Astronomy
&
Astrophysics

Automatic observation rendering (AMORE)[*]

I. On a synthetic stellar population's colour-magnitude diagram

Y. K. Ng[1,2], E. Brogt[1,3], C. Chiosi[4], and G. Bertelli[2,5]

[1] TNO TPD Space, PO Box 155, 2600 AD Delft, The Netherlands
[2] Padova Astronomical Observatory, Vicolo dell'Osservatorio 5, 35122 Padua, Italy
[3] Kapteyn Astronomical Institute, PO Box 800, 9700 AV Groningen, The Netherlands
[4] Padova Department of Astronomy, Vicolo dell'Osservatorio 2, 35122 Padua, Italy
[5] National Council of Research, IAS, Rome Italy

Abstract. A new method, AMORE – based on a genetic algorithm optimizer, is presented for the automated study of colour-magnitude diagrams. The method combines several stellar population synthesis tools developed in the last decade by or in collaboration with the Padova group. Our method is able to recover, within the uncertainties, the parameters – distance, extinction, age, metallicity, index of a power-law initial mass function and the index of an exponential star formation rate – from a reference synthetic stellar population. No a priori information is inserted to recover the parameters, which is done simultaneously and not one at a time. Examples are given to demonstrate and to better understand biases in the results, if one of the input parameters is deliberately set fixed to a non-optimum value.

Key words. methods: data analysis, numerical – stars: HR-diagram, statistics

1. Introduction

The present data flow of many ongoing surveys – such as 2MASS (Beichman et al. 1998; Skrutskie 1999), DeNIS (Epchtein et al. 1997, 1999), EIS (Renzini & da Costa 1997; da Costa 1997; da Costa et al. 1998), OGLE-II (Udalski et al. 1997; Paczyński et al. 1999), SDSS (Fan 1999 and references cited therein), and even upcoming surveys as GAIA (Gilmore et al. 1998; Perryman et al. 2001) – is so large that one requires either a semi-automated or a fully automated method to analyse the colour-magnitude diagrams (CMDs) in the resulting databases. In this paper we discuss the development and the tests of an automated analysis method, which fully employs the colour and magnitude information available about the stars populating the CMD. Our method is based on an implementation of the CMD diagnostics suggested by Ng (1998). The method uses, in contrast to other techniques (see Bertelli et al. 1992; Gallart et al. 1996 & 1999; Geha et al. 1998 Harris & Zaritsky 2001; Hernandez et al. 1999;

Holtzman et al. 1997 & 1999) the full, unbinned distribution of magnitudes and colours of the stars populating the CMD.

The purpose of this paper is to verify that astrophysical parameters for a synthetic single stellar population can be reliably retrieved with the so-called "AutoMatic Observation REnderer" AMORE. In Sect. 2 an outline of AMORE is given together with its individual building blocks. In Sect. 3 we outline the method we use and in Sect. 4 we describe the tests performed with synthetic stellar populations. The results are given in Sect. 5 and we discuss in Sect. 6 the practical limits on the convergence, which is imposed by some degeneracy of the parameter space. We end with prospects on forthcoming tests, recommendations for improvements, and an outlook on future developments.

2. AMORE

2.1. Project outline

AMORE tries to find the best matching synthetic CMD to an observed CMD. Such a synthetic CMD contains for stellar aggregates the contribution of one or more stellar populations at the same distance. In the case of a CMD along a particular line of sight in our Galaxy the synthetic CMD can moreover contain the contribution of various populations with stars distributed at different distances.

In this paper we focus on the implementation and the performance of AMORE for the fitting of CMDs. For sake of argument only one, single stellar population has been considered. The implementation of automatic fitting CMDs with multiple stellar populations with stars at the same or different distances will be subject of forthcoming papers.

2.2. Building blocks

AMORE, see Fig. 1 and Sect. 3.1 for details, combines various analysis tools developed and improved at Padova during the last decade. Conceptually, it is made up out of the following building blocks:

- a synthetic Hertzsprung-Russell Diagram generator (hereafter referred to as HRD-ZVAR see Sect. 2.3),
- a mockup version of the HRD Galactic Software Telescope (HRD-GST, see Sect. 2.4),
- a statistical diagnostic tool to compare the observed and synthetic CMDs with each other (Ng 1998), and
- the PIKAIA (extended version) and POWELL optimizers (see respectively Sects. 2.6 and 2.8), which search for the best fit between an observed and a synthetic CMD in a multi-parameter space.

In the following subsections a description is given of the various building blocks.

2.3. HRD-ZVAR

In the late-eighties Bertelli developed a code to generate synthetic Hertzsprung-Russell diagrams (HRDs) from the isochrones computed by the Padova group (cf. Chiosi et al. 1989). Initially the synthetic colour-magnitude diagram (CMD) technique was applied mainly in the studies of LMC open clusters[1] (see for example Bertelli et al. 1985, 1990 or Chiosi et al. 1989), through which the amount of convective overshoot was calibrated for the computation of a new generation of stellar evolutionary tracks. Successive improvements were gradually applied when new sets of evolutionary tracks (see Bertelli et al. 1994 for details) were computed with improved radiative opacities (Iglesias et al. 1992).

The backbone of HRD-ZVAR, the extended version of the HRD generator, is formed by the evolutionary tracks computed by Bertelli et al. (1990; $Z = 0.001$), Bressan et al. (1993; $Z = 0.020$), and Fagotto et al. (1994a,b,c; $Z = 0.0004, 0.004, 0.008, 0.050, 0.10$). The metallicities of the tracks follow the enrichment law $\Delta Y/\Delta Z = 2.5$ (see references cited in Chiosi 1996 and Pagel & Portinari 1998).

HRD-ZVAR indicates that the metallicity Z is not limited to the fixed values for which the evolutionary tracks have been computed, but is variable through interpolation between

the metal-poorest and metal-richest tracks available inside the database of evolutionary tracks. In this way one is able to generate synthetic stellar populations with a smooth metallicity coverage. A prerequisite however is to use a complete and homogeneous library of evolutionary tracks and some improvements are expected if one adopts a grid of tracks with a smoother metallicity coverage.

HRD-ZVAR has been distributed (privately) to various research groups and is also referred to as ZVAR[2]. The version distributed, modified, and used by for example Aparicio (1999), Gallart (1998), and Ng et al. (1996, 1997) is from now on referred to as V1.0. Version V1.6 is used for the simulations and results presented in this paper. This version contains a number of modifications and improvements which speeds up the code and fixes some (rarely encountered) bugs which interfered with the automatic minimization process. Although various analysis methods are available, we limit ourselves here to the description of the parameters related to the HRD-ZVAR as adopted for AMORE.

After selection of a set of tracks with fixed metallicity and the choice of the parameters η_{RGB} and η_{AGB}, (the mass loss along the Red Giant Branch (RGB) and the Thermally Pulsing Asymptotic Giant Branch (TP-AGB) phases respectively), the major input to be specified for HRD-ZVAR are:

- the metallicity range, Z ranges from Z_{min} to Z_{max};
- the age range, the age ranges from t_{min} to t_{max};
- the slope α for the power-law IMF (Initial Mass Function); and
- the index β for the exponential[3] SFR (Star Formation Rate).

The luminosity and effective temperature for each synthetic star of arbitrary metallicity is transformed to an absolute magnitude in a photometric passband with the method outlined by Bertelli et al. (1994) and Bressan et al. (1994). Default setup for HRD-ZVAR is the $UBVRI\,JHKLMN$[4] broadband photometric system. The setup can be altered to mimic any system, given the description of the spectral response of the filter and the detector of the photometric system.

2.4. HRD-GST

HRD-ZVAR was integrated in a galactic model by Ng (1994, 1997ab). The HRD-GST (Galactic Software Telescope) has been applied in the studies of the galactic structure towards the Galactic Centre (Ng et al. 1995, 1996 and Bertelli et al. 1995, 1996) and other regions in our Galaxy (Ng et al. 1997). In this paper we do not require the full complexity of the structural

---

[1] The analysis technique with stellar ratios was employed. The reason for this is that ratios are less sensitive to uncertainties in certain regions in the CMD, which might not be reproduced properly due to various reasons such as the input physics used for the calculations of the stellar evolutionary tracks or the transformations from the theoretical to the observational plane.

[2] Note that an older and modified version of HRD-ZVAR is actually the program used to generate the Bertelli et al. (1994) isochrones. HRD-ZVAR is free of the interpolation difficulties as reported by Olsen (1999).

[3] The SFR is only fixed for the time being to the adopted exponential shape.

[4] Note that the IR photometric system is based on an "average" photometric system as described by Bessell & Brett (1988). Proper transformations ought to be applied to the actual photometric system prior to any astrophysical interpretation of the results.
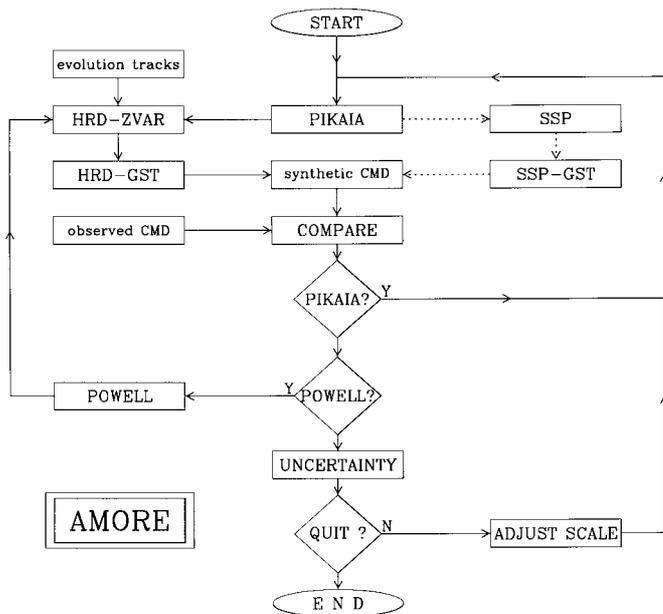
Fig. 1. AMORE flowchart: schematic diagram of the individual building blocks. PIKAIA outlines the direction of the evolution paths to be investigated. Input for the stellar population synthesis engine HRD-ZVAR is the Padova library of evolutionary tracks. The luminosity and effective temperature for each synthetic star of arbitrary metallicity is then transformed to an absolute magnitude in a photometric passband with the method outlined by Bertelli et al. (1994) and Bressan et al. (1994). The synthetic HRD is then "observed" and "detected" through a Monte Carlo "observing run" with the HRD-GST. Note that an alternative route is possible with single stellar populations (SSPs). The synthetic CMD is compared with the observed diagram and a fitness parameter is subsequently communicated to PIKAIA, which suggests a new set of parameters for each trial. The iteration lasts for a user defined, fixed number of trials. POWELL's method of minimization is subsequently applied to get closer to the local or global minimum. After computation of the uncertainties for each parameter the evolutionary run is either aborted or a new PIKAIA cycle is started after shrinking the limits of the parameter space (see Sect. 2.9).

properties from the GST model. We only use a limited number of options to "observe" a synthetic HRD at the suggested distance and to simulate the photometric errors, extinction and crowding.

A table of the photometric errors, covering a specific magnitude interval per passband, is used and the program interpolates linearly to obtain the intermediate values. We assume for the simulation that the photometric errors are Gaussian distributed. A different description of the photometric errors will be used when published artificial star tests (Stetson & Harris 1988; Gallart et al. 1999) on an observational data set are indicative for a significant deviation from a Gaussian behaviour. The visual extinction is simulated with the average value provided and appropriately scaled to a value in different passbands. In the *UBVRI* passbands we adopted the scaling according to van de Hulst (1949; curve No. 15) and for the *JHKLMN* passbands we follow the scaling laws provided by Rieke & Lebofsky (1985). We do allow for some random scatter around the average extinction. However, we do not consider (yet) the effects due to patchiness of the extinction along the line of

sight. Ng & Bertelli (1996) demonstrated that this is in first approximation, visually almost indistinguishable from a random scatter around an average extinction.

