

Statistical calibrations of trigonometric parallaxes

T. Tsujimoto¹, Y. Yamada² and N. Gouda¹

¹National Astronomical Observatory, Mitaka-shi, Tokyo 181-8588, Japan

email: taku.tsujimoto@nao.ac.jp, naoteru.gouda@nao.ac.jp

²Department of Physics, Kyoto University, Kyoto 606-8502, Japan

email: yamada@scphys.kyoto-u.ac.jp

Abstract. We examine statistical methods for calibrating trigonometric parallaxes to retrieve the absolute magnitudes of stars, using Monte Carlo simulations. Here we consider the case of the zero-point of the period-luminosity relation for Cepheid variables. The method originally proposed by Ratnatunga & Casertano was revisited by introducing a *realistic* density distribution of sample stars belonging to the catalogue through prior calculations of the photometric distance for each star. It is found that our method gives an unbiased estimate, regardless of any dispersions in their absolute magnitude. We further investigate the reliability of results which depend on the accuracy of parallax. Our finding is that the accuracy (~ 1 mas) of Hipparcos parallaxes is not enough to obtain a reliable result due to a large variation among different ensembles of stars. More precise determination of parallaxes to an accuracy of $200 \mu\text{as}$ at least, which will be easily realized by the ongoing astrometric space satellites, will give a precise zero-point together with a dispersion in absolute magnitude.

1. Introduction

The Hipparcos satellite has brought us a new era for the distance determination using the trigonometric parallaxes, π , of stars. Before Hipparcos, the ground-based observations offered π only for stars within a few tens pc. We have now reached the stage where we can obtain π for stars with their distances extending to ~ 12 kpc. As a result, the Hipparcos catalogue involves many valuable stars, in which Cepheid variables and RR Lyrae stars which serve as primary distance indicators are included.

It is well known that Cepheids obey the period-luminosity (PL) relation. The zero-point for PL relation gives a distance modulus μ_{LMC} to the Large Magellanic Cloud (LMC), which is no doubt an important step for determining the distances to distant galaxies and thus for the Hubble constant (H_0). Instead of the indirect calibrations of the zero-points so far, Hipparcos has enabled us to derive them directly from π of Galactic Cepheids for the first time. However, in fact, direct calibrations have confronted a serious problem which comes from the fact that almost all of π for these stars are measured with very large errors, including negative parallaxes. In such cases, the distance d for each star cannot be calculated from $d(\text{pc}) = 1/\pi$. We therefore need to deduce the physical quantities (e.g., the zero-points) statistically from an ensemble of Hipparcos parallaxes (e.g., Smith 1987, 1988).

To retrieve statistically an *unbiased* estimate is indeed a difficult task (e.g., Smith 2003). Many studies have been done since Roman (1952) and Jung (1971) (see the review by Arenou & Luri 1999). Essentially, two different approaches have been investigated since then, and both were applied to the calibrations with Hipparcos parallaxes. Ratnatunga &

Casertano (1991) have proposed a maximum likelihood method (hereafter, ML method) for correcting the notable Lutz-Kelker bias (Lutz & Kelker 1973), allowing full use of low-accuracy and negative parallaxes. Tsujimoto, Miyamoto & Yoshii (1998) applied their method to Hipparcos RR Lyraes, and found the absolute magnitude $M_V(\text{RR})$ of RR Lyraes at $[\text{Fe}/\text{H}] = -1.6$ is 0.59, which corresponds to $\mu_{\text{LMC}} = 18.41$ mag with the data of LMC RR Lyraes by Walker (1992), given the slope 0.20 of the $M_V(\text{RR}) - [\text{Fe}/\text{H}]$ relation. Arenou et al. (1995) also investigated extensively the statistical properties of the errors of Hipparcos parallaxes by a similar algorithm to the ML method.

Feast & Catchpole (1997) have proposed another method, the so-called reduced parallax method (hereafter, FC method), to estimate the zero-point for the Cepheid PL relation from the weighted mean of the formula free from biases such as the Lutz-Kelker one. Using Hipparcos Cepheids, they found a 0.2 mag brighter zero-point than the previous value (Laney & Stobie 1994), which results in $\mu_{\text{LMC}} = 18.70$ mag. This value gives an upper bound for μ_{LMC} among numerous determinations of the distance to the LMC (Gibson 2000), and it seems a bit too large (e.g., Freedman et al. 2001). However, Pont (1999) concluded this method to be the most rigorous one, and Lanoix, Paturel & Garnier (1999) further confirmed it by Monte Carlo simulations. In any case, even confined to the determinations of μ_{LMC} from Hipparcos parallaxes, there exists a large uncertainty in μ_{LMC} , which is one of the largest remaining uncertainties in the overall error budget for the determination of H_0 .

