The star formation history of the Milky Way disc from Gaia DR3 and BGM FASt

Author: Marc del Alcázar i Julià
*Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain**

Advisor: Francesca Figueras Siñol

Abstract: We use the Besançon Galaxy Model FASt Approximate Simulations (BGM FASt) framework together with the Approximate Bayesian Computation (ABC) algorithm to derive the posterior probability distribution function of the parameters defining the initial mass function (IMF) and the star formation history (SFH) in the Solar neighbourhood. We propose new strategies to unveil the influence of some BGM Std model ingredients from the Poissonian distance metric and the posterior distribution of the ratio between pseudo-simulation and data in the Hess diagrams. Gaia DR3 up to G=13 and a set of consolidated executions of the BGM FASt + ABC code on a Cloud Environment, using both different Mother Simulations and priors, allow us to confirm the existence of the star formation burst in the Galactic disc 2-4 Gyr ago proposed by [Mor et al. 2019]. Furthermore, for the composite IMF, we obtain the slopes of α_2 and α_3 constrained to the range [1.7, 2.8] and [1.8, 2.4], respectively, in agreement with recent values from the literature. The wide range of values obtained for the total stellar surface mass density of the thin disc at the Solar neighbourhood, between 30-50 $M_{\odot}\mathrm{pc}^{-2}$, demonstrates that, before concluding on the set of best ingredients for the Galactic stellar population model, we shall loop the process to fit again both, the Galactic potential as described in Robin et al. [2022], and the BGM FASt IMF and SFH parameters.

I. INTRODUCTION

The star formation history (SFH) of the thin disc is one of the magnitudes that, at present, better characterizes the evolution of the Milky Way in the last 10 Gyr, giving account to the internal and external processes responsible for the variation in the star formation rate (SFR) of the Galaxy. Completely entangled with the SFH, the initial mass function (IMF) of the thin disc plays a key role in the chemical evolution of this stellar system. These are the reasons that explain why the precise determination of the IMF and the SFH has centred the efforts of the Galactic research community from the middle of the 20th century to the present Gaia era (Mor et al. [2019], Sysoliatina and Just [2021], Ruiz-Lara et al. [2020], Dickson et al. [2023], Bernard [2018], Cignoni et al. [2007], Kroupa [2008], Haghi et al. [2020], Li et al. [2023], among others).

For this purpose, a wide range of tools and techniques have been implemented. Mor et al. [2019] use the Besançon Galaxy Model Fast Approximate Simulations (BGM FASt) together with the Approximate Bayesian Computation (ABC) to derive the SFH and the IMF of the thin disc of the Milky Way in the Solar neighbourhood. For the former, they report an exponential decrease with a burst of star formation peaking 2-3 Gyr ago. The origin of this SFR enhancement is still unknown. In Mor et al. [2019], they propose that it comes from a recent merger between the Milky Way and a gas-rich satellite galaxy. This hypothesis is in concordance with the results found in Stewart et al. [2008] performing cosmological simulations within the ΛCMD model, where $\sim 95\%$ of Milky Way-sized halos have accreted at least one object in the last 10 Gyr. The minor merger scenario is also coherent with the discovery of an external perturbation of the Milky Way due to the presence of the Sagittarius dwarf galaxy made by Antoja et al. [2018].

More recently, Sysoliatina and Just [2021] propose an alternative strategy to constrain 22 Galactic parameters of the Justin-Jahreiβ model using a Bayesian technique to infer their posterior probability distribution function (PDF) from Gaia DR2 and APOGEE data for stars of the local neighbourhood with apparent magnitude $7 < G < 17$. They find two bursts in the SFH in the last ~ 4 Gyr which they relate to gas infalls. Ruiz-Lara et al. [2020] modelled Gaia DR2 observed colour-magnitude diagrams with local stars inside a bubble of radius ~ 2 kpc around the Sun to infer the impact of the pericentric passages of the Sagittarius dwarf galaxy in the SFH.

For the derivation of the initial mass function, most of the studies analyze the star population on clusters and associations to infer some clues on a more general IMF [Dickson et al. 2023, Kroupa 2008, Haghi et al. 2020]. One of the most recognized is Kroupa [2008], which presents a *canonical stellar IMF* arising from the correction of the classical Salpeter slope in the low-mass regime using a multi-dimensional optimization technique on unresolved multiple stellar systems. More recently, Haghi et al. [2020] performed a series of N-body simulations to study the evolution of star clusters using the `NBODY6` code. A different approach is taken in Dickson et al. [2023], where they present the results of the best-fitting multimass models for 37 Milky Way globular clusters, obtaining values compatible with the classical Salpeter IMF in the high-mass regime.

Other studies make use of alternative techniques to derive the IMF. In Mor et al. [2019], the BGM FASt and ABC method is used to simultaneously fit the IMF and the

*Electronic address: marcdelalcazar@gmail.com

SFH. For the IMF they use a three-truncated power-law applied to Gaia DR2 up to $G < 12$. By using a completely different approach, Pelkonen et al. [2021] perform a high-resolution star-formation simulation with realistic physical conditions in clouds to infer the IMF from the core mass function under the idea that there is a one-to-one relation between stars and cores. More recently, Li et al. [2023] used the LAMOST spectrum of ~ 93000 M-dwarf stars in the Solar neighbourhood (100-300 pc) to report a variable IMF in the low-mass range which depends on the metallicity and the stellar age. This non-exhaustive list shows only some of the examples of interest in determining this fundamental function.

Within the different explored Galactic models, the Besançon Galaxy Model (BGM; Robin et al. [2003]) is a holistic population synthesis approach that has been widely used in the last decades for the statistical study of the formation and evolution of the Milky Way. From its new strategy [Czekaj et al. 2014], it was possible to directly use the IMF and the SFH to generate a full-sky thin disc population to be compared with data from *Tycho-2*, constituting the starting point of the performance of Galactic parameters inference with BGM. More recently, the outcomes of the Gaia DR2 astrometry opened a new opportunity for the improvement of the robustness of the BGM modelling. The attempt to include the dynamical self-consistency was tackled by fitting the gravitational potential of the Milky Way to the stellar kinematics and densities of Gaia DR2 [Robin et al. 2022]. At this point, it becomes essential the re-determination of the thin disc SFH and IMF, a task that is addressed in this work.

At present, the execution of a full-sky catalogue of simulated stars using the standard BGM model (hereinafter BGM Std) implies a computational cost that cannot be assumed for the inference of Galactic parameters. This process demands the use of statistically robust approximate models and powerful computational resources. To respond to these requirements, we present in this work the implementation of BGM FASt + ABC in a Cloud Environment and, for the first time, its execution considering the full-sky Gaia DR3 up to $G < 13$ [Gaia Collaboration et al. 2016, 2022]. With these new executions, we are doubling the number of stars with respect to previous works [Mor et al. 2019], as well as multiplying by a significant factor the computational resources.

In Sect. II we set the theoretical background of BGM FASt and some upgrades on its new implementation in the Cloud. In Sect. III we explain how ABC works and we describe the statistical parameters to be used to analyze the fitting process together with a set of new tools to test and validate the obtained results. In Sect. IV we show the main characteristics of the executions run in the Cloud Environment. In Sect. V we present the physical results obtained for the SFH and the IMF, comparing them with recent data from the literature. Finally, in Sect. VI, we give the conclusions of this work and what we consider are future tasks to be developed during the next months.

II. GALACTIC PHYSICS INSIDE BGM FAST

A. Basic concepts of BGM FASt

The Besançon Galaxy Model Fast Approximate Simulations is based on the idea that it is possible to generate pseudo-simulations (PS) of the Besançon Galaxy Model weighting the stars of a pre-sampled simulation (so-called Mother Simulation and labelled MS) obtained from BGM Std [Mor et al. 2018]. Then, since BGM FASt simulations are computationally cheap ($\sim 10^4$ times faster than in BGM Std), the Galactic parameters inference with tools such as the Approximate Bayesian Computation becomes feasible.

Following the scheme described in Fig. 1, a BGM FASt PS departs from two main ingredients: a MS and a set of the parameters to infer $\bar{\theta}_k$. The former is the result of the application of BGM Std [Czekaj et al. 2014], a population synthesis model that uses fundamental functions such as the space density distributions, the initial mass function, the star formation history and the age-metallicity relation of each Galactic component (e.g. thin disc, thick disc, halo, bulge and bar), among others, to generate stars with given masses M, ages τ and metallicities Z in each volume element. Then, it makes them evolve using a set of stellar evolutionary tracks. The additional intrinsic magnitudes of each star $T_{eff}, \log g$ and L are obtained after taking into account a binarity treatment. Finally, BGM Std converts these magnitudes into the observed ones following two steps: first, it applies atmosphere models to get the absolute magnitudes M_V and the intrinsic colours $BpRp_{int}$; and then, it includes the effects of reddening and extinction of the interstellar medium to get the apparent magnitudes m_V and the observed colours $BpRp_{obs}$, which constitute the observables that can be compared with the catalogue data. For the latter, the set of input parameters $\bar{\theta}_k$ is obtained from a prior probability distribution function $\bar{\rho}_t(\bar{\theta})$ that is imposed by the user in the first step ($t = 0$) and automatically build by ABC in the following ones ($t > 0$) using the accepted sets of parameters in the previous steps. Therefore, BGM FASt can be considered a function of these two ingredients. It gives rise as output to a modified MS, a PS, in which each star is associated with a weight. In the computation of these weights resides the core of BGM FASt.

We can define the i^{th} parameter space as the N-dimensional space of the i-component of the Galaxy, which contains all the parameters involved in the distribution function of the generated stars of that Galactic component. Mathematically the i^{th} parameter space is defined as $\P^i = \tau \times M \times Z \times \bar{x} \times \bar{v} \times \bar{p}$, where τ, M, Z, $\bar{x}$, $\bar{v}$ are the age, the initial mass, the metallicity, the position and the velocity of the stars of the i-component, respectively, and $\bar{p}$ accounts for other independent parameters.

Then, we can consider a general distribution function of the generated stars for the i^{th} Galactic component, $\mathcal{D}_i(\tau, M, Z, \bar{x}, \bar{v}, \alpha, \bar{p}')$, where now $\alpha \equiv [\alpha/Fe]$ is written explicitly and $\bar{p}'$ concentrates the rest of the parameters. Marginalizing all the parameters except $\bar{p}'$, we obtain the distribution function of the generated stars for the i-component in the reduced space $\P_r^i$, $\mathcal{G}_i(\tau, M, Z, \bar{x}, \bar{v}, \alpha)$. BGM FASt is based on the idea that the number of stars generated in an interval of the parameter space $\P^i$ is proportional to the mass dedicated to generate stars in that interval [Mor et al. 2018]. Therefore, the weight of the j^{th} interval of the reduced space $\Delta\P_r^{i,j}$ is proportional to the integral along that interval of the distribution function of the generated stars for the i^{th} Galactic component in the PS weighted by the mass. The normalized weight ω_j is finally obtained by dividing the resultant value by the corresponding integral of the original MS,

$$\omega_j = \frac{\int_{\Delta\P_r^{i,j}} \mathcal{G}_i^{PS}(\tau, M, Z, \bar{x}, \bar{v}, \alpha) \cdot M d\P_r^i}{\int_{\Delta\P_r^{i,j}} \mathcal{G}_i^{MS}(\tau, M, Z, \bar{x}, \bar{v}, \alpha) \cdot M d\P_r^i}. \quad (1)$$

The rigorous mathematical development of the concept as well as the practical implementation shown in Fig. 1 are explained in detail in Mor et al. [2018].

BGM FASt ends when the PS is generated from the application of the weights computed in Eq. 1 to the stars of the MS. At this point enters ABC, which uses a summary statistic and a distance metric to quantify the resemblance between the PS and the catalogue data (see Subsect. III A and III B for more details on the considered summary statistic and distance metric). Then, it compares the obtained distance δ^k with the threshold of the t^{th} ABC step. If it is lower, the set of proposed parameters $\bar{\theta}_k$ is accepted. In this case, after a given number of accepted simulations per step $N_{acc.sim}$, $\bar{\theta}_k$ becomes part of the new prior or posterior PDF $\bar{\rho}_t(\bar{\theta})$. If the distance δ^k is greater than the threshold δ^t, the set of parameters is rejected.

