Universität Bielefeld

Outlier Rejection Scheme for EPTA Frequency-resolved Timing

Jun Wang

Fakultät für Physik
Universität Bielefeld
https://www.AstronDog.com/

May 15, 2019

Outline

Introduction

Outliers influence the Frequency-resolved timing

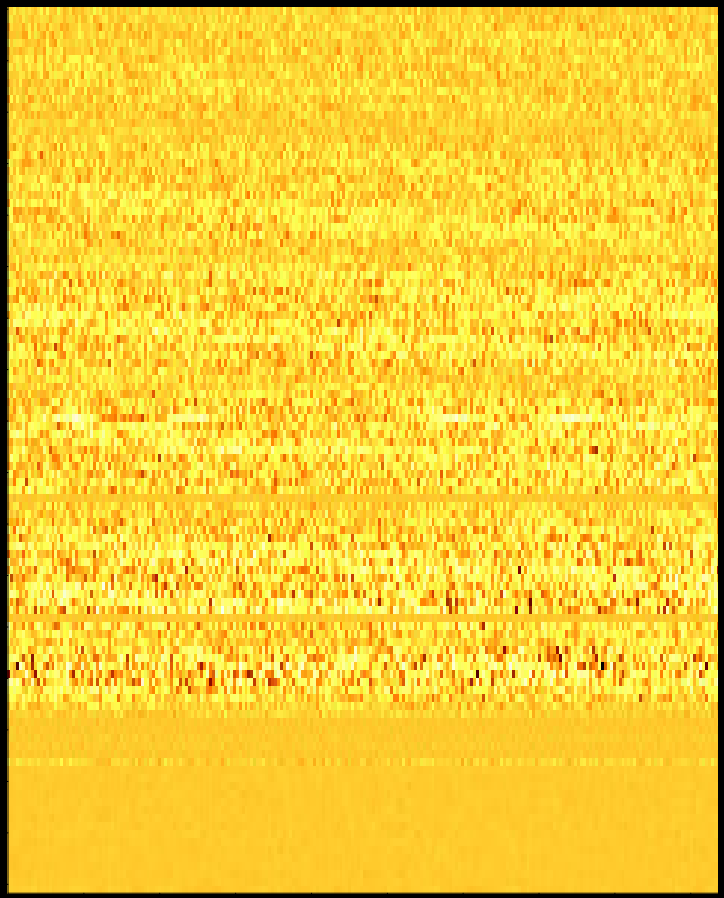


Figure: A frequency series of PSR J0218+4232, Lovell.

Outliers

1. Incomplete RFI Removal.
2. Low S/N.
3. Bad Channels.
4. Incorrect Calibration.
5. Instrumental failure.
6. ...

How to determine if the Scheme is optimal

Comprehensively analyze the following facors:

1. Minimise RMS.
2. Chisq is slightly larger than 1.
3. Maximize the Total Information Content $\frac{\sqrt{N}}{\sigma}$.
4. More ToAs are preserved.

Outlier Rejection Scheme

Scheme Evolution

Previous thoughts:

1. Goodness-of-fit criterion in the template fit process (0.5-1.5).
2. Median Standard Deviation(MAD).
3. Fitting the residuals with Gaussian distribution model and remove non-Gaussian ToAs.
4. S/N criterion(S/N above 8).

Conclusion:

1. Median standard deviation.($k = 3$)
2. Fitting the Normalized residuals with Gaussian distribution model.($k = 3$)

Outlier Rejection Scheme
1. Median Absolute Deviation(MAD)

What is MAD

The Commonly used Standard Deviation(SD) is the square root of the variance.
For a univariate data set X_1, X_2, ..., X_n, the MAD is defined as the median of the absolute deviations from the data's median:

