

The DOHA algorithm: a new recipe for cotrending large-scale transiting exoplanet survey light curves.

Mislis D.^{1*}, Pyrzas S.¹, Alsubai K.A.¹, Tsvetanov Z.I.¹, Vilchez N.P.E.¹

¹*Qatar Environment and Energy Research Institute, LAS Building, Education City, Qatar Foundation, P.O. Box 5825, Doha, Qatar*

Accepted ??, Received ??; in original form ??

ABSTRACT

We present DOHA, a new algorithm for cotrending photometric light curves obtained by transiting exoplanet surveys. The algorithm employs a novel approach to the traditional “differential photometry” technique, by selecting the most suitable comparison star for each target light curve, using a two-step correlation search. Extensive tests on real data reveal that DOHA corrects both intra-night variations and long-term systematics affecting the data. Statistical studies conducted on a sample of $\sim 9\,500$ light curves from the Qatar Exoplanet Survey reveal that DOHA-corrected light curves show an RMS improvement of a factor of ~ 2 , compared to the raw light curves. In addition, we show that the transit detection probability in our sample can increase considerably, even up to a factor of 7, after applying DOHA.

Key words: Extrasolar planets – transits – survey – algorithm.

1 INTRODUCTION

In the last decade, a significant portion of the hunt for transiting extrasolar planets has been conducted by various ground-based, large-scale surveys, such as SuperWASP (Pollaco et al. 2006), HatNet (Bakos et al. 2004), TrES (Alonso et al. 2004) and QES (Alsubai et al. 2013). A common, defining characteristic of these surveys is that they were designed to cover as large a field of view as possible.

Data obtained by these surveys tend to suffer from a, more or less, common problem: the presence of unwanted flux variations that can either mask or mimic real (astrophysical) variations. A significant part of these variations is introduced by fixed, ordered trends in the data, collectively referred to as “systematics”. The list of systematics is rather long including, among others, variations due to airmass and seeing, colour-dependent extinction, object merging etc. The imprint of systematics on the data can be viewed as components leading to common-mode behaviour among the light curves of observed stars.

In addition, unwanted flux variations can also be introduced by random events. By definition these are not systematic, or in other words, they are events that do not have a distinct, common mode imprint on the data (see e.g. Pinheiro da Silva et al. (2008)).

As the photometric accuracy for ground-based exoplanet detection is required to be of the order of 1% or better, it became readily apparent that all these variations, with amplitudes that often exceed a few percent and “signatures” that can easily mimic a transit event, can severely reduce the transit detection probab-

ity and, therefore, need to be accounted for and corrected. This has led to detrending algorithms such as TFA (Kovács et al. 2005) and SysRem (Tamuz et al. 2005). Recently, similar work has been done for space missions such as CoRoT (Mislis et al. 2010; Ofir et al. 2010) and Kepler (Still et al. 2012).

While they differ in their implementation, the core idea of these algorithms remains the same: they try to identify and correct systematic patterns, by exploiting their common mode behaviour. A crucial factor in this exercise is the actual commonness of the patterns, in the (statistical) sense of what percentage of stars are affected by them, or in other words, how representative the patterns are of the entire sample. An additional consideration is the quantitative contribution of each pattern on the overall variations and whether specific patterns can be viewed as driving the variations. We maintain the distinction between common and uncommon dominant patterns throughout the manuscript.

In this paper we present DOHA, an algorithm conceived to correct for both systematic variations (regardless of commonness) and assorted data irregularities. The structure of the paper is as follows: in Section 2 we briefly describe the data set used in testing the algorithm; in Section 3 we present the algorithm itself, while Section 4 contains the results after applying DOHA to our sample light curves. Section 5 shows a test for signal detection efficiency and Section 6 summarises our work.

2 THE SAMPLE DATA

For testing our algorithm we used data from the Qatar Exoplanet Survey (QES). The QES uses six 4k \times 4k FLI ProLine PL6801 cam-

