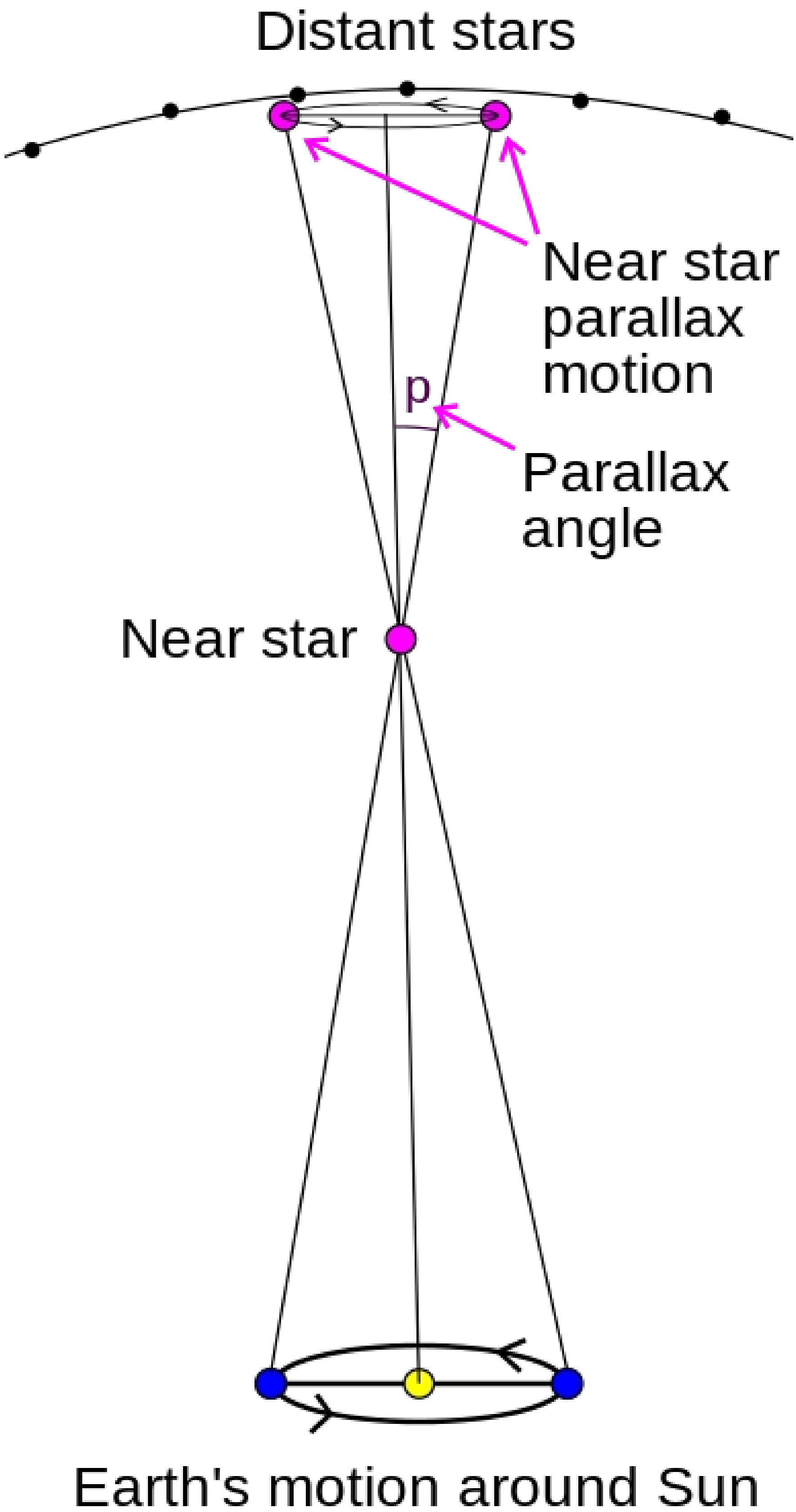
PROJECT:
Monte Carlo simulations of biases associated to absolute magnitude calibrations using parallax measurements

Background:

A purely geometric method to determine a star distance is the so-called trigonometric parallax,
that depends on the apparent motion of nearby stars compared to more distant stars, using observations made six months apart.