$$MAD = median(|X_i - median(X_i)|) \tag{1}$$

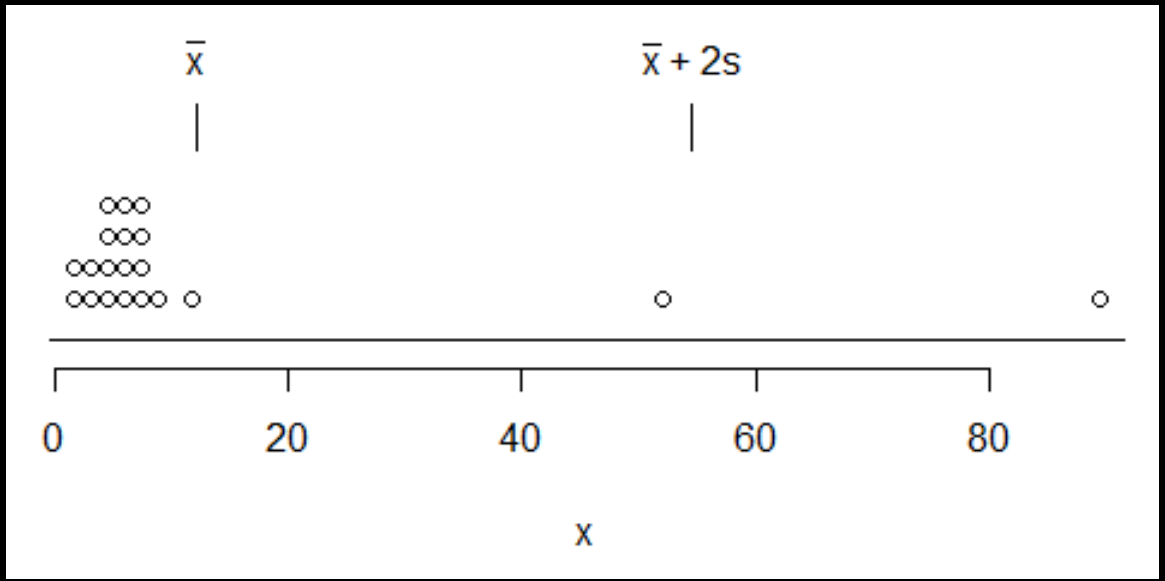


Figure: Weakness of SD in outlier detection.

Why MAD

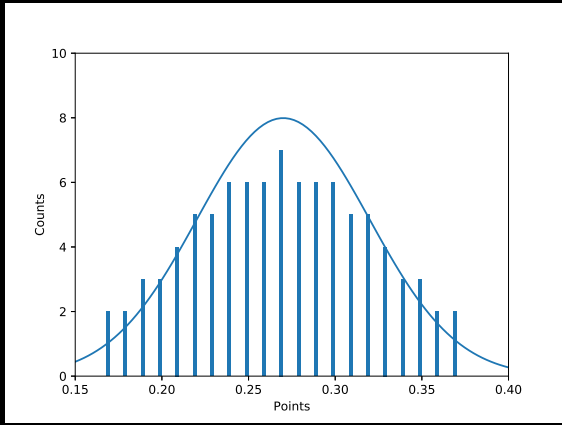


Figure: a) Normal distribution, n = 91, mean = 0.27, median = 0.27, standard deviation = 0.06.

