Cluster Difference Imaging Photometric Survey. III. The Coeval Halo and Core of NGC 2516

L. G. BOUMA, J. D. HARTMAN, J. N. WINN, AND G. Á. BAKOS

¹Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08540, USA

(Received —; Revised —; Accepted —)
Submitted to ApJL.

ABSTRACT

Recent analyses of the Gaia data have reported the existence of diffuse stellar populations ("halos") surrounding nearby open clusters. The stars in these halos could have escaped from the cores of the cluster, or they could be remnants of a larger and lower-density star formation complex that has since dispersed. They could also be field stars—false positives given by overly aggressive clustering algorithms. In this study, we focus on a single reported 250 pc-long halo around the ≈ 120 Myr open cluster NGC 2516. The classical tidal radius of this cluster is \sim XX pc. Combining photometry from Gaia, rotation periods from TESS, and lithium measurements from the Gaia-ESO and GALAH surveys, we find that: i) the halo of NGC 2516 is real and coeval with the core; ii) at fixed stellar mass and age, rapidly rotating stars have greater lithium abundances than slowly rotating stars; and iii) given Gaia-selected halo members, the contamination rate of field stars in the halo (XX%) is larger than that in the core (YY%). This work expands the set of confirmed NGC 2516 members by a factor of \approx 2, and quantifies the degree to which Gaia-based analyses can currently identify dispersed stellar populations against a background of field star contaminants. Implications for spectroscopic survey targeting, open cluster dispersal, and planet searches around young stars are discussed.

Keywords: stellar ages (1581), young star clusters (1833), open clusters (1160), stellar rotation (1629)

1. INTRODUCTION

In the traditional picture of star formation molecular clouds go gravitationally unstable, and collapse into little knots. The knots produce groups of stars close to each other. The resulting "open star clusters" are, as with any first-order guess, spherical.

How does the cluster evolve? The answer depends on the total stellar mass. The smallest 90% of star clusters disperse during the "embedded phase" (CITE). This is driven mostly by *process A* (CITE). The more massive clusters make it maybe 100 Myr. Maybe a bit longer. Their evaporation is thought to be driven by collisions with molecular clouds (CITE, Spitzer 1958), and the galactic tide (CITE), and *process A*, and *process B*. Of course, the expected evaporation time should depend on factors including the mass of the cluster itself, and how many high-mass (O and B) stars form, since their winds [and maybe supernovae] clear out most of the cloud, and the initial density of the gas cloud to begin with.

Assuming that the cluster achieves virial equilibrium, stellar fly-bys then conspire to segregate the stellar mass distri-

bution within the cluster, evaporating the lowest mass stars soonest (2T+U=0, U=-GM/r, assuming equipartition of the specific "thermal energy" per star, then $kT = 0.5mv^2$ at lower masses requires higher velocities).

Identification of the stars that disperse into the galactic field is an important task for understanding the conditions under which stars and star clusters form, and for understanding how they subsequently evolve. For instance, how does the process of radial migration across the galactic disk affect cluster dispersal? Did the Sun form in an open cluster? If so, how massive was its cluster, and is there any hope at identifying the stars that formed near the Sun? Qualitatively, the dispersal of open clusters also provides perhaps the best "test case" for the concept of "chemical tagging", also referred to as "galactic archaeology" (CITE).

Outside of the issue of star formation, identifying the remnant halos of open clusters is important for a separate project: that of discovering young transiting planets. Young transiting planets are hard to find because young stars are rare (CITE), and often reside in crowded regions of the sky (CITE). If the halos of nearby star clusters can be reliably identified, this could expand the census of nearby sub-Gyr stars by a factor of 2 or even 3 (CITE Meingast). A fortuitous benefit of searching for planets in cluster environments is that issues with stellar crowding are also alleviated.