In this paper, we perform Monte Carlo simulations as done by Lanoix, Paturel & Garnier (1999) to analyze the zero-point for Cepheid PL relation with the ML method. The ML method allows us to incorporate the density distribution of stars into the model arbitrarily. Without knowing the zero-point of the Cepheid luminosity in advance, *relative* approximate distances of individual Cepheids can be deduced from their photometric information assuming an arbitrary zero-point. Then from these relative distances, a realistic density distribution of stars belonging to the catalogue can be obtained. We show that the ML method combined with a density distribution of stars thus obtained leads to an unbiased estimate of the zero-point for the Cepheid PL relation, regardless of any values for the dispersion of the absolute magnitude. In fact, the intrinsic dispersion for Cepheids is estimated to be ~ 0.2 mag (e.g., Ngeow & Kanbur 2004). However, it could possibly be broadened by a large reddening (Luri et al. 1998).

Furthermore, the dependence of the reliability of the results on the accuracy of parallaxes is investigated. We will show how not only the absolute magnitude precision, but also its dispersion precision will improve according to the parallax accuracy. These results are also compared with those obtained by the FC method. The precise determination of biases and variances involved in the final results for the statistical calibrations is certainly demanded for not only the further study using Hipparcos data, but also the future work using the highly-precise astrometric data to be obtained by the ongoing space satellite projects such as Gaia (Perryman 2002) and JASMINE (Japan Astrometry Satellite Mission for INfrared Exploration, Gouda et al. 2003; Gouda et al., these proceedings).

2. The method

2.1. Construction of pseudo catalogues

We make pseudo catalogues based on the Hipparcos data for Cepheids. To investigate the dependence of the end results on the accuracy of parallaxes, we prepare two kinds of pseudo catalogues: One is totally based on the Hipparcos data, which is identical to simulated catalogues made by Lanoix, Paturel & Garnier (1999); another is constructed

with some changes to be done in order to be compatible with forthcoming catalogues with a high accuracy of parallax determination. Here we follow the way of DIVA, which was an astrometric satellite project once planned to be launched by the German Space Agency.

The absolute magnitude of Cepheids is expressed as $M_{V0} = \delta \log P + \rho$. An absolute magnitude M_{V0} given here is the value of the center of distribution with intrinsic magnitude dispersion. We assume that the slope δ of the PL relation is $\delta = -2.77$ (e.g., Madore & Freeman 1991). Using our pseudo catalogues, we will investigate how precisely the zero-point ρ can be determined by the improved ML method and its dependence on the accuracy of parallax. For individual stars in our pseudo catalog, the distance r , the logarithm of the period $\log P$, the absolute magnitudes in the V band M_V and the B band M_B , the apparent B and V magnitudes m_v and m_b , and the observed parallax π and its error σ_π are given by the following procedure.

Tentative stellar distribution

The spatial distribution of stars is first assumed to be uniform and then expressed as $n(r) \propto r^2$ for $0 < r < r_{\max}$. Here we introduce the parameter r_{\max} corresponding to the maximum distance where stars are distributed. We adopt different values of r_{\max} for two kinds of catalogues as shown in Table 1. The final stellar distribution will be changed so as to be compatible with the actual distribution of the observed stars.

Table 1. List of parameters and their input values.

parameter	
r_{\max} (kpc)	2.1/8.4
$\langle \log P \rangle$	0.8554
$\sigma_{\log P}$	0.2865
δ	-2.77
ρ	-1.33
N_{star}	250/15700
N_{try}	100
σ_{m0}	0.21
σ_{obs}	0.005

Distribution of log P

Then, we give the value of $\log P$ to each star, following the distribution of $\log P$ expressed by the log-normal form; $P(\log P) = G(\log P, \langle \log P \rangle, \sigma_{\log P})$, where the notation $G(x, \langle x \rangle, \sigma)$ indicates a Gaussian distribution function of x with an average $\langle x \rangle$ and dispersion σ ;

$$G(x, \langle x \rangle, \sigma) \equiv \frac{1}{\sqrt{2\Pi}\sigma} \exp \left[-\frac{1}{2} \left(\frac{x - \langle x \rangle}{\sigma} \right)^2 \right]. \tag{2.1}$$

Here Π denotes a circular constant. In our analysis, the details of this distribution form do not affect the results. Adopted values of $\langle \log P \rangle$ and $\sigma_{\log P}$ are listed in Table 1, together with other input values of the parameters.