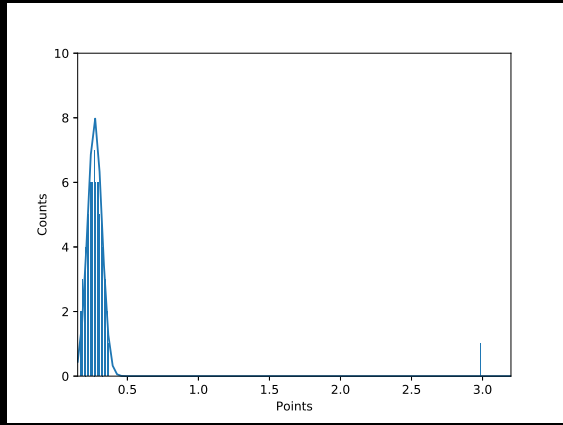
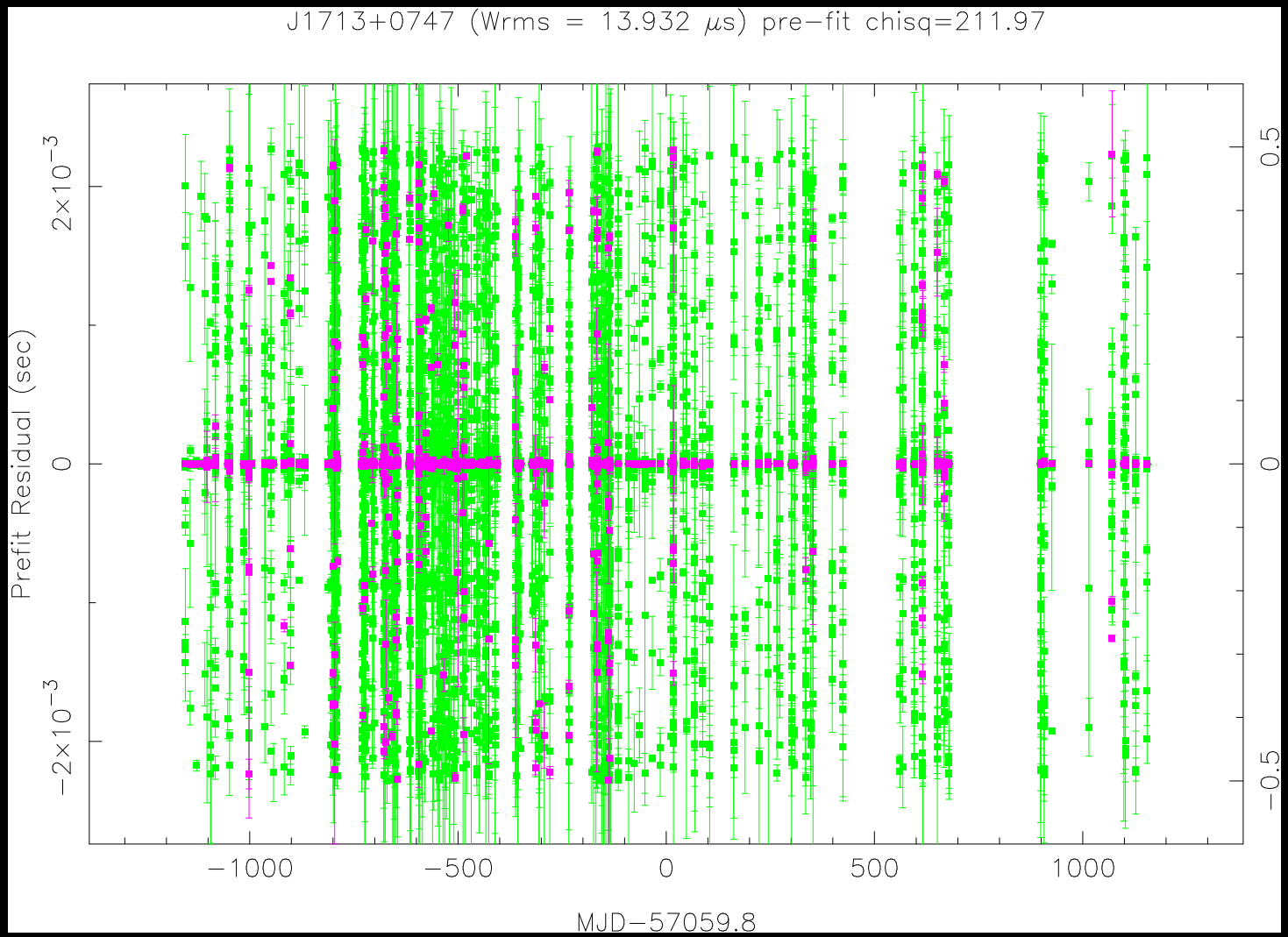


Figure: b) Asymmetry due to an outlier, n = 91, mean = 0.39, median = 0.27, standard deviation = 0.59.

