Efficient Frequency Identification in Coherent Variable Stars

Ádám SÓDOR [1], Zsófia BOGNÁR [1], Marek SKARKA [1, 2]

[1]*Konkoly Observatory, MTA CSFK, Budapest, Hungary;* [2]*Astronomical Institue, Czech Academy of Sciences*

XXXth General Assembly of the International Astronomical Union – Vienna, August 20–31, 2018

1. Background

We looked for the most efficient frequency-identification method in terms of precision, reliability and speed, in order to identify the optimal way of extracting periodicities from massive amount of extended time series of coherent variables observed by recent space- and ground-based missions.

Test sample: *Kepler* [1] light curves of δ Scuti and γ Doradus pulsators (Sódor et al. in prep.)

2. Methods we compared

We investigated Fourier-based methods:

SEQ – hybrid time- and frequency-domain method, fitting one frequency in each cycle. In-house developed implementation [2].

SIM – hybrid method, fitting all previously extracted frequencies in each cycle. In-house developed implementation [2].

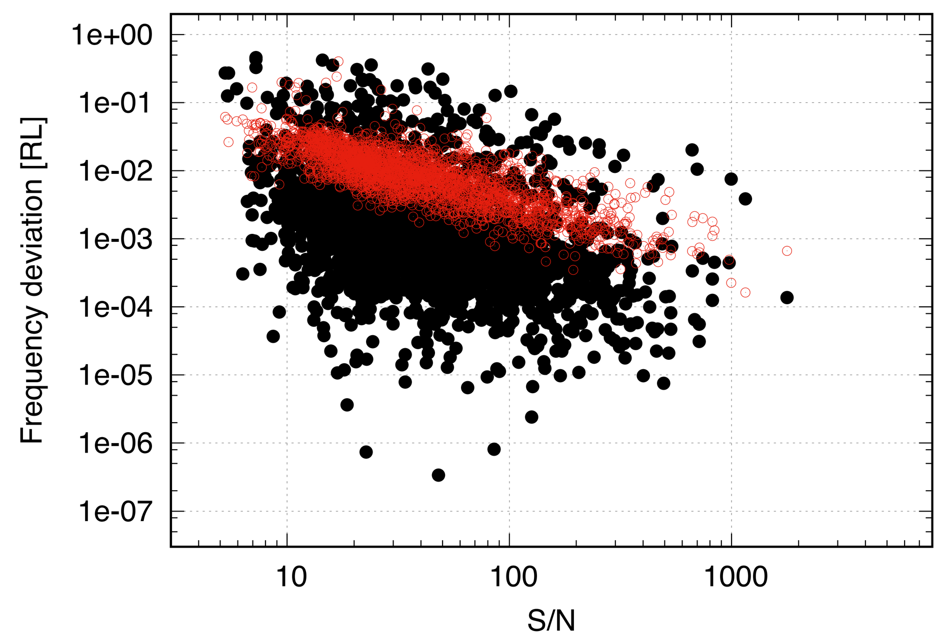
SIG – SigSpec [3] hybrid method.

CLN – Clean [4] pure frequency-domain method, in-house developed implementation.