Absolute magnitude

Since the value of $\log P$ is given to each star, we can give individual absolute magnitudes. It is assumed that absolute magnitude M_V in V band has a Gaussian distribution with mean M_{V0} and dispersion σ_{m0} , following a formula $P(M_V) = G(M_V, M_{V0}, \sigma_{m0})$. Using the PL relation, individual absolute magnitudes in V band are expressed as $M_V = \delta_V \log P + \rho_V + \Delta_1$, where Δ_1 is a deviation from M_{V0} . Its mean and dispersion

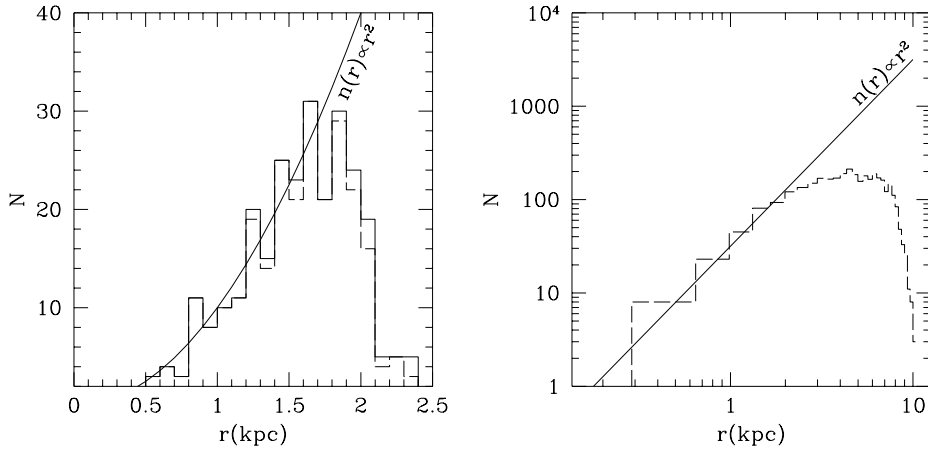


Figure 1. *left panel:* The approximate distribution of Hipparcos Cepheids deduced by giving the photometric distance to each star with the help of the PL relation (solid histogram) together with one in a pseudo catalogue that we made (dashed histogram). The solid line denotes a uniform distribution. *right panel:* The stellar distribution in a pseudo catalogue with an accuracy of parallax $\sim 200 \mu\text{s}$ (dashed histogram) compared with a uniform distribution (solid line).

are $\langle \Delta_1 \rangle = 0$ and $\sigma(\Delta_1) = \sigma_{m0}$. In the same manner, absolute magnitude M_B in B band can be expressed as $M_B = \delta_B \log P + \rho_B + \alpha \Delta_1 + \Delta_2$.

We analyze the observed PL and period-color relations in the LMC (Laney & Stobie 1994), and obtain 0.322, 0.236, and 0.0974 for dispersions of B , V , and $B - V$, which lead to the values; $\sigma(\Delta_1) = 0.236$, $\sigma(\Delta_2) = 0.0521$, and $\alpha = 1.35$. Although the value of $\sigma(\Delta_2)$ is small, it results in a rather large effective intrinsic dispersion as mentioned in the next section.

Apparent magnitude

In order to define apparent magnitude for each star, it is necessary to estimate the extinction A_V . As shown by Lanoix, Patrel & Garnier (1999), the color excesses $E(B - V)$ for Hipparcos Cepheids are uniformly distributed against the distance r in the range of $0.05 < E(B - V)/r < 0.5$. We take such a uniform distribution of $E(B - V)$. The relation between A_V and $E(B - V)$ is then used, i.e., $A_V = R_V E(B - V)$, $A_B = R_B E(B - V)$ with $R_V = 3.3$ and $R_B = 4.3$. Actual apparent magnitudes are given by the formulae: $P(m_V) = G(m_V, M_V + 5 \log r + 10 + A_V, \sigma_{mobs})$ and $P(m_B) = G(m_B, M_B + 5 \log r + 10 + A_B, \sigma_{mobs})$, where σ_{mobs} means observational error, and the value is set to be 0.005.