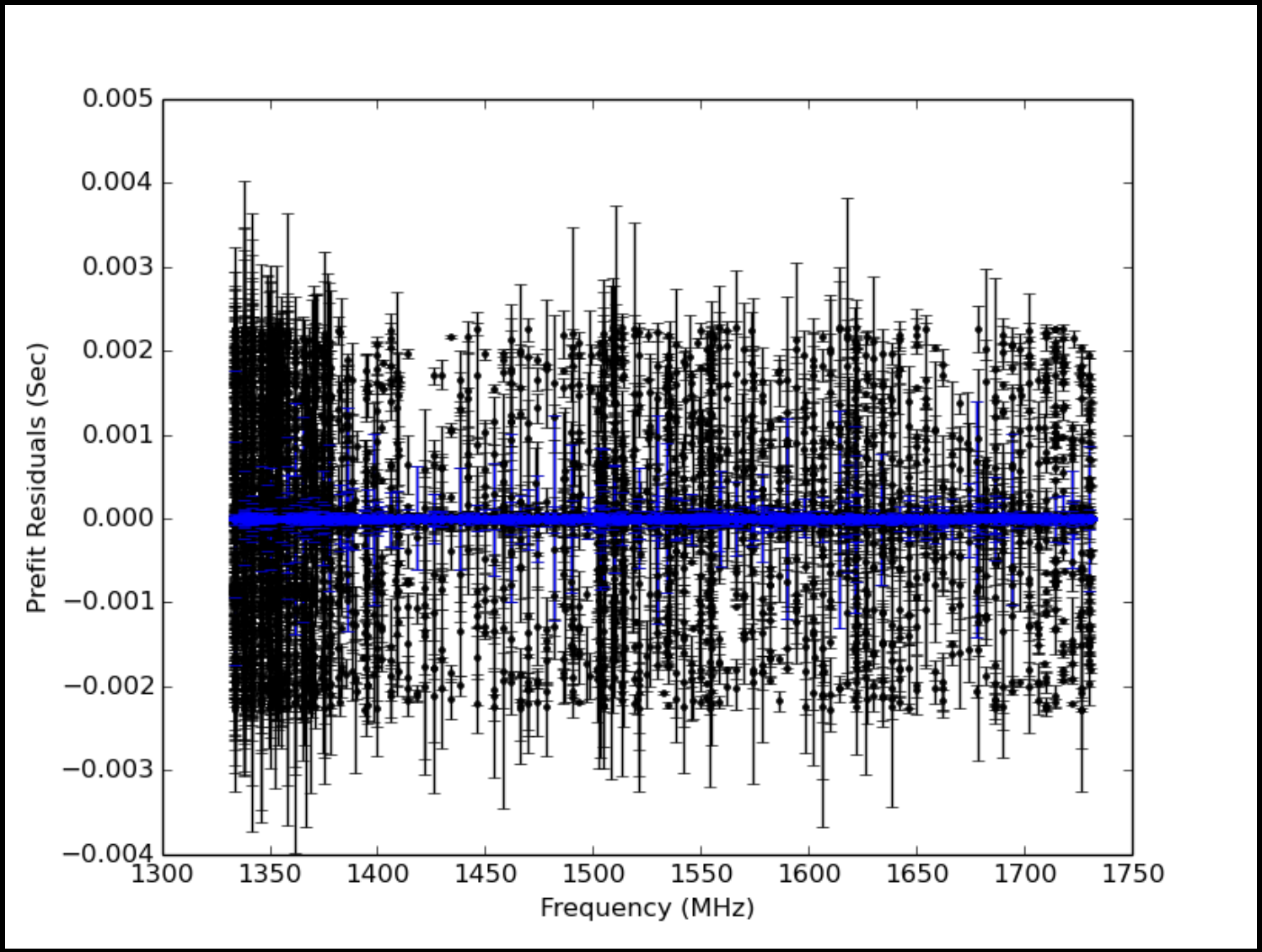
Advantage of Median Standard Deviation

1. Aimed at symmetric distribution.
2. MAD is more robust than Standard Deviation (SD).
3. MAD is totally immune to the sample size.