This process is carried out along two loops. On the one hand, the distance threshold starts at a determined upper limit that diminishes at each step t until the established lower limit is achieved or the imposed maximum number of steps is reached. On the other hand, at each step the incorporation of sets of parameters $\bar{\theta}_k$ is repeated until reaching the number of accepted simulations per step necessary to obtain a statistically robust PDF $\bar{\rho}(\bar{\theta})$. While the distance threshold of the step is above the lower limit, the derived $\bar{\rho}_t(\bar{\theta})$ becomes part of the prior PDF of the next step. Once the lower limit is achieved, $\bar{\rho}_t(\bar{\theta}) \equiv \bar{\rho}(\bar{\theta})$ becomes the approximate posterior PDF of the BGM FASt parameters.

B. The new BGM FASt implementation

For the last published results arising from the application of BGM FASt [Mor et al. 2019] it was chosen a 15-dimensional space for the parameters inference, including the three slopes of a three-truncated power-law IMF, the nine parameters of a non-parametric SFH covering the range of ages of the thin disc (0-0.1 Gyr, 0.1-1 Gyr, 1-2 Gyr, 2-3 Gyr, 3-5 Gyr, 5-7 Gyr, 7-8 Gyr, 8-9 Gyr, 9-10 Gyr), the volume stellar mass density of the young and old thick discs, and the radial scale length[1]. In this work, the inference is made within a 14-dimensional parameters space composed of the aforementioned three slopes of the IMF, and 11 parameters of a non-parametric SFH. The pass from nine to 11 SFH parameters arises from splitting the ages intervals 3-5 Gyr and 5-7 Gyr into pieces of one Gyr to get a better resolution. We consider in this case that the contribution of the young and old thick discs is well-described by the MS, so we do not fit them. Finally, we assume the MS to be dynamically self-consistent, which lets us not include the radial scale length in the inference parameters space. The possible consequences and limitations of these assumptions are discussed in Sect. V and VI.

These changes imply some modifications of the original equations of BGM FASt. For instance, the computations regarding the density laws with the Einasto profiles are no longer in use since we assume the dynamical self-consistency of the MS. Under this consideration, we suppress the spatial dependencies in the new BGM FASt implementation, keeping only the IMF and the SFH as the BGM FASt parameters to infer. The weight equation, therefore, is extremely simplified to

$$\omega_j = \frac{\int_{\Delta\tau_{i,j}} \sum_{\odot}^{i,PS}(\tau) d\tau \int_{\Delta M_{i,j}} \xi_i^{PS}(M) M dM}{\int_{\Delta\tau_{i,j}} \sum_{\odot}^{i,MS}(\tau) d\tau \int_{\Delta M_{i,j}} \xi_i^{MS}(M) M dM}, \quad (2)$$

where $\sum_{\odot,i}(\tau)$ is the star formation rate of the i^{th} Galactic component at the Solar neighbourhood at time τ, $\xi_i(M)$ is the value of the IMF at M, we consider $d\tau = \Delta\tau_{i,j}$ equal to the age intervals given in Tab. IV, $dM = 0.025$ and $\Delta M_{i,j}$ corresponds to the ranges of mass of the IMF indicated in Mor et al. [2018].

On the other hand, in Mor et al. [2019] they use a MS generated assuming a parametric SFH. In this case, as can be seen in Subsect. IV A, we set as input two different MS: one with a non-parametric SFH and another with a parametric SFH. For the former, we have slightly modified the code to use a non-parametric SFH both for the PS and the MS (numerator and denominator of Eq. (2), respectively). A schematic description of the content of the new `bgmfast` Python package can be found in Appendix C.

[1] Note that all the parameters correspond to the Solar neighbourhood.

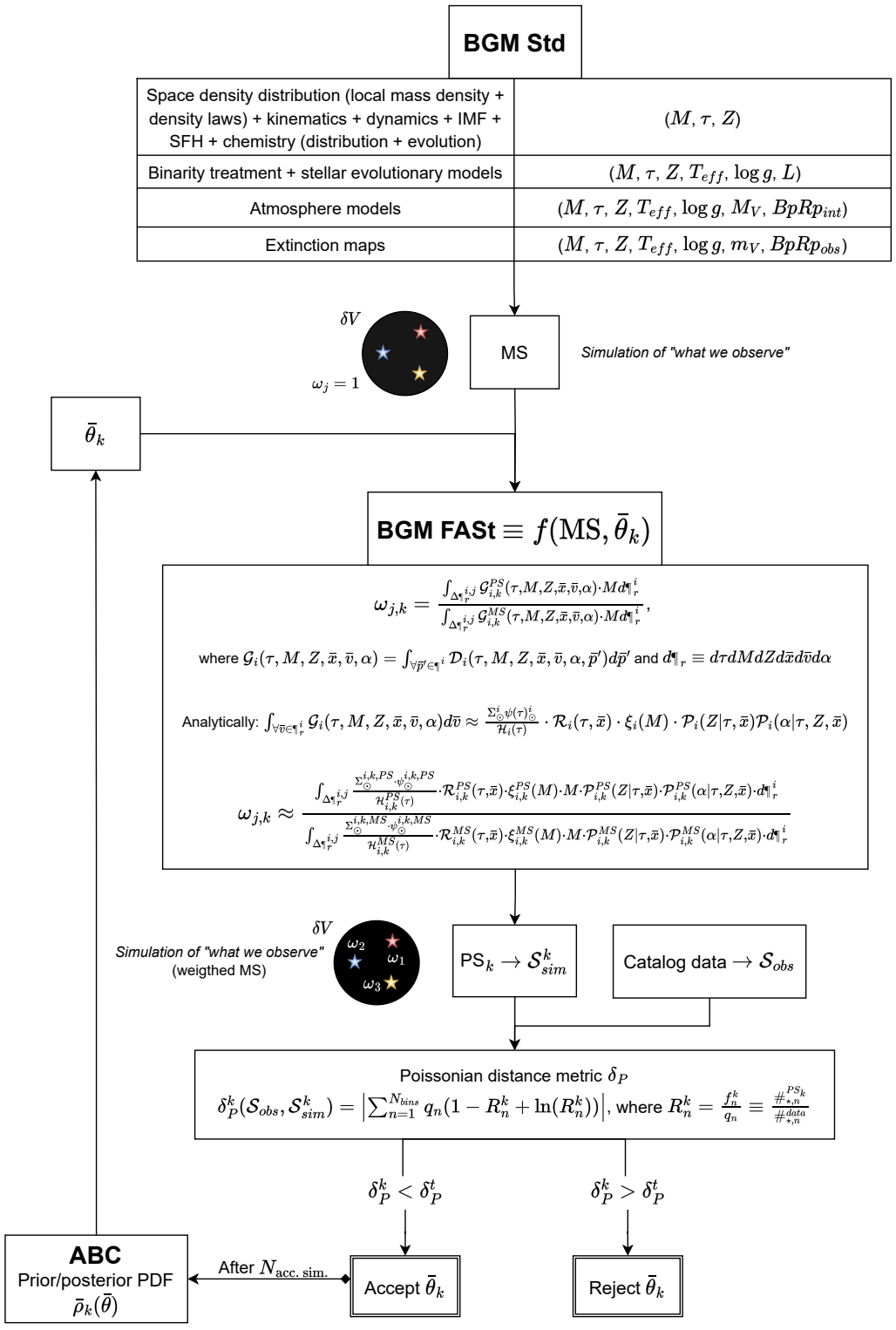


FIG. 1: Flux diagram of the process involving BGM Std, BGM FASt and ABC. Subscript and superscript legend: i^{th} Galactic component, t^{th} ABC step, k^{th} set of BGM FASt parameters, n^{th} bin of the CMD/LD, and j^{th} interval of the reduced parameter space. See in Subsect. II A a detailed explanation of this scheme.

III. TAKING ADVANTAGE OF BGM FAST: THE APPROXIMATE BAYESIAN COMPUTATION

The possibility of inferring Galactic parameters by performing iterative PS is the most remarkable opportunity arising from BGM FASt. However, it is crucial the choice of a statistically robust and computationally cheap iterative mechanism to get a final set of parameters with mathematical and physical sense. In this work, we analyze ABC's performance and give some ideas on how to understand its results within the framework of BGM FASt.
From the Bayes' theorem,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{3}$$

it is possible to compute the exact PDF of a parameter A from the likelihood function, $P(B|A)$, the prior PDF of A, $P(A)$, and the total probability of B, $P(B)$, which behaves as a normalization constant. In the framework of parameters inference, $A \equiv$model and $B \equiv$data. For very complex models such as BGM FASt (a large number of parameters and extensive data sets), it is mathematically impossible or computationally prohibitive to obtain the exact likelihood function. To surpass this inconvenient and following the approach presented in Mor et al. [2018], in BGM FASt we use a Python implementation of a sequential Monte Carlo Approximate Bayesian Computation algorithm (SMC-ABC) [Jennings and Madigan 2017], which carries out the iterative process described in Subsect. II A to estimate the posterior PDF of the parameters to infer.
According to that description, there are several elements to adjust when we work with SMC-ABC: 1) the summary statistics, 2) the distance metric, 3) the distance thresholds, 4) the number of accepted simulations per step, and 5) the maximum number of steps.

A. Summary statistics

The summary statistics $\mathcal{S}$ is the responsible for giving a proper shape to the pseudo-simulated data $\mathcal{D}_{sim}$ and the observed data $\mathcal{D}_{obs}$ to be able to compare them. Therefore, in the case of parameters inference, it has to capture the maximum information on the effects of the set of parameters $\bar{\theta}$. If the chosen summary statistics $\mathcal{S}$ is a sufficient statistic for $\mathcal{D}$, then the posterior PDF derived from the summary statistics $P(\bar{\theta}|\mathcal{S}_{obs})$ is equivalent to the posterior PDF obtained from the complete set of data $P(\bar{\theta}|\mathcal{D}_{obs})$. Considering this statement, in this work we use $\mathcal{S}$ instead of $\mathcal{D}$ to represent the sufficient statistics as it is done in Mor et al. [2018], Jennings and Madigan [2017].
In BGM FASt, we take as a sufficient statistics a binned modified Hess diagram M'_G vs. $Bp - Rp$ (hereinafter CMD), where $M'_G = G + 5\log_{10}(\bar{\omega}/1000) + 5$ and $Bp - Rp \equiv BpRp$ is the observed colour index of the filters B_p and R_p. Note that the computed M'_G is different to the intrinsic M_G because the former does not take into account the extinction due to the interstellar medium. The assumption behind the use of M'_G is that the extinction maps in the MS are good enough to characterize the extinction suffered by the light captured by Gaia. We assume it is true for stars within the colour range $-0.42 < (Bp - Rp) < 2.73$, where the photometric transformation of Evans et al. [2018] is valid. Therefore, we only fit the CMD within these limits.
We build three CMDs characterizing different Galactic latitudes ($|b| < 10$, $10 < |b| < 30$ and $|b| > 30$) for stars with $-1 < M'_G < 5$ to avoid losing valuable information on the effects of a given IMF and SFH in different regions. Stars with an absolute magnitude $M'_G > 5$ are injected into a colour-integrated modified Hess diagram (modified luminosity distribution, hereinafter LD), owing to the lack of stars with $M'_G > 5$ in the ranges of G explored (see Subsect. IV A). We take 63 $BpRp$ bins ($\Delta BpRp = 0.05$ mag) and 24 M'_G bins ($\Delta M'_G = 0.25$ mag) for the CMD and 40 bins for the LD ($\Delta M'_G = 0.25$ mag), which give us a reasonable relation between the conservation of the stars information and a statistically sufficient number of stars per bin.