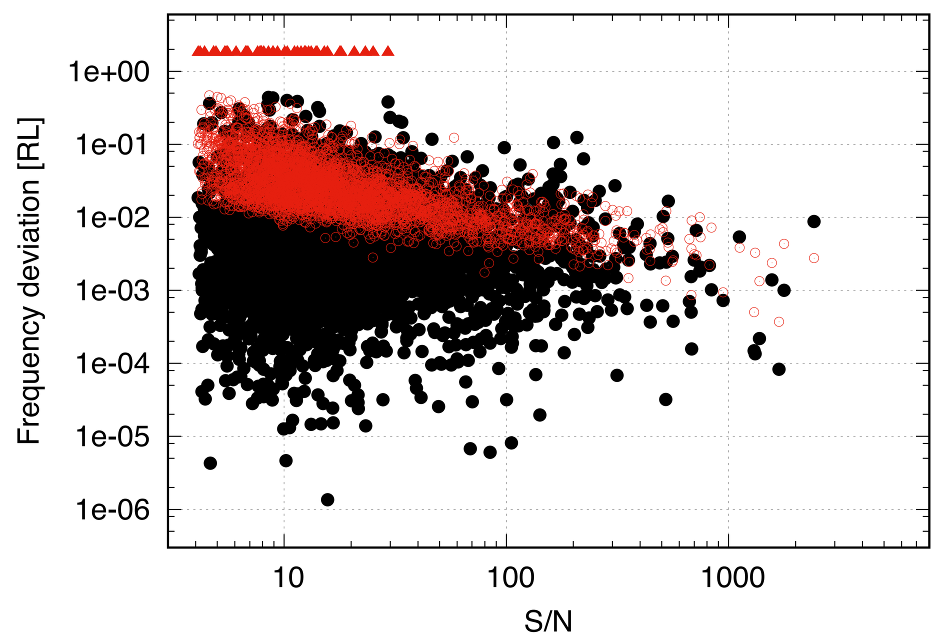
3. Precision and completeness

Precision and completeness was investigated by analysing the same data with different methods. The set of extracted frequencies were cross-matched.

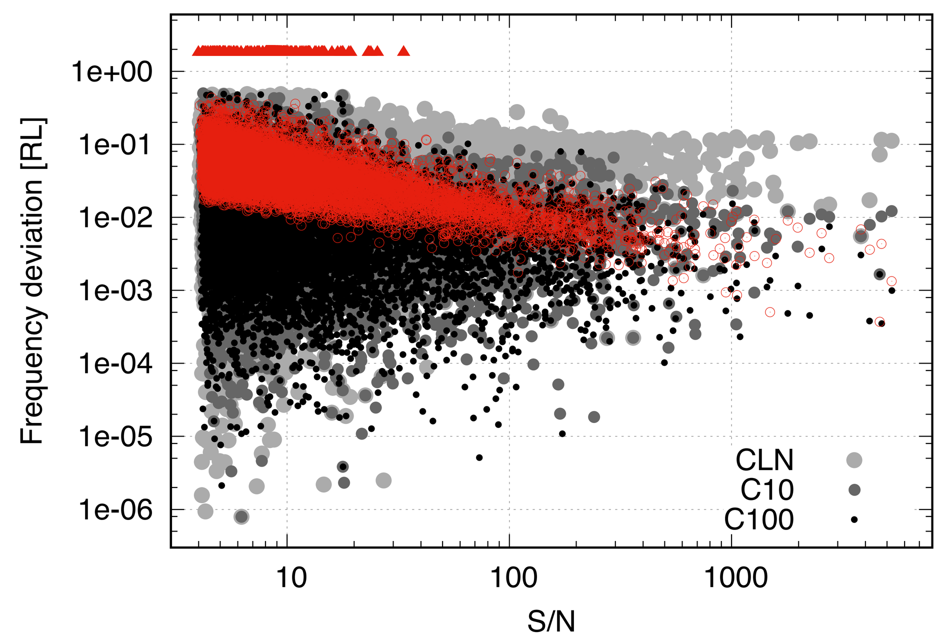
Unmatched frequency components (indicated by red triangles) were only found at low amplitudes, at low S/N. This difference is attributed to the different ways of calculating S/N, when the same component was probably identified as significant by one method while insignificant by another one.



Frequency deviations between **SIM** *and* **SIG** *against* $(S/N)_{SIM}$. *Red circles indicate standard errors of* **SIM** *fit.*



Frequency deviations between **SIM** *and* **SEQ** *against* $(S/N)_{SEQ}$. *Red triangles in the top part indicate unmatched frequencies. Red circles indicate standard errors of* **SEQ** *fit.*