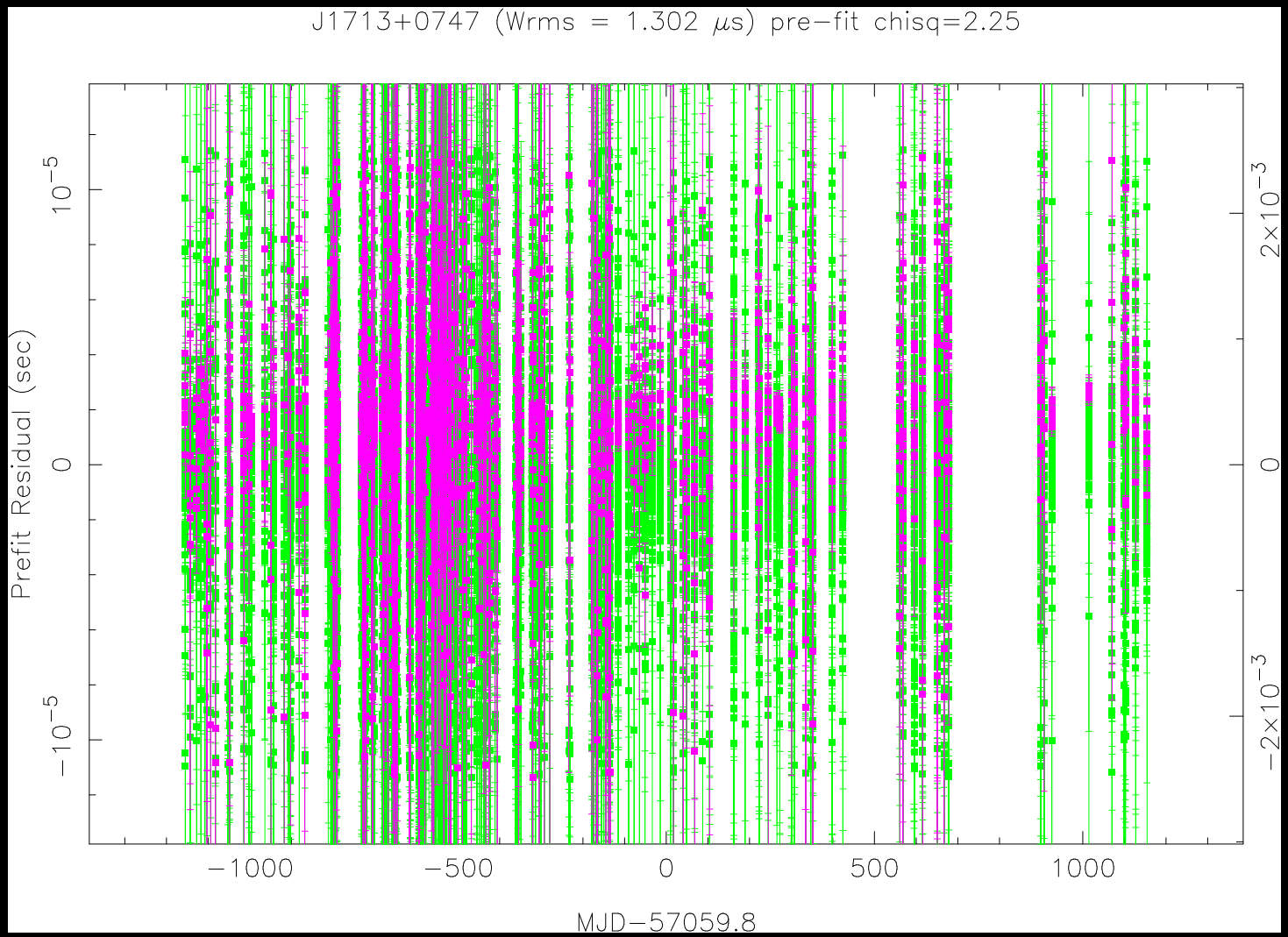
Raw ToAs from J1713+0747, JBO ($N_{TOA} = 28376$)



Example on MAD, MAD



Example on MAD,after MAD($N_{TOA} = 22588$)