B. Distance metric

Once we have the sufficient statistics, we need to define a distance δ to quantify how similar $\mathcal{S}_{sim}$ and $\mathcal{S}_{obs}$ are. For this work and following the initial idea of BGM FASt [Mor et al. 2018], we use the so-called Poissonian distance metric

$$\delta_P(\mathcal{S}_{obs}, \mathcal{S}_{sim}) = \left| \sum_{n=1}^{N_{bins}} q_n[1 - R_n + \ln(R_n)] \right|, \tag{4}$$

where $R_n = f_n/q_n$ is defined as the quotient between the number of stars in the n^{th} bin of the CMD/LD in the model (f_n) and in the catalogue data (q_n), and N_{bins} is the total number of bins. When $f_n = 0$ or $q_n = 0$, we substitute f_n by f_n+1 and q_n by q_n+1 to avoid problems with the quotient and the logarithm. When the PS fits perfectly the data, $R_n = 1\ \forall n$ and $\delta_P(\mathcal{S}_{obs}, \mathcal{S}_{sim}) = 0$.
There are three elements that play a key role in the determination of the distance: the limiting apparent magnitude G (the number of stars in the MS and in the catalogue), the input MS, and the set of BGM FASt parameters $\bar{\theta}$. In order to better understand their effects on the resulting distance, it is good to express Eq. (4) as follows:

$$\delta_P(\mathcal{S}_{obs}, \mathcal{S}_{sim}) = N_\star \left| \sum_{n=1}^{N_{bins}} s_n[1 - R_n + \ln(R_n)] \right|, \tag{5}$$

where $N_\star$ is the total number of stars in the data and s_n the fraction of $N_\star$ contained in the n^{th} bin. From

this equation, we can clearly separate the effects on the distance into two parts:

- The limiting magnitude G, characterized by the number of stars in the catalogue $N_\star$.
- The model behind, including the physics of the MS and the BGM FASt set of parameters $\bar{\theta}$, defined by R_n or, more extensively, by the absolute value of the summation in Eq. (5). We call this right-most part of Eq. (5) the intrinsic physical distance.

For the former, one can think from Eq. (5) that increasing the number of stars in a given factor will also increase the distance between the PS and the catalogue data in the same factor. That is partially true. On one hand, if we consider a sample up to a given magnitude and then we artificially duplicate each star both in the sample and in the data, we will obtain two times the original distance since the effects of the model are exactly the same in the two cases. But this is not the case if the increasing number of stars is due to a larger limiting apparent magnitude G. In this case, the physics behind the model changes, which has a direct impact on the computation of the distance δ_P.
In general, it will be larger than the one expected for the increasing number of stars, since more complex physics is involved. For instance, the contribution of the halo or the thick disc will be more important, since far distances and weak stars will be reached. Taking into account that the modelling of the Milky Way beyond the thin disc is still precisely unknown, the correspondence between the PS (which is a weighted MS) and the observed data will be worse. In that case, the distribution of R in Eq. 4 will be broader and far from the ideal $R = 1$, and the resultant distance probably higher due to the effect of the intrinsic physical distance[2].
Also, sometimes a relatively small increase in the magnitude complements regions with low statistics, reducing the effects of outliers and, therefore, implying an increasing distance smaller than the expected one. See an example of the behaviour of the intrinsic physical distance in Appendix B.

C. Distance thresholds

The third key point is the choice of thresholds for the ABC process. The idea resides in the selection of intelligent upper and lower limits to obtain a final approximate PDF of $\bar{\theta}$ statistically robust that represents a scientific improvement with respect to previous studies. This concept gives some clues on the way to choose the distance thresholds.
For the upper limit δ_{max}, we select the distance obtained from imposing the parameters of the MS22 in BGM FASt (see Subsect. IV A for a detailed description of the MS). Behind this selection, we accept that the MS presented in Robin et al. [2022] is currently the best modelling of the Milky Way within the BGM Std framework. For the lower limit δ_{min}, we run exploratory executions to find the smallest possible value reachable before accomplishing the maximum number of steps, which we set equal to 100 steps for all executions. Finally, the number of accepted PS per step $N_{acc.sim.}$ is chosen as a balance between the computation time and a good sampling of $\bar{\theta}$ for the proper evolution of the PDF. In this work, we choose a value of $N_{acc.sim.} = 200$.

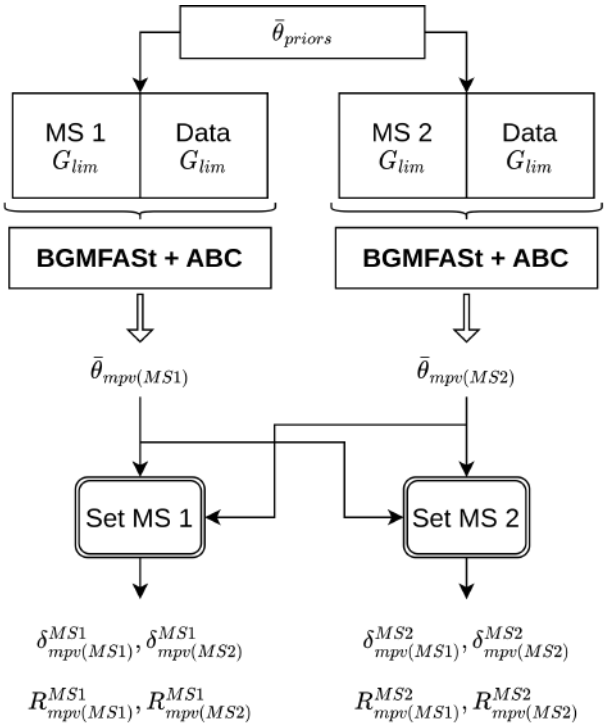


FIG. 2: Flux diagram of the approximate cross-checking. See Subsect. III D for a detailed description of this scheme.

D. The new approximate cross-checking strategy

We propose a first approximate solution to give a general idea of both the resemblance between the MS and the catalogue data, and the fineness of the derived BGM FASt parameters given the distance δ_P and the R distribution arising from a BGM FASt + ABC execution. In Subsect. III B we explained how from Eq. (5) it is derived a triple contribution to the total Poissonian distance δ_P: the total number of stars in the data $N_\star$, the

[2] It is true that a broader distribution of R can induce lower distances if there is a compensation between the values < 1 and > 1. In this case, will also be important to analyze the distribution itself, its broadness and dispersion.

closeness of the MS to the catalogue data, and the fineness of the most probable values (mpv) of the BGM FASt inferred parameters posterior PDF, $\bar{\theta}_{mpv}$. This triple degeneracy is reduced to two when using as an evaluator the mpv of the R distribution, R_{mpv}, which is independent of the total number of stars if we introduce the physical complexity arising from going to upper magnitudes in the analysis of the MS. Under this assumption, there is still a degeneracy between the MS and the mpv of the BGM FASt set of parameters distributions, $\bar{\theta}_{mpv}$, when performing the analysis of R_{mpv}.
The best procedure to break this degeneracy is to loop again the process by fitting the Galactic potential [Robin et al. 2022] to the inferred BGM FASt parameters and computing the distance between the new MS and the catalogue data. The repetition of this process should give us, in the end, the best fit of all the parameters, not only the ones inferred in BGM FASt. Considering that this is a long and complex process, we propose an alternative strategy. In the approximate cross-checking, schematically shown in Fig. 2, the sets of BGM FASt parameters $\bar{\theta}_{mpv(MS1)}$ and $\bar{\theta}_{mpv(MS2)}$ obtained with BGM FASt + ABC applied to two different MS, are introduced in the other MS to compute the resultant distances δ_P and the R distributions. The idea is that the best MS should give the smallest distances and the best values of R for any set of BGM FASt parameters, and the most performing $\bar{\theta}$ should imply obtaining the minimum distances and best values of R in any MS. However, this statement must be taken carefully, and the subtleties explained in Appendix B 1 must be considered when applying the approximate cross-checking strategy.

IV. THE SET OF EXECUTIONS IN THE CLOUD

BGM FASt works with two different samples: the Mother Simulation and the catalogue data. The role of each one is explained in detail in Subsect. II A.

A. Input data: Mother Simulations and Gaia DR3

We chose two different MS for this work:

- MS22: it is a dynamically self-consistent MS up to $G < 13$ and $V < 13$[3] generated following the outcomes of Robin et al. [2022]. It has been generated using a non-parametric SFH.
- MS18: it corresponds to the MS of the fiducial case of Mor et al. [2019]. The ingredients of this MS, including a parametric SFH, are based on the results of Czekaj et al. [2014]. In this case, the full-sky catalogue of pseudo-stars was generated up to a limiting magnitude $G < 13$.

In both MS, the astrometric and photometric magnitudes of each star were affected by errors using the indications of Gaia-DPAC consortium for Gaia DR3 (see the Gaia instrument model webpage).
For the catalogue data, we used the full-sky Gaia DR3 up to $G < 13$. Distances were directly computed as the inverse of the parallax and the photometric quantities used are the G magnitude and the observed $Bp - Rp$ colour. From these ingredients, we set three different configurations that are presented in Tab. I: 1) G12-MS22, which uses MS22 cut at magnitude $G < 12$ combined with Gaia DR3 with the same constrain; 2) GV13-MS22, which is formed by the same MS22 up to $G < 13$ and $V < 13$ and the analogous Gaia DR3 cut at the same limiting magnitudes[4]; and 3) G13-MS18, that works with MS18 and Gaia DR3 up to $G < 13$.
As can be seen in Tab. I, we observe a non-negligible disagreement between the number of stars in the MS and the Gaia DR3, with relative discrepancies of 17%, 4% and 7% for G12-MS22, GV13-MS22 and G13-MS22, respectively. In addition, these discrepancies are not homogeneously distributed along the two ranges of magnitudes described in Subsect. III A for the fitting of the CMD and the LD. Is at this point where enters BGM FASt, which is thought to reduce these "bin-to-bin" differences between the MS and the Gaia DR3.

B. BGM FASt-OCRE executions

The BGM FASt executions presented in this work have been carried out in a Cloud Environment within the European OCRE project *RainClouds*[5]. The details on the computational infrastructure, its performance and accountability are explained in Appendix A. We performed more than 12 executions of BGM FASt + ABC. The most statistically robust and scientifically useful of them are presented in Tab. II. As it is explained in Subsect. III C, the distance threshold for ABC must be determined following a coherent criterion. In this case, we have done two different kinds of executions. On one hand, we performed one *exploratory* execution per each MS + Gaia DR3 configuration with a wide range of threshold distances to determine a reasonable value for δ_{min}, which we took as the minimum distance reached by the pseudo-simulations during the explorations. On the other hand,

[3] This second constrain in V is due to the fact that BGM Std works with V magnitudes instead of G.

[4] For the cut in V in the Gaia data we used the equation $G - V = -0.02704 + 0.01424(G_{Bp} - G_{Rp}) - 0.2156(G_{Bp} - G_{Rp})^2 + 0.01426(G_{Bp} - G_{Rp})^3$ from the Gaia DR3 documentation.

[5] In 2020 the Gaia-UB team was awarded a grant from the European Open Clouds for Research Environments initiative (OCRE) for the execution of several Gaia key scientific cases, BGM FASt among them.

TABLE I: In the first row, it is found the total number of stars $N_\star$ in the Mother Simulation and in the Gaia DR3 for each of the three configurations. The second and third rows show the absolute and relative contribution of the two ranges of absolute magnitudes M'_G to the total $N_\star$.

	G12-MS22		**GV13-MS22**		**G13-MS18**	
	MS	**Gaia DR3**	**MS**	**Gaia DR3**	**MS**	**Gaia DR3**
Total $N_\star$	3071444	3685605	5805759	5552883	6912989	7369632
$-1 < M'_G < 5$	2800626 (91%)	3107952 (84%)	5304932 (91%)	4878410 (88%)	5871180 (85%)	6197481 (84%)
$5 < M'_G < 15$	121996 (4%)	174330 (5%)	323316 (6%)	336912 (6%)	630947 (9%)	484358 (7%)

Notes. Stars within the range $-1 < M'_G < 5$ are also limited to have a colour $-0.42 < (Bp - Rp) < 2.73$ to avoid the extinction discrepancies explained in Subsect. III A. Also, some stars do not fit any of the magnitude and colour criteria considered in BGM FASt, which explains why the sums of the percentages do not reach 100% of the total $N_\star$.

once carefully selected δ_{min} and δ_{max}, we performed an ABC execution that we take as the one with statistical robustness.

