

# Using Galactic Cepheids to verify Gaia parallaxes

F. Windmark<sup>1,2</sup>, L. Lindegren<sup>2</sup>, and D. Hobbs<sup>2</sup>

<sup>1</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany  
e-mail: windmark@mpia.de

<sup>2</sup> Lund Observatory, Box 43, 221 00 Lund, Sweden

Received 20 Mars 2011 / Accepted 8 April 2011

## ABSTRACT

**Context.** The Gaia satellite will measure highly accurate absolute parallaxes of hundreds of millions of stars by comparing the parallactic displacements in the two fields of view of the optical instrument. The requirements on the stability of the “basic angle” between the two fields are correspondingly strict, and possible variations (on the microarcsec level) are therefore monitored by an on-board metrology system. Nevertheless, since even very small periodic variations of the basic angle might cause a global offset of the measured parallaxes, it is important to find independent verification methods.

**Aims.** We investigate the potential use of Galactic Cepheids as standard candles for verifying the Gaia parallax zero point.

**Methods.** We simulate the complete population of Galactic Cepheids and their observations by Gaia. Using the simulated data, simultaneous fits are made of the parameters of the period–luminosity relation and a global parallax zero point.

**Results.** The total number of Galactic Cepheids is estimated at about 20 000, of which nearly half could be observed by Gaia. In the most favourable circumstances, including negligible intrinsic scatter and extinction errors, the determined parallax zero point has an uncertainty of 0.2 microarcsec. With more realistic assumptions the uncertainty is several times larger, and the result is very sensitive to errors in the applied extinction corrections.

**Conclusions.** The use of Galactic Cepheids alone will not be sufficient to determine a possible parallax zero-point error to the full potential systematic accuracy of Gaia. The global verification of Gaia parallaxes will most likely depend on a combination of many different methods, including this one.

**Key words.** stars: variables: Cepheids – space vehicles: instruments – parallaxes

## 1. Introduction

The Gaia satellite, due for launch in 2013, will measure the trigonometric parallaxes of roughly a billion objects in the Galaxy and beyond with accuracies reaching  $10 \mu\text{as}$  (microarcsec; Lindegren 2010). This huge improvement over its predecessor Hipparcos, in terms of accuracy, limiting magnitude and number of objects, will revolutionize many areas of stellar and Galactic astrophysics (Perryman et al. 2001). Moreover, it will allow entirely new kinds of investigations that depend on the statistical combination of very large data sets. One such example is the determination of the distance to the Large Magellanic Cloud (LMC). The LMC distance is fundamental for the extragalactic distance scale, and current estimates put it at 50 kpc (or  $20 \mu\text{as}$  parallax) with a relative uncertainty of 5% (Freedman et al. 2001; Schaefer 2008). Gaia should be able to observe some  $N \approx 10^7$  stars brighter than 20th magnitude in the LMC, with a standard error in the individual parallaxes of about  $200 \mu\text{as}$  or better. Potentially, therefore, the mean LMC parallax as estimated from Gaia data could have an accuracy of  $200 N^{-0.5} \approx 0.06 \mu\text{as}$ , equivalent to a relative error in distance of 0.3%, or 0.006 in distance modulus. Needless to say, such a result will be extremely interesting. However, to achieve the  $N^{-0.5}$  improvement for large  $N$  requires (1) that the individual parallax errors are effectively uncorrelated, which may be the case (Holl et al. 2010); and (2) that there is no significant global zero-point error (bias) in the measured parallaxes. This and many similar examples show that even a bias  $< 0.1 \mu\text{as}$  must be considered significant in comparison with the potential capabilities of the mission.

Achieving the desired parallax accuracy requires an exceedingly stable optical instrument in the Gaia satellite. Even extremely small periodic variations in the so-called basic angle between the two fields of view could lead to an undesirable global offset of the measured parallaxes (Mignard 2011). One well-known cause of such variations is the variable heating from solar radiation as the satellite rotates. Although the satellite has been carefully designed to minimize these effects, the resulting variations in the basic angle cannot be completely eliminated. With the use of an on-board laser interferometer they will however be continuously measured and taken into account in the instrument calibration model. However, it is obviously of great importance to be able to verify the resulting parallax zero point by independent, astrophysical means.

A large number of methods are in principle available for astrophysical verification of the Gaia parallaxes. We may distinguish three main classes of methods: (1) a priori knowledge of the parallax, using for example quasars that are so distant that their parallaxes for the present purpose can be considered to be zero. (2) Distances determined by geometric principles not relying on trigonometric parallax. Several of these methods depend on a combination of doppler velocity, time and angular measurements, for example orbital parallaxes for spectroscopic binaries with an astrometric orbit (e.g., Torres et al. 1997), expansion parallaxes for supernova remnants and planetary nebulae (e.g., Trimble 1973; Li et al. 2002), a geometric variant of the Baade-Wesselink method for pulsating stars (e.g., Lane et al. 2002), kinematic distances to globular clusters (e.g., van de Ven et al. 2006; McLaughlin et al. 2006), and rotational parallaxes

for external galaxies (Olling 2007). (3) Distance ratio methods: the ratio of the distances to two or more objects equals the inverse ratio of their parallaxes. This equality is violated in the presence of a parallax zero point error, which can therefore be derived, e.g., from photometric distance ratios established by means of standard candles, provided that extinction effects can be mastered.

Quasars (belonging to class 1 above) are among the more promising candidates. It is estimated that Gaia will observe around 500 000 quasars with individual parallax uncertainties around 250–300  $\mu\text{as}$  due to their faintness in the Gaia  $G$  band (Lindgren et al. 2008). This would lead to an uncertainty in the mean zero-point around 0.4  $\mu\text{as}$ , with a possible additional bias from foreground stars contaminating the sample.

