A Bayesian approach to open cluster distance determination

Max Palmer
Universitat de Barcelona

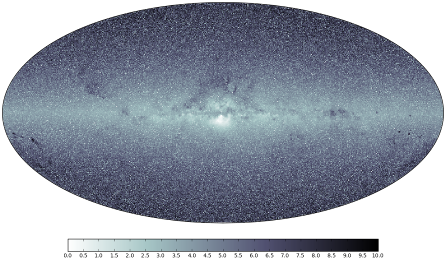
September 6, 2013

INTRODUCTION

PROJECT GOALS

- To improve on current methods for luminosity calibration in preparation for the Gaia era.
- Open clusters have been highly useful in studying luminosity calibration. A tool for precise distance estimation should be created, taking advantage of modern techniques and computing resources.
- Statistical methods should be used in order to:
  - Utilize the huge datasets becoming available in the Gaia era
  - Avoid statistical biases and selection effects

Why Gaia?



- 1 billion star catalogue containing large number of open clusters
- Micro arcsecond parallax and position precision, millimagnitude photometry, proper motions, radial velocities (many more with Gaia-ESO)
- Secondary parameters such as extinction estimates, chemical composition, temperatures, etc.

MAXIMUM LIKELIHOOD ESTIMATION

- Maximum Likelihood Estimation (MLE) has a long history in studies of open clusters. physical mathematical model, given a set of data.
- By constructing a joint probability density function, all the available information can be utilised simultaneously, providing strong constraints on the parameters to be estimated.

MATHEMATICAL FORMULATION

The likelihood function can be defined as:

$$\mathcal{V}(\theta) = \prod_{i=1}^{n} \mathcal{O}(y_i|\theta) \tag{1}$$

where $\mathcal{O}(y_i|\theta)$ is the joint density function.

MODELS

- $(y = m, l, b, \pi, \mu_l, \mu_\delta, v_r)$ is a vector describing the observed properties of each object, and $(y_0 = m_0, l_0, b_0, r_0, \mu_{\alpha 0}, \mu_{\delta 0}, v_{r0})$ is the vector describing the 'true' underlying object properties.
- We can then define the unnormalised density function such that:

$$\mathcal{D}(y_i|\theta) = \mathcal{S}(y_i) \int_{\forall y_0} \varphi_{M_0} \varphi_{rlb} \varphi_v \mathcal{E}(y_i|y_0, \epsilon)\, dy_0 \tag{2}$$

MODELS

Absolute magnitude distribution is Gaussian around isochrone:

$$\varphi_M = e^{-0.5\left(\frac{M - M_{mean}(V-I)}{\sigma_M}\right)^2} \tag{3}$$

Spatial distribution is spherical Gaussian:

$$\varphi_{r'l'b'} = r^2 \cos(b') e^{-\frac{0.5}{\sigma_S^2}\left(R^2 + r^2 - 2rR\cos(b')\cos(l')\right)} \tag{4}$$

Velocity distribution follows velocity ellipsoid:

$$\varphi_v = e^{-0.5\left(\frac{U - U_{mean}}{\sigma_U}\right)^2 - 0.5\left(\frac{V - V_{mean}}{\sigma_V}\right)^2 - 0.5\left(\frac{W - W_{mean}}{\sigma_W}\right)^2} \tag{5}$$

MODELS

Errors are Gaussian around the true value:

$$\mathcal{E}(y|y_0) = \mathcal{E}(\pi|\pi_o)\mathcal{E}(\mu_\alpha|\mu_{\alpha,0})\mathcal{E}(\mu_\delta|\mu_{\delta,0})\mathcal{E}(v_r|v_{r_0})\delta(m_g, l', b') \quad (6)$$

Step function approximates magnitude cut at observation faint limit:

$$\mathcal{S}(y) = \begin{cases} 1, & \text{if } H_p < 12.5. \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Normalisation constant:

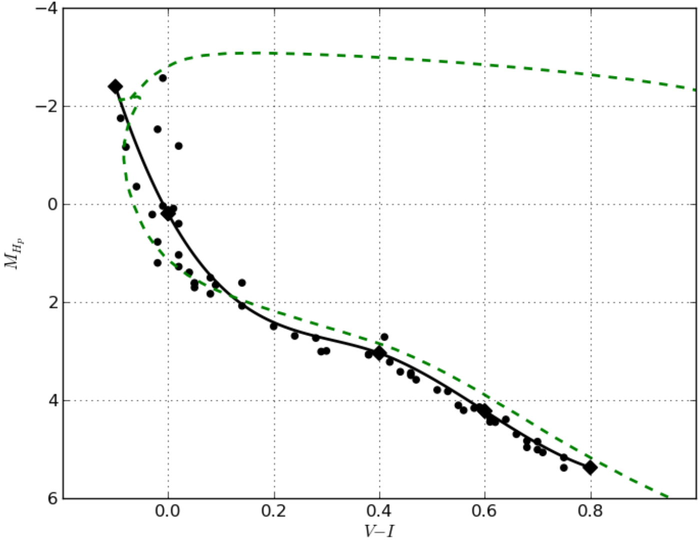
$$\mathcal{C} = \int_{\forall y_0} \int_{\forall y} \varphi_{M_{g0}} \varphi_{\pi_0 l'_0 b'_0} \varphi_v \mathcal{S}(y)\mathcal{E}(y|y_0)\, dy\, dy_0 \quad (8)$$