For the maximum threshold of distance δ_{max}, we took the value obtained by introducing the parameters of the corresponding MS in BGM FASt. The idea behind this way of proceeding is to consider the MS as the best model up to date and, consequently, to set it as the upper threshold of distance since the goal of BGM FASt is to improve the current model.

We consider executions GV13-MS22 and G13-MS18 our fiducial cases since they are the best BGM FASt performances under the theoretical background of ABC. Comparing δ_{min} and δ_{max}^{last}, GV13-MS22 almost reaches the established minimum distance in its last step while G13-MS22 gets it in 39 of the 100 planned steps. The evolution of the minimum, the mean and the maximum distances are shown in Fig. 3, where we can observe two different regimes: GV13-MS22 develops a clear asymptotic path while G13-MS18 gets a distance below the minimum threshold just at the beginning of the asymptotic trend (see in Appendix B what does it tell us the exploration of the asymptotic regime).

Apart from the comparison of the two fiducial executions in terms of their distance evolutionary regimes, we compare the distance of the set of most probable values of the inferred parameters $\bar{\theta}_{mpv}$ as well as mpv of the R distribution (both magnitudes are found in Tab. II). The final distance $\delta_{mpv}^{GV13-MS22} = 390929$ is larger than $\delta_{mpv}^{G13-MS18} = 297898$, which is unexpected if we take into account that G13-MS18 deals with 16% (MS) to 24% (Gaia DR3) more stars than GV13-MS22. As we have seen in the Subsect. III B, in some cases an increasing number of stars gives rise to a better statistical representation of the CMD/LD. In addition, since we are using a different MS, it exists the possibility that the combination of MS18 and the derived BGM FASt parameters $\bar{\theta}_{mpv}^{G13-MS18}$ fit the Gaia data better than the tandem formed by MS22 and $\bar{\theta}_{mpv}^{GV13-MS22}$.

This last hypothesis seems to be confirmed if we take a look at the distributions of R shown in Fig. 4 (black solid histogram) and, especially, the resulting mpv of these distributions, which give us valuable information on the resemblance between the final PS and Gaia DR3, as well as a quantification of it. As can be seen in Tab. II, the resulting $R_{mpv}^{G13-MS18} = 0.91^{+0.43}_{-0.66}$ has a value closer to the ideal $R = 1$ than $R_{mpv}^{GV13-MS22} = 0.84^{+1.88}_{-0.35}$. However, it is true that the dispersion is similar in both cases in

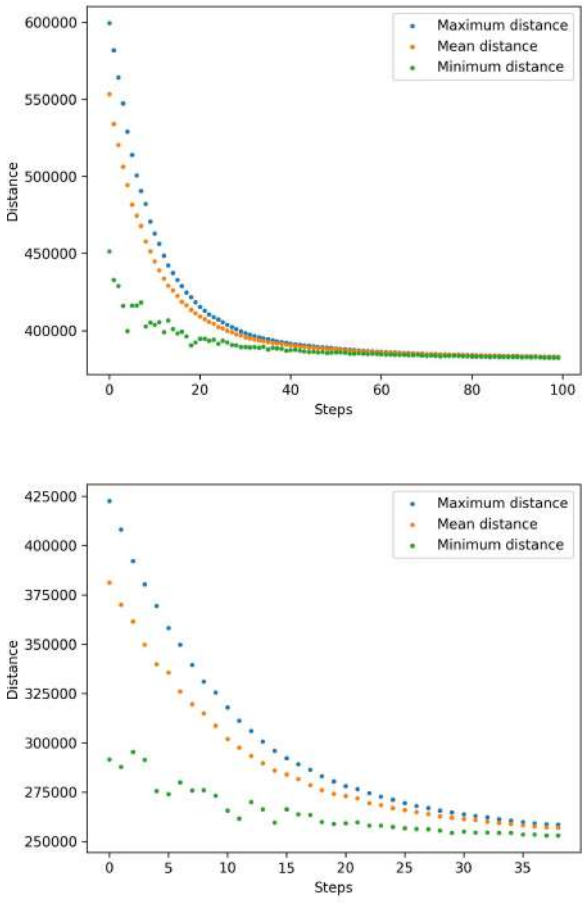


FIG. 3: Maximum, minimum and mean distance evolution over ABC steps for the two fiducial cases GV13-MS22 (top) and G13-MS18 (bottom).

TABLE II: Most reliable executions performed in this work. The name of each one is formed by the configuration of the execution (see Subsect. IV A) and the initial letter of its type when it is used for Exploratory purposes (see Subsect. IV). The only exception is G12-NP-S, which corresponds to the fiducial execution of Mor et al. [2019]. G_{lim} and V_{lim} are the limiting magnitudes, "MS" labels the Mother Simulation used for the given execution (see Subsect. IV A), "Data" corresponds to the Gaia Data Release, δ_{max} and δ_{min} are the upper and lower limits of the ABC process, δ_{max}^{last} is the maximum value of distance in the last ABC step, $N_{steps}^{exec}/N_{steps}^{plan}$ is the number of ABC steps executed compared to the planned ones, and δ_{mpv} and R_{mpv} are the distance and the mpv ratio between the MS and the catalogue of the final set of BGM FASt parameters most probable values (see Sect. III). The uncertainties of R_{mpv} correspond to the 16^{th} and 84^{th} percentiles of its distribution.

Name	**Type**	G_{lim}	V_{lim}	**MS**	**Data**	δ_{max}	δ_{min}	δ_{max}^{last}	$N_{steps}^{exec}/N_{steps}^{plan}$	δ_{mpv}	R_{mpv}
G12NP-S	ABC	12	-	2018	Gaia DR2	-	-	-	-	$5.6 \cdot 10^5$	-
G12-MS22-E	E	12	-	2022	Gaia DR3	500000	150000	275951	100/100	282344	$0.85^{+1.42}_{-0.37}$
GV13-MS22-E	E	13	13	2022	Gaia DR3	10^6	300000	383437	100/100	386330	$0.85^{+1.96}_{-0.35}$
G13-MS18-E	ABC/E	13	-	2018	Gaia DR3	422777	279011	279003	19/100	379455	$0.84^{+0.41}_{-0.59}$
GV13-MS22	ABC	13	13	2022	Gaia DR3	600000	382782	382993	100/100	390929	$0.84^{+1.88}_{-0.35}$
G13-MS18	ABC	13	-	2018	Gaia DR3	422777	258533	258530	39/100	297898	$0.91^{+0.43}_{-0.66}$

Notes. G13-MS18-E appears as an ABC and an exploratory execution at the same time because, while both δ_{min} and δ_{max} for this execution were determined following the ABC criteria (using for δ_{min} the minimum value of distance obtained in a first step of execution), G13-MS18-E was later treated as an exploratory execution to set the value of δ_{min} for G13-MS18.

absolute terms, though it follows opposite trends. This is reflected in the complementary shapes of the R distributions for GV13-MS22 and G13-MS18 observed in Fig. 4.

Also from Fig. 4, we can say that the R distribution does not depend on the latitude in any of the fiducial executions, showing similar trends for stars with $0 < |b| < 10$, $10 < |b| < 30$ and $30 < |b| < 90$. That is an indirect indication that the contribution of the Galactic components not considered in this work, especially the thick disc and the halo, are not affecting so much the distribution of R.

Finally, in order to have a first idea of the reason behind the better results found for G13-MS18–if they are due to a more reliable MS or a better set of final BGM FASt parameters than in GV13-MS22 execution–, we apply the approximate cross-checking strategy proposed in Subsect. III D. We take the set of mpv arising from the GV13-MS22 execution, $\bar{\theta}_{mpv(GV13-MS22)}$, and we put them inside the G13-MS18 configuration to obtain the cross-distance $\delta^{G13-MS18}_{mpv(GV13-MS22)}$ and the cross-ratio $R^{G13-MS18}_{mpv(GV13-MS22)}$. Then we do the same process in the other direction with $\bar{\theta}_{mpv(G13-MS18)}$. The matrices resulting from the approximate cross-checking are shown in Tab. III.

According to the criterion proposed in Subsect. III D, the MS with the best performance is the one used in the configuration G13-MS18. It gives significantly lower distances not only for the set of BGM FASt parameters inferred from applying ABC to itself, $\bar{\theta}_{mpv(G13-MS18)}$, but also for the ones derived from GV13-MS22, $\bar{\theta}_{mpv(GV13-MS22)}$. The same situation occurs with R, which appears closer to one when it is computed applying any of the sets of BGM FASt parameters to G13-MS18 than when we do the same to GV13-MS22. We cannot reach a conclusion as strong as for the MS when we try to choose the best set of inferred parameters from the approximate cross-checking. We do not find a $\bar{\theta}_{mpv}$ that behaves significantly better than the other one when applying it to both MS. For the distance, each set of inferred parameters obtain a lower distance than the other one in its own MS. On the other hand, it is true that the R is slightly better for $\bar{\theta}_{mpv(GV13-MS22)}$ in any MS but the difference is not sufficiently significant to consider this set of parameters the best of this work. Nonetheless, we cannot give a strong conclusion on the true resemblance of each MS with the Gaia DR3 or the fineness of the inferred sets of BGM FASt parameters without looping again the process by fitting the obtained

TABLE III: Poissonian distance δ_P and mpv of the ratio between the MS and the Gaia DR3 R_{mpv} matrices arising from the approximate cross-checking strategy (see Subsect. III D) using the set of parameters inferred with the GV13-MS22 (shortened to GV22) and the G13-MS18 (labelled G18) executions. The uncertainties of R_{mpv} correspond to the 16^{th} and 84^{th} percentiles of its distribution.

		$\bar{\theta}_{mpv(GV22)}$	$\bar{\theta}_{mpv(G18)}$
δ_P	**GV13-MS22**	390929	428162
	G13-MS18	331657	297898
R	**GV13-MS22**	$0.84^{+1.88}_{-0.35}$	$0.82^{+2.17}_{-0.33}$
	G13-MS18	$0.93^{+0.40}_{-0.68}$	$0.91^{+0.43}_{-0.66}$

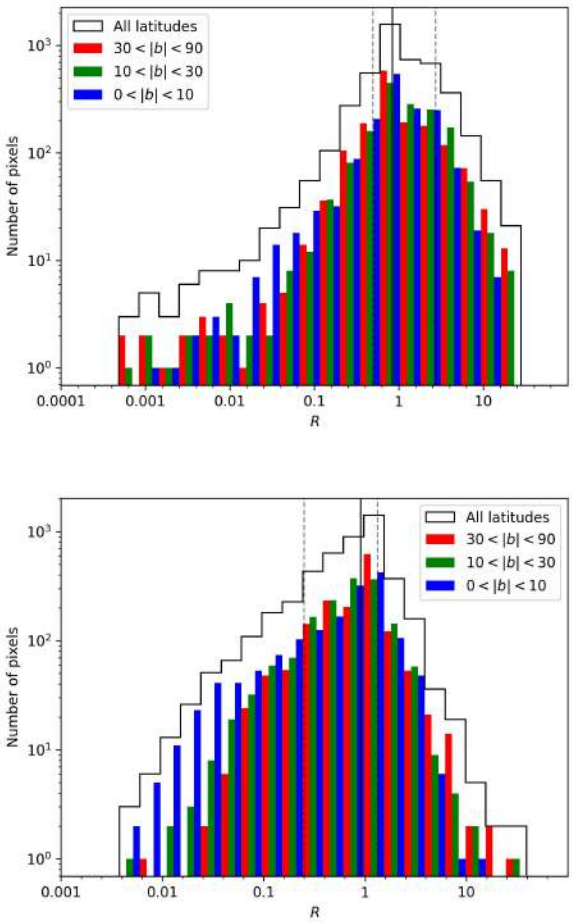


FIG. 4: R distribution obtained from the $\bar{\theta}_{mpv}$ given by the two fiducial executions GV13-MS22 (top) and G13-MS18 (bottom) taking into account all the stars (black solid histogram), stars at latitudes $30 < |b| < 90$ (red bars), stars with $10 < |b| < 30$ (green bars) and the remaining stars between $0 < |b| < 10$ (blue bars). Also, it is shown R_{mpv} (vertical black solid line) and its 16^{th} and 84^{th} percentiles (black dashed lines).

results to the potentials of Robin et al. [2022]. It is crucial to remember that ABC gives an approximate posterior PDF of the parameters. It is a distance estimator instead of a mechanism to infer the true likelihood between the catalogue data and the model. Although the Hess diagram and the luminosity distribution are chosen carefully to be summary statistics capturing all the essential ingredients of the MS and the Gaia DR3, the use of other summary statistics such as a spatial representation of the Milky Way, as well as different distance metrics, may give rise to different values of distance not compatible with the results presented in this section.