* E-mail: dmislis@qf.org.qa

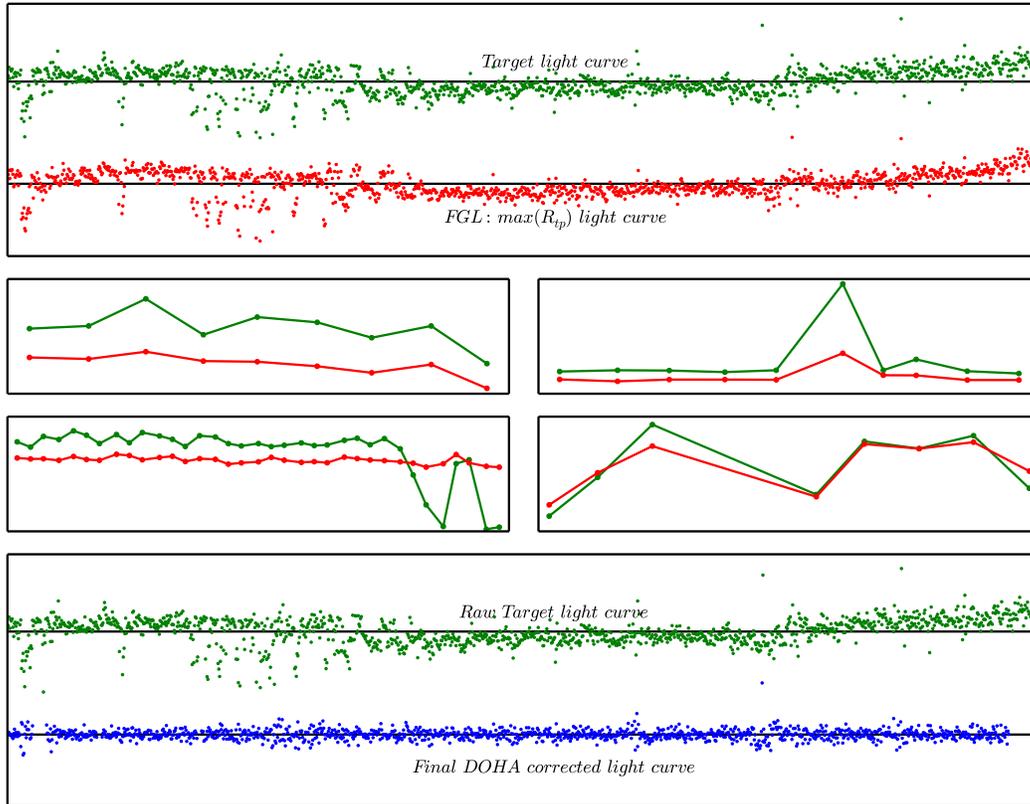


Figure 1. **Top:** raw target light curve (green), and light curve of the *highest correlated* comparison star from the FGL (red). For clarity, light curves are plotted as consecutive points and *not* according to their timestamps. **Middle:** raw target light curve (green) and master comparison light curve (red) for four individual nights. **Bottom:** raw target light curve (green) and final, DOHA-corrected light curve (blue). Note: light curves have been normalized to unity by their respective mean flux. See text for additional details.

eras, equipped with 4x400mm, 1x200mm and 1x135mm lenses, mosaiced to image an $11^\circ \times 11^\circ$ field on the sky in the magnitude range of $10 < m_V < 17$. For our purposes, we selected data from a single field (RA= $3^h 50^m$, DEC= $-3^\circ 00'$), obtained with one of the cameras equipped with a 400mm lens (FOV= $5.24^\circ \times 5.24^\circ$). The data were collected over a three-month period (end Mar-Jun 2010) and consist of $\sim 9\,500$ stars, with an average of $\sim 1\,300$ data points each and an exposure time of 60 sec. The data were reduced using the QES pipeline as described in Alsubai et al. (2013).

3 THE ALGORITHM

Let us assume that our data set consists of light curves $f(n)$, with n number of data points each, and that the total number of light curves is N . We wish to correct the light curve of the t -th *target* star $f_t(n)$. DOHA achieves correction using a two-step correlation search approach.

In the first step, the algorithm calculates the correlation coefficient R_{ip} between the target star $f_t(n)$ and a *potential* comparison star $f_p(n)$. Only stars with $\text{RMS}^P < 1.5 * \text{RMS}^T$ are considered

as potential comparison stars, with RMS^P and RMS^T the RMS values of the raw comparison and target light curves respectively. The correlation coefficient is calculated as

$$R_{ip} = \frac{1}{n-1} \sum_{i=1}^n \frac{(f_t(i) - \bar{f}_t) \cdot (f_p(i) - \bar{f}_p)}{\sigma_{f_t} \cdot \sigma_{f_p}}, \quad (1)$$

where $f_t(i)$ and $f_p(i)$ are the i -th data points of the target and potential comparison star light curves respectively; \bar{f}_t and \bar{f}_p are the mean flux values of the light curves; and σ_{f_t} and σ_{f_p} are the standard deviations of the target and comparison light curves respectively. We note that R_{ip} is calculated on the *common set of points* of $f_t(n)$ and $f_p(n)$; missing points are *not* substituted.

At the end of this first step, N^* values of R_{ip} correlation coefficients have been calculated in total. Subsequently, the mean and standard deviation of all R_{ip} values are derived, and those stars that have corresponding R_{ip} values larger than $2\text{-}\sigma$ from the mean are selected. In this fashion, we create a “family” of *definite* comparison stars, of size D , with light curves $f_d(n)$, $d = 1, 2, \dots, D$. Henceforth, we will denote this as *Family Group Light curves* (FGL).

We should note that the $2\text{-}\sigma$ cutoff limit is not cast in stone. It

is a “middle-ground” balance between selecting an adequate number of stars for the FGL on the one hand; and selecting only those stars that are strongly correlated with the target on the other. The limit can be adjusted to better suit the given data set, depending e.g. on the number of stars with sufficient data points and on the severity of the systematic and non-systematic trends.

In the second step, the algorithm splits the light curve of the target star $f_t(n)$ to its individual-night segments. Working in each segment separately, the algorithm re-calculates R_{td} correlation coefficients, but this time, only the $f_d(n)$ light curves from the FGL are taken into consideration. We use single-night segments to better account for airmass and colour-extinction variations.

As before, the mean and standard deviation of the values of R_{td} are calculated and those stars with R_{td} larger than 1.5σ from the mean are selected to create the *New Family Group Light curves* (NFGL). Assuming that the NFGL is of size M , we will denote its light curves as $f_m(n)$, $m = 1, 2, \dots, M$. We re-iterate that the NFGL is created for *each* individual night segment; there are as many NFGLs as there are individual nights in the data. Also note that the constituent light curves $f_m(n)$ of one NFGL are not necessarily the same as those of another.

Again, the 1.5σ cutoff is a reasonable “default” value, adjusted to account for the much smaller-sized FGL (compared to the original number of light curves), following the considerations described previously.

