Overview of semi-sinusoidal stellar variability with the CoRoT satellite


1 Departamento de Física, Universidade Federal do Rio Grande do Norte, 59072-970 Natal, RN, Brazil
e-mail: renan@dfte.ufrn.br
2 Departamento de Astronomia y Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile
3 The Milky Way Millennium Nucleus, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile
4 LESIA, UMR 8109 CNRS, Observatoire de Paris, UVSQ, Université Paris-Diderot, 5 place J. Janssen, 92195 Meudon, France
5 Departamento de Física, Facultad de Ciencias Básicas, Universidad Metropolitana de la Educacion, Av. José Pedro Alessandri 774, 7760197 Ñuñoa, Santiago, Chile
6 Universidade de São Paulo/IAG-USP, rua do Matão, 1226, Cidade Universitária, 05508-900 São Paulo, SP, Brazil
7 IFPB – Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, Av. Primeiro de Maio 720, 58015-430 João Pessoa, PB, Brazil
8 European Southern Observatory, Casilla 19001, Santiago, Chile
9 Department of Physics and Astronomy Uppsala University, Box 516, 751 20 Uppsala, Sweden
10 Universidade Federal do Rio de Janeiro, Observatório do Valongo, Ladeira do Pedro Antonio, 43, 20080-090 Rio de Janeiro, Brazil
11 Center for Radio Astronomy and Astrophysics Mackenzie, Universidade Presbiteriana Mackenzie, Rua da Consolação, 896 São Paulo, SP, Brazil

Received 13 April 2012 / Accepted 24 April 2013

ABSTRACT

Context. To date, the CoRoT space mission has produced more than 124 471 light curves. Classifying these curves in terms of unambiguous variability behavior is mandatory for obtaining an unbiased statistical view on their controlling root-causes.

Aims. The present study provides an overview of semi-sinusoidal light curves observed by the CoRoT exo-field CCDs.

Methods. We selected a sample of 4206 light curves presenting well-defined semi-sinusoidal signatures. The variability periods were computed based on Lomb-Scargle periodograms, harmonic fits, and visual inspection.

Results. Color–period diagrams for the present sample show the trend of an increase of the variability periods as long as the stars evolve. This evolutionary behavior is also noticed when comparing the period distribution in the Galactic center and anti-center directions. These aspects indicate a compatibility with stellar rotation, although more information is needed to confirm their root-causes. Considering this possibility, we identified a subset of three Sun-like candidates by their photometric period. Finally, the variability period versus color diagram behavior was found to be highly dependent on the reddening correction.

Key words. stars: variables: general – stars: rotation – techniques: photometric

1. Introduction

The CoRoT space mission has been operational for more than three years (Baglin et al. 2009). Its main science goals are asteroseismology and the search for exoplanets based on transit detection. Thanks to this space mission, a unique set of light curves (LCs) is now available for about 140 000 stars, with excellent time-sampling and unprecedented photometric precision. The photometric data obtained are a rich source for different astrophysical studies. For instance, the luminosity of stars can vary for a number of reasons, including gravitational deformation and eclipses due to binarity, as well as surface oscillations and rotation resulting from star spots. The variability induced by each phenomenon has a characteristic range of time scales and amplitudes. Open questions of high interest in this area include the behavior of stellar rotation periods, differential rotation as a function of latitude, distribution of spot areal coverage, the spot distributions in longitude and latitude on different stars, the presence and distribution of active longitudes, the timescale for evolution of different-sized spots, spot contrasts, and the evolutionary behavior of all these as a function of stellar mass, age, and metallicity.

An initial overview of stellar variability in CoRoT data was described in Debosscher et al. (2007, 2009), based on automated supervised classification methods for variable stars. The authors described a significant fraction of (quasi-) monoperiodic variables with low amplitude in the first four measured fields of the CoRoT exoplanet program. The majority are most likely rotationally modulated variables, with some low-amplitude Cepheids. Nevertheless, as reported by these authors, automatic procedures offer variability classification that is sensitive to different artifacts. Therefore, misclassification always...
occurs and its incidence depends significantly on the variability class considered. This is particularly true for variability caused by star spots, which are highly dynamic, and reflects the effects of surface rotation, differential rotation, spot lifetime, and transient phenomena such as flares, primarily for lower mass main-sequence stars (e.g., Lanza et al. 2007; Hartman et al. 2009). For example, star spots and photometric rotational modulation have long been studied using photometry or spectroscopy (Strassmeier 2009; Hartman et al. 2010; Meibom et al. 2009, 2011; Irwin et al. 2011). However, ground-based observations result in a number of time gaps, and the coverage area of spots on star surfaces must be several times larger than that on the Sun to achieve a robust signal. More recently, Affer et al. (2012) presented rotation period measurements for 1727 CoRoT field stars and claimed to have identified a sample of young stars (<600 Myr).

The present study describes the first results of our effort to determine variability periods from the LCs of stars in the exoplanet fields observed with CoRoT. Of the 124 471 LCs produced to date from all CoRoT observing runs, we selected those displaying semi-sinusoidal variations with particular patterns to be described in Sect. 2.2.4. The first portion of this paper is devoted to describing the stellar sample, the observations, and the procedure for determining the variability period. Our results are presented together with global analyses based on available stellar parameters, such as color index, luminosity classes, and spectral types. Here, we describe our unprecedented contribution to the treatment of CoRoT LCs, which was the correction of reddening effects on the stellar colors. In fact, CoRoT was designed to observe stellar samples in the Galactic center and anti-center directions. The exo-fields typically observe relatively faint stars with $V \sim 11$ to 16 mag. This suggests that CoRoT targets may be subject to considerable interstellar extinction effects. The reddening effect on CoRoT targets has not been explored yet. Careful analysis of these effects is mandatory, for example, to minimize possible bias on color–period behavior of variability distributions, as well as the location of stars in color–period diagrams. In particular, evolutionary scenarios of stellar variability parameters can certainly be better understood when a reddening correction is performed. Finally, our main conclusions are presented, in addition to the primary goals for future studies.