V. RESULTS & DISCUSSION

The most probable values of the BGM FASt parameters approximate posterior PDF of the two fiducial executions GV13-MS22 and G13-MS18 are found in Tab. IV and Fig. 6[6]. These values derive from fitting a Gaussian kernel density estimator to the approximate posterior PDF of the BGM FASt inferred parameters. The fits performed on the distributions and the cross-dependencies (degeneracies) for G13-MS18 are shown in Fig. 5.
We do not observe a strong dependence of the final results on the priors. This conclusion, arising from the analysis of Fig. 6, is confirmed after performing an additional exploratory execution with a set of priors significantly modified. On one hand, we take for the SFH an exponential decay with $K_\psi = 11 M_\odot \mathrm{pc}^{-2}\mathrm{Gyr}^{-1}$ and $\gamma = 0.14$, far from the $K_\psi = 3.3 M_\odot \mathrm{pc}^{-2}\mathrm{Gyr}^{-1}$ and $\gamma = 0.09$ used in GV13-MS22 and G13-MS22. On the other hand, we set the classical Salpeter slope 2.3 for the entire range of the IMF, displaced from the values $\alpha_1 = 0.8$, $\alpha_2 = 1.8$ and $\alpha_3 = 4.8$ used in our fiducial executions. While the general trend of both the IMF and the SFH is maintained in this additional execution, the only magnitude that is not conserved is the total surface density, which increases substantially in concordance with the high value of K_ψ in the additional execution. That indicates that BGM FASt and ABC are able to fit the relative star formation rate of each period of time but not its absolute value. Considering so, it is confirmed as a consistent decision the choice a SFH set of priors for the fiducial executions that maintains the total surface density of the dynamically self-consistent MS of Robin et al. [2022].

A. The bump in the non-parametric SFH

We find for the SFH a similar behaviour for GV13-MS22 and G13-MS18. The most prominent result is the enhancement of the SFR peaking at ~ 2.5 Gyr ago. This bump, which is clearly shown in all the executions performed in this work, is in complete agreement with the star formation burst with a maximum 2-3 Gyr ago reported by Mor et al. [2019]. Sysoliatina and Just [2021] also discovers two bursts of star formation centered ~ 0.5 Gyr and ~ 3 Gyr ago with a relative enhancement of $\sim 30\%$ and $\sim 55\%$, respectively. Taking into account the low resolution in the BGM FASt SFH, the observed bump is also compatible with the most significant one found in Sysoliatina and Just [2021] ~ 3 Gyr ago. However, this burst is not found by Ruiz-Lara et al. [2020], which unlike us, find three main enhancements of the SFR ~ 5.7, ~ 1.9 and ~ 1.0 Gyr ago with a significance of 3-4σ and above. The first one also appears in the G13-MS18 exe-

[6] Additional plots of the results arising from the executions presented in this work can be found in `bgmfast` GitHub.

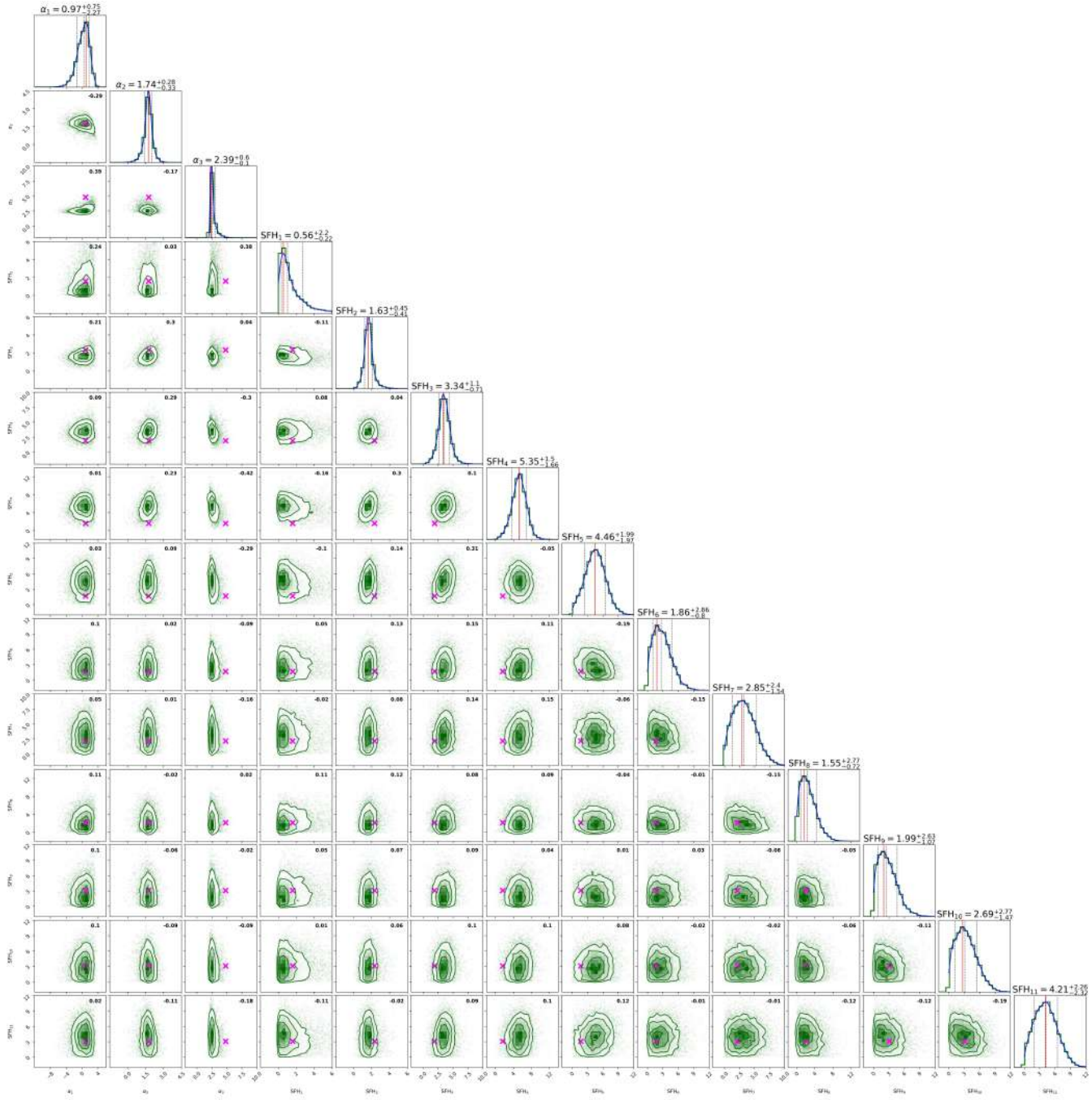


FIG. 5: Corner plot of the BGM FASt inferred parameters in the G13-MS18 execution, including the three slopes of the IMF and 11 parameters for the SFH. At the top of the columns, it is shown the projected approximate posterior PDF of each parameter with its Gaussian fit, as well as the resulting most probable value and percentiles 16^{th} and 84^{th}, which are in addition marked with solid and dashed lines, respectively. It is also indicated with a dashed line the median of the distribution. In black at the top right of each plot, it is shown Pearson's correlation coefficient. Also in the plots, the parameters adopted by the MS are marked with a magenta cross.

cution around ~ 5.5 Gyr ago. However, it is displaced to older ages in GV13-MS22, peaking ~ 6.5 Gyr ago. In addition, the error bars in the range of the old thin disc are too large to consider deriving strong conclusions on this star formation burst. Other recent studies also observe that the SFH deviates from the classical unperturbed exponential decay within the last ~ 4 Gyr. Bernard [2018] use Gaia DR1, *Tycho-2* and APASS to fit the colour-magnitude diagram for stars in the Solar neighbourhood within 250 pc, finding an enhancement of the SFR ~ 4

TABLE IV: Parameter name, age (Gyr), prior and posterior PDF of the 14 inferred parameters, three IMF slopes (dimensionless) and 11 SFH surface densities ($M_\odot \mathrm{pc}^{-2}\mathrm{Gyr}^{-1}$), and the resulting total surface density ($M_\odot \mathrm{pc}^{-2}$) of our fiducial cases GV13-MS22 and G13-MS18.

Parameter	**Age**	μ_{prior}	$\bar{\theta}_{GV13-MS22}$	$\bar{\theta}_{G13-MS18}$
α_1	all	0.8	$-1.88^{+1.96}_{-1.81}$	$0.97^{+0.75}_{-2.27}$
α_2	all	1.8	$2.76^{+0.09}_{-0.19}$	$1.74^{+0.28}_{-0.33}$
α_3	all	4.8	$1.85^{+0.62}_{-0.05}$	$2.39^{+0.60}_{-0.10}$
$\Sigma^1_\odot$	0-0.15	1.4	$0.84^{+0.73}_{-0.24}$	$0.56^{+2.20}_{-0.22}$
$\Sigma^2_\odot$	0.15-1	1.4	$3.23^{+0.16}_{-1.22}$	$1.63^{+0.45}_{-0.41}$
$\Sigma^3_\odot$	1-2	1.5	$6.06^{+0.48}_{-2.38}$	$3.34^{+1.10}_{-0.71}$
$\Sigma^4_\odot$	2-3	1.7	$9.79^{+0.57}_{-5.06}$	$5.35^{+1.50}_{-1.66}$
$\Sigma^5_\odot$	3-4	1.8	$6.74^{+0.81}_{-3.31}$	$4.46^{+1.99}_{-1.97}$
$\Sigma^6_\odot$	4-5	2.0	$6.33^{+1.48}_{-3.40}$	$1.86^{+2.86}_{-0.80}$
$\Sigma^7_\odot$	5-6	2.2	$2.51^{+2.20}_{-1.13}$	$2.85^{+2.40}_{-1.54}$
$\Sigma^8_\odot$	6-7	2.4	$4.52^{+4.40}_{-1.98}$	$1.55^{+2.77}_{-0.72}$
$\Sigma^9_\odot$	7-8	2.6	$3.92^{+2.34}_{-2.07}$	$1.99^{+2.63}_{-1.07}$
$\Sigma^{10}_\odot$	8-9	2.9	$1.87^{+2.54}_{-0.96}$	$2.69^{+2.77}_{-1.47}$
$\Sigma^{11}_\odot$	9-10	3.2	$6.44^{+1.55}_{-3.71}$	$4.21^{+2.26}_{-2.32}$
$\Sigma^T_\odot$	all	21.7	51.05	29.77

Notes. The prior PDF for both fiducial cases are the same: Gaussians centred at μ_S with variance $\sigma_S = 2$ (with the corresponding units). These PDF of the SFH are truncated at zero. The μ_S values of the IMF are exactly the ones used in the original MS22. The μ_S of the 11 $\Sigma^j_\odot$ are obtained from considering an exponential SFH of the form $K_\psi \cdot e^{-\gamma\tau}$, with $K_\psi = 3.3 M_\odot \mathrm{pc}^{-2}\mathrm{Gyr}^{-1}$ the normalization constant, $\gamma = 0.09$ Gyr^{-1} the inverse of the characteristic timescale, and τ the time.

Gyr ago followed by a decrease in the past 2-3 Gyr. Similar results are found in Cignoni et al. [2007], where they report a local maximum of the SFH $\sim$ 3 Gyr ago that they attribute to the accretion of a satellite galaxy.