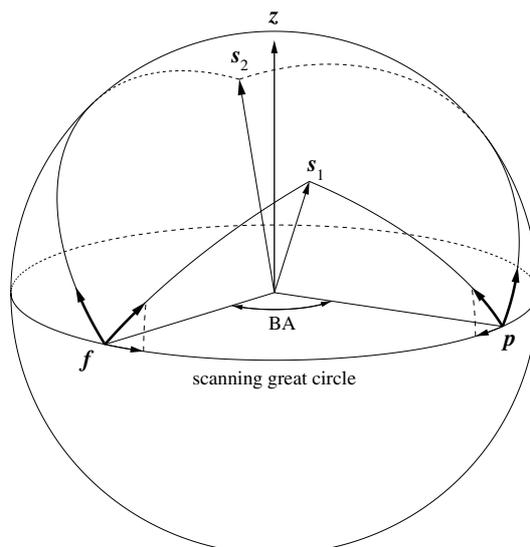
Distance determinations using various geometric principles (class 2 above) will be very important for checking the consistency and reliability of Gaia parallaxes. However, they are not entirely free of model assumptions, and are therefore in many respects problematic as a means of verifying the Gaia data. Moreover, it is doubtful if the number of objects available and the achievable accuracies are high enough for the present purpose.

In this paper we concentrate on the distance ratio method, based on Classical Cepheids as one of the most reliable standard candles. In particular, we focus on the Galactic Cepheids, which are fewer than the quasars and extragalactic Cepheids but with individually more accurate parallaxes. In this method, to avoid a circular argument, we make a simultaneous calibration of the period-luminosity ( $P - L$ ) relation and the parallax zero-point. Since Gaia will observe at least ten times the  $\sim 800$  currently known Galactic Cepheids, we need to create a synthetic population of Galactic Cepheids with the appropriate properties before we can simulate the Gaia observations and make a statistical investigation of the expected errors. Since it is difficult to quantify how much of the final errors will depend on various modelling errors (say, from a possible non-linearity of the  $P - L$  relation), our strategy is to consider first a best-case (i.e., optimistic) scenario, where all such effects are negligible, and then investigate how sensitive the results are to the various assumptions.

The method of absolute parallax measurements with Gaia is closely related to the physical origin of a possible parallax bias, and these aspects of the mission are therefore briefly explained in Sect. 2. Our modelling of the Galactic Cepheids and their observations by Gaia are described in Sect. 3. The results of the simulation and parameter fitting experiments are discussed in Sect. 4, followed by the conclusions in Sect. 5.

## 2. Cause of a possible parallax bias in Gaia

In contrast to ground-based parallaxes, which are always measured relative to background objects, astrometric satellites such as Hipparcos and Gaia in principle allow to determine *absolute* parallaxes thanks to the large difference in parallax factor between the two widely separated fields of view (Lindgren 2005; Lindgren & de Bruijne 2005). However, this capability depends critically on the short-term ( $\sim$ few hours) stability of the so-called basic angle between the two fields of view (van Leeuwen 2005; Mignard 2011). The optical instrument of Gaia is designed to be stable on these time scales to within a few  $\mu\text{as}$ , and an on-board interferometric metrology system, the Basic Angle Monitor (BAM), will moreover measure short-term variations of the basic angle with a precision  $< 1 \mu\text{as}$  every few minutes (Lindgren et al. 2008). The technical design thus guarantees that measurement biases related to basic-angle variations



**Fig. 1.** Illustrating the effect of parallax on stellar images in the preceding ( $p$ ) and following ( $f$ ) field of view of Gaia, when the Sun is in two different positions ( $s_1$  and  $s_2$ ) relative to the stars.  $z$  is the spin axis of Gaia, and the basic angle is marked between the two fields of view. The displacement due to parallax is always directed towards the Sun, as indicated by the thick arrows at  $p$  and  $f$ , and is measured by Gaia as projected along the scanning great circle. See Sect. 2 for further explanation.

are practically negligible in comparison to the random measurement errors ( $\geq 30 \mu\text{as}$  per field-of-view crossing). Nevertheless, when averaging over many stars we should be concerned about systematic effects that are much smaller than the random errors.

The geometry of the observations with respect to the Sun is important both for the determination of parallax and for possible thermal perturbations of the instrument. Indeed, as discussed by van Leeuwen (2005), a certain systematic variation of the basic angle, depending on the satellite spin phase relative to the Sun, has almost the same effect on the measurements as a global zero-point shift of the parallaxes. Thus, if such a variation exists in the real Gaia instrument, and is not recognized by the on-board metrology, then the result will be a global bias of the derived parallaxes. The cause of this shift can be understood by means of Fig. 1. Stellar parallax causes an apparent shift of the star along a great circle towards the Sun. If all stars have a constant positive parallax, the apparent shifts will be as indicated by the short arrows in the figure, depending on the position of the Sun relative to the stars. Thanks to its one-dimensional measurement principle, Gaia is only sensitive to the relative shift along the scanning great circle through  $p$  and  $f$ . Thus, with the Sun at  $s_1$  in the diagram, i.e., closest to the fields of view, the stellar parallax shifts on the detectors will be indistinguishable from a slight enlargement of the basic angle. With the Sun at  $s_2$ , furthest away from the fields, the stellar shifts will be indistinguishable from a slight reduction of the basic angle. Thus, a temporal variation of the basic angle, caused for example by the solar heating, would mimic a global shift of the parallaxes.

According to the technical specifications of Gaia, the expected parallax bias from this effect is at most a few  $\mu\text{as}$ , and most of it will be removed in the data processing by means of the BAM metrology data. Indeed, the BAM specifications are such that, theoretically, the remaining parallax bias should be much less than 0.1  $\mu\text{as}$ . Nevertheless, due to the subtle nature of the effect, it is important to verify the parallax zero point by some

independent method. Since the relevant instrumental effects are almost completely correlated with a parallax shift, an independent verification must be based on astrophysical considerations.