### 2. Working sample, observations, and data analysis

Raw LCs are collected by the CoRoT satellite as N0 data and after they are processed on the ground by the CoRoT pipeline (Samadi et al. 2007) in two levels. In the first level some electronic, background, and jitter effects are corrected, and data taken during on the South Atlantic Anomaly (SAA) passage are flagged, producing the N1 data. Subsequent treatments are processed in the second level and include sampling combination, calculation of heliocentric date, and flagging of hot pixels. The results are the N2 data, which are provided to the public for science analysis. For this investigation we selected the calibrated LCs measured with CoRoT exoplanet CCDs during 3 years of operation, with stars exhibiting visual magnitudes ranging from about 12 to 16. Time sampling for the LCs is 32 s, but for most data an average is calculated over 16 such measurements, resulting in an effective time resolution of 512 s. For a fraction of the LCs (or in some cases, parts of them), the original 32 s sampling was retained. These LCs correspond to high-priority targets measured in oversampling mode, totaling approximately 124471 CoRoT LCs from the initial run (IR), long runs (LR), and short runs (SR), with a time window of between 20 and 157 days. Additional basic properties of the data are listed in Table 1.

#### Table 1. Basic properties of the dataset analyzed by CoRoT, indicating the number of LCs in each observing run and the respective total span time, totaling 124471 LCs.

<table>
<thead>
<tr>
<th>CoRoT run</th>
<th>Total LCs</th>
<th>Total span time (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRa01</td>
<td>9880</td>
<td>54–57</td>
</tr>
<tr>
<td>LRa01</td>
<td>11408</td>
<td>131</td>
</tr>
<tr>
<td>LRa02</td>
<td>11408</td>
<td>111–114</td>
</tr>
<tr>
<td>LRa03</td>
<td>5289</td>
<td>148</td>
</tr>
<tr>
<td>LRc01</td>
<td>11407</td>
<td>142–152</td>
</tr>
<tr>
<td>LRc02</td>
<td>11408</td>
<td>144</td>
</tr>
<tr>
<td>LRc03</td>
<td>5661</td>
<td>89</td>
</tr>
<tr>
<td>LRc04</td>
<td>5716</td>
<td>84</td>
</tr>
<tr>
<td>LRc05</td>
<td>5683</td>
<td>87</td>
</tr>
<tr>
<td>LRc06</td>
<td>5683</td>
<td>77</td>
</tr>
<tr>
<td>SRa01</td>
<td>8150</td>
<td>23</td>
</tr>
<tr>
<td>SRa02</td>
<td>10265</td>
<td>31</td>
</tr>
<tr>
<td>SRa03</td>
<td>4130</td>
<td>24</td>
</tr>
<tr>
<td>SRc01</td>
<td>6975</td>
<td>25</td>
</tr>
<tr>
<td>SRc02</td>
<td>11408</td>
<td>20</td>
</tr>
</tbody>
</table>

Notes. The lower-case “a” means the Galactic anti-center and “c” means the Galactic center direction.

#### 2.1. Data treatment

The CoRoT pipeline provides N2 LCs corrected for several effects, but still with some problems that need additional treatment before the science analysis. In particular, CoRoT N2 LCs may have jumps (discontinuities) produced by hot pixels, long-term trends produced by CCD temperature variations, and outliers. There is no standard method for those post-treatments and different works have their methods according to their objectives (e.g., Renner et al. 2008; Basri et al. 2011; Affer et al. 2012). We describe below our procedure performed for data treatment, selection, and analysis. Our procedure is basically a set of steps and rules that were mainly performed manually. In particular, a simplified automatic version of this procedure was run in the beginning for a preliminary sample selection, then the selected sub-sample was re-treated and re-analyzed manually step by step.

We considered the LCs in normalized flux units, namely $F$, dividing each LC by its whole flux average. We then defined the noise level $\sigma$ of each LC as being simply a high-frequency contribution obtained from the standard deviation of the difference between the nearest-neighbor flux measurements, which yields

$$
\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (F_i - F_{i-1})^2},
$$

(1)

where $F_i$ is the flux value corresponding to the observing time $t_i$, and we considered $F_0 \equiv F_N$. For a homogeneous noise calculation the LCs were resampled to a bin of 864 s (0.01 days)\(^1\). The next steps were the data treatment, where we first performed a

---

\(^1\) This bin also saved computation time and did not affect the frequencies considered in the data analysis (Sect. 2.2).
jump correction, followed by a long-term detrend and, finally, the removal of outliers. These steps are explained below.

2.1. Jump correction

The CoRoT pipeline corrects for some jumps, but some still remain in the N2 LCs. These discontinuities may be caused by a sudden change of the mean LC level in a single (or few) time step(s). Therefore a LC may have several jumps. There is no unique method for detecting and correcting these discontinuities. Different algorithms were proposed to correct for them, usually in the search of planetary transits (e.g. Mislis et al. 2010).

For some other types of stellar variabilities, these corrections may be more challenging because we must keep the information of smoother and more irregular variations than those caused by transits. We developed a method for detecting and correcting jumps that combined with visual inspection can be used in the correction of most cases. Figure 1 shows an example of an LC with several jumps that hide any physical information. After the jump correction with this method, here applied automatically, the variability signature becomes noticeable. This method worked well automatically in this example and in several cases, although it produces miscorrections in some cases. Nevertheless, we checked in practice that for the large majority of LCs any mismatch produces a lower power in the periodogram than the main variation (e.g., in Fig. 1) and does not strongly affect the determination of the variability parameters. After a first automatic filtering, a manual correction was applied by testing different levels of corrections when analyzing the LCs, periodograms, and phase diagrams. Doubtful cases were simply rejected.