Recent clustering studies using Gaia have begun to report the identification of structures that could correspond to low-density halos of stars that may have evaporated from open clusters (e.g., Kounkel+19, Kounkel+20, Meingast+21). However, different clustering methods on the Gaia data tend to give different results (CITE: Hunt & Reffert 2020). Employing say Gaussian Mixture Modelling, or any analogous method that requires "clusters" to be ellipses in spatial/velocity phase-space unsurpsingly yields open clusters that are roughly elliptical (CITE: Cantat-Gaudin+18). Unsupervised clustering methods such as HDBScan (e.g., Kounkel+19) have been found to yield additional structures, particularly in lower density regions such as the Psc-Eri stream (CITE: Meingast+18, Curtis+19, Newton+21), but also more generally around many open clusters (Kounkel+19). Unsupervised approaches that incorporate physical constraints (e.g., imposing a maximum velocity dispersion on putative members) yield similar results, potentially with higher purity (Meingast+21).

We've recently been making TESS light curves of age-dated stars across the sky, as part of a Cluster Difference Imaging Photometry Survey (CDIPS, CITE Bouma+19). Our analysis of Cycle 1 (Sectors 1-13) yielded light curves of 483,407 candidate cluster members in the Southern Ecliptic hemisphere, available on MAST 1 . Based on rotation periods, $\approx 25\%$ appear to be bonafide cluster members.

As part of a broader project of identifying a large and clean sample of young stars for a transit search, we focus in his paper on a rather modest question: in just a single rich southern open cluster, is the cluster halo coeval with the core? The cluster we chose for this analysis was NGC 2516, since it was young (\sim 112 Myr) and close (d = 400 pc) enough to facilitate rotation measurements using TESS, and some degree of spectroscopic analysis. We want to know: is the halo real? To what extent can we use Gaia alone to reliably identify agedated needles in the haystack of boring field stars? And more generally, what are the implications for the evolution of open clusters if they do have halos?

Section 2 presents the astrometric and photometric data from Gaia, and clarifies our usage of the terms "core" and "halo". Section 3 age-dates the halo of NGC 2516, based on Gaia photometry (Section 3.1), TESS gyrochronology (Section 3.2), and lithium depletion (Section 3.3). In Section 4 we discuss the implications of this analysis for NGC 2516 specifically and stellar spin-down and open cluster evolution generally. Section 5 presents our conclusions.

2. A 250 PC HALO AROUND A CORE?

2.1. Gaia Astrometry

Figure 1 shows the problem we would like to address. In this figure, the "core" comprises NGC 2516members that Cantat-Gaudin et al. (2018) reported to have "membership probability" exceeding 10%, based on the Gaia DR2 astro-

metric data. The exact meaning of this probability, and details concerning their clustering algorithm, are discussed in Appendix A. The "halo" comprises NGC 2516members that Kounkel & Covey (2019) reported as members, also based on the Gaia DR2 data. That study did not report continuous membership probabilities, instead opting for the binary "member" or "not". These two different clustering methods yielded wildly different results. What is the true structure of NGC 2516? Are the core and halo truly coeval?

3. AGE-DATING THE HALO OF NGC 2516

3.1. HR Diagram from Gaia

The first check on whether this membership assignment is plausible was already performed by Kounkel & Covey (2019) and more recently by Meingast et al. (2021). That check is to see whether the HR diagrams of these cluster components that were selected based on positions and velocities photometrically support the claim that they are coeval.

Figure 2 presents similar results to what these investigators have already found. The core members of the cluster show a clean sequence consistent with stars with a fixed age and metallicity, and varying mass. The halo members are roughly consistent with this, but they do show greater scatter. One possible explanation for this scatter is that the halo is more contaminated by field stars. Another, explored in the other panel of the figure, is differential reddening. The halo is reported to span 20° on-sky, and varies in position from about $b = -12^{\circ}$ to $b = -20^{\circ}$, with the stars closest to the galactic plane also being further from the Sun by up to $200 \,\mathrm{pc}$ (Figure 1).