### 3. Modelling the Cepheids observed by Gaia

We create a synthetic population of Galactic Cepheids with each object assigned a number of properties generated from the relevant probability distribution models. In order to study the Cepheid  $P - L$  relation, each object is given a period ( $P$ ) and an absolute visual magnitude ( $M_V$ ), as well as a  $V - I$  colour as required for the transformation to the Gaia wide-band  $G$  magnitude. Finally, each Cepheid is given a position in the Galaxy from a given distribution model. We base our modelling on the observed distribution of the 455 Cepheids in the [Berdnikov et al. \(2000\)](#) catalogue.

#### 3.1. Cepheid properties

It is now over 100 years since the discovery of the  $P - L$  relation ([Leavitt 1908](#); [Leavitt & Pickering 1912](#)). It has since played an important part in the extragalactic distance ladder, and its calibration is still of great interest today ([Sandage et al. 2004](#); [Fouqué et al. 2007](#); [Ngeow et al. 2009](#)). Because of the difficulties involved with observations in the Galactic plane and the larger number of Cepheids known in the Magellanic Clouds, the majority of the calibrations have used Cepheids in the LMC and SMC at least to determine the slope of the  $P - L$  relation. The calibrations show a discrepancy between the  $P - L$  relations in the different galaxies, which implies that the luminosities also depend on some other property than the pulsation period. With measured metallicities becoming available for an increasing number of Cepheids, recent studies have shown the metallicity to be a likely candidate, although the results are far from conclusive ([Groenewegen 2008](#); [Romaniello et al. 2008](#)). In this work, we generally assume all Cepheids to follow the same standard  $P - L$  relation, disregarding a possible metallicity effect. This gives the best-case scenario for the use of the Galactic Cepheids for the parallax verification. In Sect. 4.2 we briefly consider how a metallicity-dependent  $P - L$  relation would affect the results.

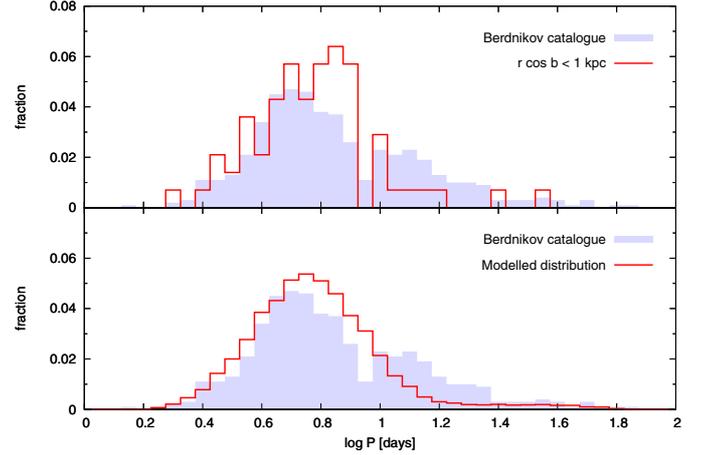
Assuming a period distribution model (see below), the period of each Cepheid is first randomly generated and the true visual magnitude then follows from the assumed  $P - L$  relation. As the  $P - L$  relation will be calibrated simultaneously with the parallax zero-point, the precise  $P - L$  relation adopted for the simulations will affect the end results only minimally. We use the relation from [Sandage et al. \(2004\)](#),

$$M_V = -3.087 \log P - 0.914, \quad (1)$$

where  $P$  is given in days. The period-colour relation by [Tammann et al. \(2003\)](#) is then used in the same way to generate the intrinsic  $V - I$  colour,

$$(V - I)_0 = 0.256 \log P + 0.497. \quad (2)$$

In this process we have neglected the intrinsic dispersion of the  $P - L$  relation due to the finite width of the instability strip in the underlying period-luminosity-colour relation ([Madore & Freedman 1991](#)). From LMC data ([Udalski et al. 1999](#)) the dispersion is found to be about 0.16 mag in  $M_V$  and 0.11 mag in  $M_I$ . The Gaia  $G$  magnitude being intermediate between  $V$  and  $I$  for typical Cepheid colours (cf. Eq. (6)), the dispersion in  $M_G$  is



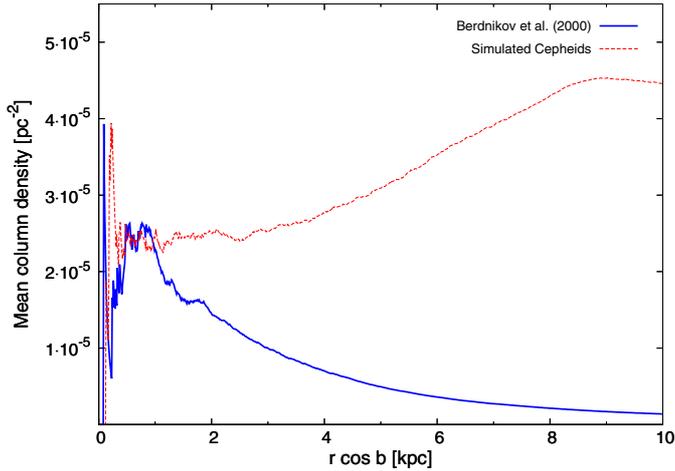
**Fig. 2.** The upper panel shows the period distribution for the Cepheids in the full Berdnikov catalogue (455 Cepheids) as well as the volume-complete sample within the assumed completeness limit of  $r \cos b < 1$  kpc (71 Cepheids). Note the larger fraction of long-period Cepheids in the full Berdnikov catalogue due to selection effects. The line histogram in lower panel shows our modelled period distribution consisting of two overlapping Gaussians, which ideally should represent the volume-complete sample in the upper panel. In the model distribution, 96% of the population has a normal distribution of  $\log P$  with mean value 0.75 and standard deviation 0.18, and 4% has a normal distribution with mean value 1.48 and standard deviation 0.20.

presumably intermediate as well. When using the reddening-free Wesenheit index, [Udalski et al. \(1999\)](#) found a considerably smaller dispersion of 0.076 mag. For our best-case scenario we ignore the dispersion in the  $P - L$  relation, assuming (optimistically) that it can be accounted for by appropriate modelling of the full period-luminosity-colour-(metallicity)-(other factors) relation for the nearby Cepheids, using distances from Gaia. The effects of an intrinsic dispersion are similar to those of an uncertainty in the correction for extinction, which we do however investigate (Sect. 4.1).

