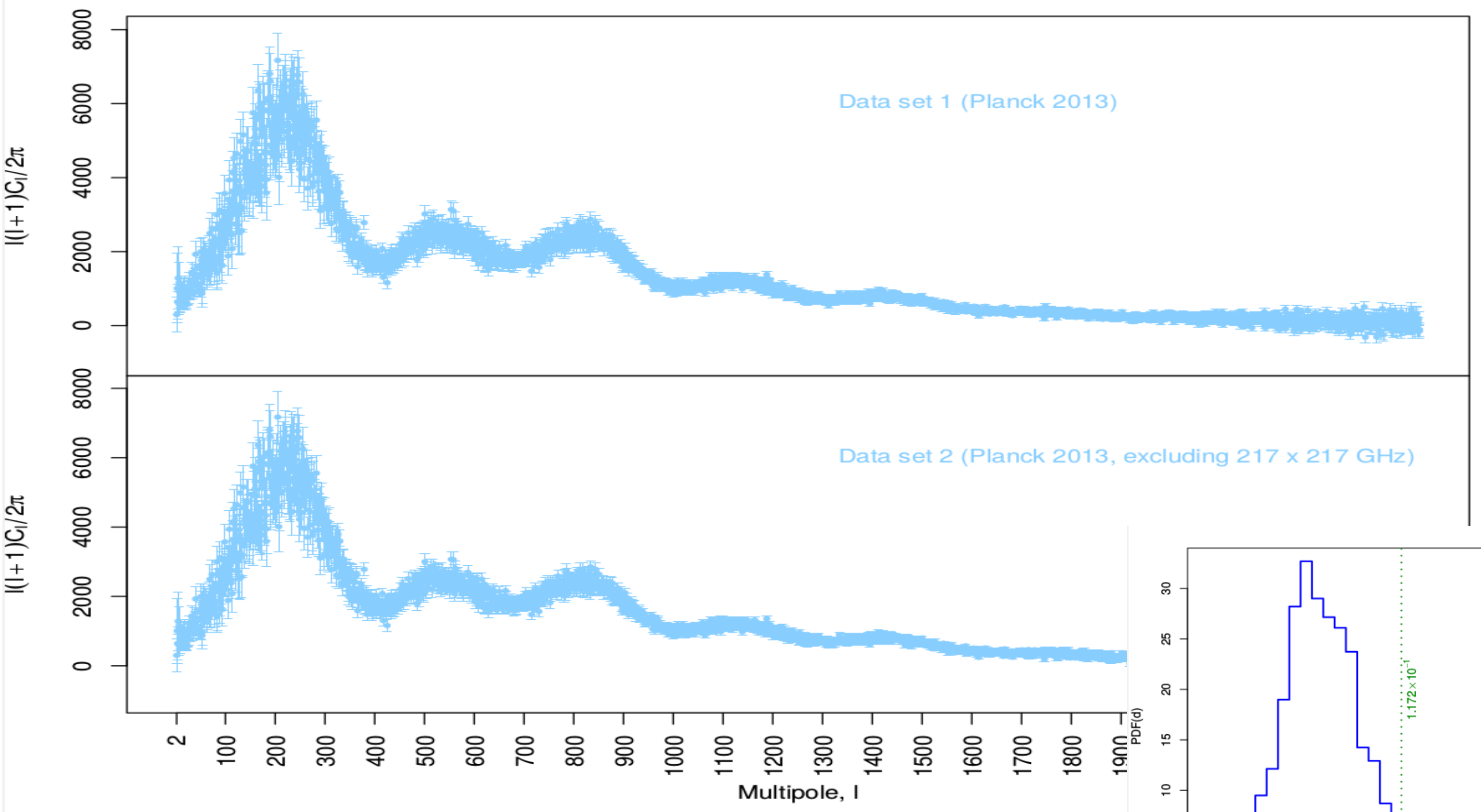
Validating Cosmological Models

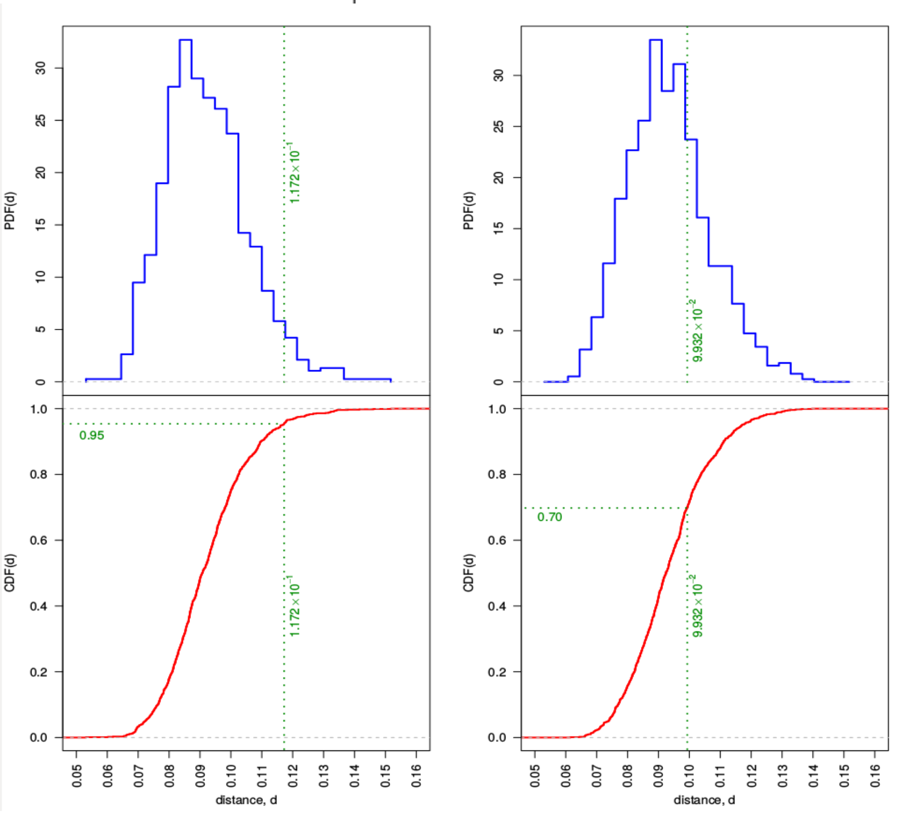


Calibrating Confidence distances



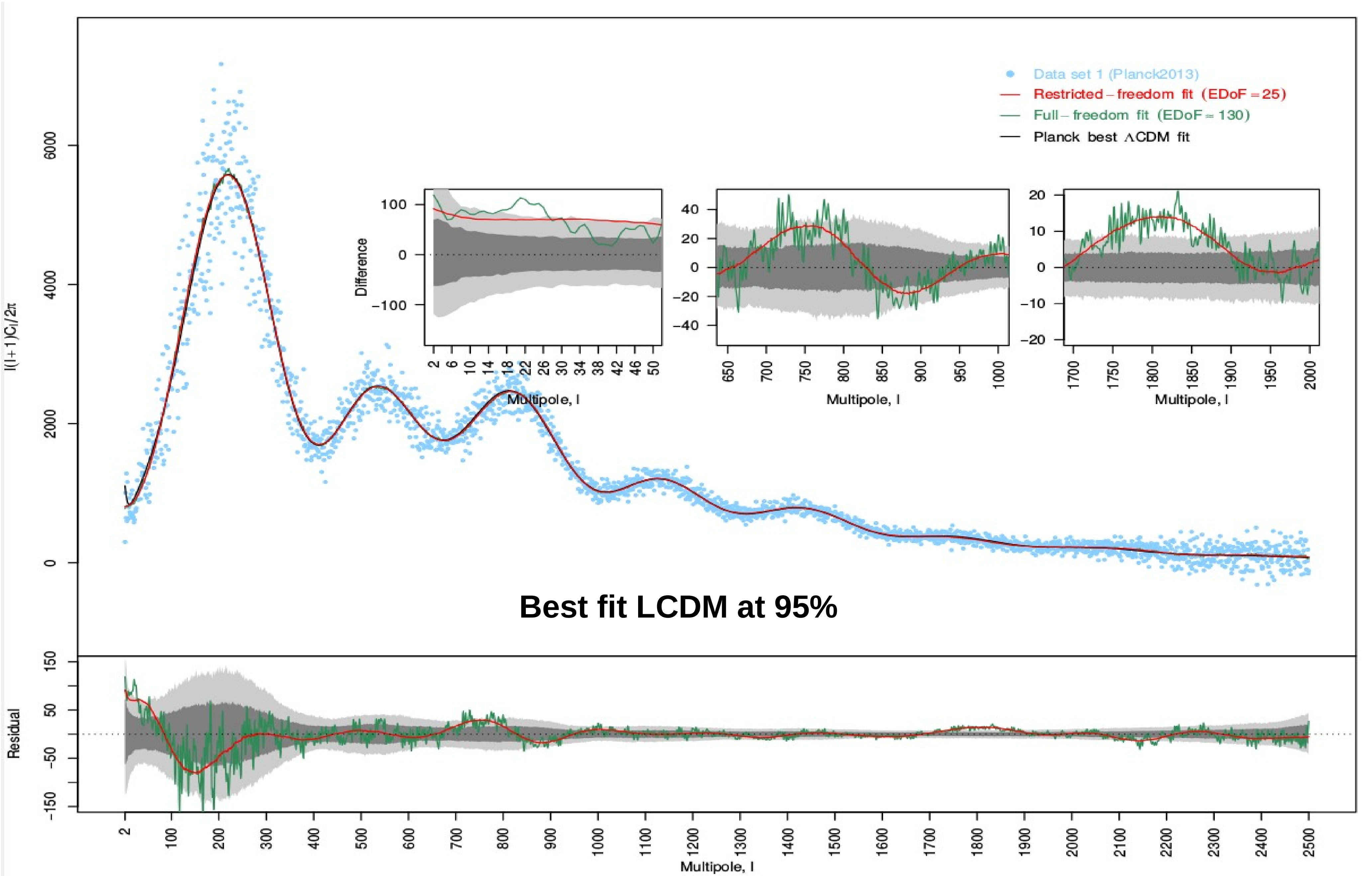
Calibrating Confidence distances





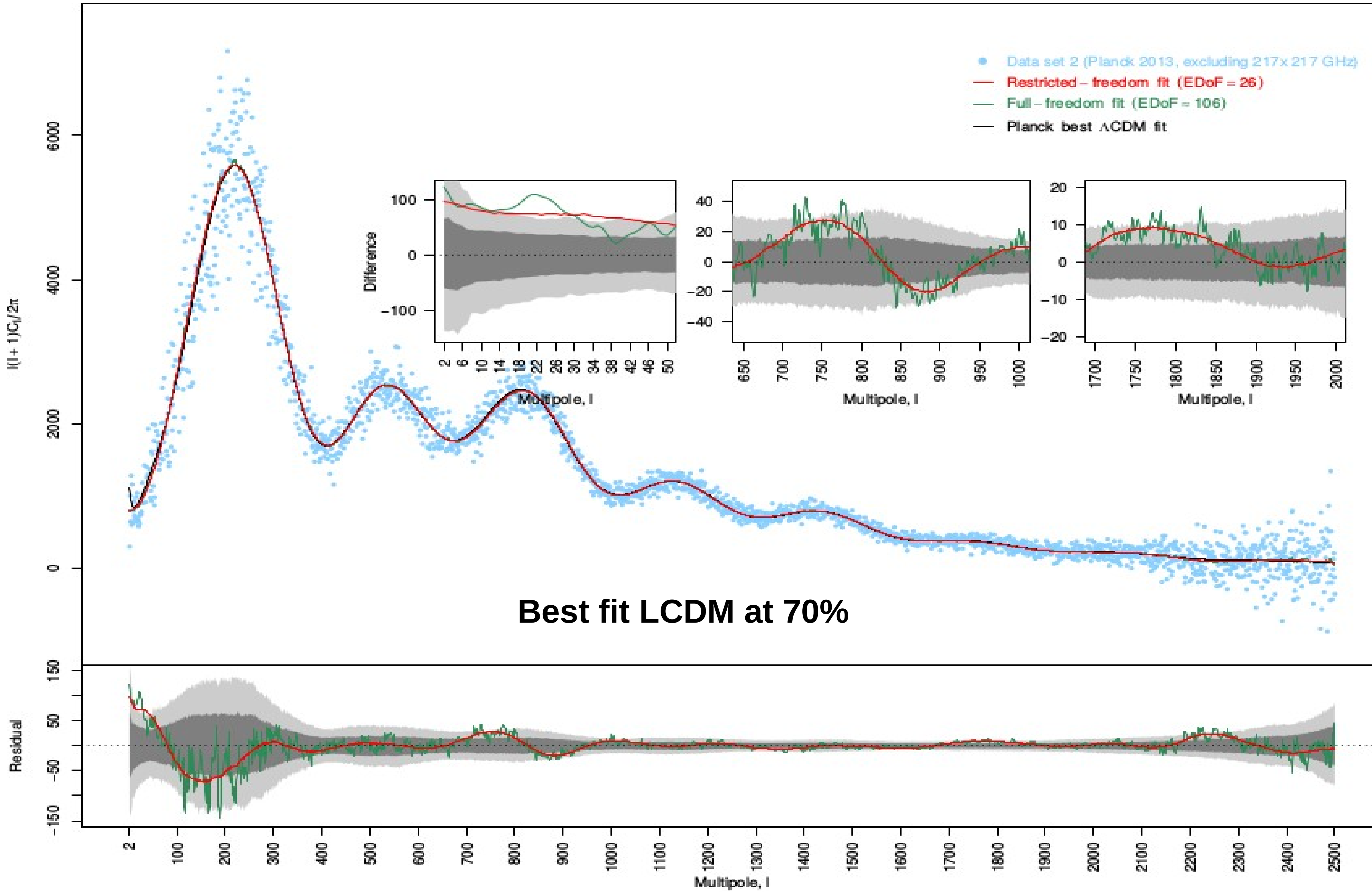
(Aghamousa, Shafieloo, JCAP, 2015)

Bias control



(Aghamousa, Shafieloo, JCAP, 2015)

Bias control



(Aghamousa, Shafieloo, JCAP, 2015)

Conclusion

- Strong Lens Time Delay Challenge II: designing an algorithm for time delay estimation based on the smoothing and cross correlation.

- Time delay estimation of SDSS J1001+5027: Applying improved algorithm on real data. Using weighted correlation and number of data criterion.

- Supernova Type Ia classification: Using Gaussian process and 2-D Kernel density to find a unique pattern for supernova type Ia.

- Model-independent estimation of CMB angular power spectrum: estimation power spectrum, uncertainties of peaks and dips, harmonicity of peaks and a proper estimation of likelihood.

- Nonparametric test of consistency between cosmological models and CMB data: consistency check of Best LCDM power spectrum with Planck 2013 data with and without 217 channel data.

- Finally: using advanced statistics and modified techniques enable us to infer more accurate and robust results from the data.