In many studies the observations are crowding limited, due to the increasing number of stars towards fainter magnitudes. Crowding gives rise to star blends which affects the magnitude and the colour of the stars. The group of stars will be detected as a single star with a magnitude equal to the sum of the stellar flux of the stars involved in the stellar blend. The remaining stars are "hidden" from detection.

The blends are well described as unresolved, apparent binaries. The simulation of apparent binaries is made with an iteratively improved blending probability, which is defined as the probability that a star within a given ensemble of stars might blend with another star from the same population. Each synthetic star within a stellar population is tested against the blending probability.

The percentage of artificial binaries is with this scheme about twice the blending factor. The blending factor in different passbands is not necessarily the same and the occurrence of star blends is furthermore not necessarily correlated, due to possible differences in the exposure time or seeing conditions.

The fainter companion stars of artificial binaries will give rise to incompleteness of the synthetic stellar sample. This allows us to map the synthetic stellar completeness function, which can be compared directly with the completeness function obtained from artificial star tests from the observed stellar sample.

In the CMD to be analysed we assume implicitly that, between the observed and synthetic photometric system, the uncertainty in the magnitude zeropoint is smaller than the uncertainty in the zeropoint of the colour, see Carraro et al. (1999), and references cited therein. We allow for this reason the possibility that a small zeropoint shift might be present between the colours of these systems.

2.5. SSPs

Although straightforward, the generation of a large number of synthetic HRDs with each their own specific age-metallicity range, star formation history and initial mass function can be a time consuming task, because of the repetition of many calculations to generate one diagram. Figure 1 indicates that an alternative route with single stellar populations (SSPs) is available for the automated analysis. However, this method requires the computation of a large, regular grid of SSPs for different age-metallicity ranges and IMFs. The star formation history and age-metallicity range are the result of the linear combination of the SSPs. The time spent on computations of new CMDs thus can be greatly reduced through the use of probability density diagrams. However, this is not our prime objective. Our present goal is to develop an automated fitting method, which compares on a star by star basis, and to demonstrate its potential. Optimization for speed is not yet our primary concern. Moreover, the generation of one synthetic CMD with 5000 stars takes about 0.75 s on a PC equipped with a 200 MHz Pentium processor. This is a good indication that our present version of the software tool is performing at an
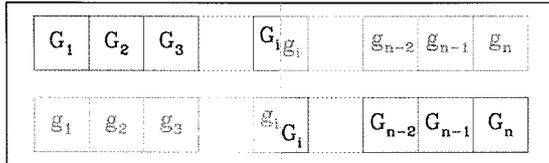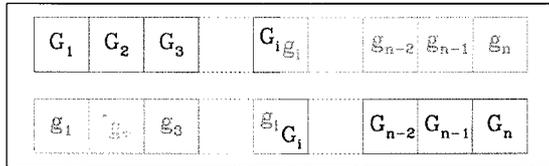
1-POINT CROSSOVER
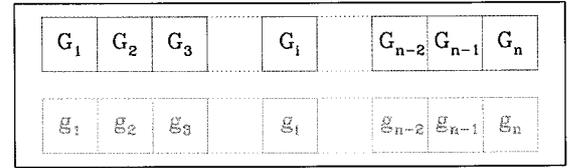
- breeding
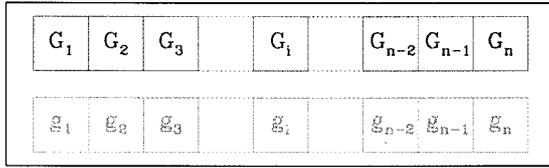
- crossover

- mutation

2-POINT CROSSOVER
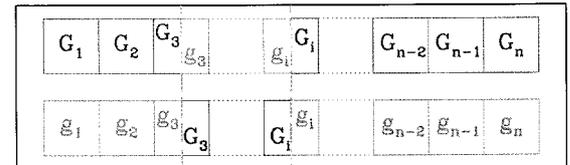
- breeding

- 2 point crossover
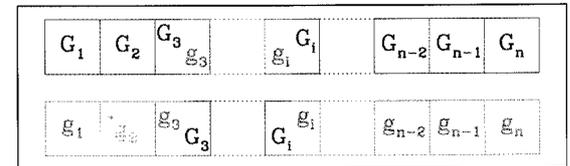

- mutation

Fig. 2. Schematic view of the main genetic operators acting on two sets of parameter strings in order to procreate two new sets (one chromosome set for each individual) of test strings: natural selection of two individuals G and g for the breeding of offspring, modification of the genetic content by means of a homologous crossover – at gene G_i/g_i, and finally the mutation of gene g_2 to $^*g_2$. Each gene on a chromosome represent a parameter to be optimized. The schematics for a 2-point crossover scheme are shown for comparison

Table 1. PIKAIA control parameters. These parameters were kept fixed at the values listed throughout the evolutionary runs.

initial value	PIKAIA default	parameter identifier	comment
100	100	np	the number of individuals per population
20	500	ngen	number of generations
2	5	nd	number of digits encoding accuracy
2	2	imut	mutation mode, imut = 2 then pmut = [pmutmn,pmutmx]
0.005	0.005	pmut	initial mutation rate
0.005	0.005	pmutmn	minimum mutation rate
0.35	0.25	pmutmx	maximum mutation rate
0.95	1.0	fdif	fitness differential
3	1	irep	reproduction plan
0	1	ielite	elitism
0	0	ivrb	verbose mode

acceptable speed. We refer to Dolphin (1997, 2001, 2002) and Olsen (1999) for a description of a method using SSPs.

2.6. PIKAIA

The PIKAIA optimization package (V1.0; public domain) was developed by Charbonneau (1995) and a full description of this package is given by Charbonneau & Knapp (1996). PIKAIA has been used successfully in a wide range of astrophysical applications (e.g. Bobinger 2000; Charbonneau et al. 1998;

Gibson & Charbonneau 1998; Kaastra et al. 1996; Kennelly et al. 1996; Lamontagne et al. 1996; McIntosh et al. 1998; Metcalfe 1999; Mewe et al. 1996; Noyes et al. 1997; Saha 1998; Wahde 1998).

PIKAIA is based on a genetic algorithm (Holland 1975; Goldberg 1989; Davis 1991; De Jong 1993), and is in principle not a function optimizer, but it does this extremely well. It searches for, locks on to, and pins down an optimal solution in a way, which is conceptually comparable to biological evolution through natural selection. Genetic algorithms are capable

to explore and find in a robust way an optimum, but not necessarily the best, setting for a particular problem. In our case this comes down to minimizing the difference between a synthetic and an observed CMD by evolving the astrophysical parameters that define the shape of the CMD (see Table 2 and Sects. 2.10 and 3). We follow the generally accepted biological terminology for the description of a genetic algorithm.

A genetic algorithm makes use of a reduced version of the evolutionary process. The gene pool, i.e. the set of parameters to be optimized, and its associated phenotypic population evolves in response to

- the reproductive success of the population, following its reproduction plan;
- genetic recombination (crossover, see Fig. 2) at breeding;
- random mutations (see Fig. 2) during breeding, which affect a subset of the events.

2.6.1. Flow control

PIKAIA has 12 flow control parameters which are discussed by Charbonneau (1995) to whom we refer for a detailed description. We limit ourselves to a short summary and the purposes of these control parameters: np defines the number of individuals in a population; ngen specifies the number of generations that the population is evolving; nd is number of digits encoding accuracy[5] used for the parameters; pcross is the probability that a crossover occurs between the chromosomes of the parents; imut, pmut, pmutmn, and pmutmx specify the mutation mode, the mutation rate (the initial mutation rate if the rate spans the range from pmutmn to pmutmx for imut = 2), and the minimum and maximum mutation rate; fdif the fitness differential controls the selection of the individuals for breeding through their fitness; irep defines the reproduction plan to be followed; ielite defines if the fittest individual can or cannot be selected for replacement; and ivrb specifies verbose mode for extra on screen information during the evolutionary run.

Table 1 holds a list of the PIKAIA flow parameters which were kept constant during all the simulations described in this paper. In the following subsections we describe the extensions added to the 12 flow parameters of PIKAIA.

2.6.2. Crossover

The crossover operator is very effective in a global exploration of the full parameter space and is in a way comparable to a variational calculus method. A one-point crossover scheme, see Fig. 2, is sometimes inadequate to combine and pass on certain features encoded on the chromosomes (Michalewicz 1996) to its offspring. In some cases only a correlated modification of a number of genes, say 2, will result in a fitter offspring. This can, for example, be mimicked through the application of a two- or multi-point crossover scheme. PIKAIA has been extended with the control parameter rcross to handle a multi-point crossover

operation. For example: rcross = 1 represents the default one-point crossover, while rcross = 2.3 represents a "2.3-point" crossover: i.e. a two-point and three-point crossover for respectively 70% and 30% of the cases ($2.3 = 0.7 \times 2 + 0.3 \times 3$).

2.6.3. Brood recombination

An important drawback of genetic algorithms is that the crossover operator is for about 75% of the time lethal to its offspring, i.e. it produces children which are not as fit as their parents (Banzhaf et al. 1998). To avoid missing, potentially, fitter offspring and to reduce the destructive effect of the crossover operator we have incorporated brood recombination in PIKAIA through a new control parameter rbrood. The default PIKAIA reproduction scheme is obtained with rbrood = 1, i.e. two parents breed once and produce two new individuals. With rbrood = 3.5 the parents are allowed to breed on average 3.5 times to produce a larger offspring (in this case 7 on average). For rbrood > 1 one ends up with more than two offspring. PIKAIA on the other hand expects from each pair of parents only two children. This constraint was obeyed in order not to significantly alter the global behaviour of PIKAIA. Therefore, in order to avoid an exponential growth of the population for rbrood ≠ 1, only the fittest two individuals of the local offspring survive[6] and enter the global population for a fitness evaluation from which a selection is made for further breeding.

The extra breeding increases the computational effort considerably, due to a larger number of function evaluations. The advantage is a more rapid increase of fitter individuals through the selection of effective crossovers from good recombinations. On the other hand, a rapid increase of fitter individuals might lead to a premature convergence to a local minimum, due to a smaller variance in the genetic pool.

2.6.4. Creep mutations

We further introduced the "creep" mutation (Charbonneau & Knapp 1996) in order to overcome the so-called "Hamming Wall" problem, i.e. the inability to cross in a decimal encoding scheme certain boundaries with a one-point mutation operator. The creep parameter pcreep defines the probability that a gene in the pool undergoes a "standard" mutation (change digit randomly in the range 0–9) or the "creep" mutation (add or subtract one from the current value of the digit). We adopted as default equal weight for the occurrence of a "creep" or "standard" mutation.

2.6.5. Correlated mutations

In general mutations occur in the optimization process to avoid premature convergence. A low mutation rate is sufficient for this purpose. However, a high mutation rate can be used as

---

[5] The 1 or 2 number of digits encoding accuracy maps the parameter via an integer value, either 0–9 or 0–99, to a range of floating point values by a controlling "normalization" function, see Charbonneau (1995) for additional details.

[6] This is irrespective if the local, fittest individuals are weaker than the weakest individual in the global population or if the locally remaining offspring are fitter than the fittest individual in the global population.

an additional way to explore the parameter space like a virus. Although one would prefer to use a (multi-point) crossover operator, we do allow that mutations can be used instead, simply because the two operators also co-exist in nature.

We modified PIKAIA's uniform mutation mode. In the majority of the cases we require a (anti-)correlated change between two or more parameters (see Sect. 6.2). In a standard mutation scheme convergence might be slow if one has to wait for the simultaneous occurrence of a favourable (anti-)correlated mutation of two specific parameters in order to improve the fitness. We introduced an extra parameter pcorr which defines the probability that a correlated mutation occurs. If this is not the case the standard mutation scheme is chosen. Otherwise we allowed that in 50% of the cases the mutations of two genes (igen1 and igen2 are extra input parameters added to the modified version of PIKAIA) are more relevant than the mutations occurring in other parameters. For the remaining 50% of the cases the two genes are determined stochastically. Additional details about the adopted values of the control parameters are given in Sect. 2.10 and Table A.1.