For the modelling of the Cepheid periods, we use the [Berdnikov et al. \(2000\)](#) catalogue to obtain a likely period distribution. Because the period is related to the luminosity, and in order to avoid observational biases, we base our model on the 71 Cepheids within the completeness limit discussed in Sect. 3.2. In the upper panel of Fig. 2 we plot the normalized distribution of periods both for the full catalogue of 455 Cepheids (shaded histogram) and for the volume-complete sample of 71 Cepheids (line histogram). We note that the full catalogue contains a larger fraction of long-period (i.e., high-luminosity) Cepheids, as can be expected for a sample that is at least partially limited in apparent magnitude. The full catalogue suggests a bimodal distribution, less evident in the volume-complete sample. In order to reproduce both the strong short-period peak and the long-period tail of the distribution, we fit two overlapping Gaussian functions to the volume-complete sample. The resulting model period distribution is shown as the line histogram in the lower panel of Fig. 2. Although the fit to the volume-complete distribution is far from perfect, the model reproduces the gross distribution reasonably well with a simple continuous function.

#### 3.2. Spatial distribution of Galactic Cepheids

We assume the Cepheids to be distributed axisymmetrically around the Galactic Center, meaning that the number density is only a function of the galactocentric (cylindrical) radius  $R$  and



**Fig. 3.** Mean column density of Cepheids within a projected distance  $r \cos b$  from the Sun. The Berdnikov Cepheids are given in the solid (blue) curve, and the dashed (red) curve represents simulated data using the radial distribution model. Note the plateau of constant density between 0.5 and 1 kpc indicating that the Berdnikov catalogue is complete within 1 kpc from the Sun.

the vertical distance  $z$  to the Galactic plane:

$$N(R, z) = \Sigma(R) \times f(R, z), \quad (3)$$

where  $\Sigma(R)$  is the surface density at distance  $R$  from the axis, and  $f(R, z)$  is the density distribution perpendicular to the disk. Since the Berdnikov et al. (2000) catalogue is only complete within  $\sim 1$  kpc around the Sun, it is necessary to estimate the radial distribution by some other means. Since classical Cepheids are young and massive stars, we can expect them to follow roughly the same distribution as other young and bright stars. McKee & Williams (1997) and Williams & McKee (1997) found the radial distribution of OB associations to be best described by an exponential function  $\Sigma(R) \propto \exp(R/R_0)$  with  $R_0 = 3.5$  kpc, and we choose this also for the Galactic Cepheids. From the galactocentric radius, the position is then generated from  $x = R \sin \theta$  and  $y = R \cos \theta$ , where  $\theta$  is randomly picked between 0 and  $2\pi$ .

The vertical distribution of Cepheids in the Berdnikov catalogue is found to be well fitted by a hyperbolic secant law (van der Kruit 1988),

$$f(R, z) = \frac{1}{\pi z_0} \operatorname{sech}\left(\frac{z}{z_0}\right) \quad (4)$$

where  $z_0(R)$  is the radius-dependent scale height. From the Berdnikov catalogue we obtain for the solar neighbourhood  $z_0(R = 8.0 \text{ kpc}) = 75 \text{ pc}$ . Amôres & Lépine (2005) found a scale height  $z_0 \propto \exp(R/12.5 \text{ kpc})$  for the distribution of Galactic HI and H<sub>2</sub>, and we therefore adopt  $z_0(R) = (40 \text{ pc}) \times \exp(R/12.5 \text{ kpc})$  for the Galactic Cepheids.

The surface density of Cepheids at the Sun's distance from the Galactic Centre,  $\Sigma(R = 8 \text{ kpc})$ , can also be estimated from the Berdnikov et al. (2000) catalogue, if we assume that the Solar neighbourhood is representative and that all Cepheids at high galactic latitudes have been included in the catalogue. In Fig. 3 we plot the mean surface density of the Berdnikov Cepheids within a projected distance  $r \cos b$  from the Sun ( $r$  is the distance from the Sun and  $b$  the Galactic latitude). We note a plateau of roughly constant column density  $\approx 2.5 \times 10^{-5} \text{ pc}^{-2}$  between 0.5 and 1 kpc, before it falls off at larger distances. The plateau is

believed to be real, and arises because of the relatively large radial scale length of the Cepheid number density. The fall-off occurs at the distance where the sample is no longer complete. We therefore conclude that the Berdnikov catalogue is complete to a projected distance of about 1 kpc.

To estimate the total number of Galactic Cepheids, we use the radial distribution described above and keep on generating Cepheids until the surface density of the generated sample agrees with the surface density of the Berdnikov sample within the completeness limit. The discrepancy between the dashed (model) and solid (observed) curves in Fig. 3 represents all the Galactic Cepheids that have yet to be detected. This extrapolation results in a total of about 20 000 Galactic Cepheids. Changing the radial scale length to  $R_0 = 2.5$  kpc, as presented by Binney & Merrifield (1998) for the Galactic thin disk, only changes the total number of Cepheids by 10% up to 22 000. These numbers are slightly larger than the 15 000 estimated by Majaess et al. (2009), who did not take the radial gradient into account. Our numbers are in reasonable agreement with expectations from models of star formation and stellar evolution. E.g., assuming that all stars with masses above  $5 M_\odot$  become Cepheids with a mean lifetime of 2 Myr, a total star formation rate of  $3 M_\odot \text{ yr}^{-1}$ , and the Salpeter (1955) IMF, we expect  $\sim 20$  000 Cepheids in the Galaxy.