Comparing these HR diagrams to the PARSEC isochrone models, we find that X, Y, and Z. In particular, "the faintest M dwarfs in the core and halo are brighter than in the field star comparison sample, consistent with these stars having not yet reached the ZAMS." This is consistent with the mainsequence turn-off being at $Bp-Rp\approx 0.05$, which implies an age of XXX. The resulting photometric age we calculate for the core is XXX. For the halo, the claimed age from photometry is YYY. Applying the same procedure to the field star comparison sample, we get an age of ZZZ.

3.2. Rotation from TESS

Duly motivated, we collected light curves from TESS photometry for YYYY of the ZZZZ cluster members.

First, we collected the XXXX Gaia DR2 source_ids corresponding to the core and halo members. For each source, we first retrieved all available CDIPS light curves, on a per-sector basis. Then, we detrended the systematics in each sector individually, and stitched together the resulting light curves before searching for the periodicity. Some details regarding our detrending approach are discussed in Appendix B. After applying a detrending step aimed at removing systematic trends, we proceeding with a few small cleaning steps aimed at improving the purity of the rotation period measurements: we masked 0.7 days at the beginning and end of each spacecraft orbit, and ran a sliding standard-deviation rejection window over the light curve, which removed any

¹ https://archive.stsci.edu/hlsp/cdips

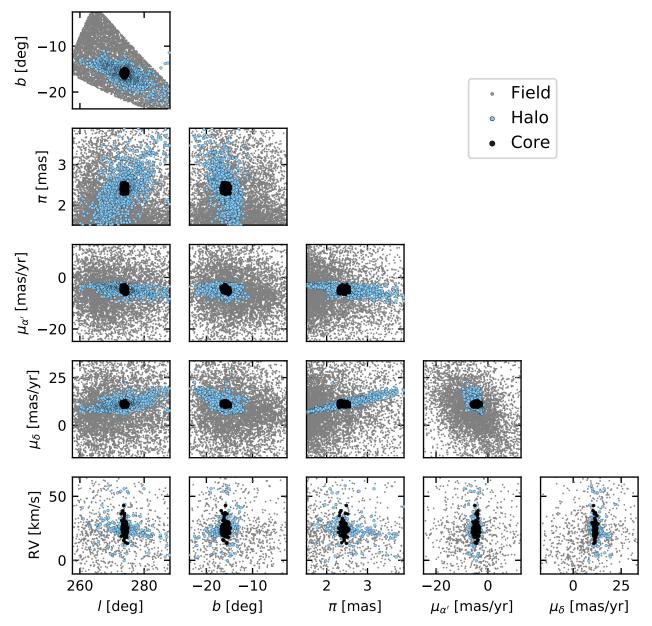


Figure 1. Reported components of NGC 2516 in position and velocity space. The "core", identified by Cantat-Gaudin et al. (2018) using Gaia DR2, is visually coincident with where you would think the cluster is if you looked at it through a pair of binoculars. The "halo" was identified by Kounkel & Covey (2019) using a less restrictive membership assignment algorithm (discussed in the appendices). The "field" is a set of randomly drawn and non-overlapping stars within a (α, δ, π) cone centered on the cluster.

outlying points within $\pm 3 \times MAD$ of the median in each window.

We then measured the rotation period of the resulting light curve. We used the the aperture radius that, based on theoretical expectations, was expected to give the optimal balance between light from the target and background-light (CITE Sullivan15). This typically resulted in an aperture radius of either 1 or 1.5 pixels. To measure the periods, we used the periodogram implementations in astrobase, in particular the Stellingwerf PDM periodogram (CITE), along with the more traditional Lomb-Scargle (CITE). We recorded the top

five periodogram peaks from each method, and their corresponding powers. Finally, as a check on crowding, we also recorded the number of stars within the aperture of equal brightness, and of brightness with 1.25 and 2.5 TESS magnitudes of the target star.