Still working on an individual night basis (inb), from the corresponding NFGL, we create a “master” comparison light curve $f_{c.inb}(n)$, which is the mean of all the $f_m(n)$ light curves in the given NFGL,

$$f_{c.inb}(n) = \frac{1}{M} \sum_{m=1}^M f_m(n) \quad (2)$$

Once the master comparison curve is calculated, correction of the target light curve $f_{t.inb}$ (for the given individual night) is achieved via a double-iterative, “global” RMS minimisation technique, as follows:

- We construct an array of *scaling correction factors* S_i , of size I ($i = 1, \dots, I$), with $S_i \in (0, 1)$ and arbitrary step; for our tests, we chose a step of 0.01^1
- We define RMS^0 , the RMS of the raw target light curve segment, $f_{t.inb}^0$, as the reference, starting point
- An I-steps iteration over all S_i begins
- Subsequently, a j -steps iteration begins, $j = 1, 2, \dots$. For the given S_i , at the j -th step, a *temporary* corrected target light curve $f_{t.inb(i)}^j$ is constructed by

$$f_{t.inb(i)}^j = f_{t.inb(i)}^{j-1} - S_i \cdot f_{c.inb} \quad (3)$$

- The RMS of this light curve, $RMS_{(i)}^j$, is calculated and compared to that of the previous j -step, $RMS_{(i)}^{j-1}$
- Iterations over j halt when $RMS_{(i)}^j \geq RMS_{(i)}^{j-1}$
- The I-steps iteration continues with S_{i+1} , and the entire process is repeated

The corrected target light curve segment $f_{t.inb}^{cor}$ is chosen to be the one with the minimum RMS, that is $RMS^{cor} = \text{MIN}(RMS_{(i)}^j)$. The *final* DOHA-corrected light curve of the target star f_t^{fin} is simply the concatenation of all corrected target light curve segments $f_{t.inb}^{cor}$.

¹ In which case, $I=99$; $S_1 = 0.01$ and $S_I = 0.99$

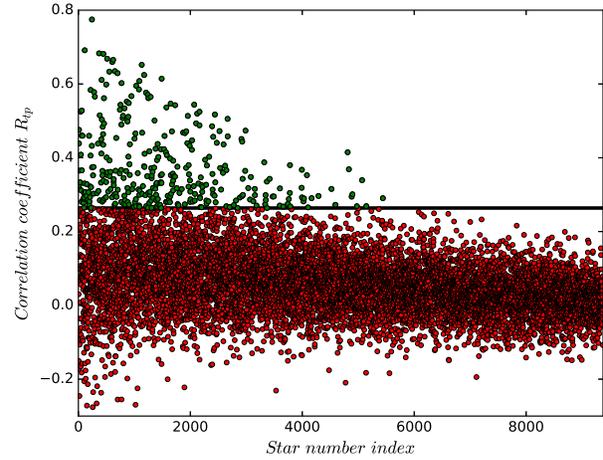


Figure 2. Correlation diagram for a randomly selected target star. The black, solid line indicates our 2σ selection criterion. The (green) points above this line form the FGL.

Finally, we note that, as with the σ -cutoffs applied in the creation of the FGL and the NFGL, the array-step for the S_i values can be adjusted to better suit a given data set.

For additional clarity, we provide representative illustrations of the algorithm’s description, as presented above, in Figures 1 and 2.

In the top panel of Figure 1 we plot a (randomly selected) raw target light curve f_t^0 (green curve), and the light curve of the *highest correlated* comparison star from the FGL (red curve). To better highlight variations around the respective mean values (indicated as solid lines), light curves are plotted as consecutive points and *not* according to their timestamps, which span a range of three months. In the four middle-panels, we show four different individual night segments and we plot the corresponding $f_{t.inb}^0$ points (again in green), along with the master comparison light curve $f_{c.inb}$ (in red) for that particular night. Finally, in the bottom panel we plot again the raw target light curve f_t^0 (green) as well as the final, DOHA-corrected light curve f_t^{fin} (blue).

Note that DOHA does not apply explicit outlier rejection. Most of the outlying points in the raw light curve (Fig. 1, bottom panel, top curve) were actually “brought-in-line” by the algorithm’s correction.

Figure 2 shows the N^* values of the R_{tp} correlation coefficients, calculated through Equation 1, for a random target star. The black, solid line corresponds to the 2σ level away from the mean. All the stars above this line (green points) are used to form the FGL.

Note that our data are magnitude-sorted (brightest to faintest). The target star used in Fig.2 is a 11.3-mag one. Given the $RMS^P < 1.5 * RMS^T$ criterion and the fact the bright stars are more likely to have high-correlation comparisons (this is explained in detail in the next Section), it is not surprising that most of the FGL members clump on the left-hand side of Figure 2. i.e. have a small star-index.

4 RESULTS

4.1 Statistical tests

For each of the ~ 9500 stars in our field, we constructed the corresponding FGL, noted the on-chip position of the target and of the

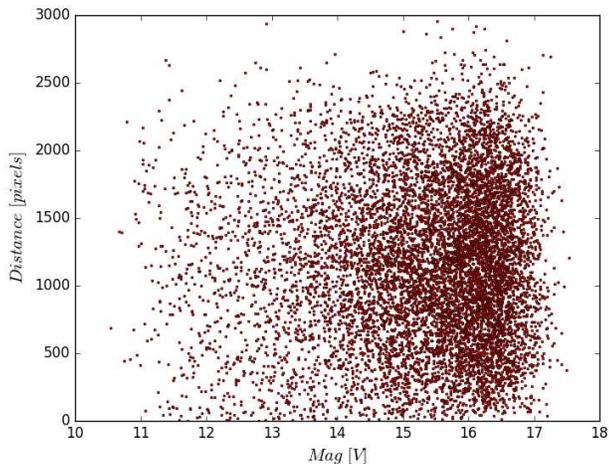


Figure 3. Target magnitude versus distance (in pixels) of the highest correlated comparison star for the given target. The distribution is totally random, and highest correlated stars can be found anywhere on chip, independent of target magnitude.