2.7. Fitness

PIKAIA searches for the optimum solution by maximizing the fitness function f. The fittest solution has a fitness $f = 1$, while the worst has $f = 0$. We use the Ng (1998) fitness function, a combination of a chi-squared and Poisson like functions. These functions minimize the differences between observed and synthetic diagrams via a star-by-star matching[7] scheme.
The Ng fitness function is defined as:

$$f = \frac{1}{1 + F} \, , \tag{1}$$

where F is

$$F = F_\chi^2 + F_P^2 \, . \tag{2}$$

F_χ is the chi-squared function of the best fitted points within a 3–$5\,\sigma$ error ellipse and F_P is the Poisson function of the residual points outside this ellipse. Both F_χ and F_P are dimensionless, but they hold information about the average uncertainty in units of σ, say σ_χ and σ_P. For example: the average uncertainty per point for a fit with F_χ is $F_\chi \times \sigma_\chi$.

F_χ and F_P are respectively defined as:

$$F_\chi = \sqrt{\overline{\chi^2}} = \sqrt{\chi^2(O, S)/N_{\text{match}}} \, , \tag{3}$$

and

$$F_P = \frac{N_{O,\text{not}} + N_{S,\text{not}}}{\sqrt{N_O} + \sqrt{N_S}}. \tag{4}$$

The intuitive motive behind this is to make a division between the synthetic points matching the observed CMD for which the errors are expected to be normally distributed and the points which do not match and are allegedly assigned to the Poisson merit function. The method actually uses F_χ as a loosely fixed

"anchor", puts the outlier points in F_P and then reduces the number of unmatched points by minimizing F_P. We refer to Ng (1998) for additional details and a discussion of these functions. Suffice to say that for an acceptable solution both F_χ and F_P are about 1 or smaller[8] which on its turn can be relaxed to the condition $F \lesssim 2$ [9] and thus yields $\frac{1}{3} \lesssim f < 1$.

The formal $1\sigma_k$ uncertainty of each parameter k, see Table 4, is obtained through variation of this parameter and by minimizing the function $|\sqrt{F_k} - \sqrt{F_{\min}} - 1|$. Conceptually, this is similar to moving the merit function in the F_P, F_χ – plane away from its optimum setting, to the nearest position on a contour $+1\sigma_k$ higher. The associated fitness function $f_{\sigma,k}$[10] is:

$$f_{\sigma,k} = \frac{1}{1 + \left| \sqrt{F_k} - \sqrt{F_{\min}} - 1 \right|} \, , \tag{5}$$

where k is the particular parameter for which the uncertainty is estimated and $F_{\min}$ is the global value obtained for the fittest population.

2.8. POWELL

We implemented a hybrid optimizer in which we use PIKAIA to explore the parameter space and then use POWELL's minimization algorithm (Powell 1964; Press et al. 1986) to pin down the nearest local or global minimum through a direction set method which produces N mutually conjugate (non-interfering) directions. For details and an excellent description of this algorithm we refer to Press et al. and references cited therein.

A hybrid minimization strategy is used, because PIKAIA is by definition not a function optimizer, but it tends to get close near a (local) optimum. POWELL is used to get even closer to the (local) optimum. If we had landed in a local optimum then we needed PIKAIA to jump out of it. The origin of our need for a hybrid search strategy is comparable to the minimization problems encountered by Harris & Zaritsky (2001).

2.9. Contracting parameter space

A full exploration of parameter space at one digit accuracy would take considerable time (in the order of weeks), even though certain forbidden combinations of parameters can be excluded à priori. A full exploration at two or more digits accuracy is nearly impossible due to present day computational limits.

We implemented a dynamic, scalable parameter range in our search for an optimum set of parameters. The parameter

---

[7] This option is feasible due to the increment of the present day computational speed.

[8] For our testcase, as described in Sect. 4, F_χ ranges from 0.7–1.0 for all observed data points with a matching synthetic point within a 3σ uncertainty ellipse. This corresponds to a goodness of fit parameter ranging from 0.49–1.0. It further indicates that it is justified to assume that the measurement errors are normally distributed.

[9] This condition is comparable but not equivalent to the results obtained by Gallart et al. (1999). They demonstrated from a comparison with colour-magnitude bins that a good agreement between the input and recovered SFR(t) required a reduced chi-squared of $\chi_\nu^2 \simeq 2.0$.

[10] For both age and metallicity ranges the associated uncertainties denote for the lower and upper values the -1σ and $+1\sigma$ boundary.

Table 2. AMORE control parameters, used as the "educated guess" for the astrophysical parameters in the first PIKAIA cycle. These values result in a fitness $f = 0.004$. The value for the parameters are constrained between the lower and upper limit.

#	parameter	value	lower limit	upper limit	description
1	$\log d$	3.7	3.6	4.0	log of distance
2	A_V	0.30	-0.001	0.5	extinction
3	$\log t_{low}$	9.5	9.0	10.3	log lower age limit
4	$[Z]_{low}$	-0.90	-1.69897	0.69897	log lower metallicity limit
5	α	1.35	1.001	3.5	IMF slope
6	$\log t_{high}$	9.6	9.0	10.3	log upper age limit
7	$[Z]_{high}$	0.30	-1.69897	0.69897	log upper metallicity limit
8	β	1.0	-2.00	5.00	SFR slope

range shrinks after each optimization cycle with POWELL. This leads to an improved accuracy in the results with a fixed number of digits encoding accuracy.

The automated re-scaling of the parameter range improves the resolution of the exploration of the search grid. In addition, due to the re-scaling one may circumvent partial degeneracy of the parameters.

The global function $F(F_\chi, F_P)$ gives the global distance to the minimum in terms of $\sigma^2(\sigma_\chi, \sigma_P)$. If we have n parameters then each parameter k is in a simple approach on average about $\sqrt{F/n}\,\sigma_k$ away from its optimum value, because we assume that

$$F = \sum_{k=1}^{n} \left(\frac{y_{k,\mathrm{sim}} - y_{k,\mathrm{true}}}{\sigma_k} \right)^2 , \qquad (6)$$

where $y_{k,\mathrm{sim}}$ is the simulated value of parameter k and $y_{k,\mathrm{true}}$ is the true value of parameter k. The average offset per parameter k is therefore $\sqrt{F/n}$. The new limits can then be set to $\pm \sqrt{F/n}\,\sigma_k$. To balance the cancellation of errors, due to a negative correlation between some parameters, we adopt a rather conservative approach by constraining the limits of each parameter k to $\pm 3 \times 1.3 \sqrt{F/n}\,\sigma_k$. In addition we add the condition that $3 \times 1.3 \sqrt{F/n} > 0.005$.

2.10. Adjustable parameters: Free and fixed

AMORE uses in essence two different sets of parameters as input. One set contains program flow control parameters and information about the observational data to be simulated: the photometric errors and crowding factors per passband, a shift of the zeropoint of the colours due to a difference between the observed and synthetic photometric system, and the spread around the average extinction value to account partly for the differential reddening in a field.

The other set of input parameters is used by PIKAIA and can be divided into two parts (see Tables 1 and 2). One part contains the PIKAIA control parameters. The other part contains the lower and upper limits of the astrophysical parameters to be optimized as well as an initial guess for the value of those parameters (first column of Table 2, resulting HRD in Fig. 6b). These parameters are a combination of the synthetic population's intrinsic properties, i.e. age, metallicity, slope of the power-law IMF and the index of the exponential SFR.
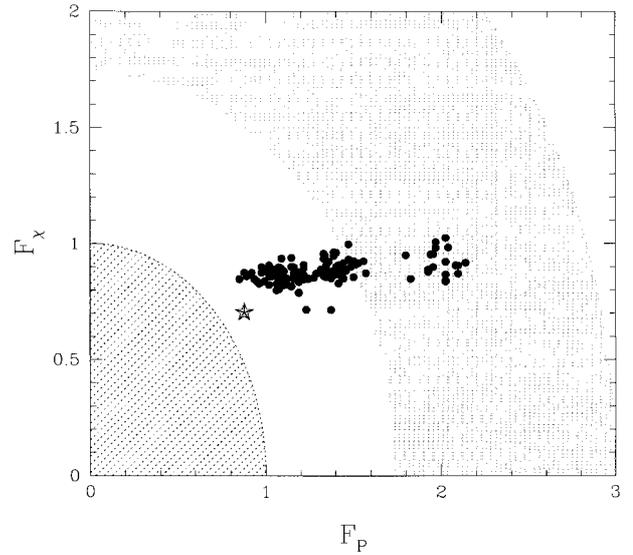

Fig. 3. Evolutionary status for all the trial models (filled dots; see Table A.1 for details) for each population after 400 generations, see also Fig. 4. The filled star ★ points to the location obtained with the original parameter settings as given in Table 4. Acceptable solutions are found in the region for which both F_χ and F_P are less than 1, see Sect. 2.7. The condition has been relaxed to $F < 2$ (solid white line). The asterisk indicates the position obtained for a different realisation of the test population by changing the random seed. The shaded regions indicate solutions for which the difference between the CMDs from the "observed" and synthetic population is on average less than respectively 3σ, $\sqrt{3}\sigma$, and 1σ. Note that rounding errors, see Sects. 5.1.1 and 6.3, give rise to a degeneration of the parameter space around $f > 0.25$ (i.e. $F < 3$).

In our case we allow for age and metallicity not to be restricted to one fixed value, but to cover a specific range (see Table 2).

In addition, there are two parameters which mimic the synthetic population's behaviour as placed in a mockup version of our Galaxy. These parameters are the distance from the Sun and the average extinction.

In total we thus have 8 free parameters which AMORE has to optimize simultaneously.

3. Method

3.1. Genetic evolution

In terms of genetic programming the objective of AMORE is to determine the genome (i.e. the set of *astrophysical parameters* described in Sect. 2.10, see also Table 2) of a specified individual (i.e. the *observed stellar population*). Note that it is not possible to directly observe the genome. The genome is determined from the phenotype of each individual (i.e. the *synthetic CMDs*, see for example Fig. 6). The genetic information is located in the genes of one chromosome[11].

A guess of the genotype of the observed CMD is obtained through comparison with a synthetic CMD, which is generated via supervised evolution and breeding (PIKAIA together with POWELL). The stars in the synthetic CMD population with a particular genotype are raised to maturity (HRD-ZVAR and HRD-GST). A group of individuals[12] is allowed to procreate (the chance of an individual procreating depends on its fitness and the selection pressure, see Charbonneau 1995 and Charbonneau & Knapp 1996) and the genetic information of the parents is passed on to their offspring (see Fig. 2).

A fitness evaluation (a comparison between the observed and synthetic CMD) provides a ranking of the resulting group of individuals. If the individual has "good genes" it survives, remains in the group and gets a chance of procreation.

The evolutionary process of breeding and fitness evaluation is repeated for a fixed number of generations. The gene pool of the resulting best individual at the end of the evolutionary run with AMORE hopefully represents a near-optimum representation of the unknown genome.

3.2. Running AMORE

Initially PIKAIA is in control (see Fig. 1) of the evolution for a fixed number of generations. Afterwards POWELL tries to improve the genome of the fittest individual communicated through PIKAIA. We then determine the uncertainty for each gene on the chromosome. Subsequently, we tighten the limits on the range of variation allowed for each gene and re-scale the parameters on the genetic print of the fittest individual accordingly. During the shrinkage of the parameter range we do not re-scale the genetic information of the remaining individuals, but preserve their former values as semi-random input for the continued optimization process. The latter addition to the hybrid scheme is most likely a significant driver in speeding up the search for a fitter individual.

After each optimization with POWELL a new cycle with PIKAIA is started with the current best parameter set as "educated next guess" for AMORE's progressive evolution. The total number of PIKAIA cycles is user defined.

---

[11] The genetic information is currently located on one chromosome. Individuals with two chromosomes might be considered as a future extension. A two-chromosome approach has the advantage that certain genetic information can remain present in a recessive form.

[12] To avoid confusion the term *group* is used instead of the biological term *population*, because each individual in the group is actually a synthetic stellar population.

In Sect. 2.9 we argued that the parameters are on average about $\sqrt{F/n}\,\sigma_k$ away from its optimum value. The convergence however is not governed by the average "distance" that each parameters is away from its optimum setting. It is mainly determined from the ability to tune the parameter which has the largest offset from its optimum value.