Figure 3 shows the resulting rotation periods. The points on this plot are light curves for which the peak Lomb Scargle periodogram period was below 15 days, had power exceeding 0.08, and had at most one equal-brightness companion within the aperture. These selection criteria are entirely heuristic.

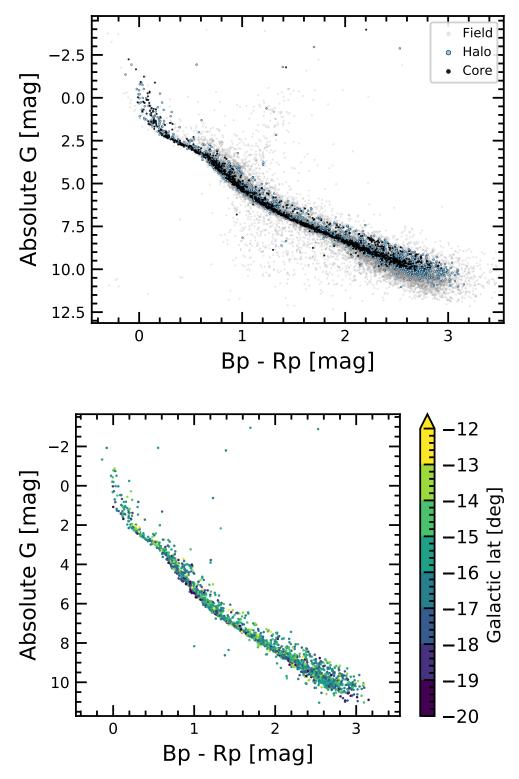


Figure 2. HR diagrams of NGC 2516, using Gaia EDR3 photometry. *Top:* The core (black) shows a clean sequence consistent with stars with a fixed age and metallicity, and varying mass. The halo (blue) is similar, but somewhat noisier. The faintest M dwarfs in the core and halo are brighter than in the field star comparison sample (gray), consistent with these stars having not yet reached the ZAMS. *Bottom:* Reported members of the halo, as a function of galactic latitude. Can the additional scatter in the halo be understood through differential reddening? **Maybe**.

But splitting the sample into the "core" and the "halo", they suggest that the halo is real.

The resulting gyrochronology age we find for the core is XXX. For the halo, the claimed age from gyrochronology is YYY. Applying the same procedure to the field star comparison sample, we get an age of ZZZ.

3.2.1. *Kinematics* ⊗ *Rotation*

How far away from the core, in position and velocity space, does the halo actually extend? We can explore this by crossmatching the rotator sample against the Gaia data. Figure 4 shows the result.

3.3. Lithium from Gaia-ESO and GALAH

Figure 5 shows the results. At fixed stellar mass (and age), it seems that the rapid rotators tend to show elevated lithium equivalent widths. Similar trends have been previously noted in the Pleiades by Soderblom et al. (1993) and Bouvier et al. (2018), in the Psc-Eri stream by Arancibia et al. (2020), and in M 35 by Jeffries et al. (2020).

The trend contradicts what might be the naive expectation that rapidly rotating stars should show more vigorous convection, leading to earlier destruction of photospheric Li. Possible physical explanations include processes both internal and external to the star (CITE see the recent review by Bouvier+20). Internal processes would be tied to the interplay between surface rotation, differential rotation in the interior, and the convective mixing efficiency (e.g., CITE Siess + Livio 1997, Baraffe+2017). Internal processes could also include the possibility that stronger magnetic fields in the stellar interior inhibit convection (READ e.g., Ventura+98, Chabrier+07, Somers + Pinsonneault 2014). An external process that could also be important is the effect of star-disk magnetic braking during the PMS phase (CITE: magnetic braking). Longer disk lifetimes would lead to the star's outer convective zone being "locked" for longer while the radiative core continues contracting. The resulting differential rotation and rotational mixing could drive the lithium depletion (CITE: Bouvier 08, Eggenberger+12).

4. DISCUSSION

- 4.1. What is the contamination fraction in the halo? Does it change vs. location?
 - 4.2. Are the "very slow rotators" bonafide members?