To determine if there was a jump within a time interval \( t_{i-1} \) and \( t_i \), we considered a box of duration \( \Delta t \) both to the left (previous) and to the right (after) of the time interval. First the mean flux was estimated within each box separately. When the difference between the left and right flux averages, \( \Delta F \), was greater than a defined threshold, \( \Delta F_t \), then a discontinuity was assumed to occur from \( t_{i-1} \) to \( t_i \). To avoid correcting for false jumps caused by a very steep flux variation within the LC, a linear fit was performed independently on the data contained in both boxes of duration \( \Delta t \). The higher of the two angular coefficients of the fits was assumed as \( |\Delta F|/\Delta t_{\text{max}} \), the estimated rate of flux variation between the boxes. The jump threshold, \( \Delta F_t \), was defined as

\[
\Delta F_t = a + b |\Delta F|/\Delta t_{\text{max}} (t_i - t_{i-1}),
\]

where \( \sigma \) is the noise level, whereas \( a \) and \( b \) are constants. To correct for the discontinuity when \( \Delta F > \Delta F_t \), we considered two boxes of a short duration \( \Delta t < \Delta t \) to the left and right of the detected jump. The flux levels of the left and the right box were adjusted so as to make the box averages equal. Our experience showed that the box durations of \( \Delta t = 1 \) day and \( \Delta t = 0.1 \) day and a threshold level with \( a \approx 4 \) and \( b \approx 2 \), were capable of detecting and correcting most jumps.

2.1.2. Long-term trends and outliers

After the jump correction, we minimized long-term trends by dividing the LC by a third-order polynomial fit, as performed in previous works (e.g., Basri et al. 2011; Affer et al. 2012). Finally, we removed some outliers with flux values that typically differed by more than about five times the standard deviation of a LC. From this point on, a LC was considered to be fully treated and its analysis could be performed.

2.2. Light curve analysis and selection

To properly analyze the LCs and select their parameters, we developed simple noise-free LC models from harmonic fits similar to those described in Debosscher et al. (2007). For each LC, the Lomb-Scargle periodogram (Lomb 1976; Scargle 1982) was computed for periods with a false alarm probability FAP < 0.01 (significance level >99%). The highest periodogram peak, named frequency \( f_1 \) or period \( P_1 \), was refined to a near frequency with the highest ratio of amplitude (calculated from a harmonic fit of the phase diagram; see below and Sect. 2.2.2) to the minimum dispersion (computed from Eq. (2) given in Dworetsky 1983) of the phase diagram \(^2\). Next, the refined frequency \( f_1 \) was used to calculate a harmonic fit with four harmonics. The fit was computed from a non-linear least-squares minimization using the Levenberg-Marquardt method (Levenberg 1944; Marquardt 1963). It was used to estimate a preliminary variability period and mean amplitude together with their errors. The final period was not necessarily \( P_1 \), as we explain in Sect. 2.2.2. Based on this fit, a mean signal-to-noise ratio (S/N) of the LC was estimated as

\[
S/N = \frac{A(\text{mag})}{\sigma(\text{mag})},
\]

where \( A(\text{mag}) \) is the mean variability amplitude in units of magnitude and \( \sigma(\text{mag}) \) is the mean LC noise defined in Eq. (1) and converted to magnitude.

For the LC, the fit was then subtracted from the time series (prewhitening) and a new Lomb-Scargle periodogram was computed. The same procedure was repeated in ten iterations, finding ten independent frequencies, each with a harmonic fit of four harmonics. These ten independent frequencies were then used to obtain a harmonic best fit with the original (trend-subtracted) time series as follows:

\[
y(t) = \sum_{j=1}^{4} \sum_{m=1}^{10} a_{ij} \sin(2\pi f_j t) + b_{ij} \cos(2\pi f_j t) + b_0,
\]

\(^2\) This adjustment reduces numerical errors such as those originating from the periodogram resolution and from the LC time window.
where $a_{ij}$ and $b_{ij}$ are Fourier coefficients, $t$ is the time and $b_0$ is the background level. The choice of four harmonics and ten iterations is based on a compromise between a good fit and computation time. This harmonic fit was used as a model to analyze some temporal variations on the amplitude and other aspects. These aspects are specifically criteria (iv) to (vi) of Sect. 2.2.3.

### 2.2.1. Selection by S/N

In the simplified automatic procedure, the methods described above were run in the beginning to estimate the mean S/N of the LCs and select a first subsample. To define a proper cut-off value for the S/N, we determined the reliability of a variability period as a function of the S/N by testing several simulations of semi-sinusoidal variabilities (300,000 simulations). These simulations were random pieces of actual CoRoT LCs with variabilities showing more than five cycles and S/N $> 5$ (that were assumed to have a good period determination, namely “true” periods), extracted from our own sample. For each simulated LC, the high-frequency signal – assumed to be the noise – was amplified or reduced by a random factor to change the S/N. Next, the Lomb-Scargle periodogram was computed and the most significant peak was compared with the “true” period from the simulation. We then counted how many times the simulation periods were correctly recovered. The recovery fraction as a function of the S/N is shown in Fig. 2.

In practice, Fig. 2 indicates the likelihood of determining the correct variability period as a function of the S/N of an LC. This probability rises with increasing S/N and usually does not reach 100% even with several cycles. The highest probability occurs for $S/N \gtrsim 1.0$, thus, only LCs with S/N above this cut-off value were selected. The reason why the recovery fraction does not reach 100% is explained below.

### 2.2.2. Period determination

An important problem in the variability period determination is the fact that the observed period may be an alias or harmonic of the actual period (e.g., Hartman et al. 2010). Aliases are seen as several discrete peaks in the periodogram of a LC. In some cases, selecting the correct period among harmonics or aliases may be ambiguous, and choosing the correct peak is generally a difficult task. Aliases may be avoided when the time window of the LC is long enough to present several cycles of the variability under analysis. Indeed, the greater the number of observed cycles, the better the determination of variability period.