### 3.3. Gaia observations

To simulate how the Cepheids are observed by Gaia we need their positions and apparent  $G$  magnitudes as seen from the Sun. The positions are immediately obtained from the simulated galactocentric  $(x, y, z)$  coordinates by subtracting the coordinates of the Sun. The apparent  $V$  magnitude is given by

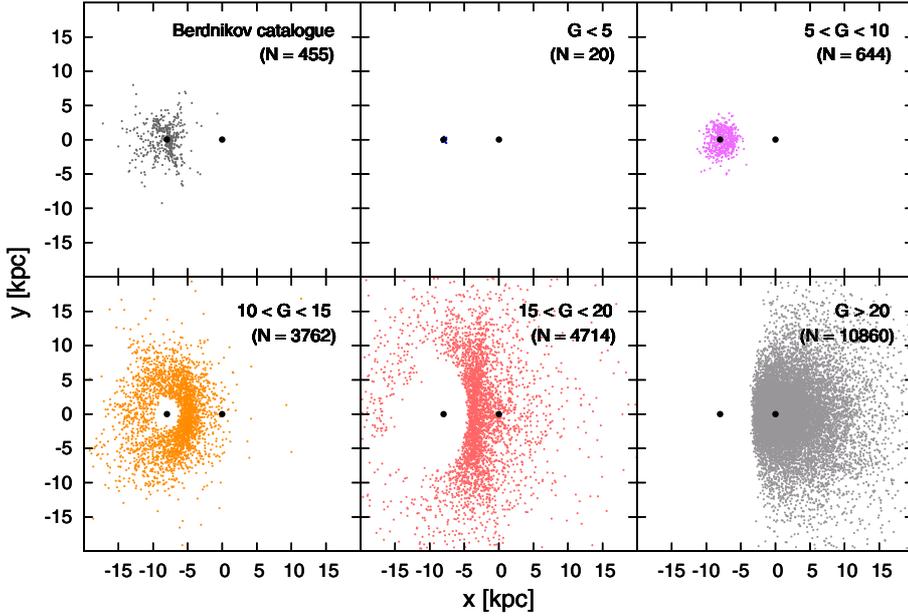
$$V = M_V + 5 \log r - 5 - A_V, \quad (5)$$

where  $r$  is the heliocentric distance in pc and  $A_V$  the total extinction. Amôres & Lépine (2005) describe an axisymmetric three-dimensional Galactic extinction model obtained from observations of the distribution of Galactic HI and H<sub>2</sub> gas. We use this model to obtain both  $A_V$  and the colour excess in  $V - I$ , assuming the ratio  $R_{V-I} \equiv A_V/E_{V-I} = 2.42$  (Cox 2000; Tammann et al. 2003). The Gaia  $G$  band is similar to  $V$  for blue objects but brighter for red objects. The  $V$  magnitude is transformed into  $G$  by the following relation from Lindegren (2010):

$$G = V - 0.017 - 0.088(V - I) - 0.163(V - I)^2 + 0.009(V - I)^3 \quad (6)$$

(cf. Jordi et al. 2010). The synthetic population is then observed according to a standard model developed by the Gaia Data Processing and Analysis Consortium (DPAC; Mignard et al. 2008). Gaia has a bright limit of  $G = 6$ , where the detectors saturate, and a faint limit of  $G = 20$ , corresponding to  $V \approx 20 - 25$ . The Gaia observational model predicts the standard error in the measured parallax,  $\sigma_\varpi$ , for each object in this range, taking into account the standard error per scan across the object, depending on the  $G$  magnitude, and the number and geometry of the scans over the five year mission, depending on the object's position on the sky. The observed parallax ( $\varpi_G$ ) is then obtained by adding a normally distributed random measurement error, with the calculated  $\sigma_\varpi$ , to the true parallax.

In Fig. 4 the Cepheids of the Berdnikov catalogue are compared with the synthetic Cepheid population, divided into five bins in apparent  $G$  magnitude. We estimate that Gaia will observe roughly 9000 Galactic Cepheids, or almost half of the total population. This value is found to be relatively insensitive to



**Fig. 4.** The Galactic Cepheids as observed from the Sun, with the Galaxy seen face on. The Sun is positioned at  $(x, y) = (-8, 0)$  kpc, and the Galactic center at  $(0, 0)$ . The upper left panel shows the Berdnikov sample, and the other five panels show the synthetic Cepheid population divided into different  $G$  magnitude bins. The Cepheids in the lower right panel and those brighter than  $G = 6$  would not be observed by Gaia.

variations in the Cepheid distribution parameters. We note that even though the extinction towards the Galactic Centre is believed to be very large, several hundred Cepheids are still visible near and even behind it. These stars are all found at  $|z|$  of hundreds of pc, where the total extinction is relatively small compared to in the plane. This can also be seen in Fig. 5, where the vertical distribution of Cepheids towards the Galactic Centre ( $|l| < 5^\circ$ ) is plotted. At distances larger than 5 kpc, there are no Cepheids visible in the Galactic plane.

#### 4. Analysis of simulated observations

With the models described in the previous section it is possible to generate a list of observed Galactic Cepheids together with the simulated data (e.g.,  $P$ ,  $G$ ,  $V - I$ ,  $\varpi_G$ ,  $\sigma_\varpi$ ). In this section we describe the tools used to analyse the simulated data and the results of the parameter fitting.