Probably not. They are not isotropically distributed around the cluster.

4.3. How did the halo form?

- 4.4. Mass differences between center and outer reaches?
- 4.5. Fast rotators: are we going faster than other clusters?

5. CONCLUSION

ACKNOWLEDGMENTS

The authors thank X and Y for fruitful discussions. L.G.B. and J.H. acknowledge support by the TESS GI Program, program NUMBER, through NASA grant NUMBER. This study was based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2020A-0146; 2020B-NUMBER PI: L. Bouma), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. ACKNOWLEDGE PFS / CAMPANAS. This paper includes data collected by the TESS mission, which are publicly available from the Mikulski Archive for Space Telescopes (MAST). Funding for the TESS mission is provided by NASA's Science Mission directorate. We thank the TESS Architects (George Ricker, Roland Vanderspek, Dave Latham, Sara Seager, Josh Winn, Jon Jenkins) and the many TESS team members for their efforts to make the mission a continued success.

Software: astrobase (Bhatti et al. 2018), astropy (Astropy Collaboration et al. 2018), astroquery (Ginsburg et al. 2018), cdips-pipeline (Bhatti et al. 2019), corner (Foreman-Mackey 2016), IPython (Pérez & Granger 2007), matplotlib (Hunter 2007), numpy (Walt et al. 2011), pandas (McKinney 2010), scipy (Jones et al. 2001), wotan (Hippke et al. 2019).

Facilities: Astrometry: Gaia (Gaia Collaboration et al. 2016, 2018). Imaging: Second Generation Digitized Sky Survey, SOAR (HRCam; Tokovinin 2018). Spectroscopy: CTIO1.5 m (CHIRON; Tokovinin et al. 2013), Photometry: TESS (Ricker et al. 2015).

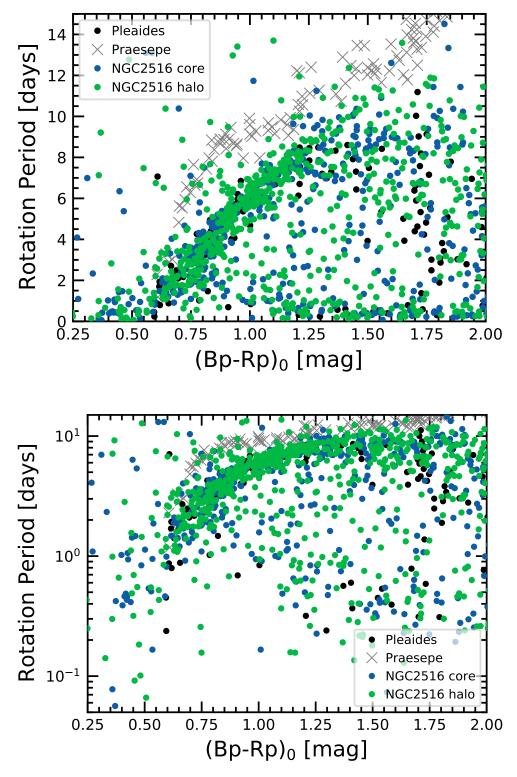


Figure 3. The core and halo of NGC 2516in the space of rotation period and Gaia color. The *top* plot shows periods against a linear scale, while the *bottom* shows them against a logarithmic scale.

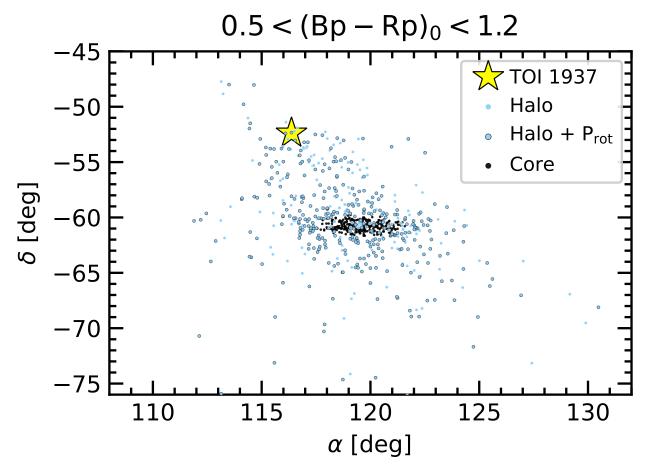


Figure 4. Gaia-based components of NGC 2516 in position and velocity space, cross-matched against the rotators.