Parallax and its error

It is assumed that the observed parallaxes π_o are distributed following the formula, $P(\pi_o) = G(\pi, 1/r, \sigma_\pi)$. The observational error of parallaxes, σ_π , depends on an apparent magnitude m_V . For the pseudo catalogues with the Hipparcos accuracy, we give σ_π to each star to reproduce the observed $m_V - \sigma_\pi$ distribution for the Hipparcos Cepheids. For cases with high accuracy of parallaxes, we assume the following formulae for reference, which are in good agreement with the case of DIVA.

$$\sigma_\pi = \begin{cases} \sigma_0 & m_V < 8 \\ \sigma_0 \times 10^{0.0146(V-8)^2 + 0.00036(V-8)^3} & 8 < m_V < 18 \\ 6.6225\sigma_0 10^{0.4(V-15.5)} & m_V > 18 \end{cases} \quad (2.2)$$

Selection of stars – reconstruction of the stellar distribution –

The actual stellar distribution in the catalogue should deviate from an uniform one mainly due to the Malmquist bias. The approximate distribution of Hipparcos Cepheids can be deduced by giving the photometric distance to each star with the help of the Cepheid PL relation. Given $\rho = -1.33$ and $\delta = -2.77$ together with the apparent magnitudes rectified with absorption, we obtain the distribution shown in the left panel of Fig. 1 (solid histogram), which is indeed different from the first assumed $n(r) \propto r^2$ (solid line) at larger distance. According to the derived distribution, we should reconstruct the stellar distribution for the pseudo catalogues, which is realized by the following procedure introducing the Malmquist bias adopted by Lanoix, Paturel & Garnier (1999). To select stars belonging to the final pseudo catalogues, we generate uniform random numbers in the range of $0 \leq t \leq 1$, and calculate t_0 from $t_0 = \frac{1}{1 + \exp[\gamma(V - V_{\text{lim}})]}$. When the value t of the obtained random number is smaller than t_0 , we use a star for a member of a sample, regarding the star as observed. Otherwise, we throw the star away from the sample. In our calculation, $\gamma = 1$ is adopted. In the Hipparcos case, we adopt $V_{\text{lim}} = 12.5$ and $V_{\text{lim}} = 15.5$ in higher accuracy cases. We use the congruent algorithm of 48-bit alignment in order to generate a uniform random number.

The resultant distributions are shown by dashed histograms in Fig. 1 for two kinds of pseudo catalogues. As shown in the left panel of Fig. 1, the final distribution is in good accord with the one for Hipparcos Cepheids (solid histogram).

2.2. *A maximum likelihood method*

Here we analyze our pseudo catalogues by the maximum likelihood method. This method was first proposed by Ratnatunga & Casertano (1991) to obtain the period-color relation for disk dwarf stars, and further applied to the metallicity-luminosity relation for Hipparcos RR Lyrae stars by Tsujimoto, Miyamoto & Yoshii (1998). The merit of this method to be noted is a simultaneous determination of the zero-point of a linear-relation such as the Cepheid PL relation and the intrinsic dispersion of absolute magnitude.

In this method, the values of parameters to be determined are obtained as those which maximize the likelihood function \mathcal{L} , i.e., $\mathcal{L}(\rho, \sigma_m) = \sum^{N_{\text{star}}} \ln [P(\pi_o | \rho, \sigma_m)]$. The probability distribution P can be expressed as

$$P(\pi_o | \rho, \sigma_m) \propto \int_{-\infty}^{\infty} d\Delta_V p(\pi_o | \pi) m(\Delta_V, \sigma_m) \nu(r) 10^{0.6\Delta_V} \tag{2.3}$$

where

$$p(\pi_o | \pi) = \begin{cases} G(\pi_o, \pi, \sigma_\pi) & \pi_o \in (\pi_-, \pi_+) \\ 0 & \text{otherwise} \end{cases}$$

$$m(\Delta_V, \sigma_m) = G(\Delta_V, 0, \sigma_m)$$

$$\Delta_V = \rho + \delta \log P + 10 + 5 \log \frac{1}{\pi} - m_V$$

Δ_V means a difference between the true absolute magnitude and the calibrated absolute magnitude, and $\nu(r)$ is the number density of stars at the position of each star. The range of allowable parallax for the analysis is denoted by (π_-, π_+) . Here we take $\pm\infty$ for these limits. A probability function $P(\pi_o | \rho, \sigma_m)$ is normalized by π_o .