In agreement with the results of Mor et al. [2019], Sysoliatina and Just [2021], we recover for G13-MS18 a general trend compatible with an exponential decreasing SFH perturbed by a Gaussian centred at $\sim$ 2.5 Gyr ago. This tendency is more difficult to observe in GV13-MS22 execution. Nonetheless, what remains fully inconsistent between different works is the total (baryonic) surface density of the thin disc, which is completely related to the dynamical consistency of the models. Even in this case, we obtain two results for GV13-MS22 and G13-MS18 which are clearly incompatible. For the former, the surface mass density grows up to $\sim 50 M_\odot \mathrm{pc}^{-2}$, while for the latter it stays much lower and similar to the priors, giving a value of $\sim 30 M_\odot \mathrm{pc}^{-2}$. It is difficult to compare these values with the literature, since not always they are explicitly presented. For instance, Ruiz-Lara et al. [2020] present a SFH normalized or in arbitrary units. We can compare our result with the total surface density of the thin disc presented in Sysoliatina and Just [2021], where it is found to have a value $\sim 29 M_\odot \mathrm{pc}^{-2}$, compatible with the results of G13-MS18.

B. The consistency of the slopes of the IMF

For the IMF, the parameter that presents the most discordance value with respect to the literature is the first slope found for GV13-MS22, $\alpha_1 = -1.88^{+1.96}_{-1.81}$. Kroupa [2008], for instance, reports an exponent $\alpha_1 = 1.3 \pm 0.3$ for stars with masses between $0.08 M_\odot < M < 0.5 M_\odot$, very similar to the value of $\alpha_1 = 1.35$ adopted in Haghi et al. [2020].

The reason that explains that very negative value is the lack of low-mass stars in the MS of this configuration, with only 94 stars falling inside the mass range of α_1 compared to the 4310763 (74%) and 1494902 (26%) used to fit α_2 and α_3, respectively. The residual number of low-mass stars in the MS of GV13-MS22 is due to the fact that it is cut at $M'_V = 7$, which can be observed in Fig. 7. This anomaly has negligible consequences in the inference of all the parameters except α_1, for which this situation is catastrophic. If we assume that we need to reach at least $M'_G = 10$ to have a significant population of low-mass stars, we can compare the 94 stars with $M < 0.53 M_\odot$ in the MS with the 404 stars with $M'_G > 10$ in the Gaia sample of GV13-MS22. Therefore, to reduce the distance between the population of low-mass stars, highly localized in the LD, BGM FASt has to give very high weights to the few stars within the α_1 range of the MS. Considering that the denominator of Eq. (2) is constant, the only way to increase the value of the weights is by making the IMF of those stars grow. Since the IMF goes as $\xi(M) \propto m^{-\gamma}$, negative values of γ will correspond to positive exponents and, therefore, increasing values of the IMF. That is what explains why in the case of GV13-MS22 the best combination of parameters is obtained with a very negative value of α_1.

On the other hand, a slightly different situation happens with G13-MS18. In this case, we have a value of $\alpha_1 = 0.97^{+0.75}_{-2.27}$ that shows a very wide distribution. The MS of the configuration G13-MS18 is cut at $M'_V = 14.6$, having a good correspondence for the number of stars up to $M'_G = 10$ with the Gaia DR3, with 2894 and 2354 stars, respectively. Note that in both cases the sample of stars within the first range of masses represents less than 1% of the total amount of stars. What happens, in this case, is that the sample is insufficient to apply a statistic method such as ABC, and we obtain a very wide distribution of α_1. Therefore, not in this case either we can consider the inferred α_1 a good indicator of its real

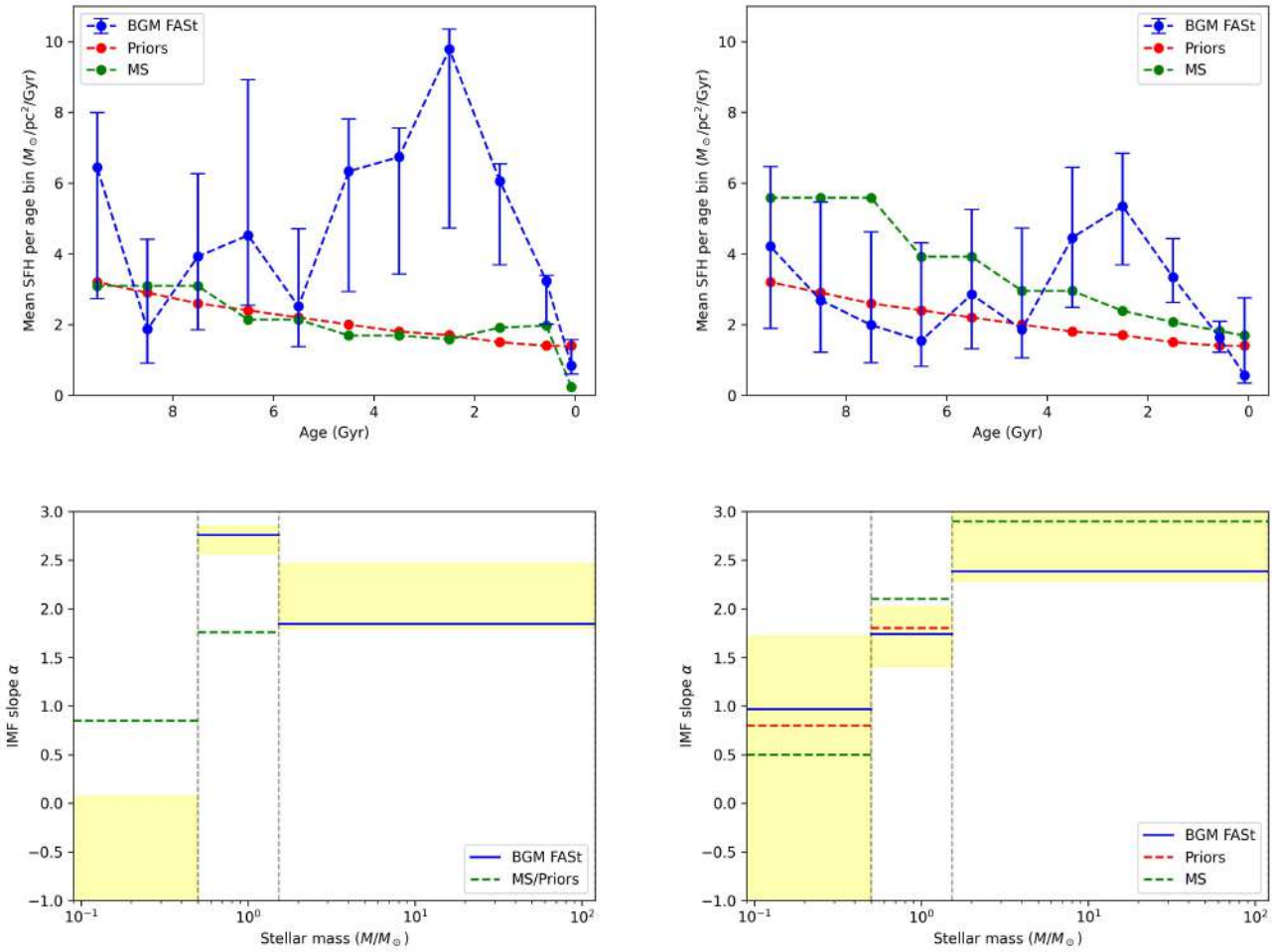


FIG. 6: In blue, it is shown the SFH (top) and the IMF (bottom) inferred with BGM FASt for the two fiducial cases GV13-MS22 (left) and G13-MS18 (right). In red and green are marked the values adopted for the priors and the MS, respectively. The BGM FASt most probable value of α_1 for the execution GV13-MS22 falls below the reasonable limits, taking a value of $\alpha_1 = -1.88^{+1.96}_{-1.81}$. The same situation happens to the prior of α_3 (which coincides with the value of α_3 in the MS in the GV13-MS22 configuration), which has a value of $\alpha_3 = 4.8$.

value.

Considering that we cannot take as valid α_1 in any of the presented cases, one may think that its incorrect fitting can affect the determination of the rest of the parameters. Actually, that is not what happens for the same reason explained above. The activity of α_1 in Eq. (2) is limited to those stars with masses between $0.09 - 0.5 M_\odot$. Since we have very few stars within this range, the consequences of a catastrophic α_1 are completely negligible. Besides, we performed an additional execution without fitting α_1 and we checked that the results do not change within the statistical fluctuations.

Regarding α_2 and α_3, we observe that both magnitudes leave their original priors to reach much lower and compatible values with the literature [Mor et al. 2019, Dickson et al. 2023, Kroupa 2008, Haghi et al. 2020], which is a positive sign that indicates a correct operation of BGM FASt and ABC. We find the values $\alpha_2 = 2.76^{+0.09}_{-0.19}$ and $\alpha_2 = 1.74^{+0.28}_{-0.33}$ for GV13-MS22 and G13-MS18, respectively. In the literature, the usual value of the second slope of the IMF is close to the classical Salpeter slope, $\alpha_2 = 2.3$. In Kroupa [2008], the second slope of a two-part power-law IMF is set at $\alpha_2 = 2.3 \pm 0.5$ for stars within a wide range of masses $0.5 M_\odot < M < 150 M_\odot$. A similar value is found in Mor et al. [2019], $\alpha_2 \approx 2$, in this case within a much narrower range comprising masses between $0.5 M_\odot$ and $1.53 M_\odot$. Considering so, the α_2 derived in this work is relatively far from the typical one in all the cases explored. Even though, the value of α_2 inferred from the G13-MS18 execution is more compatible with the literature than that obtained from G13-MS22. On the other hand, the values of the third slope of the IMF derived from GV13-MS22 and G13-MS18 are $\alpha_3 = 1.85^{+0.62}_{-0.05}$ and $\alpha_3 = 2.39^{+0.60}_{-0.10}$, respectively. Com-

paring them with the literature, we find that the one obtained with G13-MS18 fits considerably better the results reported in Dickson et al. [2023], Kroupa [2008], Haghi et al. [2020] than the α_3 from GV13-MS22. Especially the value derived in Dickson et al. [2023] for stars with $M > 1M_\odot$, $\alpha_3 = 2.37^{+0.48}_{-0.25}$, is in complete agreement with the α_3 from the G13-MS18 execution. On the other hand, all the explored values of α_3 both in this work and the aforementioned literature are far from the $\alpha_2 \approx 1.3$ given in Mor et al. [2019].

C. Hess diagrams comparison

Finally, we analyze the physical implications of the results inferred from the fiducial executions with the support of the Hess diagrams shown in Fig. 7. The first thing that draws attention are the three highlighted lines that appear in the third, four and fifth columns of the G13-MS18 execution (bottom in Fig. 7). To understand them, it must be taken into account that the stars below the solid black line in BGM FASt are not fitted in the Hess diagrams but in the luminosity distribution (see Subsect. III A). Therefore, the distance computed for the optimization of the parameters is actually the one we obtain from integrating the distances shown in the fourth column over all the *BpRp* range of colours. From this point of view, and taking into account that the two lower lines present opposite colours in the fourth column, it is derived that actually the total distance per bin of the LD is close to zero, and the two lines that are shown in Fig. 7 represent the compensation between the overcontribution of the PS (lower line, in green) and the Gaia DR3 (mid line, in red).

However, they still remain the questions of why these lines appear so highlighted and what is the origin of the third line in the upper part. These two questions have a physical origin and can be related to themselves. We propose the effect of two different possible problems in the model. First, we could have a problem with the metallicity of the stars. Variations in the metallicity generate a variety of almost parallel paths in the main sequence of the Hess diagram. This problem can be even more important for stars from the thick disc, where the content in metals is much different than that of the thin disc. The fact that in the current implementation of BGM FASt, we do not fit the stars of the thick disc, may give rise to distortions similar to the ones observed. The second possible origin of these lines is the fact that BGM Std does not contemplate stars in the pre-main sequence in the generation of the MS. That would explain the lack of stars in the PS far from the ZAMS.

