MNRAS **000**, 1–3 (2019) Preprint 30 September 2019 Compiled using MNRAS LaTeX style file v3.0

Can we predict?

L. Jetsu[⋆]
Department of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
We simulate artificial data for a sinusoid having a period $P = 1$. Then we show that this period can be detected from a short $\Delta T = 0.3P$ slice of data. We proceed to show that the slice length is irrelevant for high quality measurements. The frustrating frequency resolution limit $f_0 = 1/\Delta T$ of the power spectrum methods is pulverized. It is possible to predict the behaviour of non-linear periodic models.

Key words: Methods: statistical – Methods: data analysis

1 INTRODUCTION

Jetsu (2019, PAPER I) presented a method for solving the free parameters of non-linear models. He divided the free parameters $\bar{\beta}$ into two parts. The ones that make the model non-linear ($\bar{\beta}_I$), and those that do not ($\bar{\beta}_{II}$). If the $\bar{\beta}_I$ free parameters are fixed to constant tested values, the model becomes linear. No one, including us, has realized the full potential of this approach.

2 METHOD

Imagine that the model is a line. If there are only two points ($n = 2$), it is trivial to solve the line that goes trough them. However, if three measurements y_i do not coincide with a line, it is still possible that their distance from the line, the residuals ϵ_i, is of the same order as their measurement errors σ_i. The standard least squares fit minimizes the test-statistic $\chi^2 = \sum_i^n \epsilon_i^2/\sigma_i^2$, and gives unique solution for the free parameters of this linear model. A successful model has $\chi^2 \approx n$, because the relation $\sigma_i \approx \epsilon_i$ should be fulfilled. The sample size n must be larger than the number of free parameters of a non-linear model $p_{\beta_I} + p_{\beta_{II}}$. If $\bar{\beta}_I$ are fixed to constant values, then $n \geq p_{\beta_{II}} + 1$ is sufficient, and the model becomes *linear!*

3 SIMULATIONS

Our one period model is

$$g_1(t) = A_1 \cos(2\pi f_1 t) + B_1 \sin(2\pi f_1 t), \quad (1)$$

where the free parameters are $\bar{\beta} = [f_1, A_1, B_1]$, $\bar{\beta}_I = [f_1]$ and $\bar{\beta}_{II} = [A_1, B_1]$. Our test statistic for each tested constant frequency f_1 is

$$z(f_1) = \sqrt{\chi^2/n}. \quad (2)$$

3.1 Simulation 1

We show the result for one simulated data sample in Fig. 1. Although these data span only 30% of the full period, $\Delta T = 0.3P$, we can easily detect the period $P = 1$. This is a unique solution for our non-linear periodic model.[1]

3.2 Simulation 2

The results for the combination $n = 100, \Delta T = c_1 P_1$, $c_1 = 0.3$, $\sigma = c_2 A_1$ and $c_2 = 0.05$ in Fig. 1 do not necessarily convince the readers. The next simulation in Fig. 2 has the same frequency and the same amplitudes, but the revised values are $n = 500$, $c_1 = 0.01$ and $c_2 = 0.00001$. These results should convince even the worst sceptics. Actually, the period can be solved from an infinitesimal short slice of measurements, if σ and n are sufficient. This applies to all non-linear models, not only the periodic ones.

4 DISCUSSION

One could argue that we do not know the correct real model. This is true. However, all models can be compared with the F-test (PAPER I: Eqs. 33 and 34). For example, the comparison of the one period and the two period models showed that the latter was a better model for the light curves of FK

⋆ E-mail: lauri.jetsu@helsinki.fi

[1] We made our simulations with the Python program *Test1.py*. Our “swap-technique” is described in freely available version https://www.mv.helsinki.fi/home/jetsu/Test1.py
We will check and polish this program in the next days. This updated version will be published in Zenodo.

