Boletim da Sociedade Astronômica Brasileira, **31**, no. 1, 5-6

The SIRIUS code: Statistical Inference of physical paRameters of sIngle and mUltiple populations in Stellar clusters

S. O. Souza[1], R. A. P. Oliveira[1], L. O. Kerber[2], B. Barbuy[1], A. Pérez-Villegas[1], & F. F. S. Maia[1]

[1] Instituto de Astronomia, Geofísica e Ciências Atmosféricas – IAG-USP e-mail: stefano.souza@usp.br
[2] Universidade Estadual de Santa Cruz – UESC

Abstract. Star clusters are fundamental pieces to understand the formation and evolution of stellar systems, from a scale of few AU (binaries) to several kpc (galaxies), being at the same time unique templates to calibrate stellar evolutionary models. The physical parameters of star clusters – such as age, [Fe/H], distance, reddening, core and tidal radii, total mass, mass function, binary fraction, proper motions, and radial velocities – can be obtained by the analysis of color-magnitude diagrams (CMDs), radial density profiles, spectra of individual stars, multi-epoch astrometry, multi-band photometry, among others. Despite the wealth of information available in high quality photometric, spectroscopic and/or astrometric data from space missions (e.g., HST, Gaia), 8-10m (e.g., VLT, Gemini) and 1-4m (e.g., Blanco, SOAR, VISTA) ground-based telescopes, including those dedicated to large area surveys (e.g., OGLE, VVV, VMC, VISCACHA, DES, SMASH, APOGEE, Gaia-ESO), there is a lack of initiatives to perform a comprehensive, statistical and self-consistent analysis combining different data and techniques. In order to fulfill this gap, we are developing a python package, named SIRIUS, designed to extract as much as possible information of a stellar cluster. SIRIUS is based on a Bayesian approach using the likelihood statistics and Markov chain Monte Carlo (MCMC) method, which can be applied to determine membership probabilities and to fully characterize single or multiple stellar populations presented in a stellar cluster. In this work we present the first control experiments to attest the SIRIUS validity, as well as some successful applications of this code to analyze individual or combined HST, VLT, Gemini, VVV, and Gaia DR2 data of bulge globular clusters.

Resumo. Aglomerados estelares são peças fundamentais para o entendimento da formação e evolução de sistemas estelares, desde a escala de UA (binárias) até alguns kpc (galáxias), servindo ao mesmo tempo para calibrar modelos de evolução estelar. Os parâmetros físicos de aglomerados estelares – tais como idade, [Fe/H], distância, avermelhamento, raios de core e de maré, massa total, função de massa, fração de binarias, movimento próprio, e velocidade radial – podem ser obtidos através da análise de diagramas cor-magnitude (CMD), perfís de densidade radial, espectro de estrelas individuais, astrometria multi-época, espectrocopia multi-banda, entre outros. Apesar da riqueza de informação acessível em fotometria de alta qualidade, dados espectrocópicos e/ou astrométricos vindos de missões espaciais (HST, GAIA), 8-10m (VLT, Gemini) and 1-4m (Blanco, SOAR, VISTA), telescópios terrestres, incluindo aqueles dedicados surveys de grande área (OGLE, VVV, VMC, VISCACHA, DES, SMASH, APOGEE, Gaia-ESO), existe uma falta de iniciativas para realizar uma análise robusta, estatistica e auto-consistente combinando diferentes técnicas e dados. Com o objetivo de preencher essa lacuna, estamos desenvolvendo um código em Python, chamado SIRIUS, para extrair tantas informações quanto possível de um aglomerado estelar. SIRIUS é baseado em uma abordagem Bayesiana combinada com o método Monte Carlo de Cadeias de Markov (MCMC), que pode ser aplicado para determinar probabilidades de membership e caracterizar aglomerados estelares no contexto de multiplas populações bem como população única. Neste trabalho apresentaremos os primeiros experimentos controlados para testar a validade do SIRIUS, e também algumas aplicações bem sucedidas do código para dados individuais ou combinados do HST, VLT, Gemini, VVV, e Gaia DR2 de aglomerados globulares do bojo da Galáxia.