In the AMORE training sessions it was noted that with $F \simeq 3.0$ about three of the eight parameters are about $1\,\sigma_k$ ($\simeq \sqrt{F/3}\,\sigma_k$) away from their optimum value. AMORE has a built-in option to do a random variation from 0–$3\,\sigma_k$ of two parameters (randomly selected) from the running best solution when the fitness is less than 0.30. Above this threshold we choose a new value for two parameters according $\pm \sqrt{F/\rho}\,\sigma_k$, where ρ can be any number between 1.0 and 4.0.

4. Tests

4.1. Test objectives

We performed several tests on AMORE in order to

- verify and validate AMORE's performance in retrieving the astrophysical parameters of a synthetic single stellar population;
- determine adequate values for the parameters pcross, rcross, rbrood, pcreep, and pcorr in the extended version of PIKAIA;
- study the effects of rounding and degeneracy;
- study the effects of fixing parameters on the convergence.

4.2. Setup

The hybrid interaction between PIKAIA and POWELL, combined with a progressive shrinking/re-scaling of the parameter space, requires that a trade-off has to be made in the choice of the size of the population and the number of generations we allow this population to evolve in order to obtain results in a reasonable amount of CPU processing time.

We explored several different settings for the PIKAIA control parameters, because the tuning of those parameters is very problem dependent (Charbonneau & Knapp 1996). The values we decided to use are listed in Table 1. Four notes can be made here.

Firstly the steady-state-delete-worst reproduction plan (irep=3) we adopted, in which we replace the least-fit individual from the population when the fitness of the new individual is superior to that of the least-fit population member. Choosing this reproduction plan implies that the elitism control parameter (ielite) is non-operative, because elitism is active by default. We evaluated two other reproduction plans (Charbonneau & Knapp 1996); full generational replacement and steady-state-delete-random. The steady-state-delete-worst reproduction plan produced on average the best results.

Secondly the mutation rate of 0.35 corresponds, in case of a default 2 digit accuracy, with the on average occurrence of 2.8 mutations per astrophysical parameter.

Thirdly, the fitness differential parameter fdif, a measure for the selection pressure, would normally be chosen as high as possible (fdif=1 in this case). However, it may possible to circumvent local minima by lowering that value a bit
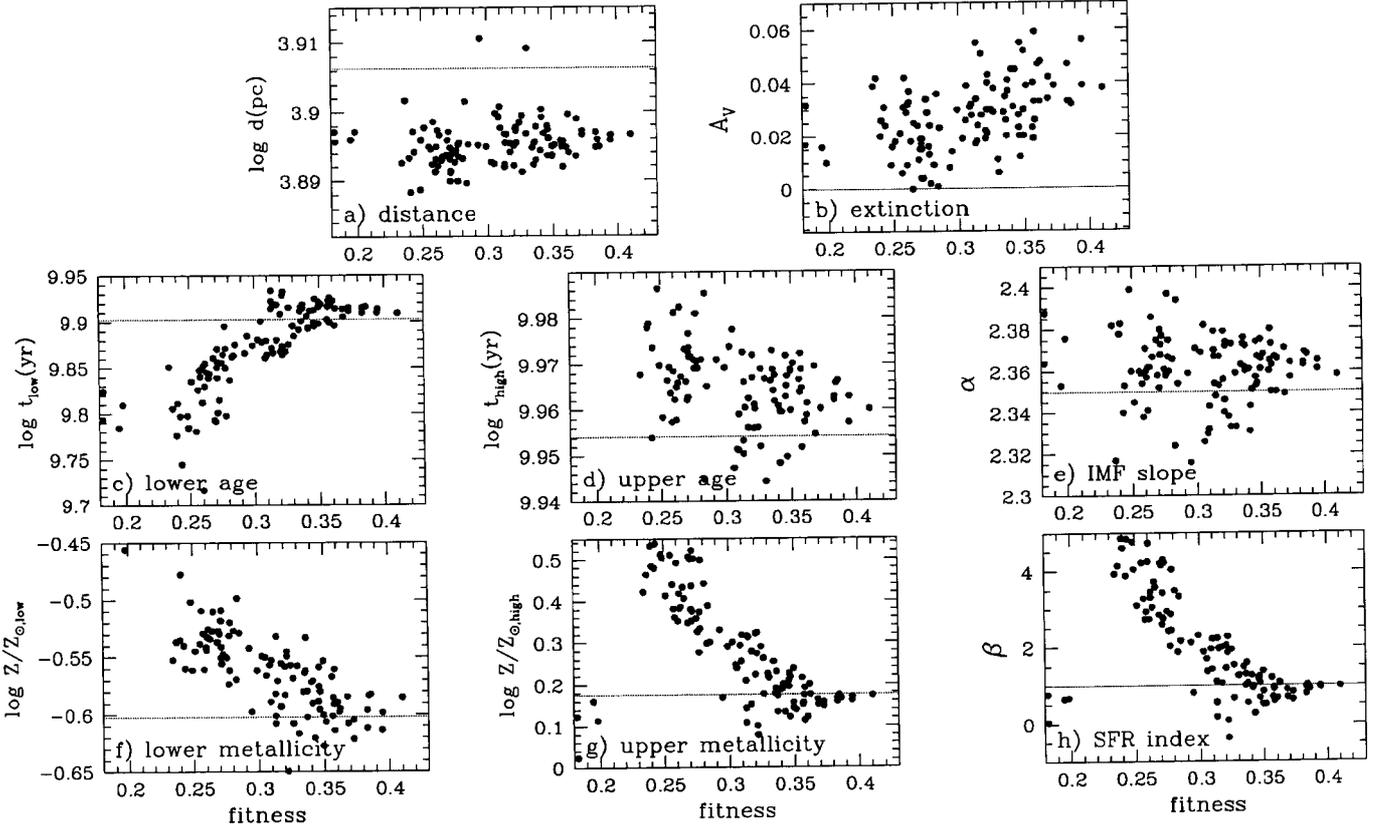
Fig. 4. The filled circles in panels **a)**–**f)** display the values obtained for the parameters using the models in Table A.1. The solid line refers to the value set for the original population.

(Charbonneau & Knapp 1996). Setting `fdif`=0.95 turned out to be a good trade-off choice.

Fourthly, we want to explore as large a fraction of the parameter space as possible at the first entry in AMORE. This is done by using only an one digit accuracy (`nd=1`). Due to the active re-scaling of the parameter space boundaries we do not require a very high precision in our exploration. A one percent accuracy (`nd=2`) of the parameter space is sufficient in the subsequent PIKAIA cycles.

In biological terms, the PIKAIA control parameters define the ecosystem in which our population evolves.

All computations presented in this paper were performed with an executable generated with the `g77` compiler. This executable was then installed on various PCs running Red Hat Linux[13] 6.X and 7.X. The PCs were equipped with Intel Pentium III or Athlon processors with clock speeds ranging from $600-1200$ MHz.

All tests, unless stated otherwise, use the synthetic population as described by Ng (1998):

- a metallicity range, spanning $Z = 0.005-0.030$;
- an age range from 8–9 Gyr;
- an initial mass function with a Salpeter slope;
- an exponential decreasing star formation rate with a characteristic time scale of 1 Gyr.

Table 3. Average fitness ($\overline{f_A}$) for different values of the parameters as obtained from Table A.1. The first column displays the parameter name which is varied, the parameter value is given in the second column, the third & fourth column show respectively $\overline{f_A}$, uncorrected for `pcorr = 0.0`, together with its standard deviation, and the fifth and sixth column display the averaged values after removal of the results of the models with `pcorr = 0.0`. See Sect. 5.1.2 for additional details.

parameter	value	$\overline{f_A}$	σ_{n-1}	$\overline{f_A}$	σ_{n-1}
pcross	0.50	0.276	0.070	0.287	0.066
	0.85	0.299	0.063	0.297	0.056
rcross	1.00	0.293	0.065	0.290	0.063
	2.00	0.285	0.068	0.297	0.065
	3.00	0.285	0.070	0.290	0.057
rbrood	1.00	0.286	0.077	0.230	0.059
	2.00	0.292	0.058	0.301	0.055
	4.00	0.285	0.066	0.276	0.067
pcreep	0.0	0.287	0.071	0.292	0.068
	0.3	0.290	0.065	0.297	0.058
	0.7	0.285	0.066	0.288	0.060
pcorr	0.0	0.278	0.077	NA	NA
	0.3	0.285	0.065	NA	NA
	0.7	0.300	0.057	NA	NA

---

[13] Red Hat® is a registered trademark of Red Hat Software, Inc. and Linux® is a registered trademark of Linus Torvalds.

The test population contains $N = 5000$ stars and is placed at 8 kpc distance. The "observational" limits are set to $V_{lim} = 22^m$ and $I_{lim} = 21^m$.

4.3. Description of the tests

4.3.1. Test 1: Determining values for `pcross`, `rcross`, `rbrood`, `pcreep`, and `pcorr`

In the first test, we evaluate the 162 models listed in Table A.1 in order to study the effect of the `PIKAIA` parameters `pcross`, `rcross`, `rbrood`, `pcreep`, and `pcorr` on the convergence and computational effort. The test has as a secondary objective to provide an understanding of the degeneracy of the parameter space.

All astrophysical parameters to be retrieved are set free, floating between reasonable minimum and maximum values (see Table 2 for details). AMORE runs for 20 iterations of 20 generations (`ngen`=20) to recover the a priori known parameters of the synthetic population. The number of iterations and generations determine the total length of an evolutionary run: $20 \times 20 = 400$ generations. Note that the range of each parameter is set within reasonable limits and not taken excessively large, because it might lead to the case that no acceptable parameter setting is found with the standard iteration loop.

4.3.2. Test 2: Rounding

The second test deals with the effects of rounding. We vary the number of significant digits in the input parameters to reveal AMORE's sensitivity to rounding. In this case we do not make an evolutionary run, because the trial set of parameters (the educated guess at the start-up of a `PIKAIA` cycle) is the correct one. For clarity we label the fitness in this test by f_{trial} instead of f_A, the fitness after a complete evolutionary run.

4.3.3. Test 3: Fixing parameters at the correct value

In the third test we take six models in which one of the parameters is set fixed at its correct value in order to study the effects on the convergence. The models chosen were two of high, two of intermediate and two of low fitness as determined from the first test. The convergence in this test basically can go two ways: either the convergence is faster, because less parameters have to be optimized. Or, due to the fact that AMORE has less maneuverability in this situation, the convergence is slower. We adjusted the limits for age and metallicity as given in Table 2 such that AMORE would not try to find solutions in forbidden regions of parameter space which might severely slow down convergence due to constant rejection by AMORE of the chosen parameter values.

For example, fixing the $[Z]_{\text{low}}$ parameter at its correct value of -0.60206 means that we have to adjust the lower limit for $[Z]_{\text{high}}$ to -0.60206 as well.

In the case of fixing the $\log t_{\text{low}}$ parameter this also implies that the initial guess has to be adjusted. We set this initial guess to 10.1.

4.3.4. Test 4: Fixing parameters at the wrong value

In the fourth test we take six models in which one of the parameters is set fixed at 1σ offset (determined from the first test) from its original value, in order to study its effect on the "second best" setting of the remaining parameters. Normally one would expect a fitness $f > \frac{1}{3}$. In this case, however, $F < F_P^2 + F_\chi^2 = 1^2 + (1 + 1)^2 = 5$ and the associated fitness constraint drops to $f > \frac{1}{6}$. However, this assessment ignores the fact that, when a parameter is offset from its optimum value, the number of matched points will decrease and F_P increases. Using Eq. (6) one has for a good fit $F = 2$. On average the offset per parameter k from the optimum value is $\sqrt{\frac{1}{4}}\,\sigma_k = \frac{1}{2}\,\sigma_k$, at best the offset is $0\,\sigma_k$, and in the worst case this is $\sqrt{2}\,\sigma_k$. So with one parameter k put at $1\,\sigma_k$ offset we distinguish the three possibilities

1   at best $F = (1 + 2)$ $= 3$ $\rightarrow$ $f = \frac{1}{4}$
2   on average $F = (1 + \frac{1}{2})^2 + \frac{7}{4}$ $= 4$ $\rightarrow$ $f = \frac{1}{5}$
3   at worst $F = (1 + \sqrt{2})^2$ $= 5.8$ $\rightarrow$ $f = \frac{1}{6.8}$.