Pleiades

Parameter	Estimated	Error
Distance (pc)	120.3	1.5
μ_α (arcsec year^{-1})	19.9	0.3
μ_δ (arcsec year^{-1})	-45.3	0.3
σ_{μ_α} (arcsec year^{-1})	1.7	0.3
σ_{μ_δ} (arcsec year^{-1})	1.5	0.2

Table: Colour independent results obtained from the method applied to the new Hipparcos reduction.

PLEIADES



Pleiades

	$(-0.1 < V - I < 0.0)$		$(0.0 < V - I < 0.4)$		$(0.4 < V - I < 0.6)$		$(0.6 < V - I < 0.8)$	
Parameter	Estimated	Error	Estimated	Error	Estimated	Error	Estimated	Error
σ_R (pc)	3.4	1.0	4.9	0.9	10.3	2.9	13.1	3.8
$\sigma_M v$	1.6	0.5	0.45	0.06	0.22	0.06	0.17	0.06
A (start point)	-2.4	0.7	0.2	0.2	3.0	0.1	4.2	0.1
B (end point)	0.2	0.2	3.0	0.1	4.2	0.1	5.4	0.3

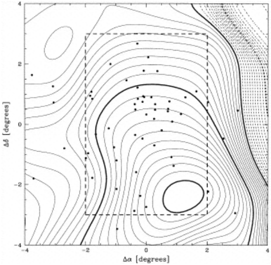
Table: Colour dependent results obtained from the method applied to the new Hipparcos reduction. In the four bins there are 9, 21, 12 and 12 stars (low $(V - I)$ to high).

Pleiades problem

- After the release of the original Hipparcos catalogue, a large discrepancy was noticed between:
  - Hipparcos $\approx$118pc
  - Everybody else $\approx$130pc
- E.g. people working on stellar evolution models say the stars should be brighter and further away.
- Many papers published on one side or the other, without a definitive conclusion.

Correlations

- Narayanan and Gould took the residual $\pi_{HIP} - \pi_{pm}$ from the original Hipparcos catalogue.
- The structure in their contour plot (left) was blamed on correlated errors



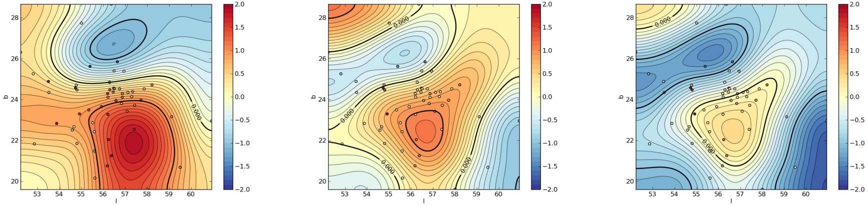


Figure: *Left*: Original, *Middle*: New Catalogue, *Right*: Correct zero point

Hyades

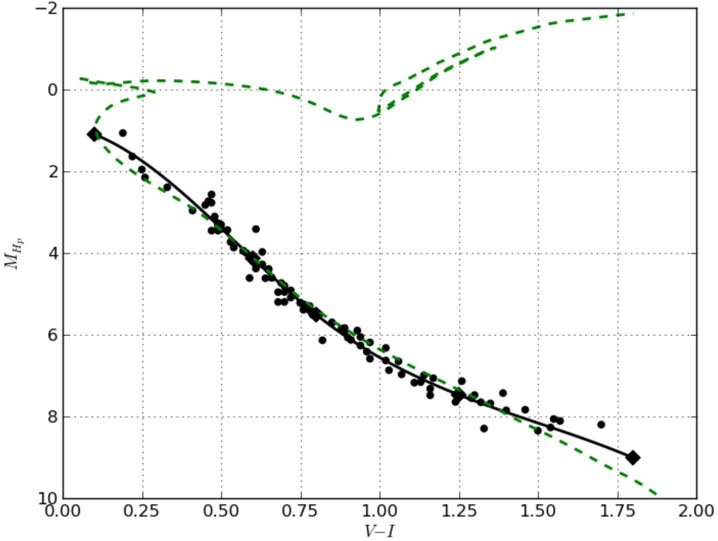
Parameter	Estimated	Error
Distance (pc)	42.8	1.5
U (kms^{-1})	-42.50	0.50
V (kms^{-1})	-19.41	0.51
W (kms^{-1})	-1.09	0.50
σ_{UVW} (kms^{-1})	4.86	0.21

Table: Colour independent results obtained from the method applied to Hyades stars in the new Hipparcos reduction.