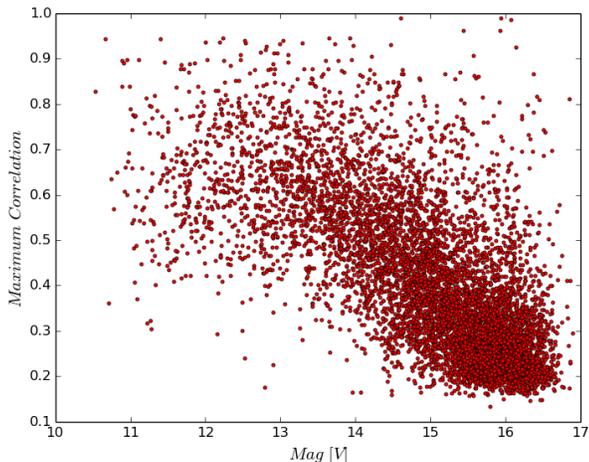


Figure 4. Target magnitude versus $\text{MAX}(R_{ip})$ correlation coefficient value. A linear trend is clearly visible, and comparison stars for the brighter targets show higher R_{ip} values.

highest correlated star in the FGL and calculated their distance in pixels. In Figure 3 we plot the target magnitude versus the distance to the highest correlated comparison star. The distance is given in pixels, with 1 pixel corresponding to $4.64''$. Figure 3 does not show any distinct pattern between the two plotted quantities. Effectively, the highest correlated star for any given target can be located anywhere on the chip.

Additionally, we noted the actual R_{ip} value of the highest correlated comparison for a given target and in Figure 4 we plot these $\text{MAX}(R_{ip})$ against target magnitude. Despite the scatter, a linear trend is clearly visible in Figure 4, showing that for bright stars, it is much more likely to find a comparison star with a high R_{ip} value. This could be an indication that red noise dominates the systematics, and bright stars are more susceptible to red noise, as they tend to occupy a larger area (more pixels) on the CCD.

Finally, we combine Figures 3 and 4 and plot the distance of

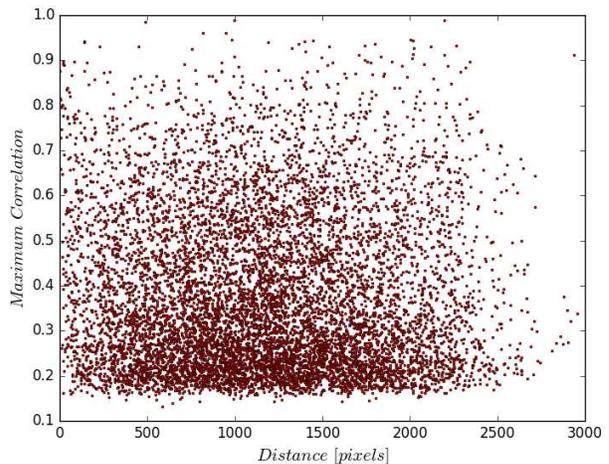


Figure 5. Distance of the highest correlated comparison star versus maximum correlation value. Once more, there is no obvious trend between the two quantities; even for very high R_{ip} values (>0.9), the corresponding comparison stars can be located anywhere on the chip.

the highest correlated comparison star versus the corresponding R_{ip} value in Figure 5. As with Fig. 3, there is again no distinct pattern between the two plotted quantities. The value of maximum correlation is independent of the distance between target and comparison, even for high R_{ip} values (>0.9).

4.2 RMS diagram

In order to better assess the performance of DOHA, we constructed the RMS diagrams of both the raw and the DOHA-corrected light curves for all stars in our sample. The resulting diagrams are shown in the left panel of Figure 6. The dashed black line indicates the theoretical noise floor curve.

To appreciate the light curve improvement, we can define the *relative RMS improvement* as $1 - \text{RMS}_{\text{DOHA}}/\text{RMS}_{\text{RAW}}$. This is plotted in the middle panel of Figure 6 against target magnitude. In general, brighter stars show larger improvement, a result of the fact that for brighter stars it is easier to find comparisons with very high R_{ip} values (see again Fig. 4). We should also note that very large relative improvement factors (>0.8) should be interpreted with some caution. As mentioned before (and see also Fig. 1 and 8) DOHA performs very well with outliers; part of the improvement can be attributed exactly to outlying points being brought to the correct level. For this reason, in the right-hand panel of Figure 6 we again plot the relative improvement, but this time using the *median absolute deviation (MAD)*, that is $1 - \text{MAD}_{\text{DOHA}}/\text{MAD}_{\text{RAW}}$.

Using statistics from the Kepler mission, Howard et al. (2012) estimate that there are 0.0066 transiting hot Jupiters per star, with orbital periods up to 10 days. The deepest, ground-observed, transiting exoplanet so far is HATS-6b, with $(R_p/R_*)^2 = 0.0323$ (Hartman et al. 2015). From Fig. 6, we calculate that in our raw sample, $\sim 17\%$ of the light curves have the required accuracy to detect such a transit. This percentage increases to $\sim 35\%$ after applying DOHA, i.e. there is a factor of 2 improvement.