Outlier Rejection Scheme
2. Normalized Residuals Gaussian Distribution Fitting

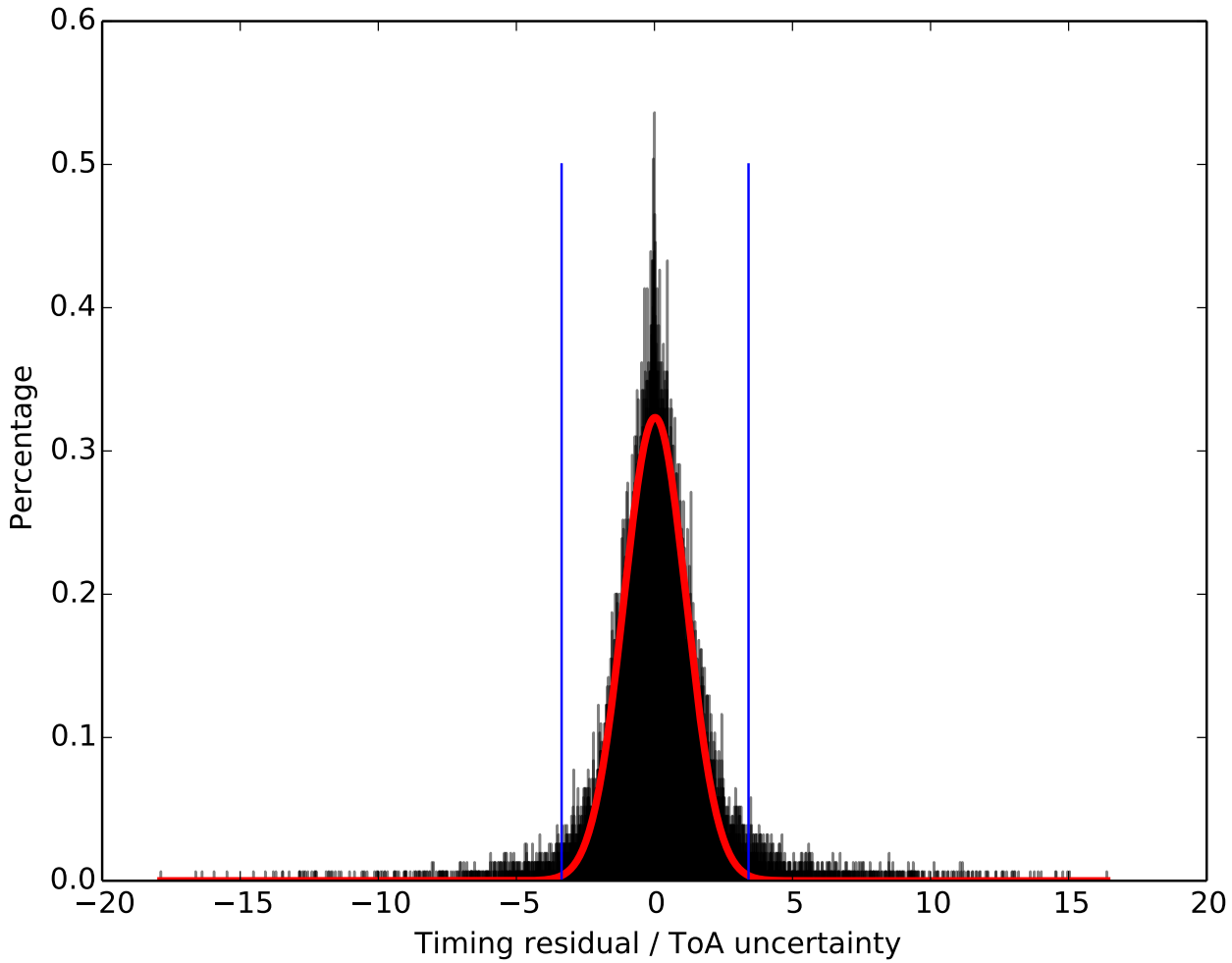
HOW

After employing the MAD scheme, the distribution of ToAs becomes more Gaussian.
We then fit the histogram of Normalized Residuals (residuals divided by ToA uncertainty) with Gaussian Distribution:

