




SHERLOCK: Scheme using Hidden markov models for Establishing Reiteratively a List Of Candidate period-spacings with liKelihood

Joey S. G. Mombarg ¹

¹ IRAP, Université de Toulouse, CNRS, UPS, CNES, 14 avenue Édouard Belin, F-31400 Toulouse, France

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#))

Summary

SHERLOCK is an open-source Python routine designed to identify the pulsation periods from a list of observed oscillation periods present in the light curve of a gravity-mode pulsator. The routine is based on Hidden Markov Models trained on a grid of theoretical stellar pulsation models to reiteratively identify pulsation periods based on the highest likelihood. The routine comes with a GUI that allows the user to provide an initial start for the automated search after which SHERLOCK will try to find all other requested pulsation modes that are present in the data.

Context

Recent space-borne missions such as the NASA *Kepler* and TESS mission have provided light curves of thousands of pulsating stars, and future missions like the ESA PLATO mission are expected to increase this number even more. The so-called gravity (g)-mode pulsators are particularly interesting to study the internal stellar structure, as the periods of g-modes excited in core-hydrogen burning stars are mostly sensitive to the physics in the deep interior. These g-mode pulsators are typically modelled with a forward-modelling approach to infer fundamental stellar properties such as the mass, age, and rotation frequency (e.g. Pedersen et al. (2021); Mombarg et al. (2021)). Starting from a processed light curve, frequencies are extracted via iterative prewhitening. This yields a list of frequencies and their amplitudes of the g-mode pulsations, but also combination frequencies and possibly spurious ones. The stellar oscillations are described by spherical harmonics with three quantum numbers; the spherical degree, azimuthal order, and radial order. In order to compare the observed frequencies with models, the pulsation modes and the corresponding quantum numbers need to be identified. For g-mode pulsators, this is typically done by exploiting a relation between the pulsation periods of consecutive radial order and equal (ℓ, m) (Miglio et al., 2008; Tassoul, 1980).

Statement of need

In period space, the so-called period-spacings (difference in period between consecutive radial orders) of g-mode pulsations should follow a pattern. The next step after iterative prewhitening is the search for period-spacing patterns in the frequency list is often non-trivial and is often done by a trained asteroseismologist. Yet, this means that the found period-spacing patterns are somewhat subjective, which can have consequences for the fundamental stellar parameters in the modelling. Automated routines to search for period-spacing patterns exist (Garcia et al., 2022; Li et al., 2020), but these do not account for any shifts in the pulsation periods due to the effect of a gradient in the mean molecular weight building up as the star evolves, or mode

39 coupling. This can result in missing pulsation modes that are particularly interesting to test the
40 theory of stellar structure. SHERLOCK is designed to automatically extract the most probable
41 period-spacing pattern based on predictions of a grid of stellar pulsation models. Moreover,
42 it provides a metric of confidence how likely a selected period is actually part of the pattern.
43 This way in the modelling a lower priority can be assigned to pulsation periods that are more
44 likely to have been mistaken.

45 Methodology

46 The tool presented here is based on Hidden Markov Models (HMMs). A HMM is defined by a
47 sequence of hidden states h_i , and visible states, where the index i indicates the time step or
48 iteration. HMM are for example used to correct typos made by the accidental pressing of an
49 adjacent key instead. In this analogy, the hidden states h_i are the intended letters and the
50 visible states v_i are the pressed keys. Given state h_i , the probability for a state h_{i+1} is given
51 by, $p(h_{i+1}|h_i)$, and is referred to as the transition probability. This transition probability in this
52 analogy can be determined by training the model on a dictionary to determine the probability
53 of the next letter given the previous one. At each time step, the HMM emits a visible state v_i
54 with probability, $P(v_i|h_i)$, which is referred to as the emission probability. This probability
55 determines how likely it is to press a specific key, given an intended letter. Translating this
56 example to the search for period-spacing patterns, we can think of the transition probability
57 as the probability of a period-spacing ΔP_i belonging to the pattern according to theory, given
58 a previous period-spacing ΔP_{i-1} . This probability is determined from a grid of stellar pulsation
59 models that provides probability distributions. The SHERLOCK algorithm works with differences
60 in period-spacings,

$$\delta(\Delta P)_i = \Delta P_i - \Delta P_{i+1} = (P_{i+1} - P_i) - (P_{i+2} - P_{i+1}), \quad (1)$$