Finally, it is important to note that the presence of these three lines is much weaker for GV13-MS22 execution (top of Fig. 7). This can be due to two different reasons: either the MS of the configuration GV13-MS22 performs a better simulation of the low-mass regions of the Hess diagram, or the difference in the number of stars between the samples of GV13-MS22 and G13-MS18 (see Subsect. IV A) implies a much lower number of stars in the low-mass region for the former than for the latter, diminishing the distances in this region. This last idea makes sense if we take into account the values presented in Tab. I, where we find almost twice as many stars within the range $5 < M'_G < 15$ in the MS of the configuration G13-MS18 (630947 stars) than in the MS of GV13-MS22 (323316 stars).

Apart from the three lines in the low-mass region of the Hess diagram, we also find some other conflicting zones in both executions. For instance, the turn-off and the giant branch are not well defined, especially in high latitudes (top rows in Fig. 7). We find in these regions a lack of stars in the PS, which can be due to the evolutionary tracks used in BGM Std. Another option is that this discordance comes from the stars of the thick disc. This possibility is compatible with the fact that these conflicting regions seem to be more spread in G13-MS18 than in GV13-MS22, which corresponds to the evidence that in the MS of the G13-MS18 configuration, the stars located at the thick disc represent 16% of the total number of stars (1138473 stars), while in the case of the MS from GV13-MS22, the percentage is reduced to an 11% (635214 stars).

In addition, the GV13-MS22 execution presents an excess of stars in the high-mass regime of the main sequence. This region is probably the most affected by the possible variations in the inferred parameters of BGM FASt, especially the value of α_3. In Subsect. V B we found a lower value of the third slope of the IMF for GV13-MS22 than for G13-MS18, corresponding to a top-light IMF in the high-mass range [Kroupa and Jerabkova 2021] that can cause an overproduction of high-mass stars. In this case, the better value of α_3 obtained in G13-MS18 may be responsible for the better performance of this execution with respect to GV13-MS22 in this region.

On the other hand, we observe a high discordance between the PS and the Gaia DR3 in the regions with the most reddened stars. The problems with the extinction map in BGM Std are not still well-resolved. Actually, this matter is so complex that studies on the performance of BGM FASt using different extinction maps were presented in Mor et al. [2019, 2018]. That is the reason why, as it is explained in Subsect. III A, we only consider stars with $-0.42 < (Bp - Rp) < 2.73$ in the range $-1 < M'_G < 5$, to avoid problems while fitting highly-reddened stars. Even though, the majority of them are still in the sample and may give rise to some distortions in the fitting process. Nevertheless, as can be seen in the third column of Fig. 7, the influence of these stars in the final total distance is negligible, since there are just a few of them.

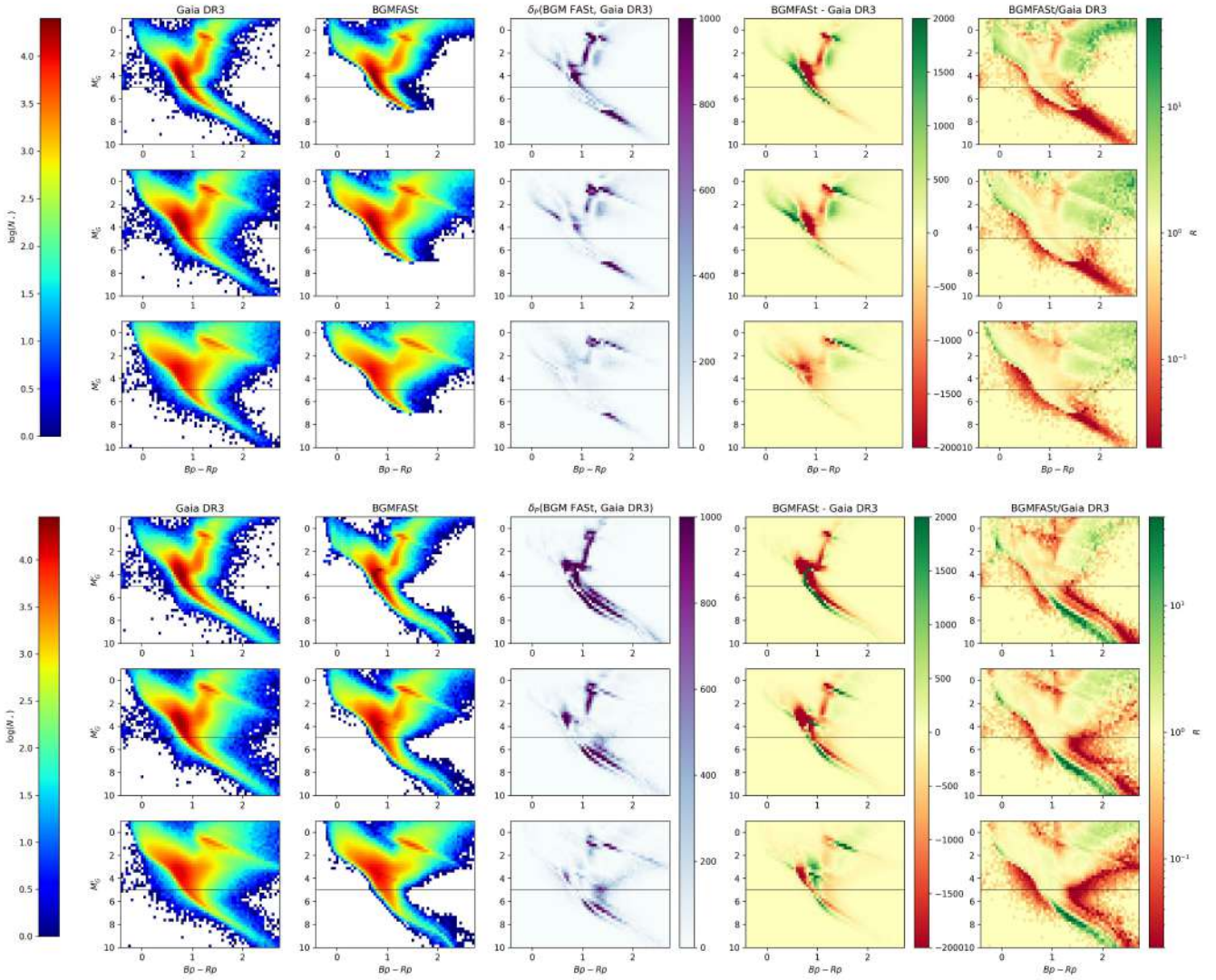


FIG. 7: Hess diagrams for the three considered latitude ranges ($30 < |b| < 90$, $10 < |b| < 30$, $0 < |b| < 10$ represented in top, mid and bottom rows, respectively) obtained from the inferred parameters in the fiducial executions GV13-MS22 (top) and G13-MS18 (bottom). From the left-most column to the right-most columns, they are shown the Hess diagrams of the Gaia DR3, the BGM FASt PS, the metric distance per pixel (each one of the elements of the summation of Eq. (III B)), the absolute difference in the number of stars and the ratio R between the PS and the Gaia DR3. In addition, it is shown with a black solid line the separation between the stars for which the distance is computed in a binned Hess diagram ($-1 < M'_G < 5$) and those for which we use the luminosity distribution ($5 < M'_G < 15$).

VI. CONCLUSIONS & FUTURE WORK

In this work, we have implemented, executed and analyzed in a Cloud Computing Environment more than 12 full-sky Besançon Galaxy Model Fast Approximate Simulations [Mor et al. 2019] for the first time up to magnitude $G < 13$, consuming $> 5 \cdot 10^5$ CPU hours of executions. We have used the Approximate Bayesian Computation, two different Mother Simulations generated with BGM Std [Robin et al. 2022, Czekaj et al. 2014] and the Gaia DR3 [Gaia Collaboration et al. 2022] to infer the star formation history and the initial mass function of the thin disc in the Solar neighbourhood.

Our results, from all the executions performed with different priors and Mother Simulations, definitively unveil the existence of a statistically significant burst of star formation peaking ~ 2.5 Gyr ago. These results are compatible with the SFH enhancements reported in the last years by several authors using different techniques [Mor et al. 2019, Sysoliatina and Just 2021]. For the third slope of the IMF, we recover a value of $\alpha_3 = 2.39^{+0.60}_{-0.10}$ close to the classical Salpeter values and in complete agree-

ment with Dickson et al. [2023], Kroupa [2008], Haghi et al. [2020]. Although statistically robust, the values obtained for the α_2 slope at intermediate masses from our two fiducial executions remain unclear. We have obtained $\alpha_2 \in [1.7, 2.8]$ for the range of masses between $0.5 - 1.53 M_{\odot}$. More work is required to understand these discrepancies. We confirm, as expected, that for the low-mass range of the IMF ($M < 0.5 M_{\odot}$) the slope α_1 remains unreachable when using a sample limited to $G < 13$, since faint sources are needed to characterize the contribution of the low-mass main sequence stars.

Our results report a significant variation in the total surface stellar density in the Solar neighbourhood for the two fiducial and most reliable executions. As known, this quantity is crucial for the estimation of the local dark matter in the Solar neighbourhood, so further work with BGM Std and BGM FASt is required to constrain this value.

Last, as a product of the present master thesis, we have developed, implemented and tested new tools to analyze the statistical performance of the BGM FASt outputs. We have provided a new framework to characterize and quantify through the Poissonian distance the contribution of the Mother Simulation ingredients, the BGM FASt inferred parameters and the limiting magnitude of the working sample. Using these tools we have identified some of the key tasks that needed to be upgraded in future BGM FASt and ABC executions. Among others:

- To increase the limiting magnitude to explore the first range of the IMF. At this point, we will have to consider the contribution of the stars from the thick disc, incorporating the corresponding parameters in the inference process.
- To execute an iterative process to fix the total surface density required for the dynamical self-consistency of the stellar system at a Galactic radius between 5-11 kpc.
- To modify the summary statistics and the distance metric to obtain a better characterization of the resemblance between the pseudo-simulation and the catalogue data.

As a first step, to confirm the results obtained in this work it will be crucial to loop again the process by fitting the outcomes of this work (IMF and SFH) to the strategy developed by Robin et al. [2022] to fit the Galactic potential and obtain a dynamically self-consistent solution. Finally, it is important to mention that the present work has allowed the Gaia-UB team to deeply evaluate the numerous benefits and caveats of the execution of highly demanding computational simulations for the Scientific Exploitation of Gaia Data in a Cloud Computing Environment. Our work corroborates the impressive capabilities of the SPARK unified analytics engine for the use of large-scale data processing tasks as the ones required by BGM FASt. These tasks are mandatory to provide new insights into fundamental and key pending questions in Galactic Astrophysics such as the derivation of the initial mass function and/or the unveiling of the star formation history of the Galactic disc in the Solar neighbourhood.

Acknowledgments

I acknowledge Dr Francesca Figueras for the opportunity to work with her in the context of BGM FASt, as well as for her constant support and willingness. I also acknowledge Dr Annie Robin for contributing to all the scientific discussions and for providing the Mother Simulations required to run BGM FASt. I am indebted to Sergi Bartolomé for his work and advice in the configuration and execution in the Cloud Environment. I acknowledge Dr Roger Mor for his valuable advice, comments and suggestions regarding the statistical treatments presented in this work. Dr Jordi Portell has continuously supervised the work with a detailed analysis of the OCRE Galactic RainCloudS resources. I also acknowledge Dr Olivier Bienaymé for his contribution and discussions on the dynamical self-consistency of the solutions and the Galactic density laws. This work was partially supported by the OCRE awarded project Galactic Research in Cloud Services (Galactic RainCloudS). OCRE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 824079.

R. Mor, A. C. Robin, F. Figueras, S. Roca-Fàbrega, and X. Luri, A&A **624**, L1 (2019), 1901.07564.

A. C. Robin, O. Bienaymé, J. B. Salomon, C. Reylé, N. Lagarde, F. Figueras, R. Mor, J. G. Fernández-Trincado, and J. Montillaud, A&A **667**, A98 (2022), 2208.13827.