##### 4.1. Parameter fitting

The idea is to use the observed parallaxes of nearby Cepheids to determine the  $P - L$  relation, and to use this  $P - L$  relation for more distant Cepheids to determine the parallax bias. To avoid circularity, we make a simultaneous fit of the  $P - L$  relation and the parallax zero point to all the data. Following Knapp et al. (2003) the fitting is made in parallax space, where the error distribution is symmetric around the true values. Statistically, this is equivalent to the method of reduced parallaxes (Feast 2002), and allows to handle correctly that some observed parallaxes are negative due to measurement errors. Such negative observed parallaxes are statistically valid measurements, but cannot be converted to distances or absolute magnitudes. During the parameter fitting, all measured objects are therefore usable, avoiding a possible bias due to selection (Feast & Catchpole 1997). Rewriting Eq. (5) and inserting the  $P - L$  relation, we get the observation equation:

$$\varpi_G [\mu\text{as}] = 10^{5+0.2[a \log P + b - V + A_V]} + c + \text{noise}, \quad (7)$$

where  $a$  and  $b$  are the slope and zero point of the  $P - L$  relation, respectively.  $c$  represents the global parallax zero point error, and is expected to be zero if Gaia works well. We can safely

assume that Gaia will be able to measure  $P$  and  $V$  with negligible uncertainty. If we assume negligible intrinsic dispersion of the  $P - L$  relation and that  $A_V$  is known exactly, we have the best-case scenario for parallax zero-point verification using Cepheids. Each data point is then weighted entirely depending on its formal parallax uncertainty,  $\sigma_\varpi$ .

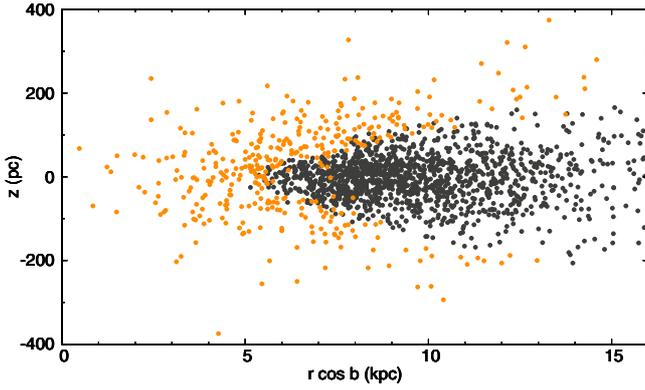
As a more realistic alternative, we introduce some uncertainty in the knowledge of the extinction value  $A_V$  by assuming a constant uncertainty  $\sigma_{A_V} = 0.05$  mag for all objects. This is pessimistic for the bright and nearby, low-extinction Cepheids, but probably optimistic for high-extinction Cepheids. We then generate assumed extinction values that are normally distributed around the true values, and take the total uncertainty in Eq. (7) to be  $\sigma'_\varpi = [\sigma_\varpi^2 + (\varpi \sigma_{A_V} / 2.17)^2]^{0.5}$ . This will lessen the importance of the nearby Cepheids, but might also introduce additional bias effects since the calculated uncertainty uses the observed parallax and not the true one. As seen from Eq. (7), an intrinsic dispersion of the absolute magnitude in the  $P - L$  relation will have the same effect as random errors in the assumed  $A_V$ . We can therefore use these experiments also to conclude on how such a dispersion would affect the results.

To avoid the uncertainties associated with determining the extinction for each Cepheid, we also investigate the use of a reddening-free method equivalent to the use of the Wesenheit function  $W = V - R_{V-I}(V - I)$  (Madore & Freedman 1991, 2009), where  $R_{V-I}$  is the ratio of the total to selective extinction. If the period-colour relation is  $V - I = d \log P + e$ , the observation equation then becomes

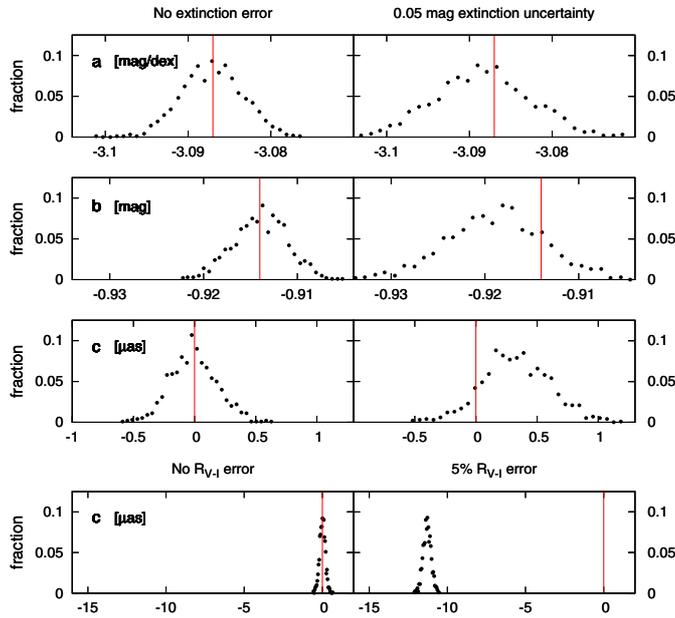
$$\varpi_G [\mu\text{as}] = 10^{5+0.2[k_1 \log P + k_2 - V + R_{V-I}(V - I)]} + c + \text{noise}, \quad (8)$$

where  $k_1 = a - dR_{V-I}$  and  $k_2 = b - eR_{V-I}$ . It is necessary to introduce  $k_1$  and  $k_2$  as the new unknowns since it is not possible to solve simultaneously for all four parameters  $a$ ,  $b$ ,  $d$  and  $e$ . This method requires however that  $R_{V-I}$  is known, and we investigate the effect of assuming a value of  $R_{V-I}$  that is too large by 5%.

Least-squares fitting using the Newton-Raphson iterative method gives the parameters  $a$ ,  $b$  (or  $k_1$  and  $k_2$ ) and  $c$  along with their formal uncertainties arising from the known uncertainties in the measured parallaxes. Biases and the total uncertainties resulting from all modelled effects can be obtained after multiple realisations of the Cepheid data and parameter fitting.



**Fig. 5.** The simulated inner Galaxy ( $|l| < 5^\circ$ ), with projected distance to the Sun plotted versus height above the Galactic plane. The light dots correspond to Cepheids that will be observable by Gaia ( $G < 20$ ) and the dark dots correspond to Cepheids that are too faint to be observed by Gaia ( $G > 20$ ).