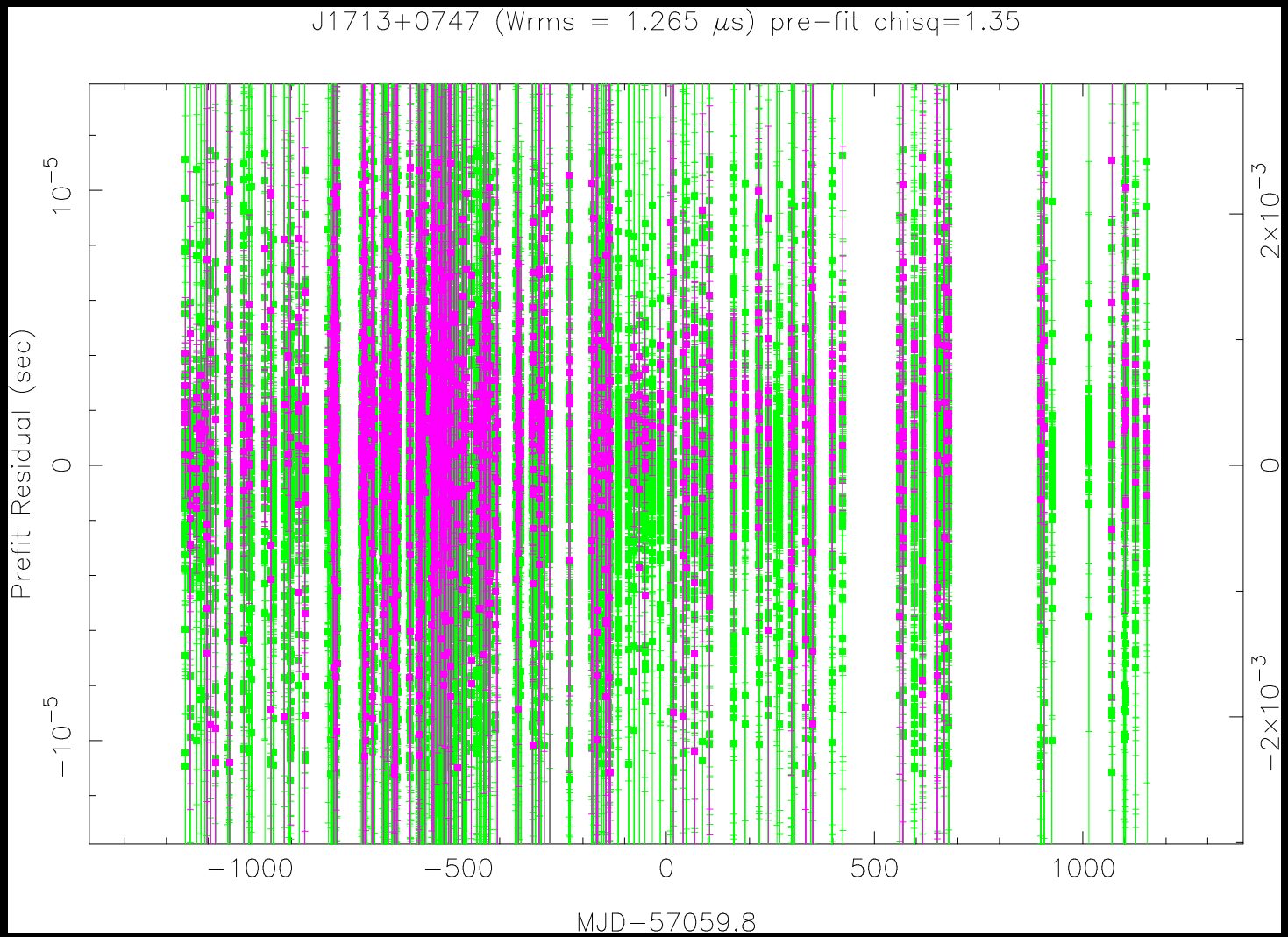
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2}$$

and remove residuals which are 3σ away.

Examples on Gaussian Fitting



After Gaussian Fitting($N_{TOA} = 20978$)



Conclusion about the scheme

1. The Wrms is much lower than before, but it's still not great.
2. The reduced chi-squared is good and very comparable to the best results people typically get.
3. The ToAs still look scattered, even though Wrms and reduced chisq are fine.

Summary & future work

Summary & In Future

Summary

The present optimal outlier rejection scheme is first use MAD, then Normalised Residuals fitting.

In Near Future

1. decide which CCA is best.
2. check ToA bandwidth.
3. ask for more recent data.
4. ask for more pulsars to be included.
5. finish writing paper!

The End

Thank You!
Danke Schön!
Dank u wel!
Merci beaucoup!
Molte Grazie!