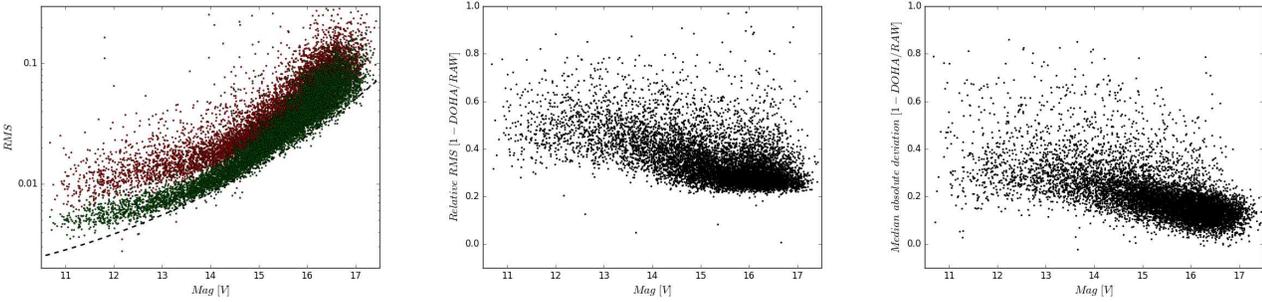


Figure 6. **Left:** The RMS diagram of the sample before (red dots), and after DOHA (green dots). The dashed black line indicates the theoretical noise floor curve. **Middle:** Relative improvement between raw data and DOHA-corrected data versus target magnitude, using RMS-statistics. **Right:** Relative improvement using MAD-statistics.

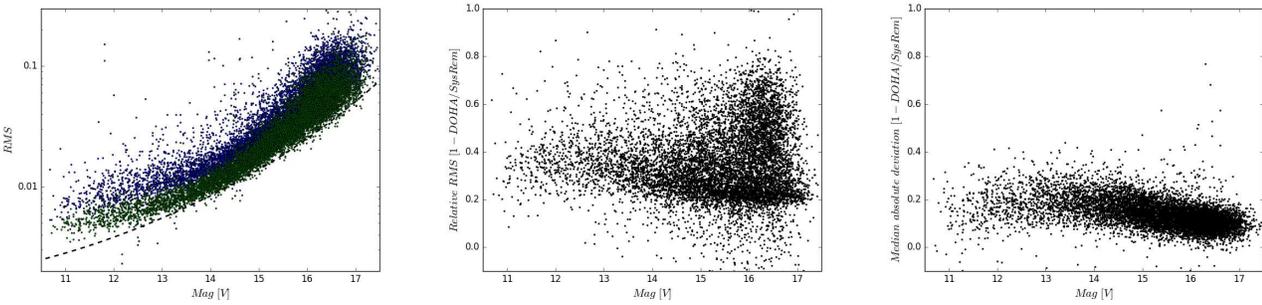


Figure 7. Same as Fig. 6, but this time for SysRem (blue dots) and DOHA (green dots).

4.3 Comparison with SysRem

As a further performance test, we subjected our sample to detrending using the SysRem algorithm, and compare it with the DOHA results from the previous section. Figure 7 shows the resulting RMS diagrams in the left panel, and the relative RMS and MAD improvements in the middle- and right panel respectively. There is a rather small percentage ($\sim 2\%$ of the total sample) of stars where SysRem yields a better RMS. Also, the same considerations about the treatment of outliers, mentioned previously, apply here.

The main difference between SysRem and DOHA is the assumption on the nature of the patterns affecting the data. The implicit assumption of SysRem is that the common-mode patterns are dominant and can be expressed as linearly varying components, calculated from the *entire* sample and, furthermore, that these calculated components are representative of the entire sample and can therefore be used to correct it. On the other hand, DOHA makes no assumption on the nature or the dominance of the patterns (common or uncommon, as described in the introduction) and, moreover, DOHA tries to find the representative components for each star individually, without being based on whole sample statistics. To illustrate the point of uncommon dominant patterns, we refer the reader to Fig. 2 again, where it can be seen that (a) the majority of stars actually show very little correlation with the target, (b) practically a quarter of the stars shows, in fact, negative correlation and (c) there are indeed stars which show high correlation.

4.4 Individual light curves

4.4.1 General examples

To illustrate the performance of DOHA more accurately, we present here three individual light curves selected from our sample. These light curves are representative examples of the patterns affecting our data. Table 1 gathers some basic information on these three stars. The reported periods come after running the “Box Least Squares” (BLS) algorithm of Kovács et al. (2002).

- 3UC171-131243 (Fig. 8, top panels): this is a constant star, but systematics introduce non-real variability, which is picked-up by the BLS search. DOHA, not only corrects the systematics creating the variability, but also corrects almost all the outliers between phase $0.1 < \phi < 0.3$.

- 3UC175-129698 (Fig. 8 middle panels): this is a typical short period variable star, ($P = 0.391388$ [d]) Note that the amplitude of the variability remains unaffected after applying DOHA. Furthermore, similar to the previous example, the algorithm manages to correct almost all the outlying points.

- 3UC177-129206 (Fig. 8 bottom panels): the light curve of this system seems to contain a “transit-like” signature at phase $\phi \sim 0.8$. The corrected light curve is much cleaner without outliers, but, most importantly, the “transit-like” signal disappears. This is a case where DOHA successfully corrects a false-positive identification. The fact that this star is, indeed, constant is supported by radial velocity measurements which show no RV variations, to a level of $K < 40$ msec $^{-1}$, corresponding to a mass of smaller than $0.2M_J$.