Keywords. Methods: statistical – globular clusters: individual: NGC 6558, HP1 – Stars: evolution

1. Introduction

SIRIUS is a Python code based on a Bayesian approach of Isochrone fitting which can be applied to determine the density of probabilities of the cluster physical parameters using the Metropolis-Hastings algorithm of Markov chain Monte Carlo (MCMC). Also, SIRIUS makes use of Machine Learning Classification methods to isolate the multiple generations when it is available. SIRIUS is designed to analyze different photometric systems with diverse stellar evolutionary models (BaSTI, DSED, etc). Furthermore, this code possesses the advantage of analyzing both synthetic and real CMDs (Kerber et al. 2007).

2. Method

SIRIUS' statistical approach came from Bayes' theorem:

$$P(\varphi|D) = \frac{P(D|\varphi) \times P(\varphi)}{P(D)}$$

where φ and D represent the parameter space (model) and the data (observational), respectively. The term $P(D|\varphi)$ represents the Likelihood distribution, $P(\varphi)$ are the *a priori* distributions about parameter space and they are called Priors. Once does not depend on the parameter space, the marginal distribution $P(D)$ sometimes is used as normalization constant for the likelihood. The Posterior distributions $P(\varphi|D)$ are distributions *a posteriori* of the observational data. The Bayes' theorem can be written, for convenience, in logarithm as:

$$\ln P(\varphi|D) = \ln P(D|\varphi) + \ln P(\varphi)$$

We assumed that the data are spread following a Gaussian distribution to establish our Likelihood function:

$$\ln P(D|\varphi) = \chi^2_{Color} + \chi^2_{Magnitude} \rightarrow \chi^2_{\xi} = \sum_{i=1}^{N} \frac{-1}{2} \left(\frac{\xi_i^{OBS} - \xi_i^{Iso}}{\sigma_i^{\xi}} \right)^2$$

where ξ represents *Color* or *Magnitude*. In this case, the maximum value of likelihood will be given when all isochrone points

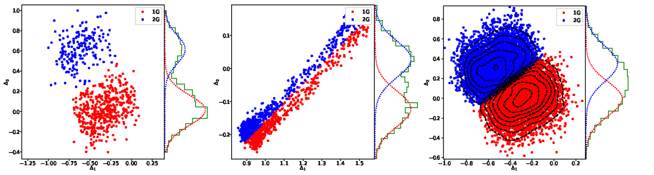

Figure 1. The upper panels represent the separation method for the RGB and SGB, respectively, and for the MS is shown in the lower-left panel.

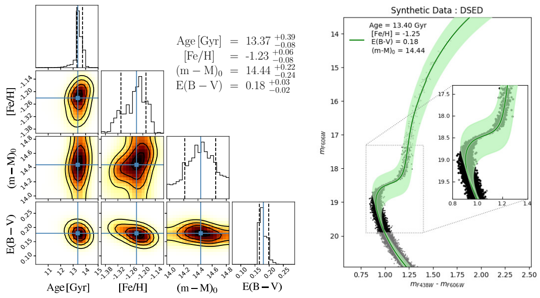

Figure 2. The left panel shows the 2D projection of parameter space and the posteriors distributions. The right one is the region of the best solutions over the CMD.

are closest to data. In order to get the probability distributions of each free parameters, we are applying the MCMC, that is a tool capable to explore the parameter space looking for the best probability distribution of a problem.

3. Control Experiments

Control experiments are used to prove the validity of a tool. In the SIRIUS' case we built synthetic CMDs in order to get the input parameters after the isochrone fitting. The synthetic CMD was made from the HST UV-Legacy Survey catalogue allowing the multiple populations analysis. The input parameters of the synthetic CMD were: Age = 13.5Gyr, [Fe/H] = -1.25, E(B-V) = 0.18, (m-M)$_0$ = 14.4 and fraction of first generation stars (N_{1G}/N_{total}) = 0.69.

We used the methods described in Milone et al. 2017 (RGB and MS – using Chromosome Map) and Nardiello et al. 2015 (SGB – using a conventional color-color diagram with UV-filters) to separate the multiple populations. An algorithm of Classification based on k-Nearest Neighbors (kNN) method using Gaussian Mixture Models (GMM) was applied in this method. The Figure 1 shows the result of the multiple populations. The output value of the fraction of first generation stars was 0.70 ± 0.05.