	$(0.1 < V - I < 0.6)$		$(0.6 < V - I < 0.8)$		$(0.8 < V - I < 1.25)$		$(1.25 < V - I < 1.8)$	
Parameter	Estimated	Error	Estimated	Error	Estimated	Error	Estimated	Error
σ_R (pc)	10.4	1.3	15.9	1.9	10.0	1.3	8.2	1.8
$\sigma_{M}v$	0.24	0.03	0.26	0.04	0.23	0.04	0.33	0.12
A (start point)	1.10	0.27	4.13	0.04	5.51	0.05	7.47	0.08
B (end point)	4.13	0.04	5.51	0.05	7.47	0.08	9.01	0.61

Table: Colour dependent results obtained from the method applied to the new Hipparcos reduction for the Hyades open cluster.In the four bins there are 24, 27, 28 and 14 stars (low $(V - I)$ to high).

HYADES



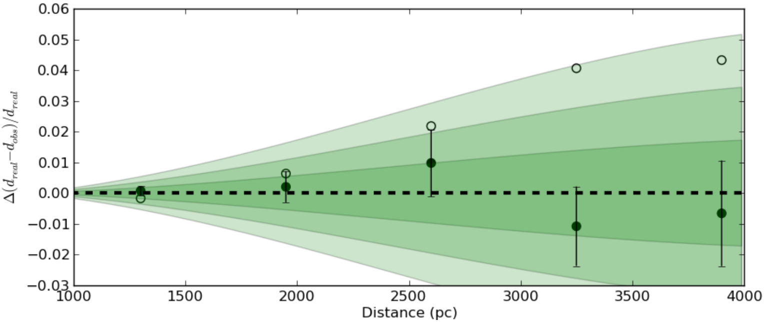
Pleiades with Gaia

- A realistic Pleiades-like open cluster has been simulated in GaiaSimu using a Chabrier/Salpeter IMF and Padova isochrones.
- The data has been processed with the Gaia Object Generator to produce simulated Gaia observations.
- The catalogue contains over 1000 members and due to its close proximity has very small relative parallax errors.

Pleiades with Gaia

- No significant improvement can be made in the mean distance for very close clusters by using this method over the mean of the parallax.
- Individual members parallaxes can be improved through posteriori estimates.
- Fitting of structure using only parallax information results in large bias and distortion of the cluster.

DISTANT CLUSTERS WITH GAIA



Conclusions

- Created an improved method for open cluster distance determination including parametrisation of:
  - Spatial structure (and mass segregation)
  - Kinematics in three dimensions (including internal motions)
  - Fitting a model free isochrone like sequence
- Implemented the derived method and tested with two well studied open clusters, including:
  - Finding evidence for mass segregation in the Plieades
  - Re-checking and refuting a claim of error correlation problems in Hipparcos
- Tested with simulations the applicability of the method to Gaia data, with promising results.

Future work

- Improvement of models (e.g. King spatial distribution instead of exponential)
- Extension to other object types:
  - An extended stay in Bologna is planned for early 2014 in order to apply the same methods to Cepheid and RR-Lyrae stars.

Conclusions

Astronomy & Astrophysics manuscript no. 3
August 5, 2013

A Bayesian approach to open cluster distance determination

M. Palmer[1], F. Arenou[2], X. Luri[1], E. Masana[1]

[1]Departament dAstronomia i Meteorologia, Universitat de Barcelona, Avda. Diagonal 647, E-08028, Barcelona, Spain.
[2]Observatoire de Paris-Meudon, D.A.S.G.A.L., URA CNRS 335, F-92195 Meudon CEDEX, France

Preprint online version: August 5, 2013

ABSTRACT

Aims. An improved method for estimating the distance to open clusters is presented, and applied to Hipparcos data for the Pleiades and the Hyades. The method is applied in the context of the historic Pleiades distance problem, with a discussion of previously made criticisms of Hipparcos parallaxes. This is followed by an outlook for Gaia, where the improved method could be especially useful.
Methods. The method, based on Maximum Likelihood Estimation, combines parallax, position, apparent magnitude, colour, proper motion and radial velocity information to estimate the parameters describing an open cluster precisely and without bias.
Results. We find the distance to the Pleiades to be 120.3 ± 1.5pc, in accordance with previously published work by F. van Leewen using the same dataset. We find that error correlations can not be responsible for the still present discrepancy between Hipparcos and photometric based methods. Additionally, the three dimensional space velocity and physical structure of Pleiades is determined, where we find strong evidence for mass segregation. The distance to the Hyades is found to be $42.8 \pm .1.5$pc. Through the use of simulations, we confirm that the method is unbiased, and will be useful for accurate open cluster parameter estimation with Gaia at distances up to several thousand parsec.

Key words. Keywords should be given

1. Introduction

Open clusters have long been used as a testing ground for a large number of astronomical theories. Determining the distances to nearby open clusters is critical, as they have historically formed the first step in the calibration of the

shorter than previous works based on photometric methods such as (Pinsonneault et al. (1998), Robichon et al. (1999b), Stello & Nissen (2001)) (see section 6).

With the launch of Gaia, the ESA's second major as-