Note that the worst case limit is in agreement with the results presented in Table A.3.

The effect of the 1σ offset of one of the parameters will partly be canceled by forcing other parameters away from the optimum value. For example, the effect of an increased extinction can be masked partially by generating a bluer stellar population with a lower metallicity and a younger age. The effect will be such that the fitness will not be around $f \simeq \frac{1}{6.8}$, but somewhere in the range $\frac{1}{6.8} < f < \frac{1}{3}$.

We fixed the parameters both at one sigma above and one sigma below the original value, because the evolutionary effects do not have to be symmetric. The only exception is the extinction, which we only fix at one sigma above the original value of $A_V = 0.0$.

Again we adjusted the limits for the upper and lower limit for age and metallicity.

5. Results

AMORE has been tested for a wide range of setups. The results in Table A.1 indicate that AMORE give both acceptable and less acceptable solutions. They are displayed in Figs. 3 and 4.

To better understand what goes on during the genetic evolution we display the results from model A.1-40. Figure 5 displays an example of the evolution of the merit function F for a number of generations. It shows how the initially dispersed individuals gradually find their way, start to cluster together around generation 10, and penetrate the region with acceptable solutions after about 50 generations. After 100 generations the improvements become marginal for this model.

Figure 6 displays for the same model A.1-40 the phenotypical changes of the CMD for several fitnesses during the genetic evolution. The various panels show that the synthetic CMD resembles better and better the "observed" CMD when the fitness improves. Note that at fitness $f = 0.05$ one already gets for the eye appealing solutions.

Figure 7 shows the improvements of the astrophysical parameters as a function of increasing fitness for the models A.1-40 and A.1-51. The panels for distance and extinction show that the distance is systematically underestimated, while the extinction is overestimated. But in general one notices that

Table 4. The effect of rounding in the parameters on the fitness evaluation; f_{trial} is the fitness obtained with AMORE for the given values of the parameters. The *average value* and σ are based on the 162 models in Table A.1. The σ is obtained from an unweighted average for the fitness of the models involved. The $\widehat{\sigma}$ is the error as estimated by Ng 1998.

| parameter	$\log d$(pc)	A_V	$\log t_{low}$	$\log t_{high}$	$[Z]_{low}$	$[Z]_{high}$	α	β	f_{trial}
original	3.906335	0.^{m}00	9.90309	9.95424	−0.60206	0.17609	2.35	1.0	0.44597
round-v1	3.906	0.^{m}00	9.903	9.954	−0.60	0.18	2.35	1.0	0.28595
round-v2	3.906	0.^{m}00	9.903	9.954	−0.602	0.176	2.35	1.0	0.30812
round-v3	3.9063	0.^{m}00	9.9031	9.9542	−0.602	0.176	2.35	1.0	0.42439
average value	3.8958	0.^{m}027	9.866	9.984	−0.554	0.244	2.358	1.574	
σ	0.0033	0.^{m}014	0.049	0.047	0.050	0.13	0.034	1.40	
$\widehat{\sigma}$	0.012	0.^{m}06	0.043	0.023	0.18	0.08	0.03	1.4	
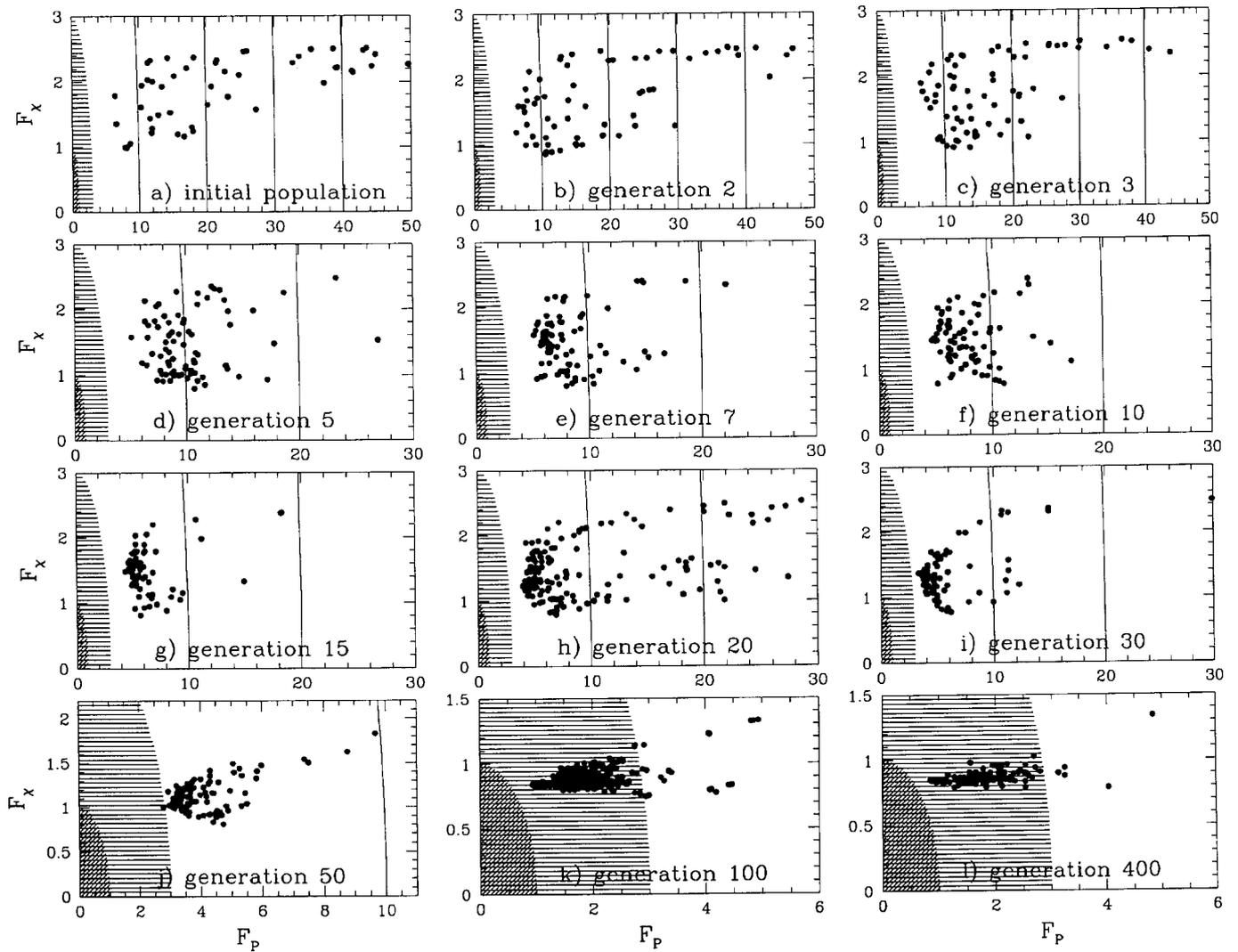
Fig. 5. Conception diagram of the evolution of the genetic population of model A.1-40 during the optimization process displayed in Fig. 7. Frame **a)** shows the initial population and the frames **b)**–**l)** show the population after several generations up to generation = 400. The outer shaded region indicates solutions for which the difference between the CMDs from the "observed" and synthetic population is on average between 1–3 σ. The inner shaded regions marks the region with solutions for which the difference between the "observed" and synthetic CMDs are less than 1 σ. Such solutions are close to perfect matches between the "observed" and synthetic data and are considered to belong to a group of solutions for which one may say "too good to be true". The solid lines indicate the 10 σ, 20 σ, 30 σ and 40 σ contours.

the astrophysical parameters obtained from model A.1-40 get quite close to the parameters of the CMD to be matched.

5.1. Test 1: Parameter values and degeneracy

5.1.1. Degeneracy of the parameter space

Table A.1 is displayed in Fig. 3. The clustering in the figure provides an indication that a degeneracy of the parameter space is present near $f > 0.25$ (i.e. $F < 3$, see Eq. (1)). Without a major computational effort it will be difficult to obtain a significant improvement of the parameters once $f > 0.25$. However, $F < 3$ indicates a region in the (F_χ, F_P)-plane for which the systematic offset of the individual parameters from its true value are on average less than $\sqrt{F/n}\,\sigma_k = 0.6\,\sigma_k$, see Sect. 2.9 for details. In practice it turns out that a strong correlation between three of the eight parameters has the culprit; at least two of them have to change simultaneously in the proper direction in order to improve the fitness (see also Sect. 6.3). They have an average offset of $\sim \sqrt{F/3}\,\sigma_k = 1\,\sigma_k$, while for the remaining parameters this is $\ll 1\,\sigma_k$.

In addition, Fig. 4 displays the retrieved parameters for all models as a function of fitness. Note, that AMORE systematically underestimates the distance of the test population. On the other hand, the effect of this underestimation is in its turn partially canceled by overestimating the extinction, the upper age limit and the slope of the power-law IMF slightly (see also Fig. 7). Another clue we get from Fig. 4 is that the slope of the SFR β is very poorly constrained.

5.1.2. Determining values for `pcross`, `rcross`, `rbrood`, `pcreep`, and `pcorr`

Looking at Table A.1, a result that is immediately obvious is that `pcorr = 0.0` has a strong stabilizing effect on the simulation. Its effects overwhelm the effects of the `pcreep` parameter and lock f_A at a certain value, which may be quite good (e.g. models 10, 11 and 12) or quite poor (e.g. models 28, 29 and 30). Setting `pcorr > 0` is thus preferred to avoid getting locked in a low value for f_A.

In order to compensate for this strong stabilizing effect, we also evaluate in Table 3 the average fitness of the models when we exclude all models which have `pcorr = 0.0`.

As expected, the `rbrood` parameter has a strong influence on the amount of computational time needed. Although the models with high values of `rbrood` are somewhat better than models with low values, this effect is only marginal. Considering that a high value of `rbrood` lessens the genetic variation in the gene pool while increasing the computational time needed for a run with several factors, it is desirable to have a low value of `rbrood`.

The different parameters are not independent, as can be seen from Table A.1 and Table 3. Simply taking the best options in Table 3 yields model 134 for the case in which `pcorr = 0.0` has not been corrected for, a reasonable, but not an exceptionally good model.

Table 5. Fitness statistics when fixing one parameter at its correct value. Averaged fitness values $\overline{f_A}$ and their associated standard deviation σ_{n-1} are obtained from simulations with the setup parameters from models 9, 14, 22, 34, 40 and 52. See Table A.2 for additional details.

parameter	$\overline{f_A}$	σ_{n-1}
$\log d$	0.299	0.081
A_V	0.274	0.081
$\log t_{\mathrm{low}}$	0.332	0.021
$\log t_{\mathrm{high}}$	0.304	0.056
$[Z]_{\mathrm{low}}$	0.269	0.036
$[Z]_{\mathrm{high}}$	0.361	0.036
α	0.305	0.066
β	0.312	0.076

5.2. Test 2: Rounding

Table 4 shows the effect of the accuracy of the retrieved values of the parameters on the evaluation of the fitness. If one applies a rounding to one or two significant digits it is not possible to reproduce the expected fitness, i.e. the expected fitness drops from $f_{\mathrm{trial}} = 0.44$ to $f_{\mathrm{trial}} = 0.28$. A better agreement can be obtained by reporting the values of the parameters with the addition of one or more (apparently insignificant) digit(s). A renewed search (Table 4; round-v3) circumvents the local optimum at $f_{\mathrm{trial}} = 0.28$ and results in a near global fitness of $f_{\mathrm{trial}} = 0.42$, which is close enough to the value expected.

The *true* σ line in the table shows that both the $[Z]_{\mathrm{high}}$ and the β parameter are the weak links in the overall parameter estimation (see also Fig 4).

5.3. Test 3: Fixing parameters at the correct value

The results of fixing parameters at the correct value are listed in Table A.2 and an example of the diagnostics is listed in Table 6. Details of the individual setups for these tests are given below. In general, the results of the tests for which one of the parameters was set to the correct value were slightly better than the results for the models for which all parameters are set free, see Tables 4 and 5 for additional details. This behaviour is due to the fact that by forcing one parameter to a fixed value the evolutionary path changes. The models were selected from the results with low and intermediate fitness given in Table A.1.