The isochrone fitting experiment was performed considering the cluster as one population. Once the populations are isolated it is possible to apply the SIRIUS code in each one separately (in the contribution from Oliveira et al. in this proceedings there are applications to the context of the multiple populations). The result of the experiment is in Figure 2.

4. First Result: HP1

SIRIUS was applied to analyze the HP1, a Bulge globular cluster candidate to a relic of the Galaxy. The results are written on a paper recently accepted in MNRAS. In this paper we analyzed proper-motion-cleaned CMDs from Gemini/GSAOI+GeMS and HST multi-epoch photometry. Two CMDs were used K$_s$ vs. J-K_s and m$_{F606W}$ vs. m$_{F606W}$-K$_s$ (Figure 3). In order to obtain a

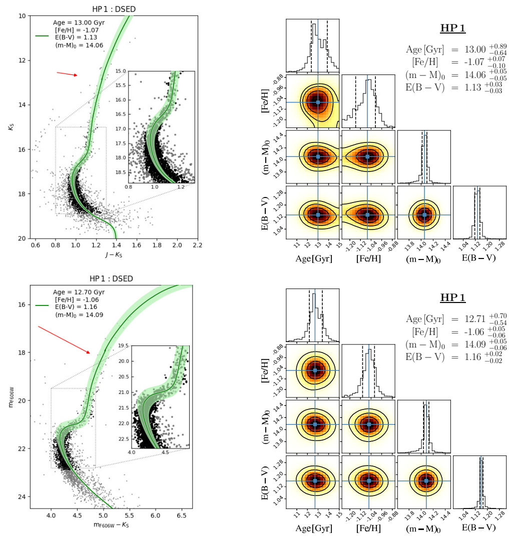

Figure 3. The upper panels show the result for the CMD K$_S$ vs. J-K$_S$ and the lower ones for the CMD m$_{F606W}$ vs. m$_{F606W}$-K$_S$.

Table 1. Average results using the NIR and NIR-Optical CMDs and two stellar evolutionary models (DSED and BaSTI* models). BaSTI* is the BaSTI results after the expected correction for the atomic diffusion ($\Delta Age \approx 0.90$ Gyr).

[Fe/H] (dex)	Age (Gyr)	(m-M)$_0$ (mag)	d$_\odot$ (kpc)	E(B-V) (mag)
$-1.09^{+0.07}_{-0.09}$	$12.75^{+0.86}_{0.81}$	$14.10^{+0.06}_{-0.05}$	$6.59^{+0.17}_{-0.15}$	$1.15^{+0.02}_{-0.02}$

better estimation of the parameters, We employed two different stellar evolutionary models BaSTI and DSED. The average of the posterior distributions and the final values of each parameter are presented in Table 1. This result confirms the idea of the HP1 is one of the oldest objects of the Galaxy.

5. Conclusions

SIRIUS was successfully tested by means of several control experiments. Even though SIRIUS was designed to analyze globular clusters in the Milky Way, it is flexible to analyze stellar clusters in general. SIRIUS has been applied to analyze clusters from the HST UV Legacy Survey and the VISCACHA Survey. Others stellar evolutionary models will be added to SIRIUS code in the future.

Acknowledgements. FAPESP, CNPq and CAPES. Based on observations obtained at the Gemini Observatory and HST

References

Hogg, D. W., & Foreman-Mackey, D. 2018, ApJ, 236, 11
Kerber, L., Santiago, B., Brocato, E. 2007, A&A, 462, 139
Milone, A. P., Piotto, G., Renzini, A., et al. 2017, MNRAS, 464, 3636
Nardiello, D., Piotto, G., Milone, A. P., et al. 2015, MNRAS, 451, 312
Kerber, L. O., Libralato, M., Souza, O. S., et al. 2019, MNRAS (in press)
Pietrinferni, A., Cassisi, S., Salaris, M., & Castelli, F., 2004, ApJ, 612, 168
Dotter, A., Chaboyer, B., Jevremovic, D., et al. 2008, ApJ, 178, 89