Therefore, we determined the reliability of a variability period as a function of the number of cycles observed by testing several simulations of semi-sinusoidal variabilities (300,000 simulations). Similarly to the S/N analysis (Sect. 2.2.1), random pieces of actual CoRoT LCs were taken, in this case with variabilities showing more than ten cycles (to consider the best period determination, namely “true” periods) in our sample with an $S/N > 1.0$. The recovery fraction as a function of the number of cycles is shown in Fig. 3. This probability, $\rho$, rises with increasing number of cycles and does not reach 100%, as in the S/N analysis. Based on this figure, variabilities were divided into two groups: a higher confidence group, with more than three observed cycles in their LCs and $\rho \gtrsim 80\%$, and a lower confidence group, exhibiting less than three observed cycles in their LCs.

We suggest that the highest recovery fraction, as obtained above automatically, does not reach 100% because for a number of LCs (~5%) the actual period is not the strongest periodogram peak, even for long-term observations. This may occur in particular when photometric variability can be modeled as two main sinusoids per cycle. This is the case, for example, for some multimode pulsators and also for many rotating stars that display active regions at opposite faces and produce two main dips per lap. Therefore, this limitation has instrumental and physical origins.

Accordingly, we developed a simple method for minimizing this problem. Figure 4 shows a CoRoT LC that was interpreted here to be such a case. To identify these cases, the phase diagram was always checked for twice the period $P_1$ of the strongest periodogram peak. When two dimmings had notably different depths in the phase diagram, the true period $P$ was taken to be $2P_1$, otherwise it was $P_1$. In this analysis, the harmonic fit of Eq. (4) was obtained for the phase diagram with a fixed period (thus without the sum on $i$) and the final period was also refined to the highest amplitude and lowest dispersion (as explained above). This method is not a final solution to avoid aliases, but it certainly helps to reduce this problem, thus, increasing the recovery fraction closer to 100%. Finally, for some LCs with more than one type of variability superposed one another (e.g., rotation + pulsation), we selected the one that better met the criteria described in Sect. 2.2.3. Thus, for some cases, another period, $P_i$ or $2P_i$, was selected instead of $P_1$ or $2P_1$.

### 2.2.3. Semi-sinusoidal signature

Based on CoRoT LCs with known rotational modulations, as, e.g., CoRoT-2 (Silva-Valio & Lanza 2011), CoRoT-4 (Lanza et al. 2009), CoRoT-6 (Lanza et al. 2011), and CoRoT-7 (Lanza et al. 2010), the semi-sinusoidal signature was defined here by six main criteria.

(i) The variability period is longer than ~0.3 days.

(ii) The mean amplitude is typically $\leq 0.5$ mag.

---

**Fig. 2.** Recovery fraction for period determination as a function of S/N. The vertical dashed line indicates the S/N (~1.0) above which the recovery fraction is close to the maximum.

**Fig. 3.** Recovery fraction for period determination as a function of the number of cycles. The vertical dashed line indicates the number of cycles (~3) above which the recovery fraction is greater than ~80%.
(iii) The periodogram shows a relatively narrow spread around the variability peak.
(iv) The flux maximum and minimum per cycle are often asymmetric with respect to the flux average per cycle.
(v) The amplitude varies randomly and smoothly, with a characteristic period of $\sim 10^{-30} \times \text{of the variability.}$
(vi) The short-term flux variation has a smooth semi-sinusoidal shape that can be superposed with a second semi-sine, near in period, varying independently and smoothly in amplitude and phase.

Such a detailed description is needed because rotational modulation may indeed present very complex patterns and these criteria avoid subjectivity in the visual inspection. The ranges in criteria (i) and (ii) are those expected for most rotating stars (e.g., Eker et al. 2008; Hartman et al. 2010) and were chosen as a compromise to save time. To apply criteria (i) and (ii) the Lomb-Scargle periodogram was calculated within the Nyquist-frequency range for all LCs and those with the main periodogram peak with a frequency greater than 0.3 c/d or a mean amplitude greater than 0.5 mag were discarded.

Criteria (iii) to (vi) were applied by visual inspection. Based on criterion (iii), variabilities showing strongly spread peaks in the periodogram (e.g. some semiregular and irregular variables) were discarded. Based on criterion (iv), several pulsators, which usually show nearly symmetric maximum and minimum flux per cycle, were rejected. Based on criterion (v), variabilities with regular amplitude variations (e.g., RR Lyrae) or with nearly constant amplitude (e.g., eclipsing binaries) were also rejected. Figure 5 shows the example of a selected LC and its amplitude over time. The figure also shows the example of a discarded LC, whose amplitude variations are regular. Finally, for the LCs we kept, criterion (vi) was used to select those with a short-term semi-sinusoidal behavior. Note that the CoRoT time window and noise limits in many cases hampered a proper analysis of criteria (iii) to (vi) altogether. In particular, the long-period variabilities were more often subject to a mis-selection (and these were often classified as the lower confidence group defined in Sect. 2.2.2), but they were still selected here because of their importance in studying stellar evolution. Therefore, considering that the final sample is a list of candidates, the selection was not very conservative.

Instrumental effects were taken into account using a procedure similar to that described in Degroote et al. (2009). We selected 1000 LCs of different runs, interpreted as constant stars, and computed the average of their Fourier periodograms to identify instrumental signatures. The variabilities found in each individual LC were visually compared with the instrumental signatures and periodicities identified as instrumental were rejected.

It is important to note that identifying LCs with semi-sinusoidal signatures as defined here is useful for selecting rotating candidates if no other information than photometry is available. However, not all semi-sinusoidal LCs are necessarily produced by rotation and not all rotating variables produce semi-sinusoidal signatures. A better selection of rotating variables can only be made with the aid of spectroscopic data. Nevertheless, selecting this particular type of variation may provide a good filtering of rotating candidates.

In some cases, a faint third semi-sine contribution may also be found. Superposed semi-sines may be due to spots but also to some pulsations. The other criteria must be analyzed carefully to validate this one.

---

3 In some cases, a faint third semi-sine contribution may also be found. Superposed semi-sines may be due to spots but also to some pulsations. The other criteria must be analyzed carefully to validate this one.