6 *Mislis D.*

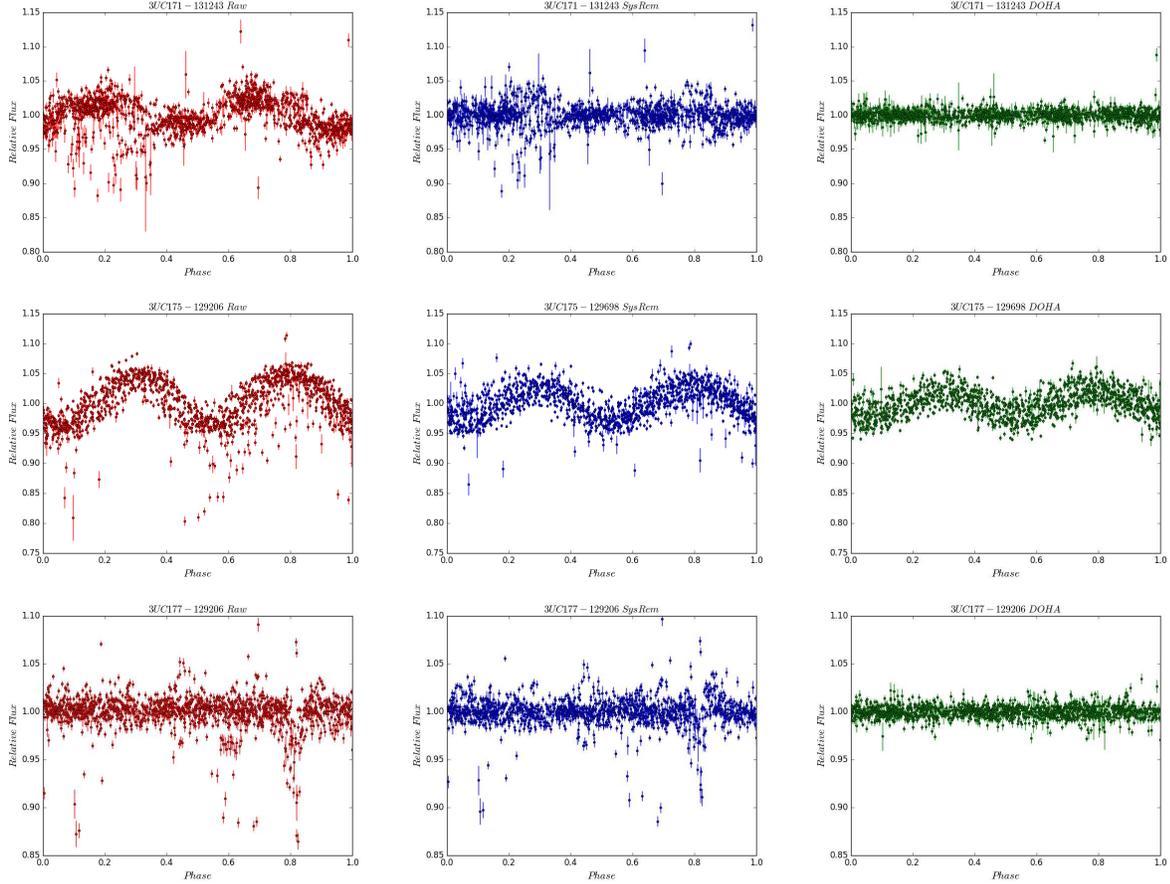


Figure 8. The light curves of (top to bottom) 3UC171-131243, 3UC175-129698 and 3UC177-129206. **Left:** raw light curve; **Middle:** light curve corrected with SysRem; **Right:** light curve corrected with DOHA. The data were phase-folded using the period obtained from running BLS. See Table 1 for more information on the stars and the periods.

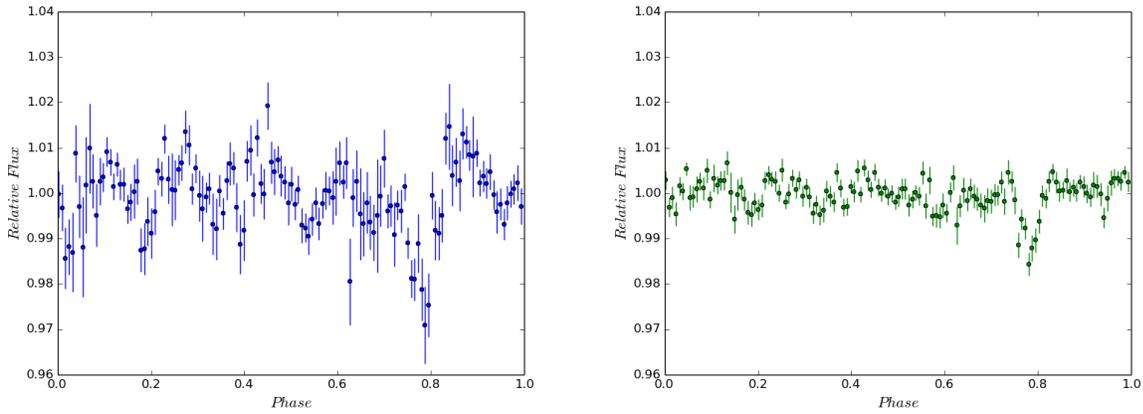


Figure 9. WASP-1b phase-folded and binned light curves using SysRem (left panel) and DOHA (right panel) for detrending.

4.4.2 *Transiting light curves*

A further test was to assess the performance of DOHA on known transiting exoplanets that have been observed with the QES. We have selected data that actually contain three known planets in the

same field²: WASP-1b (Collier Cameron et al. 2007) with period $P_W = 2.5199464$ days and magnitude $mag_W = 11.63$; HAT-P-19b (Hartman et al. 2011) with period $P_H = 4.008784$ days and magnitude $mag_H = 12.9$; and KELT-1b (Sivert et al. 2012) with period

² This field is different than the one described in Sec. 2.