A practical demonstration of this is to hold a finger up in front of your face and look at it with your left eye closed and then your right eye. The position of the finger will appear to move compared to more distant objects.

A nearby star viewed from two different positions will appear to move with respect to more distant background objects, due to the change of position of the earth along its orbit. This 'movement' is called parallax.
A six-month baseline of the observations ensures that the star is observed when the earth has covered the whole diameter of its orbit.



If the parallax angle (p - usually very small due to the large distances of stars) is measured in arcseconds (arcsec), the distance to the star (R) in parsecs (1pc=3.26 light years) is given by:

$R=1/p$

Just one star (the Sun) has a parallax greater than 1 arcsec as seen from the Earth. All other stars are at distances greater than 1 pc and parallax angles less than 1 arcsec.

Suppose now you want to determine the absolute magnitude of a class of stars that is expected to have all the same intrinsic brightness.
You first measure the parallax of a sample of these stars distributed over a certain range of distances. From each parallax you get the individual distances, hence their individual absolute magnitudes (assume extinction is negligible or well known for each star).
In principle, if there weren't measurement errors, these 'derived' absolute magnitudes would all have the same value, that provides the sought calibration.

However, parallax measurements have an error.
Each derived distance will have therefore an associated error, hence the derived absolute magnitudes won't be all the same.
Instinct will probably tell us to get the mean of these absolute magnitudes as the best estimate of the unknown true value.

If you do this, there is a very good chance that the derived calibration is biased compared to the real value of the unknown absolute magnitude.

What is the reason for this bias? Can we predict quantitatively (and correct for) this bias?
This is the topic of this project. Monte Carlo simulations are the perfect tool to study the problem and provide solutions.

The MC simulation

The fortran77 code LKbiasMSc.f provides the tool.
This code calculates first a synthetic sample of stars with a given 'true' absolute magnitude M (for example in the optical V band) and distribution of 'true' distances R drawn randomly according to given law.
To each synthetic star at a true distance R the code assigns an apparent magnitude and a true parallax $p=1/R$.

To simulate the observations, each true parallax is perturbed by a Gaussian random error with a given σ, hence providing the observed parallax p_o, that will be different from the true one due to the measurement error. This parallax error σ can be chosen to be either constant or to depend on the apparent magnitude of the synthetic star (the option GAIA-like error below, that mimics the error law of the parallaxes measured by the GAIA astrometric satellite, and is a complicated function of the apparent magnitude of the synthetic star).

From each value of p_o the code determines the retrieved distance R_o that, again, will be different from the true distance R due to the parallax measurement errors.
From R_o the code finally determines the retrieved absolute magnitude of each synthetic star M_o, that is then perturbed by a Gaussian photometric observational error σ_{phot}.
All these quantities are printed in an output file (LK.dat), to be used for the interpretation of the results.

The project

After compilation (you will need a fortran77 compiler. The code is very simple, any compiler shouldn't encounter any problem) , you can start running the code.
The following inputs are required from keyboard

i) Select number of stars (1-500000)

This is the size of the 'synthetic' sample of stars for which you will have 'synthetic' parallaxes

ii) Select space distribution (exponent n of R^n distribution with R=distance between 1 pc and 8 Kpc)
n=2 --> constant spatial density (number /unit volume) of stars
n=1 --> spatial density of stars decreasing as R^{-1}
n=0 --> spatial density of stars decreasing as R^{-2}

iii) Select type parallax error
GAIA-like error law =0 - Constant value=1

If selection 1 (Constant value) is selected the next query is

iiib) Select constant parallax error (arcsec)

In both cases the parallax error is assumed to be Gaussian
With option iiib) you prescribe the σ constant value of this error.

iv) Select true mean absolute magnitude of stars (mag)

v) Select distribution true magnitudes
Gaussian=0 - Uniform=1

These two options consider either a Gaussian or a uniform distribution of the spread of the intrinsic absolute magnitudes around the mean value selected before

vi) Select σ photometric error (mag)
on the measured apparent magnitude

The photometric error is assumed to be Gaussian

vi) Do you want to include an apparent magnitude limit?
No=0 - Yes=1

Typically observations have a magnitude limit because of the finite power of the observing instrument.