4 These amplitude variations were calculated within boxes with a duration equal to the variability period.
Fig. 6. Subsample of LCs presenting the typical variabilities considered in our analysis. The upper panel illustrates LCs of FGK stars, the middle panel shows LCs of M-type stars, and the lower panel depicts the LCs for the Sun-like candidates defined in Sect. 3.3.

2.2.4. Final selection

All CoRoT N2 LCs of the exofield were first analyzed automatically to select sources with valid flux measurements, a mean S/N greater than 1.0 (see Sect. 2.2.1), and meeting criteria (i) and (ii). Based on 2MASS infrared photometric data, sources from this sample showing contamination and confusion flags were excluded. These flags indicate that photometry and/or 2MASS position measurements of a source may be contaminated or biased by the proximity of an image artifact or a nearby source of equal or greater brightness.

With visual inspection, all methods and criteria described above yielded a final sample of 4206 targets exhibiting confident semi-sinusoidal variability, as we show in the next section, with spectral types F, G, K, and M and luminosity classes III, IV, and V as listed in the CoRoTSky database. A portion of the sample has unknown spectral types and luminosity classes, hereafter represented by a question mark (?). Table 2 displays the computed periods and amplitude of variability, and different stellar parameters (CoRoT ID, right ascension, declination, spectral type, luminosity class, B magnitude, V magnitude, CoRoT run, J magnitude, H magnitude, Ks magnitude, variability period, variability amplitude, and signal-to-noise ratio). The error average of the variability period is ∼3% and that of the amplitude is ∼2 mmag. Figure 6 shows a sample of LCs presenting the typical variabilities considered in our sample, namely a semi-sinusoidal behavior.

2.3. Sample description and biases

Considering that we had obtained a list of rotating candidates, the selection methods described above may have biased our sample, for example, by excluding some regular sinusoidal variabilities. On the other hand, the selection may have polluted the final sample with other variables that are not rotators (for instance some semi-regular pulsators) that may show variabilities somewhat similar to the semi-sinusoidal signatures. Of course, because the aim of the methods was to minimize such a sample pollution, a compromise with some bias is unavoidable.

For a general description of our final sample of 4206 stars, Fig. 7 shows their spectral type and luminosity class distributions, while Fig. 8 depicts the variability amplitudes in mag and the periods of the corresponding 4206 LCs. Thus, most of the stars in our sample exhibit variability amplitudes lower than 0.05 mag, within a range compatible with rotational modulation. However, other types of variabilities may also be found within this amplitude range. The period distribution (Fig. 8, right panel) may include physical aspects, but they mostly denote biases. This can be explained by at least two facts. First, the limited time span of CoRoT LCs of up to ∼150 days makes it more difficult to identify the criteria described in Sect. 2.2.3 the longer
the periods are. Second, the higher the frequency, the lower the number of flux measurements cycle by cycle in an LC, which also complicates identifying those criteria. Therefore, in view of the full-width at half-maximum (FWHM) of the period distribution, the best selection of semi-sinusoidal variabilities in our sample lies around 3–20 days.

Figure 9 shows the color–magnitude diagram by comparing the CoRoT parent sample of 124471 LCs (in black) with our final sample of 4206 stars (in red). This comparison indicates that some more biases were introduced by our selection procedure. Essentially, there is a cut-off region for stars fainter than ∼14 mag for \((J - H) \lesssim 0.8\) and ∼11 mag for \((J - H) \geq 0.8\), caused by the S/N selection described in Sect. 2.2.1. We verified that the distribution of the parent sample of 124471 LCs is quite similar to a random selection of 2MASS sources. This means that our sample is valid for relatively bright field stars. Considering all biases, we are aware that our final sample is not statistically complete. Nevertheless, it is large enough for a robust global analysis if one is cautious to interpret how the biases may affect the physical results. For example, restricting the sample to that region of the color–magnitude diagram may be advantageous, because it produces a stronger relation of the color with the stellar evolutionary stage (see Sect. 3).

2.4. Our sample selection versus automatic classifiers

It could be suggested that the sample provided in Debosscher et al. (2007, 2009) contains all parameters needed for the results presented in Sect. 3. However, their sample was obtained from a fully automatic classifier, which is useful for the preliminary selection of a large sample of LCs, but may present a number of problems, particularly for CoRoT LCs, as detailed below. The importance of our sample compared with that of Debosscher et al. (2007, 2009) for the study of CoRoT targets is justified below.

As mentioned in the introduction, automatic classifiers are subject to misclassifications as a result of data artifacts. Discontinuities found in CoRoT LCs may be interpreted as variabilities by producing an incorrect calculation of periods and the statistical measurements used in the classifiers. For example, Fig. 10 (top) shows a CoRoT LC classified by Debosscher et al. (2007, 2009) and Sarro et al. (2009) as displaying rotational modulation, with a period of 6.15 days, Mahalanobis distance of 1.36, and class probability of 98.8%. Although these values are typical of a good classification, visual inspection clearly shows this is a fake period caused by strong discontinuities.

Incorrect classifications as that in Fig. 10 (top) may contaminate a sample of stars and therefore hamper the identification of physical results. For instance, consider a subsample of FGKM stars classified in Debosscher et al. (2007, 2009) and Sarro et al. (2009) as having rotational modulation, to be compared with our results (see Sect. 3.1). Only the best classifications with Mahalanobis distances smaller than 1.5 and class probabilities higher than 90% were considered in that subsample, but it shows no clear behavior in a color–period diagram, as seen in the bottom panel of Fig. 10. This is not the case of our sample, which is a source of relevant physical results to be demonstrated in this study. Visual inspection was crucial in minimizing misclassifications in our sample.