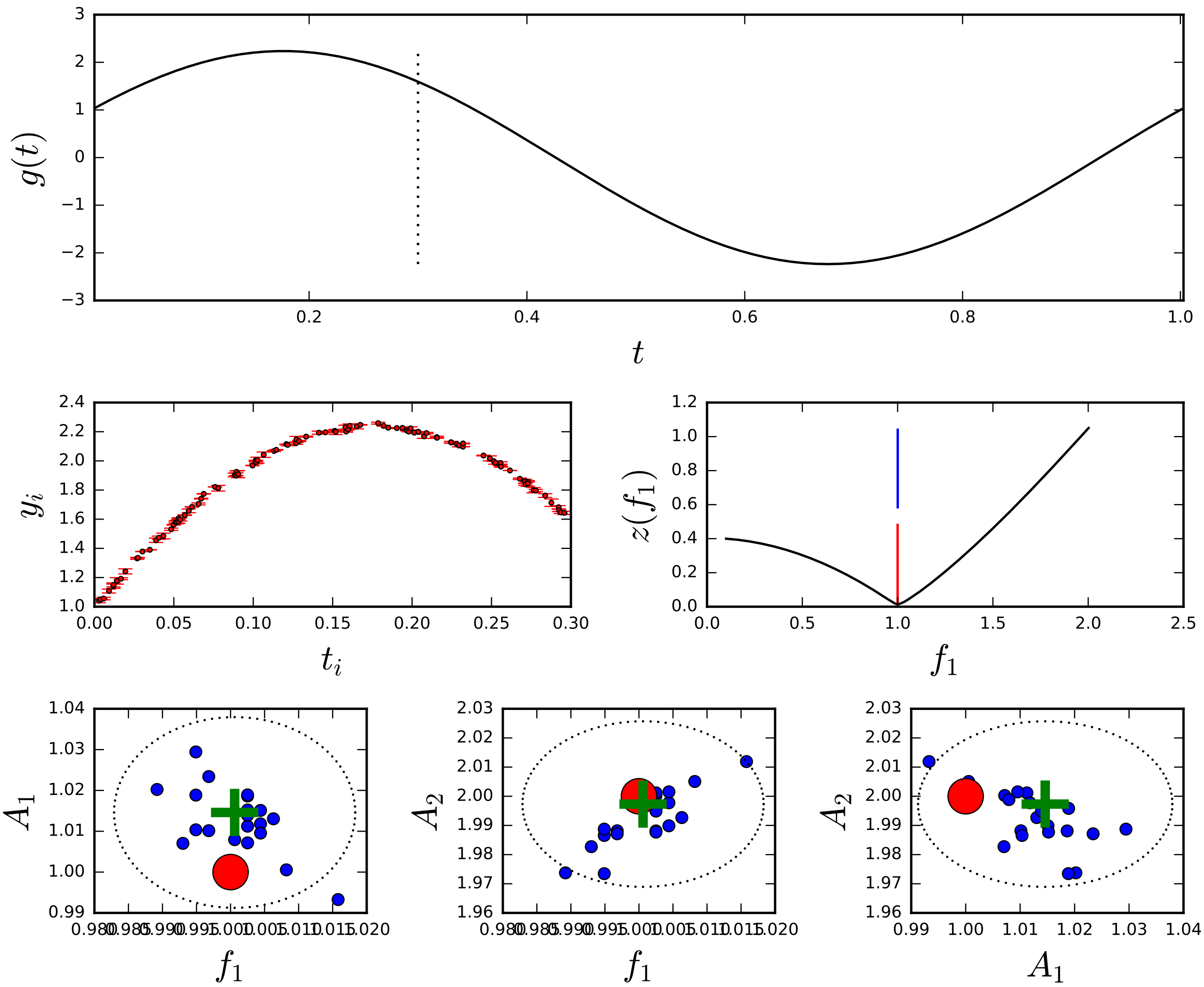

Figure 1. Our $\Delta T = 0.3P_1$ simulation. **Up**: Model $g(t)$ for $f_1 = 1$, $A_1 = 1$ and $A_2 = 2$ in Eq. 1. Simulated data extends to dotted vertical line. **Middle left**: Data y_i simulated from $g(t)$ model plus Gaussian noise $N(0, \sigma)$, where $\sigma = A_1/20$. Sample size is $n = 100$. **Middle right**: Test statistic $z(f_1)$ of Eq. 2 in period range $P_{\rm min} = 0.5$ and $P_{\rm max} = 10.0$. Vertical red and blue lines denote simulated period $P_1 = 1$ and best detected period $P_{\rm best}$, respectively. **Down**: Results for free parameters in 20 bootstrap samples (blue circles), their mean of bootstrap estimates (green cross) and simulation values used (large red circle). Dotted ellipse denotes $\pm 3\sigma$ limits for bootstrap estimates.

Com. This was typically confirmed at $Q_F < 10^{-16}$ significance level, which is a major improvement to the correct direction in finding the real model.

One could also argue that we do not know the correct test intervals in the non-linear free parameter space (NLFPS). This is only partly true, because we check, if there are other solutions, even far from the correct one. Furthermore, the z values close to each other correlate in NLFPS. For example in physics, the reasonable limits are already known in many cases. We use the bootstrap method to compute the errors for σ_{β_I} and $\sigma_{\beta_{II}}$. We have published a detailed description of how the free parameter search could be done in multi-dimensional NLFPS (PAPER I: Sect. 11).

The method relies on brute numerical force. This takes a lot of computation time. However, observing a phenomenon having a long period P requires a long observing time ΔT. Unlike the power spectrum method, our method does not suffer from the frustrating $f_0 = 1/\Delta T$ frequency resolution limit emphasized by Loumos & Deeming (1978), so why wait? Here, we *assume* that the correct model is a sinusoid, but so do also all those who apply the power spectrum method. The free parameter estimates will become more accurate as the observations continue. Although the computations take time for more complex models, the solution is at least unique (i.e. unambiguous).