**Fig. 6.** Distribution of the fitted parameters in 1000 realisations of the Cepheid data. The upper three panels show the distributions of  $a$ ,  $b$  (in the  $P - L$  relation) and  $c$  (the parallax zero-point) for the model in Eq. (7) when extinction is perfectly known (left) and when it has an uncertainty of 0.05 mag (right). The bottom panels show the distribution of  $c$  for the model in Eq. (8) when  $R_{V-I}$  is perfectly known (left) and when it has an error of 5% (right). The vertical lines indicate the true values.

#### 4.2. Numerical results

In Fig. 6 we present the distributions of the fitted parameters  $a$ ,  $b$  and  $c$  after 1000 realisations of the Cepheid observational data. In the left panels, the correct extinction is assumed during the parameter fitting, meaning that the spread arises only due to the uncertainty in the Gaia parallaxes. This (unrealistic) best-case scenario leads to very well-determined  $P - L$  relation parameters ( $\sigma_a = 0.0035 \text{ mag dex}^{-1}$ ,  $\sigma_b = 0.0027 \text{ mag}$ ) and a parallax zero-point uncertainty or  $\sigma_c = 0.20 \mu\text{as}$ . No bias is observed: the distributions are symmetric around the true parameter values. The bottom left panel shows the distribution of  $c$  when using the reddening-free model of Eq. (8) with the correct value of  $R_{V-I}$ .

Again, the results are unbiased but the parallax zero-point uncertainty is slightly larger,  $\sigma_c = 0.25 \mu\text{as}$ .

In the right panels of Fig. 6 we show how the results are affected by an imperfect knowledge of the extinction, all other factors being the same as in the left panels. The top three diagrams show the results for  $a$ ,  $b$  and  $c$  when the assumed extinction has an uncertainty of  $\sigma_{A_V} = 0.05 \text{ mag}$ . The parallax zero-point uncertainty has increased to  $\sigma_c = 0.27 \mu\text{as}$ , and in addition there is a significant bias in the  $P - L$  relation zero-point ( $b$ ) and a corresponding bias in the parallax zero-point of  $0.3 \mu\text{as}$ . The bottom right diagram shows that the reddening-free model of Eq. (8) is very sensitive to an error in the assumed  $R_{V-I}$ . A 5% error in  $R_{V-I}$  introduces a bias of about  $11 \mu\text{as}$  in  $c$ . To keep the bias below  $0.1 \mu\text{as}$  would require that  $R_{V-I}$  is known to better than 0.04%.

The intrinsic dispersion of the  $P - L$  relation is at least about 0.1 mag (Sect. 3.1), and its expected effect on  $c$  is therefore at least twice as big as the 0.05 mag uncertainty in the extinction, including a likely bias of the order of  $0.6 \mu\text{as}$ . Using the reddening-free method reduces the scatter both due to the extinction and the intrinsic dispersion, but instead the results are then very sensitive to an error in  $R_{V-I}$ , as we have seen.

We have briefly investigated the effects of a metallicity dependent  $P - L$  relation. Groenewegen (2008) found  $M_V = -2.60 \log P - 1.30 + 0.27[\text{Fe}/\text{H}]$  for the Galaxy, where the metallicity dependence has an uncertainty of  $0.30 \text{ mag dex}^{-1}$ . We assumed this metallicity dependence to be the true one, and implemented a radial metallicity gradient of  $[\text{Fe}/\text{H}](R) = 0.42 - 0.052 R$  with an internal scatter of 0.1 dex (Lemasle et al. 2008). Assuming that the metallicity of each Cepheid can be determined with an accuracy of  $\sigma_{[\text{Fe}/\text{H}]} = 0.1 \text{ dex}$  and adding a metallicity parameter in the  $P - L$  relation, the parallax zero-point uncertainty increases to  $\sigma_c = 0.52 \mu\text{as}$  in the case of a perfect knowledge of the extinction. This model also results in an increased bias of  $0.45 \mu\text{as}$  in  $c$ . If metallicity is not accounted for in the  $P - L$  model, the scatter in  $c$  is somewhat reduced but the bias is even larger ( $1.3 \mu\text{as}$ ).

The experiments described above used all the Cepheids observed by Gaia in the least-squares fitting, independent of their individual accuracies and degrees of extinction. As we have seen, the results are very sensitive to extinction errors. It is possible that this sensitivity is a consequence of including many Cepheids with large extinction in the analysis. In order to investigate this we tried various ways of removing the worst data, e.g., by only using Cepheids within certain distance or extinction limits. This gave only a slight improvement in terms of the scatter in  $c$ , but was usually found to introduce additional biases that proved difficult to avoid. One reason for this could be that the selection of “good” data depend on measured quantities, which in the case of noisy data invariably introduces selection biases. We also note that the method requires the observation of both nearby and distant Cepheids in order to separate the  $P - L$  zero point ( $b$ ) from the parallax zero point ( $c$ ); excluding either nearby or distant Cepheids from the analysed sample introduces large statistical uncertainties in both parameters.

## 5. Conclusions

In order to explore the full statistical potential of the Gaia parallaxes it is desirable that the global parallax zero point can be verified to within  $0.1 \mu\text{as}$ . We have explored the possible use of Galactic classical Cepheids for this purpose.

A model of the Galactic Cepheid population has been formulated which allows us to simulate their observation by the Gaia

satellite. From the simulated data, we have made simultaneous fits of the  $P - L$  relation and the Gaia parallax zero point under a variety of assumptions.

We find that the parameters  $a$  and  $b$  of the  $P - L$  relation can be determined with a typical precision better than 0.01 mag dex<sup>-1</sup> or 0.01 mag, respectively, which is far better than current calibrations. The results for the parallax zero point  $c$  are less encouraging. Even under optimal circumstances (accurate knowledge of extinction, metallicity, etc.), the Galactic Cepheid method cannot determine  $c$  better than to within a few tenths of a  $\mu$ as. Moreover, we find that the resulting  $c$  is very sensitive to errors in the extinction correction, or to an error in the  $R_{V-I}$  value if a reddening-free method is used. Attempts to improve the situation, e.g., by limiting the sample to low-extinction Cepheids, were largely unsuccessful due to the introduction of additional biases caused by the selection being made from observed values.