Results are written in the ASCII file LK.dat

What appears in this file is the following:

Apparent magnitude limit= xx

Results simulation
(1) true parallax (arcsec)
(2) true distance (pc)
(3) apparent magnitude
(4) observed parallax (arcsec)
(5) parallax error (arcsec)
(6) retrieved distance from observed parallax (pc)
(7) true absolute magnitude
(8) retrieved absolute magnitude
(99.0 for negative derived distances)
(9) photometric error (mag)

Negative derived distances appear when the measured parallax is negative, due to the fact that the parallax measurement error is larger than the parallax itself.
Do not consider stars with negative distances (hence parallaxes) in the analyses below.

The tasks of this project are listed below:

TASK 1

Consider 400000 stars distributed with a constant spatial density distribution (case n=2).
Use a true absolute magnitude equal to 0.6 with a Gaussian dispersion very small, 0.001 mag.
Employ a constant parallax error of 50 microarcsec (0.00005)
Use a small σ photometric error, 0.01 mag.
No magnitude limit.

In the analysis consider the stars with apparent distances within 4Kpc.

Now the analysis.

i) Consider all stars (within 4Kpc) and determine the mean value (and dispersion) of the retrieved absolute magnitude. Compare with the true absolute magnitude. Are there differences?

ii) Split the sample into different intervals according to the retrieved distance.
For example R_o=0-1000, 1001-2000, 2001-3000, 3001-4000 pc
For each of these intervals determine the mean value (and dispersion) of the retrieved absolute magnitude. Compare

with true absolute magnitude.
Are there differences? Are there trends with R_o ?

To understand and explain the results you should produce and interpret the following diagrams:

a) Make an histogram of the true number distribution of stars with distance (from column 2 of the file LK.dat) and compare with the histogram of the retrieved distribution of stars with distance (from column 6 of the same file).
Use a bin size of the order of 200 pc.

b) Make a scatter plot of all individual retrieved absolute magnitudes as a function of the retrieved distances (columns 8 and 6 of the file LK.dat).

c) Plot histograms of the distribution of the retrieved absolute magnitudes for the four retrieved distance intervals at point ii).

Taking advantage of these sets of plots, can you explain if/why you have biases in the retrieved absolute magnitude compared to the true value?

Two crucial pieces of information must be kept in mind for the analysis.
Firstly, the stellar spatial distributions described by the simulations predict a non uniform distribution of the number of stars with distance.

Also, in general, given that $R=1/p$, the modulus of the error dR on the derived distance for a given parallax error dp is (by simple error propagation)

$dR=(1/p^2)dp$

hence

$dR=R^2\,dp$

This means that for a given parallax error, more distant stars show the largest migration to shorter and longer distances, compared to stars at shorter distances.

TASK 2

Repeat TASK 1 but including a magnitude limit of 14 mag.

TASK 3

Repeat TASK 1 but changing the spatial density distribution. Use now the case n=0.

TASK 4

Repeat TASK 2 but considering a GAIA-like parallax error distribution (if you think you need to compare the parallax error with this option to the case of the previous tasks, remember that the individual parallax errors are tabulated in the output file).

Do you find differences amongst the results from these different tasks?

What are your conclusions about absolute magnitude calibration biases due to parallax errors?

Please submit a report of your results/conclusions as a pdf file.
The report should contain all plots above described and should be at most 3000 words long.