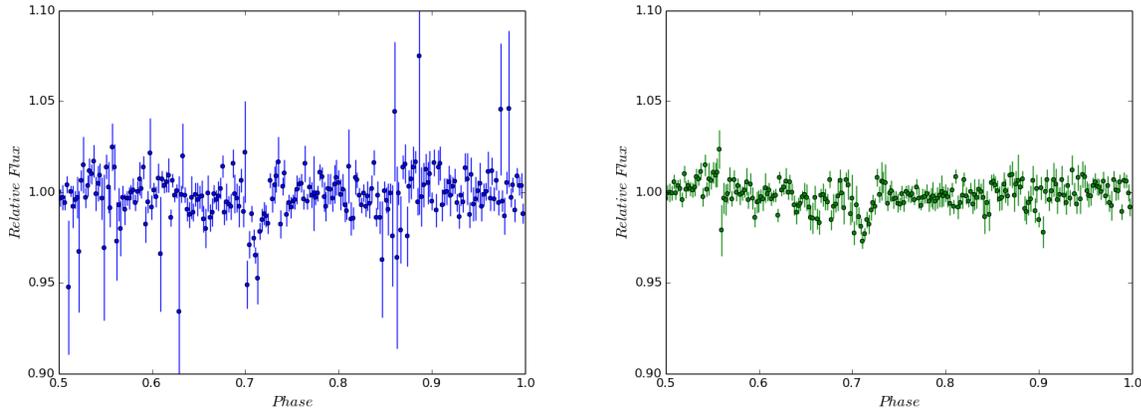


Figure 10. Same as Fig. 9, but for HAT-P-19b.

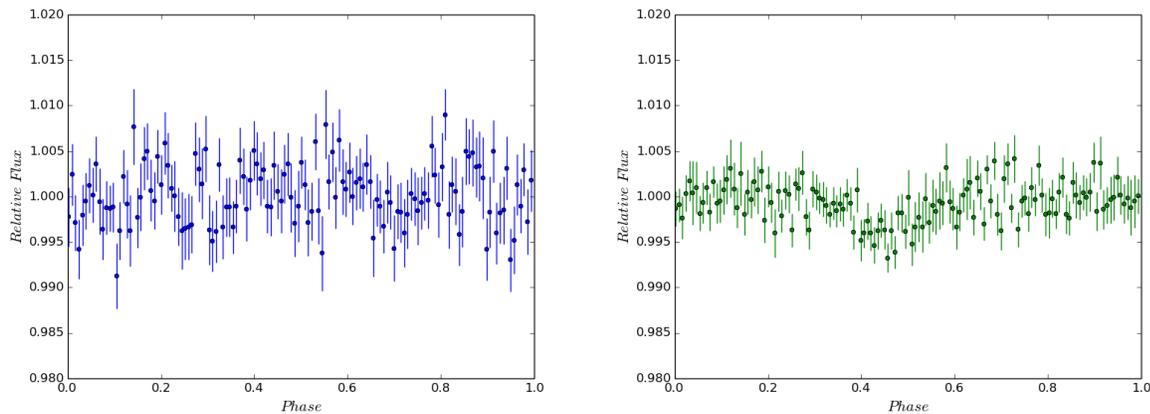


Figure 11. Same as Fig. 9 & 10, but for KELT-1b.

Table 1. Basic information for the three example stars.

UNSO4 ID	RA	Dec	Mag	BLS period [d]
3UC171-131243	13:57:16.09	-04:31:51.10	13.4	1.966354
3UC175-129698	13:52:05.31	-02:45:32.70	13.0	0.391308
3UC177-129206	13:54:36.36	-01:42:11.80	12.8	1.596665

$P_K = 1.217514$ days and magnitude $mag_K = 10.7$. KELT-1b is a very massive object ($27M_J$), but because it is orbiting a mid-F type star, the depth of the transit is very small (0.6%), making it an ideal target for our test.

We first detrended the raw data using *SysRem* and subsequently ran BLS on the resulting light curves. BLS successfully detected WASP-1b and HAT-P-19b, but failed to detect KELT-1b. We then, repeated the process, only this time we corrected the raw light curves using DOHA. In this case, all three planets were successfully detected by the BLS, with the correct parameters for orbital period and transit depth. Figures 9-11 show the *SysRem* (left panels) versus DOHA (right panels) phase-folded and binned light curves of the three planets.

5 SIGNAL DETECTION ALGORITHMS

As a final test, we investigated the effect of DOHA on the performance of signal detection algorithms. For comparative purposes, we also ran the same tests using the *SysRem*-detrended light curves, as presented in Sec. 4.3.

The tests were conducted as follows: we injected simulated transit signals, generated using the Pál (2008) model, in all the 9374 raw light curves of our sample (Sec. 2). For all transits, the stellar and planetary parameters were kept fixed to $R_\star = 1.0 R_\odot$, $M_\star = 1.0 M_\odot$ and $R_p = 1.0 R_J$, while the orbital period was randomly chosen from a uniform distribution, with $1 [d] < P_{orb} < 5 [d]$. This combination of stellar and planetary parameters was chosen to ensure a large number of detections for statistical purposes.

Subsequently, the raw light curves (now including the transit signals) were subjected to correction using both *SysRem* and DOHA. In each corrected set, we searched for transits using two separate signal detection algorithms: BLS and *SiDRA* (Mislis et al. 2016). We note that the test was not designed to compare BLS with *SiDRA*, only to assess how the probability of detecting a transit, using each detection algorithm, changes after applying DOHA.