The term $\nu(r) 10^{0.6\Delta_V} d\Delta_V$ corresponds to the number of stars per unit solid angle between r and $r+dr$, and thus equivalent to $\nu(r)r^2 dr$. The number density $\nu(r)$ is assumed to be constant in the original method of Ratnatunga & Casertano (1991). However, as shown in the left panel of Fig. 1, the distribution of Hipparcos Cepheids deviates from a uniform distribution. Furthermore, in the pseudo catalogues with higher accuracies

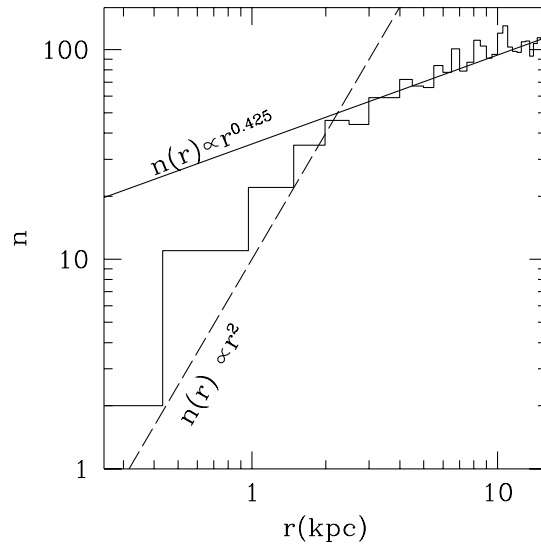


Figure 2. The stellar distribution of one sample in the pseudo catalogues with an accuracy of parallax $\sim 200 \mu\text{as}$. The solid line is obtained by the power-law fitting at larger distance. The dashed line denotes a uniform distribution.

of parallaxes, the deviation from a uniform distribution is significant, as shown in the right panel of Fig. 1. Therefore we introduce a stellar distribution close to the actual one, assuming the power-law form of $\nu(r) \propto r^{-\alpha}$. We perform power-law fitting for each catalogue, and obtain different values of the power index. Fig. 2 shows one sample with $n \propto r^{0.425}$, which gives $\nu(r) \propto r^{-1.575}$ from the relation $n(r) = \nu(r)r^2 = r^{2-\alpha}$. As in this case, we tried to fit the distribution at larger distances such as $r > 2 \text{ kpc}$ because most of stars are resided in this area.

3. Results

For one hundred pseudo catalogues with an accuracy of parallax corresponding to Hipparcos, the resultant distribution of ρ and σ_m is shown in the left figure on the left panel of Fig. 3. The dashed lines denote the input values, i.e., $\rho = -1.32$ and $\sigma_m = 0.2$. It is found that there is no bias for the returned values of ρ , regardless of the returned values of σ_m , though there exists a large scatter in the values of ρ , ranging over $-1.7 < \rho < -1$. This means that the accuracy ($\sim 1 \text{ mas}$) of Hipparcos parallaxes is not enough to obtain a reliable result for the zero-point for Cepheid PL relation. Besides, we cannot expect to obtain the right σ_m . For comparison, the results calculated with the ML method in which a stellar distribution is assumed to be uniform are shown in the right figure on the left panel of Fig. 3. For the cases which return higher values of σ_m than the true (input) value, biases are apparently seen in the returned values of ρ . These biases appear in the opposite sense to the Malmquist bias which yields a bias toward a brighter magnitude. This is caused by an overcorrection of the Malmquist bias for the assumed uniform density distribution of stars.

The results of one hundred pseudo catalogues with an accuracy of parallax corresponding to DIVA, i.e., $200 \mu\text{as}$ are shown on the right panel of Fig. 3. Here we perform calculations for three input values of σ_m , i.e., $\sigma_m = 0.2, 0.5,$ and 0.7 denoted by three