5.3.1. Fixed distance and extinction

Distance and extinction are negatively correlated. When both parameters are set free, a certain degree of degeneracy is to be expected. Fixing one of the parameters at its correct value can break this degeneracy. The result depends strongly on the evolutionary path of the other parameters.

The lower value for the average fitness $\overline{f_A} = 0.299 \pm 0.081$, when fixing the distance at the correct value, is caused by the presence of one outlier (see Table A.2), which is caused by the
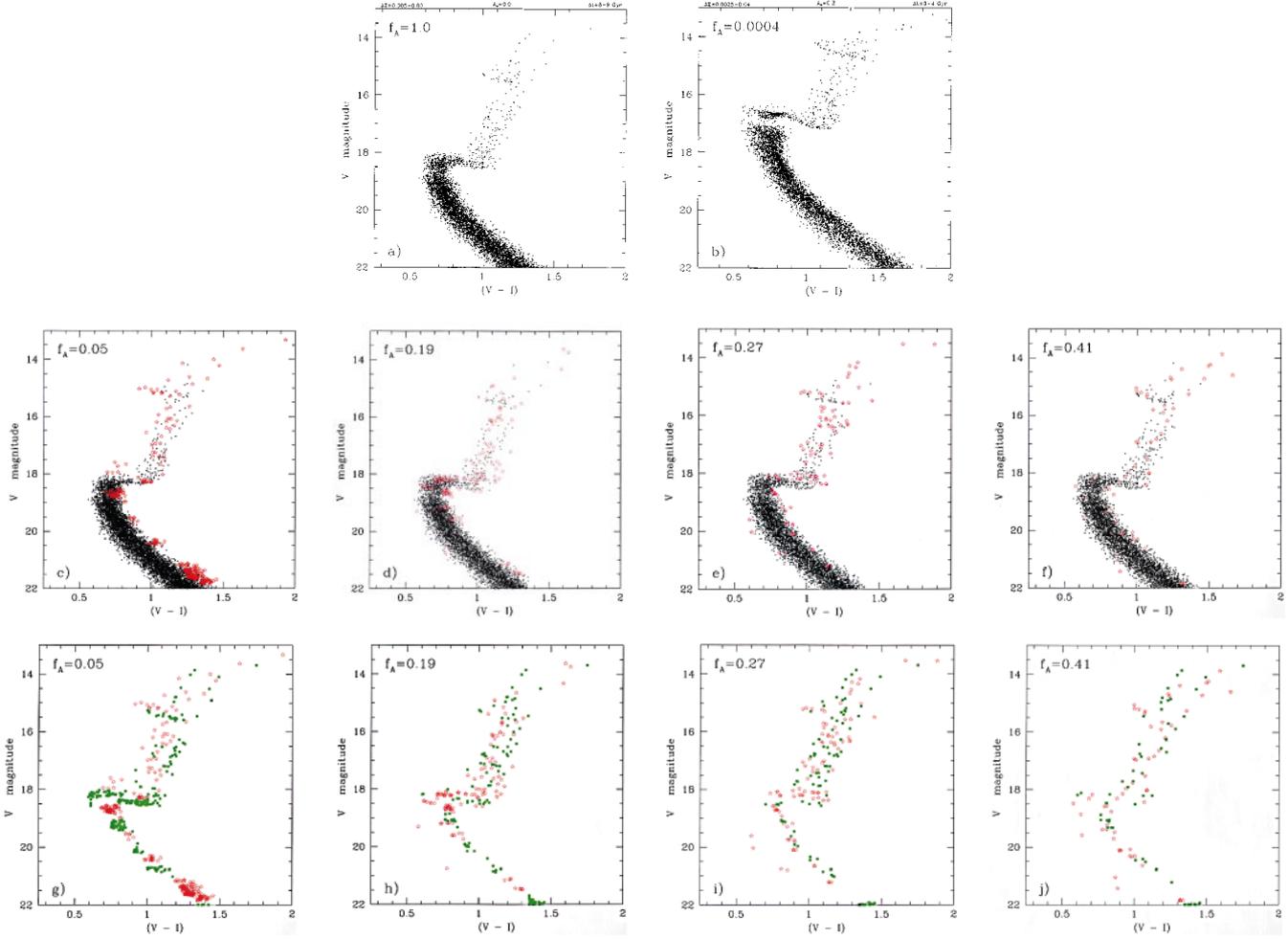
Fig. 6. Genetic evolution of the colour-magnitude diagram (CMD) from the first test population. Panel **a)** displays the original population to be matched. The physical parameters for this population are described Table 2 and Sect. 4.2. The CMD of the initial trial population is shown in panel **b)**. Panels **c)**–**e)** display the resulting CMDs obtained with setup A.1-40 for different fitnesses (see Sect 2.7). The fitnesses $f = 0.05$, $f = 0.19$ and $f = 0.27$ are respectively reached after 20, 60 and 80 generations. The dots in the panels **b)**–**e)** are used for each matching point, while the red open stars ★ in panels **c)**–**j)** are the points in the simulation which have no counterpart in the original CMD. Panel **f)** displays the fitness $f = 0.41$ as obtained after 341 generations. Note that the CMDs of panels **a)** and **f)** as well as **c)**–**e)** are visually almost indistinguishable. Panels **g)**–**j)** displays the residuals between the simulated and the original CMD (**a**; green solid squares ■) are those points in the original which have no counterpart in the simulation.

age-metallicity degeneracy. Excluding this value results in an average fitness of $\overline{f_A} = 0.328 \pm 0.039$. In general: the extinction can be reliably retrieved when fixing the distance.

When considering a fixed extinction, the results show a strong variation in both age and metallicity. It should also be noted that the average $^{10}\log(\text{distance (pc)})$ retrieved is only 3.8953 ± 0.0066. This is more than one sigma away from the optimum value for the distance (see Table 4). This is an indication that retrieval of the distance by fixing the extinction is hampered by the age-metallicity degeneracy. Therefore, the distance cannot be reliably retrieved when fixing the extinction to its correct value.

5.3.2. Fixed age and metallicity

Fixing one of the age limits results in values for both the age and metallicity which are close to the input values. This is due to the (partial) breaking of the age-metallicity degeneracy. The distance-extinction degeneracy remains. The results also suggest that the age-metallicity degeneracy has a stronger impact on the fitness than the distance-extinction degeneracy.

Fixing the upper metallicity limit to its correct value shows that the values for age and metallicity come closer to their original, input values. This is quite in contrast with the results obtained from fixing the lower metallicity to its correct value. Table A.2 shows that both the high metallicity limit and the slope of the exponential SFR are not well constrained. This behaviour can be accounted to the implicit shape of the linear age-metallicity relation adopted in the HRD-GST. The number of high metallicity stars is smaller than the number of low metallicity stars due to the adopted, exponentially decreasing ($\beta = 1$), star formation rate. The consequence is that the high metallicity
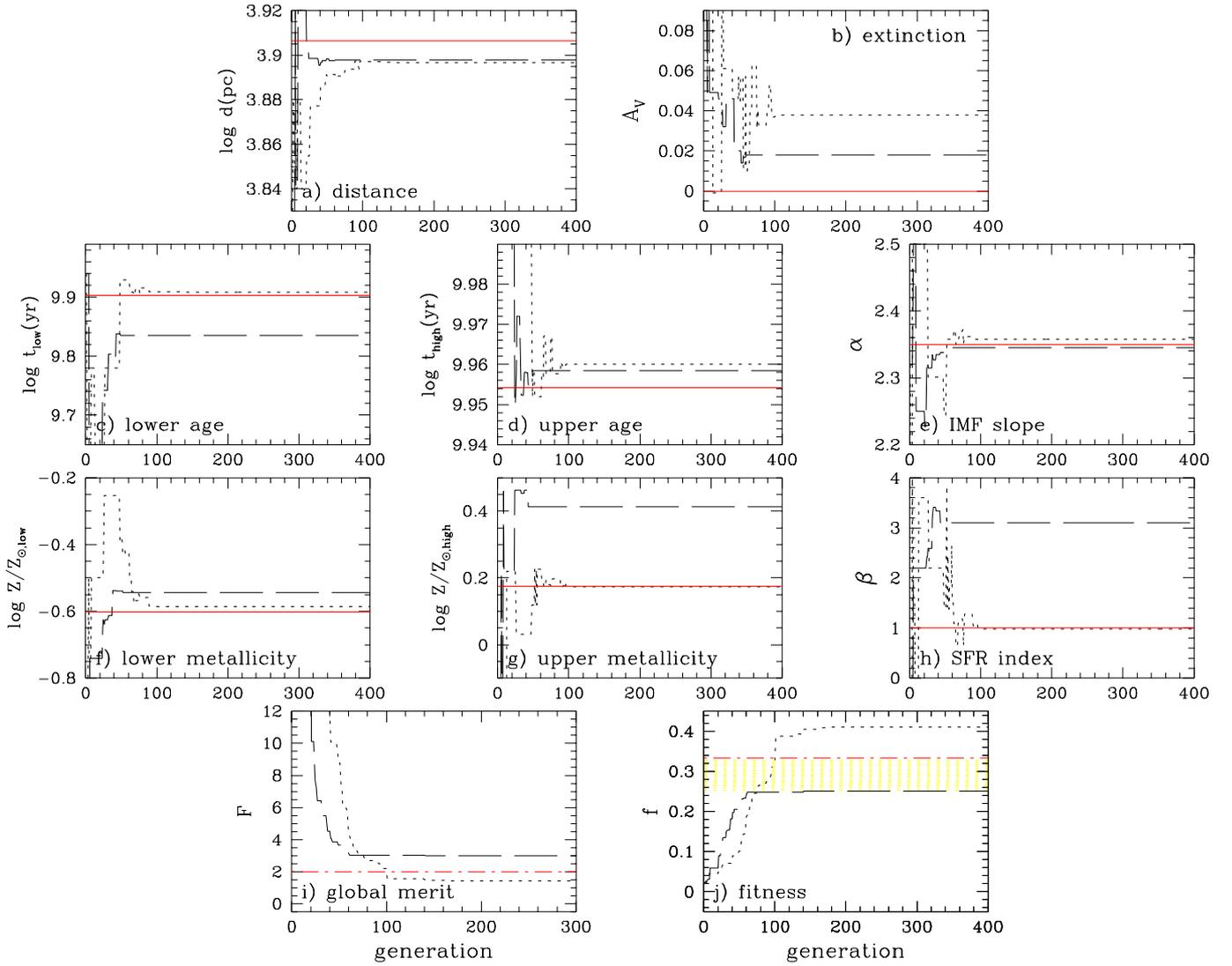
Fig. 7. Panels **a)**–**h)** display the convergence curves for the parameters of models A.1-40 (dotted line; fitness $f = 0.41$) and A.1-51 (long dashed line; fitness $f = 0.25$). The solid line in the frames **a)**–**f)** refers to the value adopted for the original population. The long, dot dashed line in frames **i)** and **j)** shows the threshold values to be crossed for acceptable solutions, i.e. $F < 2$ and $\frac{1}{3} < f < 1$. The short dashed area in frame **j)** marks the region where degeneracy of the parameter space becomes noticeable (see Sects. 6.3 for details).

limit can be better determined when both the low and high limits are determined in union.

5.3.3. Fixed slope for the power-law IMF

Fixing the slope α of the power-law IMF to its correct value ensures stability in the magnitude direction of the CMD. This implies that the degeneracy in the colour dependent parameters, like age, metallicity, and partly the star formation rate, becomes more apparent. Although the overall results are quite good, only model A.2-14 is significantly affected by this degeneracy. If we leave out model A.2-14 from the statistics the average fitness becomes $\overline{f_A} = 0.331 \pm 0.020$.

5.3.4. Fixed slope for the exponential SFR

The exponential star formation rate parameter β is tied to the age and metallicity range, see also Sect. 5.3.2. Fixing the parameter β better constrains in particular the upper metallicity limit. However, it does not avoid that the genetic evolution enters into a local age-metallicity gap, see Table A.2. Excluding model A.2-34 improves the average fitness in Table 5 to $\overline{f_A} = 0.342 \pm 0.029$.