Finally, all CoRot targets classified in Debosscher et al. (2007, 2009) as possible rotating variables were visually inspected by us, independently of their Mahalanobis distances and class probabilities. Only ∼4% of those targets show semi-sinusoidal signatures as defined in Sect. 2.2.3. Therefore only this fraction was included in our sample. These targets correspond to ∼60% of our whole sample, the remaining targets was not classified in Debosscher et al. (2007, 2009) as possible rotating variables. Therefore, our sample has a substantial number of additional candidates for the study of
stellar rotation. We emphasize that the 96% of the LCs classified in Debosscher et al. (2007, 2009) that are not included in our sample are not necessarily misclassifications. The large fraction of rejected LCs, in the context of the present study, indicates that these LCs do not fulfill the main criteria adopted by our selection procedure for a semi-sinusoidal signature.

2.5. Comparison with period measurements available in literature

The literature offers now a substantial list of 1978 period measurements computed from CoRoT LCs, 1727 of which interpreted by their authors as rotation periods (Affer et al. 2012). From this sample, 216 targets are in common with our sample, which offers the possibility for a preliminary comparison between the two sets of measurements. Figure 11 displays our period estimates versus those obtained by Affer et al. (2012).

Of the 216 targets in common, the periods agree excellently for about 95%. For the main discrepant cases, indicated in Fig. 11, the following aspects should be outlined: case A has two types of variabilities superposed, one of which was selected by us and the other by Affer et al. (2012). The variability selected by us is compatible with a semi-sinusoidal signature, while the latter has an amplitude approximately constant over time, which is more often observed in eclipsing binary LCs (see Sect. 2.2.3). For the other cases (B to I), our periods match the semi-sinusoidal signature, while the periods given by those authors correspond to long-term contributions not compatible with such a signature. A relevant aspect of this comparative analysis is that, except for case A, the disagreement is associated with the long period measurements computed by Affer et al. (2012).

2.6. Influence of reddening

We used the $(J−H)$ color index obtained from the 2MASS photometry in our analysis, which may be affected by reddening. To determine its effect on our results, we computed pseudo-colors as described in Catelan et al. (2011), which are supposed to be reddening-free. For instance, these authors considered the data collected with a set of five different broadband filters of the Vista Variables in the Vía Láctea (VVV) ESO Public Survey to estimate reddening-free indices, which can be calculated by calibrating magnitudes or colors. Accordingly, we used 2MASS magnitudes and Eq. (7) in Catelan et al. (2011) to determine the pseudo-colors. As demonstrated below, a reddening correction can dramatically affect the behavior in the period-versus-color distribution for our stellar sample.

3. Results and discussion

The aim of this pioneering investigation is to identify and quantify the level of semi-sinusoidal variability in stellar LCs produced by the CoRoT space mission. To that end, we dedicated most of our effort to identifying through visual inspection the LCs without ambiguities in their semi-sinusoidal behavior. As a result, 4206 periods of variability for stars of spectral types F, G, K and M are now available. This section presents some statistics and characteristics of the periods obtained, in particular as a function of colors.

3.1. General description of the variability behaviors

Figure 12 (top panel) shows the variability amplitude as a function of variability period for the final sample of 4206 selected stars. One also observes a slight trend of finding higher amplitudes at longer periods. Whether this has a physical contribution or is only caused by biases is not clear. In principle, the longer the period, the more difficult it is to detect a faint signal because of the fewer observed cycles—which could produce a bias in the amplitude. However, the cut-off by S/N based on Fig. 2 should have reduced this possible bias. On the other hand, the observed behavior for the period dependence on amplitude in Fig. 12 may also be affected by a color bias (see Fig. 9). Despite these biases, this behavior may still have a physical influence, as we discuss in the beginning of the next section.

Figure 12 (bottom panel) depicts variability amplitude as a function of the variability period for a subsample of main-sequence stars (selected from CoRoT luminosity class). According to Basri et al. (2011), main-sequence stars are expected to be more active for shorter periods and may be more obviously periodic or display larger variation amplitudes (unless activity was too uniformly distributed). Therefore, one could expect some trend for an amplitude decrease with increasing period for main-sequence rotating variables, which is not observed in our sample. However, according to literature data, this behavior was also absent from a sample of field stars selected as rotating candidates in Hartman et al. (2010) and in a sample of chromospherically active binaries with photometric rotation periods studied in Eker et al. (2008).

By combining period, color index, amplitude, and luminosity class, we achieved a detailed overview of the semi-sinusoidal variabilities for the sample described in Sect. 2. Figure 13 shows the color–period diagram, where the variability period is plotted as a function of color index $(J−H)$. Circle size indicates the variability amplitude in mag and colors correspond to the CoRoT luminosity class. This figure provides outline variabilities in the evolutionary context. From a global perspective, there are at least two important facts in this color–period diagram. First, it shows two distinct stellar populations: one to the left, with $(J−H) \lesssim 0.85$, and another to the right, with $(J−H) \gtrsim 0.85$. Second, these populations tend to show an increase of the period

---

3 More details about those public data are provided in Sect. 3.2.
with increasing color index, each at a different rate. These two distinct populations should be related to different evolutionary stages of the stars. In fact, there is a substantial number of giant stars in the population to the right. On the other hand, stars from classes III, IV, and V are more or less uniformly distributed to the left. This dispersion of luminosity classes may be associated with uncertainties in the parameters of the CoRoTSky database. There are also quite a few stars with low amplitude variability in the region with \((J−H)<0.55\), while those with a color index between 0.55 to 0.9 mostly have higher amplitudes.