61 where $P_i < P_{i+1} < P_{i+2}$. For each radial order in the grid, we compute $\delta(\Delta P)_1$ and $\delta(\Delta P)_2$.
62 For an observed $\delta(\Delta P)_{1, \text{obs}}$, we select the 500 values of $\delta(\Delta P)_1$ that are closest to the
63 observed one. Next, we compute distribution of the values of $\delta(\Delta P)_2$, that is, the expected
64 next expected difference in period-spacing. This distribution is interpolated to make a PDF
65 from which the transition probability is computed. The emission probability is based on the
66 number of periods in the list that falls within a confidence interval, where periods with higher
67 amplitudes are favoured over lower amplitude periods, as lower amplitude periods are more
68 likely to be spurious. This probability is defined as,

$$p_{\text{emis}, i} = \frac{A_i}{\sum_j A_j}, \quad (2)$$

69 where A_i is the amplitude of period P_i , and the index j runs over all potential periods that
70 fall within the search window. The search window (or confidence interval) is taken to be the
71 most extreme values for which $p_{\text{emis}}/\max(p_{\text{emis}}) > 0.001$. An example of the distribution of
72 this quantity is shown in Figure 1.

73 The algorithm is initiated with the periods of three consecutive radial orders, provided by the
74 user. SHERLOCK will then compute a search window, defined as the minimum and maximum
75 period for which the probability is larger than 0.001. Then, for each candidate in the search
76 window a total probability is computed,

$$p_{\text{total}, i} = p_{\text{emis}, i} p_{\text{trans}, i}, \quad (3)$$

77 and is normalised such that $\sum_i p_{\text{total}, i} = 1$. The period with the highest total probability is
78 then selected and the next iteration starts.

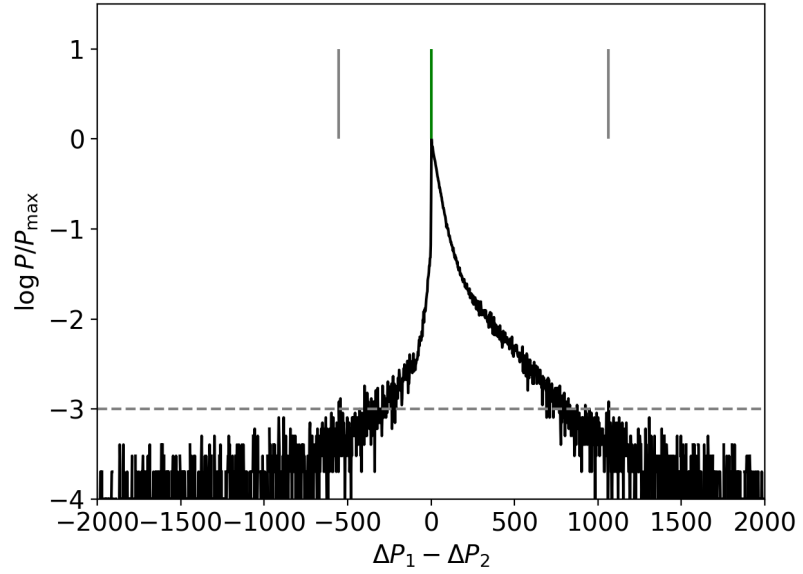


Figure 1: Normalised interpolated emission probability for a difference in period-spacings $\Delta P_1 - \Delta P_2$, computed from a grid of stellar pulsation models. The green tick indicates the expected value, the grey ticks the search window, based on the cutoff in $\log p/p_{\max}$ as shown by the horizontal grey dashed line.

Stellar models for training

In order to calculate the transmission probability, a grid of GYRE 6.0 (Townsend & Teitler, 2013) pulsation models was constructed to compute the periods of the stellar eigenmodes, specifically for dipole ($\ell = 1$) modes with $m \in [-1, 0, 1]$ and for radial orders $n_{\text{pg}} = -100$ to -1 . Since the probabilities are based on differences in periods-spacing, period-spacing patterns of higher spherical degree can also be found without needing the extent the grid, as these modes behave the same way as dipole modes with the same sign for m . A uniform rotation profile is assumed, where the a randomly picked rotation rate between 0 and 0.6 times the critical rotation rate is picked for each stellar equilibrium model. The stellar equilibrium models are computed with MESA r22.11.1 (Jermyn et al., 2023; B. Paxton et al., 2011, 2013, 2015, 2018; Bill Paxton et al., 2019), where the mass is sampled according to a Sobol sequence between 3 and $9 M_{\odot}$. The chemical diffusion coefficient due to core boundary mixing (overshoot) is described by an exponential decaying function where coefficient related to the width of the zone, f_{ov} , is sampled in a similar way between values of 0.005 and 0.035. The chemical mixing in the radiative envelope is modelled by mixing induced due to internal gravity waves (e.g. Varghese et al., 2023).