5.4. Test 4: Fixing parameters at the wrong value

The results of fixing the parameters at a $\pm 1\sigma$ offset from its original value are listed in Table A.3. An example of the diagnostics of these tests are given in Table 8. Details of individual setups are given below.

Table 6. Description of the diagnostic statistics for model A.1-40 when fixing one parameter at its correct value, see Sects. 5.3, 2.7 and 6.2.1 for additional details.

model	$N_{O,not}$	$N_{S,not}$	N_{match}	F_χ	F_P	F	f_A
ideal	0	0	5000	0.000	0.000	0.000	1.000
free	60	60	4940	0.847	0.849	1.438	0.410
fixed: $\log d$ (pc)	155	155	4845	0.903	2.192	5.621	0.151
fixed: A_V	103	103	4897	0.840	1.457	2.827	0.261
fixed: $\log t_{low}$ (yr)	76	76	4924	0.817	1.075	1.823	0.354
fixed: $\log t_{high}$ (yr)	104	104	4896	0.902	1.471	2.978	0.251
fixed: $[Z]_{low}$	94	94	4906	0.836	1.329	2.466	0.289
fixed: $[Z]_{high}$	76	76	4924	0.815	1.075	1.820	0.355
fixed: α	74	74	4926	0.834	1.047	1.799	0.357
fixed: β	80	80	4920	0.727	1.131	1.808	0.356

Table 7. Fitness statistics when fixing one parameter $\pm 1\sigma$ from its original value. Averaged fitness values $\overline{f_A}$ and their associated standard deviation σ_{n-1} are obtained from simulations with the setup parameters from models 9, 14, 22, 34, 40 and 52. See Table A.2 for additional details.

parameter	offset	$\overline{f_A}$	σ_{n-1}
$\log d$	-1σ	0.291	0.048
	$+1\sigma$	0.280	0.066
A_V	$+1\sigma$	0.284	0.020
$\log t_{low}$	-1σ	0.258	0.044
	$+1\sigma$	0.279	0.012
$\log t_{high}$	-1σ	0.229	0.022
	$+1\sigma$	0.207	0.014
$[Z]_{low}$	-1σ	0.254	0.066
	$+1\sigma$	0.315	0.027
$[Z]_{high}$	-1σ	0.270	0.054
	$+1\sigma$	0.283	0.018
α	-1σ	0.266	0.025
	$+1\sigma$	0.273	0.027
β	-1σ	0.232	0.071
	$+1\sigma$	0.276	0.011

5.4.1. Erroneous distance

A wrong assumption about the distance gives in a vertical shift in the CMD (all other parameters give a diagonal shift) and cannot be masked out through a correlated change of any of the other parameters. This results in quite a wide range in the values of the other parameters, except the value of α. The unexpected result is that irrespective if the distance is too short or too far: the power-law IMF slope flattens!

If the distance is overestimated there are more synthetic stars present at fainter magnitudes. To get relatively more synthetic stars at brighter magnitudes one needs to flatten the power-law IMF slope.

If the distance is underestimated then more synthetic stars are present at brighter magnitudes. One expects that a steeper power-law IMF slope is required as compensation. This is not always true, see Table A.3. Stars pop up at the lower end of the main sequence. They are taken away from the stars located at

brighter and brighter magnitudes. One therefore requires also in this case a flatter IMF slope.

A flatter slope of the power-law IMF can be a hint that the distance of the stellar aggregate is wrong. Or it might be a hint that the zero point of the adopted synthetic photometric system is different from the actual photometric system used.

Recognizing that the slope is indeed flatter than the majority of the other cases outlined in Table 4 one may start to explore the assumption that the distance is wrong: release the constraint during the next exploration.

The sensitivity to the distance implies that AMORE can be used to determine the distance to a stellar aggregate quite reliably. A bonus is that due to an initially wrongly assumed distance the extinction is in most cases better constrained.

5.4.2. Erroneous extinction, age and metallicity

In $(V, V-I)$ CMDs a strong correlation between extinction, age and metallicity exists (see also Ng & Bertelli 1996 and references cited therein). A higher value of the extinction can be compensated by a younger age and/or a lower metallicity. Indeed, the results in Table A.3 show that this actually occurs for the lower age and metallicity limit. The upper age and metallicity limit, however, drifts away in the opposite direction to compensate for "erroneous corrections" applied by other parameters.

A higher value of the star formation index results in a lower number of stars at the upper age metallicity limit. To get a sufficient number of high metallicity stars one has to stretch the upper metallicity limit to a slightly higher value.

We further notice that a wrong value for the age and metallicity does not affect the extinction significantly. Our findings indirectly supports the method to determine high resolution ($4' \times 4'$) extinction maps towards the Galactic bulge by Schultheis et al. (1999) with the data obtained for the DeNIS project (Epchtein et al. 1997).

5.4.3. Erroneous slope for power-law IMF

The slope of the power-law IMF is very strongly constrained (Ng 1998) for the test population. As a consequence the changes in the values of the remaining parameters are not

extremely large. The slightly larger value of the slope α pushes slightly more synthetic stars to fainter magnitudes, introducing a relative deficiency of stars at brighter magnitudes. This is compensated through a younger age, a decrease of the lower metallicity limit and an increase of the upper metallicity limit. The higher value for the upper age and metallicity limit compensates in its turn for the overestimation of the exponential star formation index.

A lower value of α is partly compensated for by lowering the upper age limit, lowering the lower metallicity limit and overestimating the upper metallicity limit.

5.4.4. Erroneous index for exponential SFR

A larger index for an exponentially decreasing SFR pushes more stars of the population to the blue edge of the CMD, resulting in a slightly bluer stellar population. AMORE compensates this by mainly increasing the upper metallicity limit, i.e. reddening the synthetic stellar population.

A lower value for the SFR index has the opposite effect. AMORE compensates for the now slightly redder population by lowering the upper metallicity limit, making the population bluer.

6. Discussion

6.1. Relative contribution of the parameters to the fitness

Figure 4 and the Tables 4, 5, 7, A.2 and A.3 provide the hint that not all parameters have an equal contribution to the fitness. It appears that the β and $[Z]_{\text{high}}$ parameter can vary considerably and still yield a decent fitness. Moreover, Tables A.2 and 5 indicate that knowledge of the value of the $[Z]_{\text{high}}$ parameter results in acceptable values for the other parameters.

The origin of this behaviour lies in the implicit definition of the exponential star formation rate (for $\beta = 1$ one has a decreasing star formation towards a younger age) attached to a linear age-metallicity relation. The latter relation will give less metal-richer stars. The small number of stars with higher metallicity induces a larger variation of the $[Z]_{\text{high}}$ parameter without affecting significantly the overall fitness.

6.2. Convergence

The fine-tuning of the genetic algorithm is a tedious task. It is not straightforward to find the optimum setting for the problem to be solved. One has to balance the exploring quality through crossovers against the variation of the parameters through (creep) mutations.

We did not want to deal with a mutation dominated search, because it tends to move farther away from an optimum parameter setting in the majority of the cases. We used therefore a relatively high crossover probability (`pcross`) and we set the mutations at a fixed rate, such that on average only 2.8 mutations occur in the gene pool of each individual.

At a certain stage however one requires the variation of other correlated parameters to obtain an improvement. This becomes particularly necessary when approaching the optimum setting of the parameters. A favourable crossover and mutation might do the trick, but it might take a while before this occurs. We introduced in Sect. 2.6.5 the possibility that two parameters might be more sensitive to mutations than others. This approach gave better results for the majority of the trial cases (see Table 3), but it failed to obtain improvements when changes of one parameter were neutralized through the variation of one or more parameters. The distance-extinction and the age-metallicity degeneracies slow down the convergence of AMORE for $f > 0.3$, see Fig. 4.

One of the modifications to consider for future implementation is a two-chromosome approach. In that case acceptable values for the parameters do not shift out of the population if the overall fitness is less, but still reside in the gene pool as a recessive quality. This however, will require a major extension to PIKAIA and a significant amount of genetic research to be done about dominant and recessive qualities in the AMORE gene pool.

Another modification to consider in order to improve the accuracy and to speed up convergence, is to replace the finite resolution of the digital encoding scheme with a genetic coding based on floating point, i.e. each gene on the chromosome is represented by one floating point number. According to Michalewicz (1996) a real encoding scheme can be superior and improve convergence. Such an encoding scheme is indeed to be included in the next release of PIKAIA 2.0 (Charbonneau; in preparation).

6.2.1. Unstable solutions

In one test (fixing A_{V} at one sigma above the original value for model 9) no convergence was achieved and the run was aborted. Because AMORE is quite sensitive to rounding these effects can be circumvented by slightly altering the input parameters. We decided in this case against such an action, because that would make the sample inhomogeneous.

6.3. Degeneracy

Isochrones for a particular age and metallicity can be mimicked with another set of isochrones of different age and metallicity (Worthey 1994; Charlot et al. 1996, and references cited therein). This degeneracy of the parameter space increases if one considers the distance and the extinction towards a stellar aggregate. There is no straight forward method to circumvent partial degeneracy of the parameters to be explored. One might consider to apply AMORE for the analysis of colour-colour diagrams in order to rule out the distance, to determine the extinction and a number of other parameters, and finally to determine the distance to the stellar population from one of the CMDs. The combined analysis of colour-magnitude and colour-colour diagrams is expected to improve the results obtained by AMORE so far. However, this requires that at least one additional colour should be available for each star considered above a certain detection and completeness threshold. Moreover, as was mentioned in Sect. 5.3.1, knowledge of the

Table 8. Description of the diagnostic statistics for simulations with setup parameters from model A.1-40. One of the parameters is forced to a value $\pm 1\sigma$ from its original value. See Sects. 5.4 and 2.7 for additional details.

model	offset	$N_{\rm O,not}$	$N_{\rm S,not}$	$N_{\rm match}$	F_χ	$F_{\rm P}$	F	$f_{\rm A}$
ideal		0	0	5000	0.000	0.000	0.000	1.000
free		60	60	4940	0.847	0.849	1.438	0.410
fixed: $\log d$ (pc)	-1σ	91	91	4909	0.788	1.287	2.278	0.305
	$+1\sigma$	85	85	4915	0.722	1.202	1.967	0.337
fixed: $A_{\rm V}$	$+1\sigma$	98	98	4902	0.932	1.386	3.790	0.264
fixed: $\log t_{\rm low}$ (yr)	-1σ	91	91	4901	0.897	1.287	2.460	0.289
	$+1\sigma$	94	94	4906	0.905	1.329	2.587	0.278
fixed: $\log t_{\rm high}$ (yr)	-1σ	106	106	4894	0.977	1.499	3.202	0.238
	$+1\sigma$	126	126	4874	1.017	1.782	4.211	0.192
fixed: $[Z]_{\rm low}$	-1σ	78	78	4922	0.907	1.103	2.043	0.329
	$+1\sigma$	80	80	4920	0.835	1.131	1.977	0.336
fixed: $[Z]_{\rm high}$	-1σ	148	148	4852	0.896	2.094	5.188	0.162
	$+1\sigma$	94	94	4906	0.867	1.329	2.519	0.284
fixed: α	-1σ	95	95	4905	0.886	1.344	2.590	0.279
	$+1\sigma$	103	103	4897	0.894	1.457	2.921	0.255
fixed: β	-1σ	145	145	4855	0.930	2.051	5.071	0.165
	$+1\sigma$	103	103	4897	0.834	1.457	2.818	0.262

extinction does not automatically imply that the distance can be retrieved accurately.