Figure 14 shows the color–period diagram, where, instead of \((J−H)\), we used a pseudo-color, \(c_3 = (J−H)−1.47(H−K_S)\), suggested to be reddening-free in Catelan et al. (2011, Sect. 2.4). In this case, the gap that was clearly seen in Fig. 13 is not evident, even though close inspection of the distribution confirms that the \(c_3\) distribution is also bimodal, with the two modes strongly tied to the same two modes that are seen in Fig. 13. There are at least three possible reasons for the less prominent gap in Fig. 14, as compared to Fig. 13. First, the combination of three filters in the case of \(c_3\), as opposed to just two in the case of \((J−H)\), leads to an increase in the propagated errors, and accordingly, the typical error bars in the \(c_3\) values are larger than the corresponding ones in \((J−H)\). Second, theoretical evolutionary tracks reveal that the extension of the Hertzsprung gap is reduced (in mag units) when going from \((J−H)\) to \(c_3\): for instance, for a \(3\ M_\odot\) star, based on evolutionary tracks from the BaSTI database (Pietrinferni et al. 2004), the interval between the main-sequence turnoff point and the base of the RGB amounts to about 0.5 mag in \(J−H\), but only about 0.35 mag in \(c_3\). Third, \((J−H)\) is found to be more tightly correlated with the spectroscopic temperatures from Gazzano et al. (2010) than \(c_3\) – an effect that may also be related to the increased errors that affect the latter quantity.
3.2. Root-cause of the semi-sinusoidal variability in CoRoT LCs

The observed color–period scenario (Figs. 13 and 14) may be reflecting some physical contribution even if there is some bias, following the discussion presented in Sect. 2.3. As observed in Fig. 9, evolved sources are rather selected at higher colors. Thus, such a bias may produce an evolutionary selection that can partially explain the behavior observed in Figs. 13 and 14. For example, if our sample is composed of rotating candidates, it is natural to expect longer periods for higher colors, based on physical reasons related to stellar evolution theories (e.g., Ekström et al. 2012). This may also be the case in Fig. 12 (top panel) for the amplitude.

To check our results, we compared our sample with ~1800 field stars available in the HATNet Pleiades Rotation Period Catalogue described in Hartman et al. (2010). The authors conducted a survey to determine stellar rotation periods in the Pleiades cluster and obtained photometric periods of non-cluster members. The non-cluster members, assumed to be field stars, show variabilities suspected to be rotational modulation, several of which may have other physical natures, as observed in the present study. For these field stars, the period distribution is very similar to those of our sample, but there are several (~35%) sources containing periods between about 0.1 and 1.0 days. In the color–period diagram, field stars observed in Hartman et al. (2010) are reasonably compatible with our sample for the ≥1.0 day period. As in our sample, there is a slight increase in amplitude with the rise in color index.

Furthermore, we analyzed the data available in the catalog of chromospherically active binaries provided in Eker et al. (2008), which contains information on brightness, colors, photometric and spectroscopic data, and physical quantities for 409 field and cluster binary stars. These data provide a basis for determining to what extent our sample exhibits photometric characteristics similar to those of stars with measured rotation periods. Indeed, some binary systems may be impacted by tidal effects; however, the overall statistics of the sample in Eker et al. (2008) can be considered for comparison with our sample. Moreover, our sample may also be composed of non-eclipsing binary systems affected by tidal interactions. The color–period diagram of the Eker et al. (2008) sample shows higher amplitudes for \((J - H) \geq 0.55\), in line with our sample. Nevertheless, the amplitude range of the sample of Eker et al. (2008) has a maximum around 0.05 mag, while the highest amplitudes for our sample occur at around 0.025 mag, possibly because CoRoT was designed to observe fainter sources. In summary, the global behavior of the variabilities in our final sample is compatible in many aspects with that expected for rotating stars, based on the literature. Nevertheless, we should be cautious with the interpretation of the whole list of periods, because physical phenomena other than rotation can produce LCs with semi-sinusoidal behavior, as discussed here.

3.3. Rotating Sun-like candidates?

The rotation period of the Sun ranges from 23 days at the equator to 33.5 days at the poles (e.g., Lanza et al. 2003). Based on solar values of \((J - H)\) color index defined in the literature, we made an additional effort to identify stars in the present sample with \((J - H)\) colors near the solar value that display variability periods close to the Sun’s rotation, namely rotating Sun-like stars. According to recent research, the solar \((J - H)\) index ranges from 0.258 (Holmberg et al. 2006) to 0.355 (Rieke et al. 2008). In addition, Zhengshi et al. (2010) computed \((J - H) = 0.288\), from Valcarce et al. (2012) we estimated \((J - H) = 0.347\), whereas Casagrande et al. (2006) give a list of different estimates of solar color indexes \((J - H)\). Based on these references, we established an average \((J - H) = 0.315 \pm 0.04\). Within the solar rotation period from 23 to 34 days and for color indices \((J - H)\) between 0.275 and 0.355, we identified two stars; however, only one source is a G-type star of luminosity class V, with an amplitude lower than 0.05 mag. Considering a \((J - H)\) range two times wider, from 0.235 to 0.395, results in a total of three stars exhibiting period and amplitude variability, as well as a spectral type and luminosity class close to the Sun. Therefore, one of the by-products of this study is a set of three rotating Sun-like candidates in the context of photometric period, namely the CoRoT IDs 104049149, 104685082, and 105290723. According to Figs. 7 and 8, the identified candidates seem to be compatible with the number of targets available in our sample, considering the distribution of spectral type, color, and variability period.

3.4. Variability of M-type stars

Rotational modulation in M-type stars can be considered possible based on the results reported by Hünsch et al. (2001). The authors examined M-type giant stars and found indications of variability in H-alpha and Ca I 6572, which may be related to chromospheric activity. Our sample contained 96 stars of spectral type MV with amplitude variability ranging between 0.004 and 0.2 mag and 416 stars of spectral type MIII with amplitudes ranging between 0.01 and 0.5 mag. Of course, follow-up is needed to check the nature of these variabilities, but this may be a substantial amount of M-type stars with rotational modulation.

Our sample of M-type stars may also be useful for future studies, based for example on the investigation conducted by Herwig et al. (2003). These authors analyzed the s-process in rotating stars of the asymptotic giant branch (AGB), but data obtained to date are insufficient to understand many aspects of stellar evolution. This is because M-type giant stars – either RGB or AGB – generally do not exhibit significant stellar activity. New results may be obtained if at least a fraction of our M-type stars are confirmed to present rotational modulation.