The combination of *SysRem*+BLS yielded 149 transits (1.6% of the total), whereas DOHA+BLS successfully identified 1226 transits (13.1% of the total). To have a clearer view, we divided our

sample in 0.5mag-wide bins and in Figure 12 we plot the BLS detection efficiency in each magnitude bin. If we now restrict the considered magnitude range to $mag < 14$ (the working magnitude range of the QES survey), then BLS correctly identifies 6.2% of the transits, using SysRem; and 58% of the transits, using DOHA.

For the test with SiDRA³, we imposed a strict 70% confidence cut-off (see Mislis et al. 2016, for details). At this level, SiDRA classified 505 systems as definite planets using the SysRem light curves (5.4% of the total); and 938 systems (10.0% of the total) using the DOHA-corrected light curves. If we again restrict the magnitude range, as before, then SiDRA returns 7% of the total number of planets, using SysRem; and 20% using DOHA.

As a by-product, using the DOHA light curves, SiDRA correctly identified three (already known) variables in the field (2 RR Lyr and 1 W Uma) which have been missed in a variable search using SysRem-detrended light curves.

It is evident that DOHA significantly increases the chances of finding transiting planets, regardless of the signal detection algorithm employed.

6 CONCLUSIONS

In this paper we have presented DOHA, a new algorithm for correcting light curves obtained with large-scale, ground-based photometric surveys, with an emphasis on those of transiting exoplanets. Adopting the reasoning of a comment made by the referee during the review process, we denote DOHA as a *cotrending*, rather than a detrending algorithm.

DOHA is based on the standard differential photometry technique of correcting a target light curve using a master comparison light curve constructed from suitable, individual comparison stars. The success of DOHA lies in its ability to optimise the way in which suitable comparison stars are selected for *each* target *separately*. DOHA looks for and corrects common-mode patterns shared by the target and a “base” of comparison stars (constituting a small subset of all stars in the field), which is built after a two-step correlation search; the first accounting for long-term trends, the second for intra-night variations. In short, DOHA exploits the defining characteristic of systematics, that is their manifestation as common-mode behaviour of the data, without making any assumptions on their nature and prevalence. As such, DOHA is able to correct data trends and patterns regardless of their commonness and/or individual contribution to the variations in the sample. Our algorithm can either be used as stand-alone on raw light curves, or as a compliment to detrending algorithms, correcting for residual uncommon patterns.

To test and assess the performance of DOHA, we have used ~9500 light curves from the QES transiting survey. The results show that DOHA is able to improve the light curve RMS by a factor of 2, doubling the probability of detecting a transit signal. Results also indicate that DOHA is particularly efficient on bright stars.

Finally, by adding simulated transits in all of our sample light curves, we showed that, using DOHA combined with two separate signal detection algorithms, the number of successful detections can increase considerably.

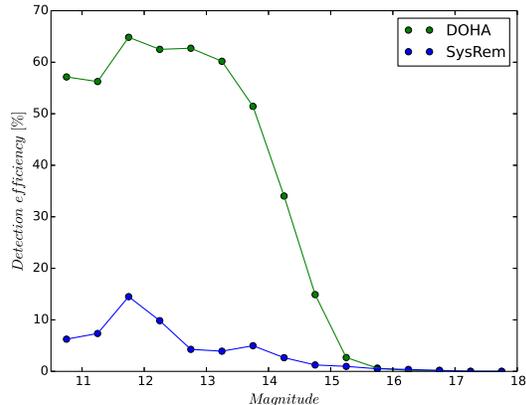


Figure 12. BLS detection efficiency using SysRem (blue curve) and DOHA (green curve) in each 0.5mag-wide magnitude bin.

ACKNOWLEDGMENTS

We would like to thank the anonymous referee for a prompt and useful report. This publication was made possible by NPRP grant # X-019-1-006 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the author.

REFERENCES

- Alonso R. et al, 2004, ApJ, 613L, 153
 Alsubai K. et al. 2013, Acta Astron., 63, 465
 Bakos G., Noyes R. W., Kovács G., Stanek K. Z., Sasselov D. D. and Domsa I., 2004, PASP, 116, 266
 Collier Cameron A., Bouchy F., Hébrard G., et al. 2007, MNRAS, 375, 951
 Hartman J., D., Bakos G., Sato B., et al. 2011, ApJ, 726, 52
 Hartman J., D., Bayliss D., Brahm R., et al. 2015, ApJ, 149, 166
 Howard A., Marcy G., Bryson S., et al. 2012, ApJS, 201, 15
 Kovács G., Zucker S. & Mazeh T., 2002, A&A, 391, 369
 Kovács G., Bakos G. and Noyes R., 2005, MNRAS, 356, 557
 Mislis D., Schmitt J. H. M. M., Carone L. et al. 2010 A&A, 522, 86
 Mislis D., Bachelet E., Alsubai K. A., et al. 2016, MNRAS, 455, 626
 Ofir A., Alonso R., Bonomo A., et al. 2010, MNRAS, 404L, 990
 Pál A., 2008, MNRAS, 390, 281
 Pinheiro da Silva L., Rolland G., Lapeyrere V. et al. 2008, MNRAS, 384, 1337
 Pollaco, D. L. et al, 2006, PASP, 118, 1407
 Siverd R., Beatty T., Pepper J., et al. 2012, ApJ, 761, 123
 Still M & Barclay T., 2012, Astrophysics Source Code Library, ascl.soft08004S
 Tamuz, O., Mazeh, T., Zucker, S., 2005, MNRAS, 356, 1466

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.

³ SiDRA is an entropy-based, random forest classification algorithm, and does not yield physical parameters, such as the orbital period.