Performance

To demonstrate to performance of SHERLOCK, we use the well-studied gravity-mode pulsator KIC7760680 (see e.g. Pápics et al., 2015). Earlier studies that have manually searched for a period-spacing pattern in pulsation spectrum of this stars found a pattern of 36 consecutive radial orders. Furthermore, the dips found in the period-spacing pattern is indicative a chemical gradient near the core. Figure 2 shows the pattern found by SHERLOCK, where the red colored periods are the initial three consecutive radial order suggested by the user. Once the user has found a possible start of a pattern, it takes SHERLOCK less than 2 seconds to complete it. For this star, initial suggestions of periods labeled (2,3,4) or (6,7,8) in the top panel of Figure 2, result in the same final pattern found. However, when starting for example from the

periods (4,3,2) around 1 day, the same pattern is not recovered. Therefore, the initial periods should ideally be three high-amplitude periods with similar spacing. In any case, an incorrect suggested start will result in an incorrect final result that is not always trivial to notice, but typically the longest possible pattern is the correct one. The bottom panel, showing the total probability for each selected period, indicates that for example mode 2 toward periods shorter than the initial ones, and modes (5,6) towards periods larger than the initial ones, could have possibly been mistaken as other candidates with high probability are present within the search window. Therefore, when performing forward modelling of this star's period-spacing pattern, these modes can be assigned less weight according to the total probability given by SHERLOCK.

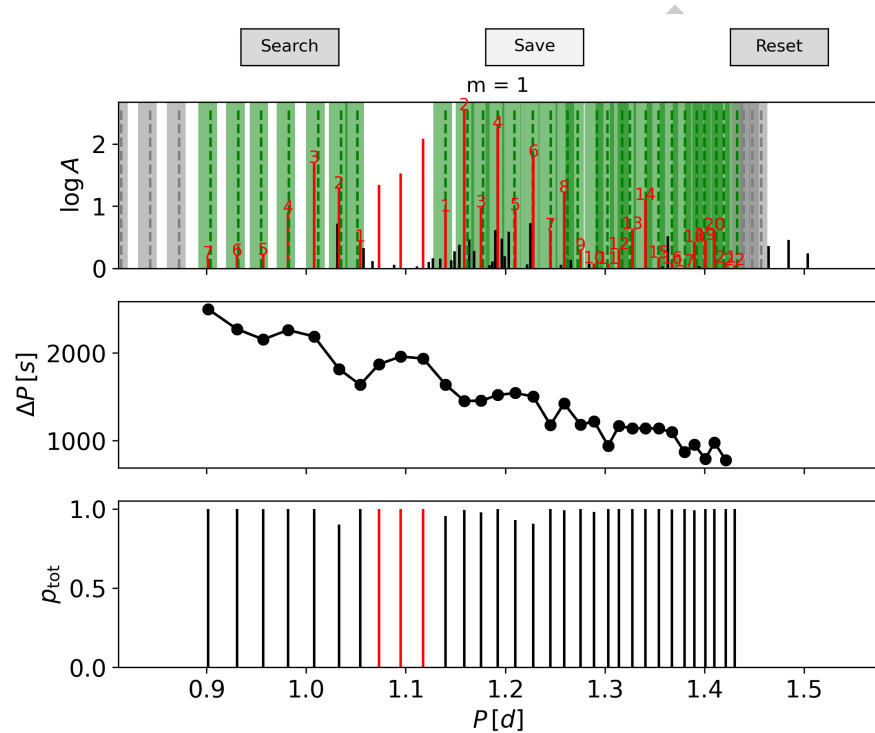


Figure 2: Example of the interface and output of SHERLOCK. Top panel: Periods and amplitudes of the inserted frequency list. Periods colored red were selected by SHERLOCK to be part of the period-spacing pattern, assuming azimuthal order $m = 1$. The green shaded areas indicate the search windows and the dashed line the expected location (in grey if no period was found). Middle panel: The period vs period-spacing of the selected modes. Bottom panel: Total probability for each selected mode. Red colored modes indicate the three initial modes suggested by the user.

Acknowledgements

The research leading to these results has received funding the French Agence Nationale de la Recherche (ANR), under grant MASSIF (ANR-21-CE31-0018-02). The author thanks Mathias Michielsen for his advice on unit tests and Poetry.

References

- Garcia, S., Van Reeth, T., De Ridder, J., Tkachenko, A., IJspeert, L., & Aerts, C. (2022). Detection of period-spacing patterns due to the gravity modes of rotating dwarfs in the TESS southern continuous viewing zone. 662, A82. <https://doi.org/10.1051/0004-6361/202141926>