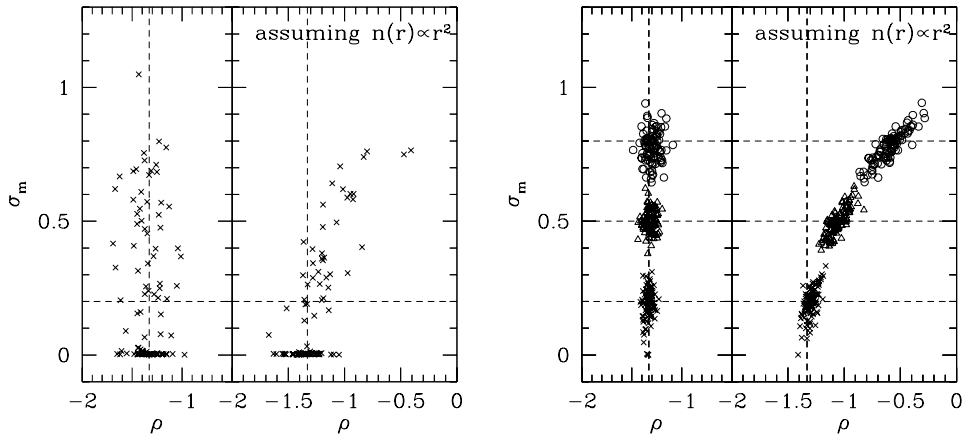


Figure 3. *left panel:* A left figure shows a distribution map of one hundred trial results for the case of the accuracy equivalent to Hipparcos parallaxes. The dashed lines correspond to the input values of ρ and σ_m . A right figure shows the results obtained by the original ML method in which a uniform stellar distribution is assumed. *right panel:* The distribution maps for one hundred pseudo catalogues with an accuracy of parallax $\sim 200 \mu\text{as}$ for three cases of $\sigma_m = 0.2, 0.5$ and 0.7 .

dashed lines. Although the values of $\sigma_m = 0.5$ and 0.7 are much larger than an intrinsic dispersion of M_V , they could possibly be due to the error in the value of the reddening. Furthermore, it should be remarked that a large value of $\sigma_m \sim 0.7$ is suggested by the statistical calibration of parallaxes and proper motions for Hipparcos Cepheids (Luri et al. 1998). For all cases, our modified ML method gives unbiased estimates of ρ with a small scatter (the left figure). However, this is not the case for the ML method with a assumption of a uniform stellar distribution (the right figure). The biases become larger according to larger input σ_m .

The left panel of Fig. 4 clearly demonstrates that the area of the distribution of (ρ, σ_m) becomes smaller according to higher accuracy of parallaxes. Providing that an accuracy of parallax better than $\sim 50 \mu\text{as}$ is reached, we will be able to obtain the precise combination of ρ and σ_m . The right panel of Fig. 4 shows the returned values of ρ with 1σ errors as a function of an accuracy of parallax (crosses with solid bars) together with the results obtained by FC method (open triangles with dashed bars). Interestingly, in the case of the FC method, the bias always remains regardless of the accuracy of parallax, though the application of the FC method to cases with high accuracies such as $10 \mu\text{as}$ is not appropriate. This bias might be ascribed to an effective intrinsic dispersion σ_{eff} of the absolute magnitude. In fact, the FC method gives a reliable value of the zero-point as long as σ_{eff} can be reduced to a small value such as $0 - 0.1$. It was then proposed by Feast & Catchpole (1997) that such a small σ_{eff} could be realized if we use a color-period relation to estimate a reddening, and combine it with the luminosity-color-period relation. However, we obtain

$$\sigma_{\text{eff}} \sim \sqrt{((\alpha - 1)R_V - 1)^2 \sigma(\Delta_1)^2 + R_V^2 \sigma(\Delta_2)^2} \sim 0.176, \quad (3.1)$$

using the values already obtained, $\sigma(\Delta_1) = 0.236, \sigma(\Delta_2) = 0.0521$, and $\alpha = 1.35$. This relatively large σ_{eff} is a result of non-zero $\sigma(\Delta_2)$, while Feast & Catchpole (1997) assumed $\sigma(\Delta_2) = 0$.