As demonstrated in Sect. 5.1.1 the degeneracy among parameters becomes noticeable for $f > 0.25$ or $F < 3$. This corresponds to systematic offset for each parameter of on average $\sim 0.6\sigma_k$ and at maximum $\sim 1\sigma_k$. The Poisson uncertainty of the original population results in a fitness of $f = 0.43$ ($F \simeq 1.33$). However, solutions with a comparable fitness do exist due to the degeneracy of the parameter space. A direct consequence is that there is an intrinsic offset present among the parameters amounting to on average $\sim 0.4\sigma_k$ and at maximum $\sim 0.7\sigma_k$. This intrinsic offset is present in the solutions obtained with AMORE and actually is responsible for slowing down the convergence in the fitness range $0.30 < f < 0.43$. It will therefore be nearly impossible to recover in one pass the original input values. However, some improvements might be obtained by averaging the parameter values obtained from AMORE runs with different initial conditions.

7. Conclusions

We demonstrate that an automatic search can be made for the astrophysical parameters of a synthetic stellar population from the analysis of colour-magnitude diagrams with an optimizer, based on a genetic algorithm. However, AMORE tends to slightly underestimate the distance. It subsequently attempts to compensate this with an higher extinction, a higher upper age and a slightly steeper slope for the power-law IMF. At $f > 0.3$ the combined effect of the age-metallicity and the distance-extinction degeneracy slows down the convergence. The data suggests that AMORE has more problems dealing with the age-metallicity than with the distance-extinction degeneracy. For general purpose, however, the retrieved values are sufficiently accurate.

8. Future work

The good results obtained so far for a single synthetic stellar population is an indication about AMORE's potential for the detailed analysis of CMDs. The next step is to improve one step at a time various aspects of AMORE before it can be used as an interpretative tool for large photometric surveys.

Despite limitations in the input physics of the underlying stellar evolutionary tracks and by the transformation from the theoretical to the observational plane, the results with real data from Gallart et al. (1999), who uses the same set of evolutionary tracks, are encouraging. It will therefore be important to verify first with, for example, well studied open clusters (see Carraro et al. 1998, 1999 and references cited therein) for which age and metallicity range we may apply AMORE safely. Extinction is also of some concern, because a high extinction may result in the MS turnoff point to fall below the detection limit. This would deprive AMORE of a clear reference point.

Another case of interest is of course the question how many different stellar populations can be distinguished with AMORE. Separating multiple, mixed populations from each other through the automated and the objective analysis of colour-magnitude diagrams could be a valuable tool for the analysis of galaxy formation and evolution.

This requires a rigorous follow up study on the separation of multiple (synthetic) populations. One further ought to verify if the automated analysis of colour-colour diagram can reduce the effects of error cancellation between distance and extinction.

Finally, after a succesful implementation, testing and validation phase, we plan to combine AMORE with the Padova spectrophotometric code (see Bressan et al. 1994, 1996; Tantalo et al. 1996, 1998a,b). A synthetic population has to be generated, containing sub-populations with different ages and metallicities. Then a synthetic spectrum must be generated for the mixed population and subsequently used as input

for a synthetic, spectral fitting program to determine the underlying stellar populations. In this way one can establish the calibration of the spectrophotometric tool in a self-consistent way. Furthermore, an implicit verification can be made that the populations are consistent with those obtained from a CMD analysis with AMORE.

Appendix A: AMORE test results, the data

The appendix contains the data of the simulations as discussed in Sect. 4. The first test, in which we explored different values for the PIKAIA control parameters, is described in Sect. 4.3.1 and the data are given in Table A.1.

The data of the third test, in which one of the astrophysical parameters was fixed at its correct value as discussed in Sect. 4.3.3, are given in Table A.2. The data on the fourth test, in which one parameter was fixed one sigma from its correct value as discussed in Sect. 4.3.4, are given in Table A.3.

The tables in the Appendix are available in electronic form at http://www.edpsciences.org and at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/392/1129.

References

Aparicio, A. 1999, Proc. IAU Symp. 192, The stellar content of the Local Group galaxies, 7–11 Sep. 1998, ed. P. Whitelock, & R. Cannon, ASP Conf. Ser., 304

Banzhaf, W., Nordin, P., Keller, R. E., & Francone, F. D. 1998, Genetic Programming (Morgan Kaufmann Publishers)

Beichman, C. A., Chester, T. J., Skrutskie, M., Low, F. J., & Gillett, F. 1998, PASP, 110, 480

Bobinger, A. 2000, A&A, 357, 1170

Bertelli, G., Bressan, A., & Chiosi, C. 1985, A&A, 150, 33

Bertelli, G., Betto, R., Bressan, A., et al. 1990, A&AS, 85, 845

Bertelli, G., Mateo, M., Chiosi, C., & Bressan, A. 1992, ApJ, 387, 320

Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&AS, 106, 275

Bertelli, G., Bressan, A., Chiosi, C., Ng, Y. K., & Ortolani, S. 1995, A&A, 301, 381

Bertelli, G., Bressan, A., Chiosi, C., & Ng, Y. K. 1996, A&A, 310, 115

Bessell, M. S., & Brett, J. M. 1988, PASP, 100, 1134

Bressan, A., Fagotto, F., Bertelli, G., & Chiosi, C. 1993, A&AS, 100, 647

Bressan, A. C., Chiosi, & Fagotto, F. 1994, ApJS, 94, 63

Bressan, A. C., Chiosi, & Tantalo, R. 1996, A&A, 311, 425

Carraro, G., Ng, Y. K., & Portinari, L. 1998, MNRAS, 296, 1045

Carraro, G., Girardi, L., & Chiosi, C. 1999, MNRAS, 309, 430

Charbonneau, P. 1995, ApJS, 101, 309

Charbonneau, P., & Knapp, B. 1996, A User's Guide to PIKAIA 1.0, NCAR Technical Note 418+IA (Boulder: National Center for Atmospheric Research)

Charbonneau, P., Tomczyk, S., Shou, J., & Thompson, M. J. 1998, ApJ, 496, 1015

Charlot, S., Worthey, G., & Bressan, A. 1996, ApJ, 457, 625

Chiosi, C. 1996, in Proc. From Stars to Galaxies, 9–13 October, 1995, Elounda (Crete; Greece), ed. C. Leitherer, U. F. von Alvensleben, & J. Huchra, ASP Conf. Ser., 98, 181

Chiosi, C., Bertelli, G., Meylan, G., & Ortolani, S. 1989, A&A, 219, 167

da Costa, L. 1997, Messenger, 88, 34

da Costa, L., Bertin, E., Deul, E., et al. 1998, Messenger, 91, 49

Davis, L. 1991, Handbook of Genetic Algorithms (New York: van Nostrand Reinhold)

De Jong, K. A. 1993, in Foundations of Genetic Algorithms 2, ed. L. D. Whitley (San Mateo; Morgan Kaufmann), 5

Dolphin, A. 1997, New Astron. 2, 397

Dolphin, A. E., Walker, A. R., Hodge, P. W., et al. 2001, ApJ, 562, 303

Dolphin, A. 2002, MNRAS, 332, 91

Epchtein, N., Batz de, B., Capoani, L., et al. 1997, ESO Messenger 87, 27

Epchtein, N., Deul, E., Derriere, S., et al. 1999, A&A, 349, 236

Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994a, A&AS, 104, 365

Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994b, A&AS, 105, 29

Fagotto, F., Bressan, A., Bertelli, G., & Chiosi, C. 1994c, A&AS, 105, 39

Fan, X. 1999, AJ, 117, 2528

Gallart, C., Aparicio, A., Bertelli, G., & Chiosi, C. 1996, AJ, 112, 1950

Gallart, C. 1998, ApJ, 495, L43

Gallart, C., Freedman, W. L., Aparicio, A., Bertelli, G., & Chiosi, C. 1999, AJ, 118, 2245

Geha, M. C., Holtzman, J. S., Mould, J. R., et al. 1998, AJ, 115, 1045

Gibson, S. E., & Charbonneau, P. 1998, JGR 103, No. A7, 14511

Gilmore, G. F., Perryman, M. A., & Lindegren, L. 1998, SPIE, 3350, 541

Goldberg, D. E. 1989, Genetic Algorithms in Search, Optimization & Machine Learning, (Reading, MA: Addison-Wesley)

Harris, J., & Zaritsky, D. 2001, ApJS, 136, 25

Hernandez, X., Valls-Gabaud, D., & Gilmore, G. 1999, MNRAS, 304, 705

Holland, J. H. 1975, Adaptation in Natural and Artificial Systems (1st ed.; Ann Harbor: University of Michigan Press – 2nd ed.; 1992, Cambridge: MIT Press)

Holtzman, J. A., Mould, J. R., Gallagher, J. S., et al. 1997, AJ, 113, 656

Holtzman, J. A., Gallagher, J. S., Cole, A. A., et al. 1999, AJ, 118, 2262

Iglesias, C. A., Rogers, F. J., & Wilson, B. G. 1992, ApJ, 397, 717

Kaastra, J. S., Mewe, R., Liedahl, D. A., et al. 1996, A&A, 314, 547

Kennelly, E. J., Walker, G. A. H., Catala, C., et al. 1996, A&A, 313, 571

Lamontagne, R., Moffat, A. F. J., Drissen, L., Robert, C., & Matthews, J. M. 1996, AJ, 112, 2227

McIntosh, S. W., Diver, D. A., Judge, P. G., et al. 1998, A&AS, 132, 145

Metcalfe, T. S. 1999, AJ, 117, 2503

Mewe, R., Kaastra, J. S., White, S. M., & Pallavich, R. 1996, A&A, 315, 170

Michalewicz, Z. 1996, Geneticalgorithms + datastructures = evolutionprograms, 3rd revised and extended edition (Springer-Verlag)

Ng, Y. K. 1994, Ph.D. Thesis, Leiden University, The Netherlands

Ng, Y. K. 1997a, in Proc. The impact of large-scale near-IR sky surveys, 24–26 April, 1996, Puerto de la Cruz (Tenerife; Spain), ed. P. Garzon-Lopez et al., 91

Ng, Y. K. 1997b, in Proc. 12th IAP Colloquium Variable stars and the astrophysical returns from microlensing surveys, 8–12 July, 1996, Paris (France), ed. R. Ferlet, J.-P. Maillard, & B. Raban, 113

Ng, Y. K. 1998, A&AS, 132, 133

Ng, Y. K., & Bertelli, G. 1996, A&A, 315, 116

Ng, Y. K., Bertelli, G., Bressan, A., Chiosi, C., & Lub, J. 1995, A&A, 295, 655 (Erratum A&A 301, 318)

Ng, Y. K., Bertelli, G., Chiosi, C., & Bressan, A. 1996, A&A, 310, 771

Ng, Y. K., Bertelli, G., Chiosi, C., & Bressan, A. 1997, A&A, 324, 65

Noyes, R. W., Lha, S., Korzennik, S. G., et al. 1997, ApJ, 483, L111

Olsen, K. A. G. 1999, AJ, 117, 2244

Paczyński, B., Udalski, A., & Szymanski, M. 1999, Acta Astron., 49, 319

Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, A&A, 369, 339

Pagel, B. E. J., & Portinari, L. 1998, MNRAS, 298, 747

Powell, M. J. D. 1964, Comput. J., 7, 155

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vettering, W. T. 1986, Numerical Recipes, 2nd edition (Cambridge University Press)

Renzini, A., & da Costa, L. 1997, Messenger, 87, 23

Rieke, G. H., & Lebofsky, M. J. 1985, ApJ, 288, 618

Saha, P. 1998, ApJ, 115, 1206

Schultheis, M., Ganesh, S., Simon, G., et al. 1999, A&A, 349, L69

Skrutskie, M. 1998, in The impact of near-infrared surveys on galactic and extragalactic astronomy, Proc. 3rd Euroconference (Kluwer), ASSL 230, 11

Stetson, P. B., & Harris, W. E. 1988, AJ, 96, 909

Tantalo, R., Chiosi, C., Bressan, A., & Fagotto, F. 1996, A&A, 311, 361

Tantalo, R., Chiosi, C., & Bressan, A. 1998, A&A, 333, 419

Tantalo, R., Chiosi, C., Bressan, A., Marigo, P., & Portinari, L. 1998, A&A, 335, 823

Udalski, A., Kubiak, M., & Szymański, M. 1997, Acta Astron., 47, 319

van de Hulst, H. C. 1949, Rech. Astr. Obs. Utrecht, Vol. 11, part 2

Wahde, M. 1998, A&AS, 132, 417

Worthey, G. 1994, ApJS, 95, 107