3.5. Is there a center versus anti-center difference in the behavior of variability period distributions?

Here we obtain some statistics on the distribution of the computed variability period. Figure 15 shows the period distribution for the entire sample of 4206 and for different luminosity classes, with stars segregated according to Galactic region, namely Galactic center and anti-center. The upper panel in Fig. 15, where the three luminosity classes V, IV, and III are combined, shows that both distributions peak at around ten days and decrease rapidly for increasing periods. Nevertheless, for stars located in the Galactic center there is an excess of long periods compared with those in the Galactic anti-center. To determine whether the present data sets for the Galactic center and anti-center are significantly different, we performed a Kolmogorov-Smirnov (KS) test (Press et al. 2009), which calculates the probability that two distributions are derived from the same parent distribution. Figure 16 shows the cumulative functions for the two variability period distributions. The probability of \(1.7 \times 10^{-48}\) obtained by the KS test indicates that the two distributions are in fact not drawn from the same population distribution function. In addition, KS analyses were conducted by comparing stars in the Galactic center and anti-center according
to luminosity class. Probability values shown in Fig. 16 indicate a scenario where the variability period distributions for stars in the referred Galactic regions, when compared by luminosity class, are in fact not derived from the same parent distribution. This result reinforces the scenario observed in Fig. 15, with a clear excess of long periods among stars located in the Galactic center compared with those in the anti-center. Of course one could question whether the difference in these distributions is produced by biases related to the LC time spans. Although long and short runs are found either in the Galactic center or in the anti-center direction, there are several long runs in the center direction with shorter time spans than usual (LRC03–06; see Table 1). This could limit the sample to shorter periods in that region, but longer periods are found in the center direction. Therefore, the difference in the period distributions does not seem to be caused by biases and may have a physical explanation with a similar discussion as in Sect. 3.1. The explanation is possibly related to the fact that more population II stars lie in the Galactic center than in the anti-center direction.

4. Conclusions and future work

This study presents an overview of stellar LCs obtained by CoRoT within a wide range of period, color, and variability amplitude. This is the first time that a homogeneous set of stellar variability measurements, obtained using only one instrument, has been analyzed for a large sample with wide ranges of period, variability amplitude, and color, taking into account the effects of reddening on the results. As such, we were able to demonstrate the global distribution of these parameters in a representation valid for field stars.

A total of 124471 LCs were analyzed, from which we selected a sample of 4206 LCs presenting well-defined semi-sinusoidal signatures. Each LC was treated individually by correcting trends, outliers, and discontinuities. Through Lomb-Scargle periodograms, harmonic fits, and visual inspection, we selected the most likely periods for each variability. Our sample shows periods ranging from $\sim 0.33$ to $\sim 92$ days, and variability amplitudes between $\sim 0.001$ and $\sim 0.5$ mag, for FGKM stars with $(J-H)$ from $\sim 0$ to $1.4$.

The color–period diagrams of the sample indicate several aspects compatible with rotational modulation. The increase in variability amplitude around $(J-H) \approx 0.55$ corroborates studies on rotating variable stars by Eker et al. (2008) and Gilliland et al. (2009). The overall behavior of the increasing period with rising color index is compatible with theoretical predictions of stellar rotation. Results from this investigation were compared with public data for field variables by Hartman et al. (2010). The distribution periods and variability amplitudes reported here are compatible with data in the corresponding color range. In addition to our overall results, we identified a subset of three Sun-like candidates in the context of photometric period and color, which may be of particular interest for future studies. Moreover, we analyzed a subsample of more than 400 M-type giant stars, whose behavior seems compatible with recent studies of rotational modulation. In addition, the distribution of variability periods for the CoRoT targets tends to be different when compared with Galactic center and anti-center directions. Finally, the behavior of the variability period distribution in the period–color diagram appears to substantially depend on reddening correction, which may significantly affect age–period analyses such as that carried out in Affer et al. (2012).

Observations of apparently bright stars generally provide information concerning intrinsically bright stars. The CoRoT
mission makes the important contribution of increasing the sample of intrinsically faint stars and accumulates a large amount of micro-variability data for the sources. This demonstrates the importance of this work for studying the general variability for a significant sample of intrinsically faint field stars. Moreover, this investigation enables future studies of the particular case of stellar rotation.

Although in many respects our results match those expected for rotating stars, photometric data alone are insufficient for identifying the physical nature of the variabilities. Therefore, additional research is necessary to confirm the root-cause of the variabilities. As part of future research, we will combine our database with a set of spectroscopic observations currently under analysis by our team. This will allow a more accurate assessment of the results, particularly with regard to stellar rotation.

Acknowledgements. CoRoT research activities at the Federal University of Rio Grande do Norte are supported by continuous grants of CNPq and FAPERN Brazilian agencies and by the INCT-INEEspaço. I.C.L. acknowledges a Post-Doctoral fellowship of CNPq; C.E.F. Lopes and S.V. acknowledge financial support of the FAPESP agency; E.J.-P. acknowledges a Post-Doctoral fellowship of the CNPq; C. Cortés, J.P.B., S.C.M. and D.B.F. acknowledge graduate fellowships of CAPES agency; E.J.-P. and A.V. acknowledge financial support of the FAPESP agency. G.F.P.M. acknowledges the financial support by CNPq (476909/2006-6 and 474972/2009-7) and FAPERJ (APQ1/26/170.687/2004) grants. M.C. and C.E.F.L. acknowledge support by the Chilean Ministry for the Economy, Development, and Tourism’s Programa Iniciativa Científica Milenio through grant P07-021-F, awarded to The Milky Way Millennium Nucleus; by the BASAL Center for Astrophysics and Associated Technologies (PFB-06); by Proyecto Fondecyt Regular #1110326; and by Proyecto Anillo ACT-86. The authors warmly thank the CoRoT Technical and Manager Staffs for the development, operation, maintenance and success of the mission. This work used the SIMBAD Astronomical Database operated at the CDS, Strasbourg, France.

References

Levenberg, K. 1944, Quarterly of Applied Mathematics, 2, 164

A&A 555, A63 (2013)