5 CONCLUSIONS

Here, we show that our numerical χ^2 method can predict the period of a non-linear model. Even a short slice of measurements can be used to solve its free parameters. In PAPER I, we showed that for the two period models it is possible to go beyond the f_0 resolution limit of the power spectrum methods (Loumos & Deeming 1978).

Imagine, what super computers, quantum computers

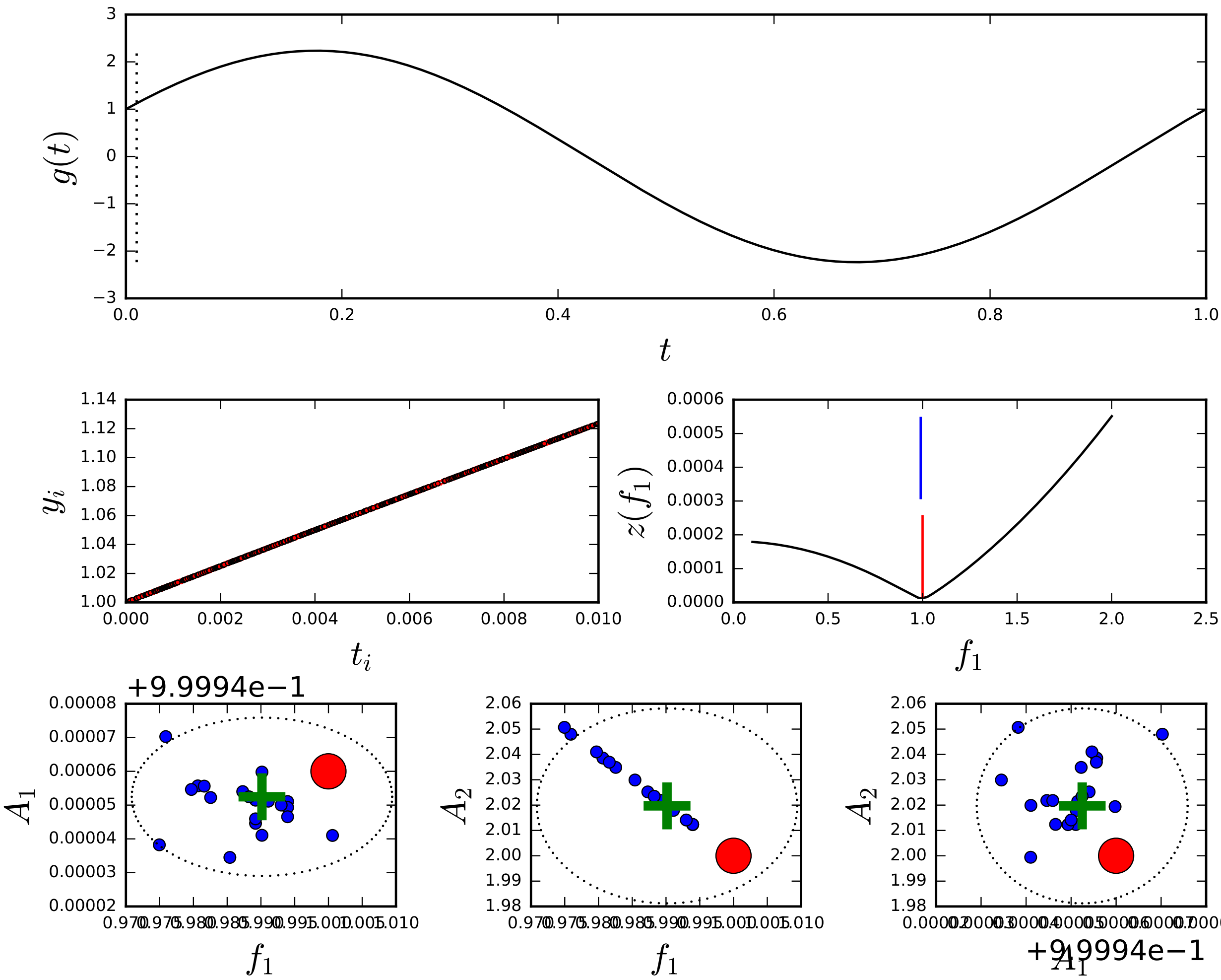

Figure 2. Our $\Delta T = 0.01 P_1$ short slice simulation. $\sigma = 0.0001 A_1$ and $n = 500$, otherwise as in Fig. 1

or artificial intelligence can achieve with our numerical approach. Frankly, we find this scary. Maybe somebody can prove us wrong.

ACKNOWLEDGEMENTS

We thank the anonymous referee and Thomas Hackman for their comments on the manuscript. This work has made use of NASA's Astrophysics Data System (ADS) services.

REFERENCES

Jetsu L., 2019, submitted to MNRAS on Sep 20, 2018,

Loumos G. L., Deeming T. J., 1978, Ap&SS, 56, 285

This paper has been typeset from a TEX/LATEX file prepared by the author.