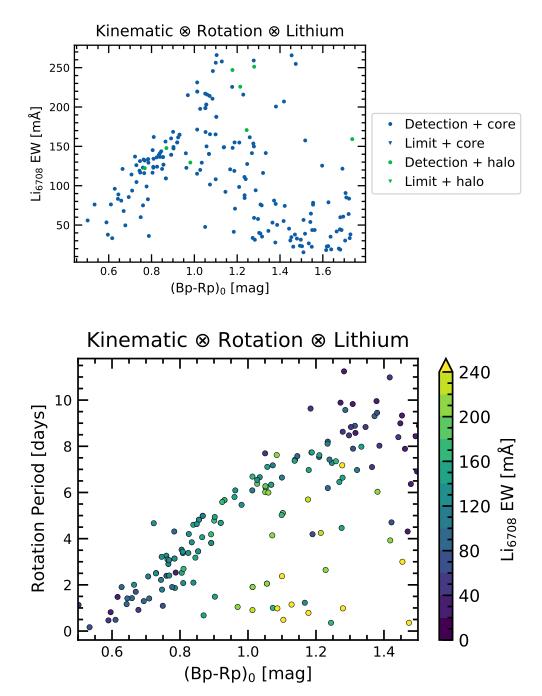


Figure 5. Lithium equivalent widths in NGC 2516, plotted against extinction-corrected color (*top*), and the combination of rotation period and extinction-corrected color (*bottom*).

REFERENCES

- Arancibia, J., Bouvier, J., Bayo, A., et al. 2020, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 61C, 81
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
- Bhatti, W., Bouma, L., & Yee, S. 2019, cdips-pipeline v0.1.0, https://doi.org/10.5281/zenodo.3370324
- Bhatti, W., Bouma, L. G., & Wallace, J. 2018, astrobase, https://doi.org/10.5281/zenodo.1469822
- Bouma, L. G., Hartman, J. D., Bhatti, W., Winn, J. N., & Bakos, G. Á. 2019, ApJS, 245, 13
- Bouvier, J., Barrado, D., Moraux, E., et al. 2018, A&A, 613, A63 Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, A&A, 618, A93
- Foreman-Mackey, D. 2016, Journal of Open Source Software, 1, 24
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016,
 A&A, 595, A1
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
- Ginsburg, A., Sipocz, B., Madhura Parikh, et al. 2018, Astropy/Astroquery: V0.3.7 Release
- Hippke, M., David, T. J., Mulders, G. D., & Heller, R. 2019, AJ, 158, 143

- Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
 Jeffries, R. D., Jackson, R. J., Sun, Q., & Deliyannis, C. P. 2020,
 MNRAS, 500, 1158, arXiv: 2010.04217
- Jones, E., Oliphant, T., Peterson, P., et al. 2001, Open source scientific tools for Python
- Kounkel, M., & Covey, K. 2019, AJ, 158, 122
- McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51
- Meingast, S., Alves, J., & Rottensteiner, A. 2021, A&A, 645, A84Pérez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
- Soderblom, D. R., Jones, B. F., Balachandran, S., et al. 1993, The Astronomical Journal, 106, 1059
- Tokovinin, A. 2018, PASP, 130, 035002
- Tokovinin, A., Fischer, D. A., Bonati, M., et al. 2013, PASP, 125, 1336
- Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22