- Jermyn, A. S., Bauer, E. B., Schwab, J., Farmer, R., Ball, W. H., Bellinger, E. P., Dotter, A., Joyce, M., Marchant, P., Mombarg, J. S. G., Wolf, W. M., Sunny Wong, T. L., Cinquegrana, G. C., Farrell, E., Smolec, R., Thoul, A., Cantiello, M., Herwig, F., Toloza, O., ... Timmes, F. X. (2023). Modules for Experiments in Stellar Astrophysics (MESA): Time-dependent Convection, Energy Conservation, Automatic Differentiation, and Infrastructure. *265*(1), 15. <https://doi.org/10.3847/1538-4365/acae8d>
- Li, G., Van Reeth, T., Bedding, T. R., Murphy, S. J., Antoci, V., Ouazzani, R.-M., & Barbara, N. H. (2020). Gravity-mode period spacings and near-core rotation rates of 611 γ Doradus stars with Kepler. *491*(3), 3586–3605. <https://doi.org/10.1093/mnras/stz2906>
- Miglio, A., Montalbán, J., Noels, A., & Eggenberger, P. (2008). Probing the properties of convective cores through g modes: high-order g modes in SPB and γ Doradus stars. *386*, 1487–1502. <https://doi.org/10.1111/j.1365-2966.2008.13112.x>
- Mombarg, J. S. G., Van Reeth, T., & Aerts, C. (2021). Constraining stellar evolution theory with asteroseismology of γ Doradus stars using deep learning. *Stellar masses, ages, and core-boundary mixing*. *650*, A58. <https://doi.org/10.1051/0004-6361/202039543>
- Pápics, P. I., Tkachenko, A., Aerts, C., Van Reeth, T., De Smedt, K., Hillen, M., Østensen, R., & Moravveji, E. (2015). Asteroseismic Fingerprints of Rotation and Mixing in the Slowly Pulsating B8 V Star KIC 7760680. *803*(2), L25. <https://doi.org/10.1088/2041-8205/803/2/L25>
- Paxton, B., Bildsten, L., Dotter, A., Herwig, F., Lesaffre, P., & Timmes, F. (2011). Modules for Experiments in Stellar Astrophysics (MESA). *192*, 3. <https://doi.org/10.1088/0067-0049/192/1/3>
- Paxton, B., Cantiello, M., Arras, P., Bildsten, L., Brown, E. F., Dotter, A., Mankovich, C., Montgomery, M. H., Stello, D., Timmes, F. X., & Townsend, R. (2013). Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars. *208*, 4. <https://doi.org/10.1088/0067-0049/208/1/4>
- Paxton, B., Marchant, P., Schwab, J., Bauer, E. B., Bildsten, L., Cantiello, M., Dessart, L., Farmer, R., Hu, H., Langer, N., Townsend, R. H. D., Townsley, D. M., & Timmes, F. X. (2015). Modules for Experiments in Stellar Astrophysics (MESA): Binaries, Pulsations, and Explosions. *220*, 15. <https://doi.org/10.1088/0067-0049/220/1/15>
- Paxton, B., Schwab, J., Bauer, E. B., Bildsten, L., Blinnikov, S., Duffell, P., Farmer, R., Goldberg, J. A., Marchant, P., Sorokina, E., Thoul, A., Townsend, R. H. D., & Timmes, F. X. (2018). Modules for Experiments in Stellar Astrophysics (MESA): Convective Boundaries, Element Diffusion, and Massive Star Explosions. *234*, 34. <https://doi.org/10.3847/1538-4365/aaa5a8>
- Paxton, B., Smolec, R., Schwab, J., Gaudy, A., Bildsten, L., Cantiello, M., Dotter, A., Farmer, R., Goldberg, J. A., Jermyn, A. S., Kanbur, S. M., Marchant, P., Thoul, A., Townsend, R. H. D., Wolf, W. M., Zhang, M., & Timmes, F. X. (2019). Modules for Experiments in Stellar Astrophysics (MESA): Pulsating Variable Stars, Rotation, Convective Boundaries, and Energy Conservation. *243*(1), 10. <https://doi.org/10.3847/1538-4365/ab2241>
- Pedersen, M. G., Aerts, C., Pápics, P. I., Michielsen, M., Gebruers, S., Rogers, T. M., Molenberghs, G., Burssens, S., Garcia, S., & Bowman, D. M. (2021). Internal mixing of rotating stars inferred from dipole gravity modes. *Nature Astronomy*, *5*, 715–722. <https://doi.org/10.1038/s41550-021-01351-x>
- Tassoul, M. (1980). Asymptotic approximations for stellar nonradial pulsations. *43*, 469–490. <https://doi.org/10.1086/190678>
- Townsend, R. H. D., & Teitler, S. A. (2013). GYRE: an open-source stellar oscillation code based on a new Magnus Multiple Shooting scheme. *435*, 3406–3418. <https://doi.org/10.3847/1538-4365/ab2241>

172 [1093/mnras/stt1533](https://doi.org/10.1093/mnras/stt1533)

173 Varghese, A., Ratnasingam, R. P., Vanon, R., Edelmann, P. V. F., & Rogers, T. M. (2023).
174 Chemical Mixing Induced by Internal Gravity Waves in Intermediate-mass Stars. *942*(1),
175 53. <https://doi.org/10.3847/1538-4357/aca092>

DRAFT