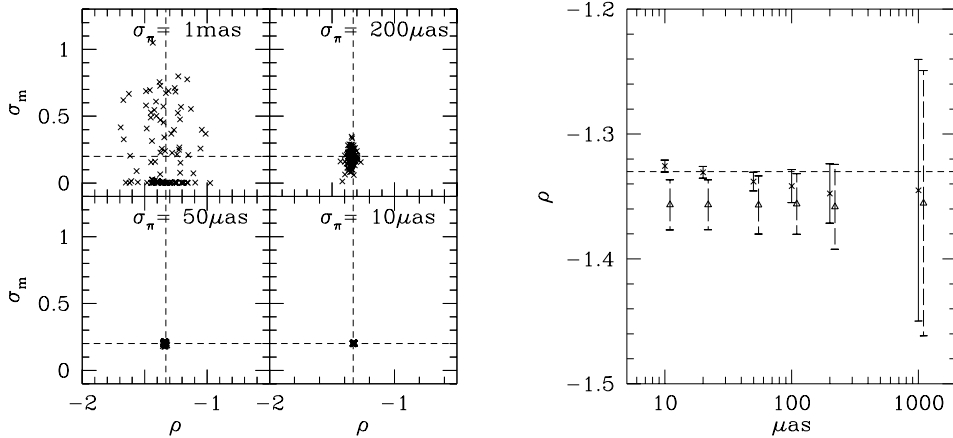


Figure 4. *left panel:* Distribution maps of one hundred trial results in the ρ - σ_m diagram for four cases with accuracies of 1 mas, 200 μ as, 50 μ as, and 10 μ as for parallax. *right panel:* The returned values of ρ as a function of an accuracy of parallax calculated by two methods, i.e., the modified ML method (crosses with solid bars) and the FC method (open triangles with dashed bars).

4. Conclusions

Here we present a maximum likelihood method in which a *realistic* density distribution of sample stars to be analyzed is introduced. Prior information on the stellar density is obtained through calculation of the photometric distance for each star instead of use of the Galactic model as done by Arenou et al. (1995). Our method gives an unbiased estimate of the absolute magnitude of stars, regardless of any dispersions in their absolute magnitude. However, an accuracy of 200 μ as, or better, for parallaxes is required to obtain a reliable result from only one ensemble of stars. If not, as in the case of Hipparcos, we cannot get a definitive conclusion due to a large variation in the estimate. Furthermore, an accuracy more than $\sim 50 \mu$ as promises to give a precise value of the dispersion in absolute magnitude together with its mean value.

Acknowledgements

This work is supported by a Grant-in-Aid for the 21st Century COE Center for Diversity and Universality in Physics.

References

- Arenou, F., Lindgren, L., Froeschlé, M., Gómez, A. Z., Turon, C., Perryman, M. A. C. & Wielen, R. 1995 *A&A* 304, 52–60.
- Arenou, F. & Luri, X. 1999 in *ASP Conf. Ser. 167, Harmonizing cosmic distance scale in a post-Hipparcos era* (ed. D. Egret & A. Heck), pp. 13–32. ASP.
- Feast, M. W. & Catchpole, R. M. 1997 *MNRAS* 286, L1–L5.
- Freedman, W. L., et al. 2001 *ApJ* 553, 47–72.
- Gibson, B. K. 2000 *Mem. Soc. Astron. Ital.* 71, 693–700.
- Gouda, N. et al. 2003 *SPIE*, vol. 4850, pp. 1161–1168.
- Jung, J. 1971 *A&A* 11, 351–358.
- Laney, C. D. & Stobie, R. S. 1994 *MNRAS* 266, 441–454.
- Lanoix, P., Paturel, G. & Garnier, R. 1999 *MNRAS* 308, 969–978.

- Luri, X., Gómez, A. E., Torra, J., Figueras, F. & Mennessier, M. O. 1998 *A&A* 335, L81–L84.
Lutz, T. E. & Kelker, D. H. 1973 *PASP* 85, 573–578.
Madore, B. F. & Freeman, W. 1991 *PASP* 103, 933–957.
Ngeow, C.-C. & Kanbur, S. M. 2004 *MNRAS* 349, 1130–1136.
Perryman, M. A. C. 2002 *Astrophys. and Space Sci.* 280, 1–10.
Pont, F. 1999, in *ASP Conf. Ser.* 167, Harmonizing cosmic distance scale in a post-Hipparcos era (ed. D. Egret & A. Heck), pp. 113–128. ASP.
Ratnatunga, K. U. & Casertano, S. 1991 *AJ* 101, 1075–1088.
Roman, N. 1952 *ApJ* 116, 122–143.
Smith, H. 1987 *A&A* 171, 336–347.
Smith, H. 1988 *A&A* 198, 365–369.
Smith, H. 2003 *MNRAS* 338, 891–902.
Tsujiimoto, T., Miyamoto, M. & Yoshii, Y. 1998 *ApJ* 492, L79–L82.
Walker, A. R. 1992 *ApJ* 390, L81–L84.



Seiji Ueda, Takuji Tsujimoto (centre), Naoteru Gouda, Taihei Yano, Vladimir Elkin