APPENDIX

A. CLUSTERING METHOD DETAILS

Cantat-Gaudin et al. (2018) applied a procedure that ultimately yielded what we call "the core". Their procedure was to first query a Gaia DR2 cone around the previously reported RA and dec of the cluster, and within ± 0.5 mas of its previously reported distance. The outer radius of their cone was either r2 from MWSC (CITE Kharchenko 2013), or twice the "cluster radius" listed by Dias. No proper motion cut was applied. They then applied an unsupervised classification scheme called UPMASK to G < 18 mag stars within this cone (CITE). The steps of UPMASK are first to perform a k-means clustering in the "astrometric space" ($\mu_{\alpha'}, \mu_{\delta}, \pi$). Then, a "veto" step is applied to assess whether the groups of stars output from the k-means clustering are or are not more concentrated than a random distribution. This is implemented by "comparing the total branch length of the minimum spanning tree connecting the stars with the expected branch length for a random uniform distribution covering the investigated field of view". "To turn this yes/no flag to a membership probability, Cantat-Gaudin et al. (2018) then redraw new values of ($\mu_{\alpha'}, \mu_{\delta}, \pi$) for each source based on the listed value, uncertainty, and covariance. After a certain number of redrawings, the final probability is the frequency with which a given star passes the veto". In the case of NGC 2516, this yielded a reported "r50" within which half of the cluster members were found to be within 0.496°.

Kounkel & Covey (2019) applied a different unsupervised clustering method to the 5-dimensional Gaia DR2 data (omitting radial velocities, due to their sparsity). Their selection function (see their Section 2) yielded $\approx 2 \times 10^7$ stars, mostly within ≈ 1 kpc and typically with G < 18 mag. Their clustering algorithm was the "hierarchical density-based spatial clustering of applications with noise" (HDBSCAN, CITE McInnes17). The classical DBSCAN algorithm "identifies clusters as overdensities in a multi-dimensional space in which the number of sources exceeds the required minimum number of points within a neighborhood of a particular linking length ϵ . HDBSCAN does not depend on ϵ ; instead it condenses the minimum spanning tree by pruning off the nodes that do not meet the minimum number of sources in a cluster and reanalyzing the nodes that do. Depending on the chosen algorithm, it would then either find the most persistent structure (through the excess of mass method), or return clusters as the leaves of the tree (which results in somewhat more homogeneous clusters). In both cases it is more effective at finding structures of varying densities in a given data set than DBSCAN." "The two main parameters that control HDBSCAN are the number of sources in a cluster and the number of samples. The former is the parameter that rejects groupings that are too small; the latter sets the threshold of how conservative the algorithm is in its considerations of the background noise (even if the resulting noisy groupings do meet the minimum cluster size). By default, the sample size is set to the same value as the cluster size, but it is possible to adjust them separately."

B. DETRENDING DETAILS

In "detrending" for our general variability search, our goal was to preserve astrophysical variability, while removing systematic variability. One particular concern for the TESS light curves is systematic variability at the timescale of the 14-day satellite orbit, mostly induced by scattered light from the Earth and Moon.

We therefore turned to the principal components (i.e., the eigenvectors) calculated following the procedure described by Bouma et al. (2019). In brief, these vectors are computed using a set of "trend stars" selected from across each CCD according to ad-hoc heuristics that (hopefully) lead them to be dominated by *systematic* variability (Sec 3.7.2).

The principal component vectors, also referred to as the eigenvectors, are rank-ordered by the degree of variance that they predict in the training set (of "trend stars").

We then posit that any given target star's light curve is described as a linear combination of the eigenvectors. Optionally, we also considered the inclusion of additional systematic vectors that could affect the light curve, such as the CCD temperature, the flux level measured in the background annulus, and the centroid positions of the stars on the CCDs. These can be treated as additional "features" in the linear model.