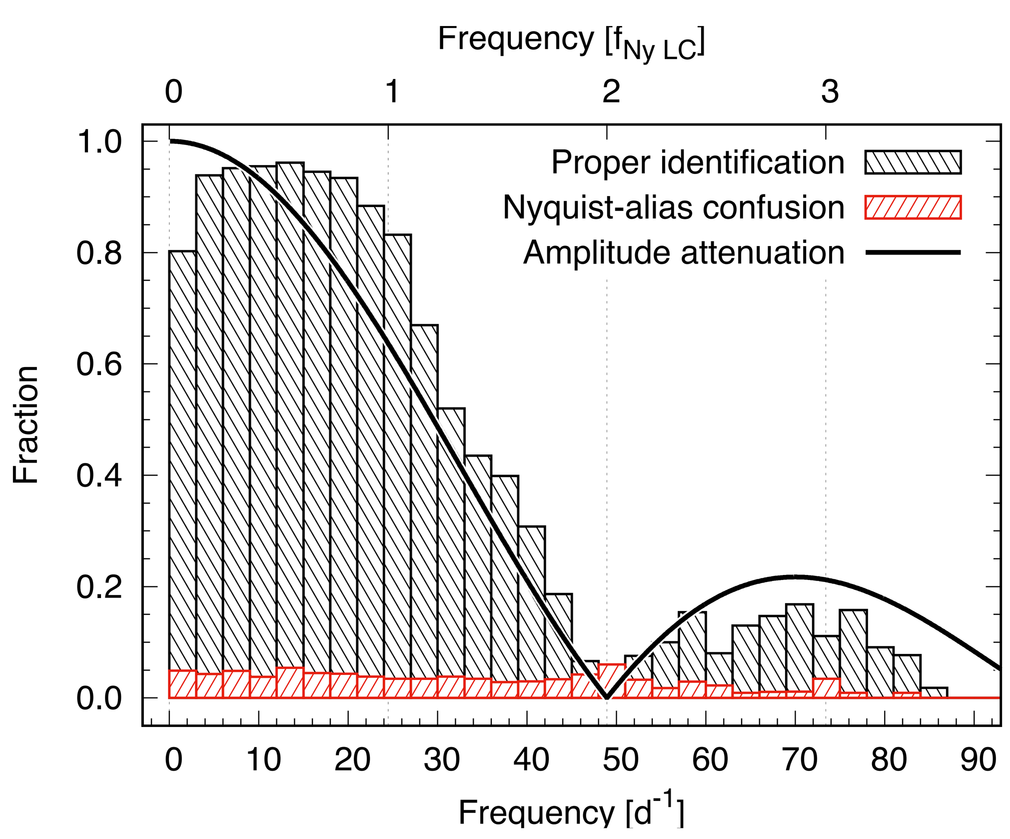


Frequency deviations between **SEQ** *and different resolution of* **CLN** *against* $(S/N)_{SEQ}$. *Red circles indicate standard errors of* **SEQ** *fit.*

4. Super-Nyquist performance

The Nyquist-limit of the 30-min (long-cadence) *Kepler* light curves is 24.5 d^{-1}, but it was shown that in several-year-long time series the real frequency components can be distinguished from the Nyquist aliases [5].

We tested the automatic distinction capabilities of Clean by running the identification on both the short-cadence (as reference) and long-cadence light curves of the same objects, then cross-matching the results.



Ádám Sódor

Download in pdf

5. Speed

Execution times strongly depend on the hardware*. We give examples only for comparison. Execution times correspond to a typical *Kepler* long-cadence light curve.

SIM and **SIG** have superlinear, while **SEQ** and **CLN** have linear time complexity:

SIM $T(n) = O(n^{2.9})$ strongly polynomial; 4.5h for 100, 1.4d for 200 components.

SIG $T(n) = O(n^{1.2})$ slightly polynomial; 6.7 s/comp. at the 100th, 8.5 s/comp. at the 200th comp.

SEQ 2.9 s/comp., linear.

CLN 0.025 s/comp., linear.

* Which, according to Moore's law [6], gets improved exponentially; the computing power is doubling every 18 months. This relation amazingly holds for the past 50 years equally well!

References

[1] Borucki W.J. et al. 2010, *Science*, **327**, 977

[2] Sódor Á. 2012, *Konkoly Observatory Occasional Technical Notes*, **15**

[3] Reegen P. 2007, *A&A*, **467**, 1353

[4] Roberts D. H. et al. 1987, *AJ*, **93**, 968

[5] Murphy S. et al. 2013, *MNRAS*, **430**, 2986

[6] Moore G.E. 1965, *Electronics*, **38**, 114

Conclusion

We find that Clean and the other methods perform equally well at identifying significant coherent periodicities, however, Clean can be orders of magnitudes faster. The only deficiency of Clean is the inability of deriving uncertainties on the extracted frequencies and phases.

Therefore, Clean is the best choice for extracting coherent periodicities from massive amounts of extended, strongly multiperiodic time series data.

Acknowledgements

The financial support of the Hungarian NKFIH Grants K-115709, K-113117 and K-119517 are acknowledged. ÁS was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. ZsB acknowledges the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the PD_17 funding scheme, project no. PD-123910. MS acknowledges the support of the postdoctoral fellowship programme of the Hungarian Academy of Sciences at the Konkoly Observatory as a host institution and Czech Grant GA CR 17-01752J. This work benefited a lot from the discussions within the SoFAR international team* supported by the International Space Science Institute.

*http://www.issi.unibe.ch/teams/sofar/