By extrapolating Cepheid statistics from the [Berdnikov et al. \(2000\)](#) catalogue, we estimate the total number of Galactic Cepheids to be  $\sim 20\,000$ . We estimate that Gaia will observe about 9000 of them, which is a factor ten larger than the currently known number. Although many of them are faint, their observation by Gaia will greatly improve our knowledge of the  $P - L$  relation and its dependence of other factors such as metallicity. A detailed global modelling of their characteristics is very worthwhile, and should take into consideration a possible parallax zero point error. However, the ultimate astrophysical verification of Gaia's parallax zero-point is likely to depend on a combination of many different methods including the presented one.

Finally, if we assume that the parallax zero point can be verified to a good accuracy without the use of Cepheids, one could use the Cepheids observed by Gaia to learn more about extinction. With the highly accurate observations that Gaia will provide, this method could yield very precise mapping of the Galactic extinction ( $R$  and  $A_G$ ) with an accuracy that has previously not been achievable.

*Acknowledgements.* We like to thank Berry Holl, Anthony Brown, Timo Prusti and the referee, Floor van Leeuwen, for helpful comments.

## References

- Amôres, E. B., & Lépine, J. R. D. 2005, *AJ*, 130, 659  
 Berdnikov, L. N., Dambis, A. K., & Vozyakova, O. V. 2000, *A&AS*, 143, 211  
 Binney, J., & Merrifield, M. 1998, *Galactic astronomy* (Princeton University Press)  
 Cox, A. N. 2000, *Allen's astrophysical quantities*, ed. A. N. Cox  
 Feast, M. 2002, *MNRAS*, 337, 1035  
 Feast, M. W., & Catchpole, R. M. 1997, *MNRAS*, 286, L1  
 Fouqué, P., Arriagada, P., Storm, J., et al. 2007, *A&A*, 476, 73  
 Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, *ApJ*, 553, 47  
 Groenewegen, M. A. T. 2008, *A&A*, 488, 25  
 Holl, B., Hobbs, D., & Lindgren, L. 2010, in *IAU Symp.*, ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel, 261, 320  
 Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, *A&A*, 523, A48  
 Knapp, G. R., Pourbaix, D., Platais, I., & Jorissen, A. 2003, *A&A*, 403, 993  
 Lane, B. F., Creech-Eakman, M. J., & Nordgren, T. E. 2002, *ApJ*, 573, 330  
 Leavitt, H. S. 1908, *Annals of Harvard College Observatory*, 60, 87  
 Leavitt, H. S., & Pickering, E. C. 1912, *Harvard College Observatory Circular*, 173, 1  
 Lemasle, B., François, P., Piersimoni, A., et al. 2008, *A&A*, 490, 613  
 Li, J., Harrington, J. P., & Borkowski, K. J. 2002, *AJ*, 123, 2676  
 Lindgren, L. 2005, in *The Three-Dimensional Universe with Gaia*, ed. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, ESA SP, 576, 29  
 Lindgren, L. 2010, in *IAU Symp.*, ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel, 261, 296  
 Lindgren, L., Babusiaux, C., Bailer-Jones, C., et al. 2008, in *IAU Symp.*, ed. W. J. Jin, I. Platais, & M. A. C. Perryman, 248, 217  
 Lindgren, L., & de Bruijne, J. H. J. 2005, in *Astrometry in the Age of the Next Generation of Large Telescopes*, ed. P. K. Seidelmann, & A. K. B. Monet, ASP Conf. Ser., 338, 25  
 Madore, B. F., & Freedman, W. L. 1991, *PASP*, 103, 933  
 Madore, B. F., & Freedman, W. L. 2009, *ApJ*, 696, 1498  
 Majaess, D. J., Turner, D. G., & Lane, D. J. 2009, *MNRAS*, 398, 263  
 McKee, C. F., & Williams, J. P. 1997, *ApJ*, 476, 144  
 McLaughlin, D. E., Anderson, J., Meylan, G., et al. 2006, *ApJS*, 166, 249  
 Mignard, F. 2011, *Adv. Space Res.*, 47, 356  
 Mignard, F., Bailer-Jones, C., Bastian, U., et al. 2008, in *IAU Symp.*, ed. W. J. Jin, I. Platais, & M. A. C. Perryman, 248, 224  
 Ngeow, C., Kanbur, S. M., Neilson, H. R., Nanthakumar, A., & Buonaccorsi, J. 2009, *ApJ*, 693, 691  
 Olling, R. P. 2007, *MNRAS*, 378, 1385  
 Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, *A&A*, 369, 339  
 Romaniello, M., Primas, F., Mottini, M., et al. 2008, *A&A*, 488, 731  
 Salpeter, E. E. 1955, *ApJ*, 121, 161  
 Sandage, A., Tammann, G. A., & Reindl, B. 2004, *A&A*, 424, 43  
 Schaefer, B. E. 2008, *AJ*, 135, 112  
 Tammann, G. A., Sandage, A., & Reindl, B. 2003, *A&A*, 404, 423  
 Torres, G., Stefanik, R. P., & Latham, D. W. 1997, *ApJ*, 485, 167  
 Trimble, V. 1973, *PASP*, 85, 579  
 Udalski, A., Szymanski, M., Kubiak, M., et al. 1999, *Acta Astron.*, 49, 201  
 van de Ven, G., van den Bosch, R. C. E., Verolme, E. K., & de Zeeuw, P. T. 2006, *A&A*, 445, 513  
 van der Kruit, P. C. 1988, *A&A*, 192, 117  
 van Leeuwen, F. 2005, *A&A*, 439, 805  
 Williams, J. P., & McKee, C. F. 1997, *ApJ*, 476, 166