There are many methods for determining the coefficients of the linear model, after the full set of eigenvectors (plus optionally "sytematic" vectors) has been asssembled. We explored two: ordinary least squares, and ridge regression. Ridge regression is the same as ordinary least squares, except it includes an L2 norm with a regularization coefficient. The regularization coefficient that best applied for any given target light curve was solved for using a cross-validation grid search, using sklearn.linear_modelRidgeCV (CITE).

Each target light curve was mean-subtracted and normalized by its standard deviation, as were the eigenvectors. The linear problem was then solved, and the light curve was reconstructed by re-adding the original mean, and re-multiplying by the standard deviation to ensure that the variance of the light curve did not change.

We found, somewhat to our surprise, that the choice of using ordinary least squares versus ridge regression did not seem to significantly affect the resulting light curves. In other words, the inclusion (or lack thereof) of a regularization did not strongly alter the best-fitting coefficients.

A few other choices seemed to be more important:

- To smooth, or to not smooth the eigenvectors. Ideally, the eigenvectors should be smooth (and not contain residuals from e.g., eclipsing binaries that snuck their way into the template set). However this does not always apply. Sometimes, the eigenvectors are also noisy, which leads them to induce extra variability into the PCA "detrended" light curves. To address this problem, we opted to smooth the eigenvectors using a windowed filter (with a "biweight" weight scheme, implemented in wotan by Hippke et al. (2019); window length 1 day, cval 6). One issue with this is that systematic sharp features (captured e.g., in "spike vectors") no longer are captured, so they end up in the "PCA detrended" light curves. They can be filtered out relatively easily though (using rolling outlier rejection), and we prefer this approach to having systematic features injected by the PCA detrending.
- How many eigenvectors to use. A larger number always leads to greater whitening. In Bouma et al. (2019), we performed a Factor Analysis cross-validation to determine the number of eigenvectors to use. The typical number adopted based on this analysis was 10−15. Even though this approach was chosen to prevent over-fitting, in our experience, for stellar rotation it still often lead to overfitting, especially for rotation signals with periodicities of ≥ 3 days. (Shorter signals typically are not distorted, since the eigenvectors generally do not contain the high-frequency content that leads to the distortions). For this analysis, we therefore impose the maximum number of eigenvectors to be 5.
- Which supplementary systematics vectors to use. We considered using the BGV, CCDTEMP, XIC, YIC, and BGV vectors, packaged with the CDIPS light curves. we found that the background value measured in an annulus centered on the aperture, BGV, tended to produce the best independent information from the PCA eigenvectors, and so we adopted it as our only "supplementary" trend vector. We opted to not smooth it (in hopes that it would provide direct complement to the smoothed PCA vectors; 1 sharp vector containing literally the background information, plus 5 smooth vectors).

For every "target star", we then decorrelated the raw (image-subtracted and background-subtracted) light curve using a linear model with ordinary least squares.

C. ROTATION \otimes RUWE

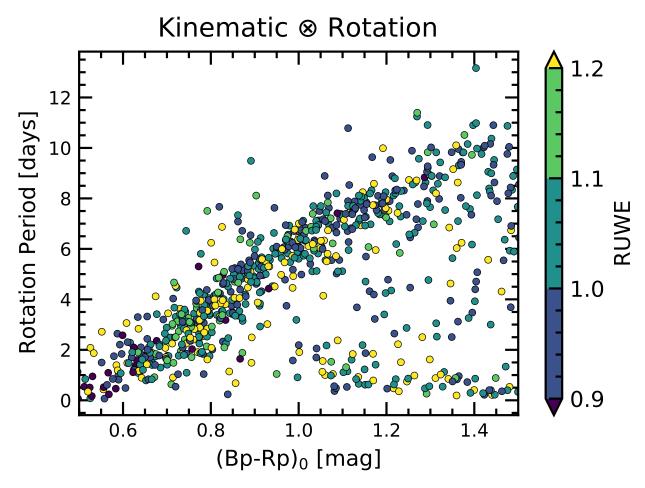


Figure 6. Rotation versus color, colored by RUWE. Looks like on the slow sequence, there's more yellow at the bottom?