K. Sysoliatina and A. Just, A&A **647**, A39 (2021), 2102.09311.

T. Ruiz-Lara, C. Gallart, E. J. Bernard, and S. Cassisi, Nature Astronomy **4**, 965 (2020), 2003.12577.

N. Dickson, V. Hénault-Brunet, H. Baumgardt, M. Gieles, and P. J. Smith, MNRAS **522**, 5320 (2023), 2303.01637.

E. J. Bernard, in *Rediscovering Our Galaxy*, edited by C. Chiappini, I. Minchev, E. Starkenburg, and M. Valentini (2018), vol. 334, pp. 158–161.

M. Cignoni, S. Degl'Innocenti, P. G. P. Moroni, and S. N. Shore, in *Statistical Challenges in Modern Astronomy IV*, edited by G. J. Babu and E. D. Feigelson (2007), vol. 371 of *Astronomical Society of the Pacific Conference Series*, p. 407.

P. Kroupa, in *Pathways Through an Eclectic Universe*, edited by J. H. Knapen, T. J. Mahoney, and A. Vazdekis (2008), vol. 390 of *Astronomical Society of the Pacific Conference Series*, p. 3, 0708.1164.

H. Haghi, G. Safaei, A. H. Zonoozi, and P. Kroupa, ApJ **904**,

43 (2020), 2012.07095.
J. Li, C. Liu, Z.-Y. Zhang, H. Tian, X. Fu, J. Li, and Z.-Q. Yan, Nature **613**, 460 (2023), 2301.07029.
K. R. Stewart, J. S. Bullock, R. H. Wechsler, A. H. Maller, and A. R. Zentner, ApJ **683**, 597 (2008), 0711.5027.
T. Antoja, A. Helmi, M. Romero-Gómez, D. Katz, C. Babusiaux, R. Drimmel, D. W. Evans, F. Figueras, E. Poggio, C. Reylé, et al., Nature **561**, 360 (2018), 1804.10196.
V. M. Pelkonen, P. Padoan, T. Haugbølle, and Å. Nordlund, MNRAS **504**, 1219 (2021), 2008.02192.
A. C. Robin, C. Reylé, S. Derrière, and S. Picaud, A&A **409**, 523 (2003).
M. A. Czekaj, A. C. Robin, F. Figueras, X. Luri, and M. Haywood, A&A **564**, A102 (2014), 1402.3257.
Gaia Collaboration et al., A&A **595**, A1 (2016), 1609.04153.
Gaia Collaboration et al., arXiv e-prints arXiv:2208.00211 (2022), 2208.00211.
R. Mor, A. C. Robin, F. Figueras, and T. Antoja, A&A **620**, A79 (2018), 1809.03511.
E. Jennings and M. Madigan, Astronomy and Computing **19**, 16 (2017), 1608.07606.
D. W. Evans, M. Riello, F. De Angeli, J. M. Carrasco, P. Montegriffo, C. Fabricius, C. Jordi, L. Palaversa, C. Diener, G. Busso, et al., A&A **616**, A4 (2018), 1804.09368.
P. Kroupa and T. Jerabkova, arXiv e-prints arXiv:2112.10788 (2021), 2112.10788.

Appendix A: Infrastructure, Performance & Accountability of the OCRE Computational Cloud

1. Infrastructure

We ran the BGM FASt Galaxy Model in Spark clusters (Dataproc) with a total maximum number of three clusters. For instance, in June 5th we had the following infrastructure:

- 784 cores (520 Intel Cascade Lake and 264 Intel Haswell).
- 2.9 TB RAM.
- 7.6 TB disc (4 TB in SSD).

Each cluster was created as follows:

- 1 Master node, with 4 vCPUs, 16GB of RAM and 500GB of disc storage.
- 4 Worker nodes, each with 64 vCPUs, 240GB of RAM and 500GB of disc storage.
- A shared Google Cloud Storage disc.

The first cluster used Intel Haswell CPUs and HDD disc storage. When we wanted to perform different runs simultaneously, we thought we could also try to improve the new clusters' performance by using a newer processor architecture and solid-state drives. That is why we moved to use Intel Cascade Lake CPUs and SSD discs for the second and third clusters. The shared disc stored the data inputs between the different machines and their computation logs.

2. Performance

We use the following command line to execute the BGM FASt code in the Cloud Environment:

```
/opt/conda/anaconda/bin/spark-submit
--num-executors 31 --executor-cores 8
master_bgmfast.py.
```

We opt for this command line after considering the following ideas. The maximum number of executors we can use per process is the total number of CPUs of the cluster, subtracting one for the master node and one for each cluster for the performing of the internal processes of each machine. Taking into account that we have a maximum of 256 CPUs, this leaves us with 251 available CPUs that we can use to assign executors to run BGM FASt. To be able to run each pseudo-simulation quickly, it is necessary to assign more than one CPU to each executor. In this way, each calculation has more resources and the time per PS is shortened. We tested by expanding the number of CPUs assigned to each executor until 8 CPUs/executor, the point at which the time of each PS

no longer decreased and remained at 1s. Using 8 CPUs per executor lowers the number of available executors which becomes 251/8 = 31.37, which we round down to 31 executors.

3. Accountability

Since the time we adequately labelled the Google Cloud resources with the "BGMFast" tag (around 9 March 2023 and until 11 June 2023), we have spent a total of 22.6k€ in the executions. Resource consumption before this date was very small, just with initial setup and configuration activities. The day with the highest cost was 4 June (1023€), when we had the three clusters up and running plus a testing notebook. The highest costs come from N1 and N2 computing cores (7.77k€ and 3.99k€, respectively), the licensing fee for Google Cloud Dataproc (4k€), the N1 and N2 RAM (3.9k€ and 2.1k€, respectively), and the storage (462€ standard plus 274€ SSD). Note that the second cluster was added at a quite late stage (23-May), and the third one just during the final days (2-June). The daily costs of the Cloud infrastructure are shown in Fig. 8.

Appendix B: Clues for the proper execution of BGM FASt and ABC

1. Warnings and suggestions for the application of the approximate cross-checking strategy (from Subsect. III D)

1. The parameters used to generate the BGM Std that are not taken into account in BGM FASt and those derived in BGM FASt are physically and computationally bounded. That means that introducing BGM FASt parameters obtained with different configurations of the MS may give rise to non-sense combinations if we analyze the whole set of ingredients. That can be given from the statistical point of view (the resulting distance is much larger) or from the physical point of view (the combination of parameters is physically non-sense).

2. It is highly recommended to compare MS up to the same limiting magnitude. The reason is that, as it is commented in Subsect. III B, the increasing of the limiting magnitude implies both a higher scale factor $N_\star$ and the involvement of more complex physics. The former can be surpassed using R instead of δ_P as an evaluator, but the latter is intrinsic to the sample and cannot be avoided (except, for instance, if we consider for the comparison only the less conflicting regions, a possibility that is not explored in this work).

3. It is also recommended to start in both cases from the same set of priors for the BGM FASt parameters. This is the least important matter, since the resulting $\bar{\theta}_{mpv}$ is not dependent on the selection of the initial priors if the distance thresholds and the number of steps are taken correctly (see Sect. V). However, it gives consistency to the results.

4. In some cases, it is possible that an overfit of the MS by BGM FASt gives rise to a set of BGM FASt parameters that only works well in the MS that generated them. This can also happen if we work with a problematic MS which is not a good approximation to the observed data. To detect this kind of problem, it is always important to analyze and compare the CMD and LD of the catalogue data and the PS resulting from $\bar{\theta}_{mpv}$, in order to find out problematic regions.

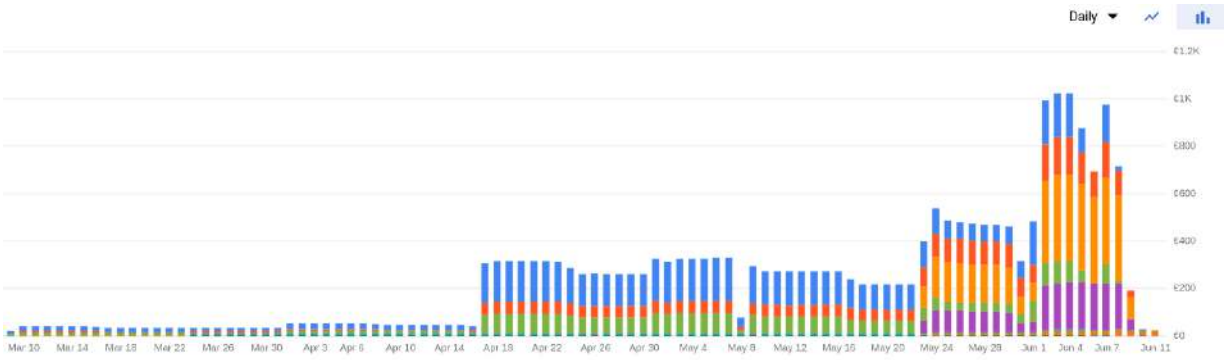
2. Definition and application of the new concepts developed in this work

- **The role of the intrinsic physical distance (from Subsect. III B):** when going from $G < 12$ to $G < 13/V < 13$ (see Subsect. IV B), we increase the number of stars in the MS and Gaia a factor 1.9 and 1.5, respectively (see Tab. I). However, if we compare the distances we obtain injecting the same set of BGM FASt parameters $\bar{\theta}_{MS}$ (the set of parameters corresponding to the MS) in both configurations, we find an increasing factor of 1.4, which is lower than the expected 1.9 and 1.5 if we only take into account the scale factor of the number of stars, $N_\star$. That demonstrates that we must take into account the role of the intrinsic physical distance when comparing Poisson distances obtained up to different limiting magnitudes.

- **The meaning of the asymptotic regime (from Subsect. IV B):** from Fig. 3 we observe that GV13-MS22 explores in detail the asymptotic regime before ending its process while G13-MS18 reaches the minimum threshold of distance before completely entering into this trend. This result is indicative that δ_{min} is better chosen for GV13-MS22 than for G13-MS18, which is in agreement with the fact that we perform for GV13-MS22 an exploratory execution to define its lower threshold of distance while in the case of G13-MS18, we use the results of G13-MS18-E, that departs from a much narrower range $[\delta_{max}, \delta_{min}]$.

Appendix C: BGM FASt Python package

We show in Fig. 9 and Fig. 10 the flux diagrams of the old and new `bgmfast` Python package. In the scheme of the old version (Fig. 9), shown in red are the functions that are no longer needed due to the fact that in the new execution of BGM FASt, we do not fit the



DOWNLOAD CSV

SKU	Service	SKU ID	Usage	Cost	Discounts	Promotions and others	Subtotal
N1 Predefined Instance Core running in EMEA	Compute Engine	9431-52B1-2C4F	351,554.08 hour	€10,053.47	-€2,280.11	–	€7,773.36
Licensing Fee for Google Cloud Dataproc (CPU cost)	Compute Engine	DC33-D3C7-25CA	485,451.81 hour	€4,002.74	€0.00	–	€4,002.74
N2 Instance Core running in EMEA	Compute Engine	9F61-45D7-D4FB	142,974.04 hour	€4,127.37	-€133.82	–	€3,993.55
N1 Predefined Instance Ram running in EMEA	Compute Engine	39F4-0112-6F39	1,318,327.89 gibibyte hour	€5,053.41	-€1,146.11	–	€3,907.30
N2 Instance Ram running in EMEA	Compute Engine	A109-54C1-7CB0	571,896.2 gibibyte hour	€2,212.94	-€69.87	–	€2,143.07
Storage PD Capacity	Compute Engine	D973-5D65-BAB2	13,808.89 gibibyte month	€462.29	€0.00	–	€462.29
SSD backed PD Capacity	Compute Engine	B188-61DD-52E4	1,897.39 gibibyte month	€274.32	€0.00	–	€274.32
External IP Charge on a Standard VM	Compute Engine	C054-7F72-A02E	15,577.27 hour	€45.44	€0.00	–	€45.44
Standard Storage Belgium	Cloud Storage	A703-5CB6-E0BF	525.31 gibibyte month	€8.81	€0.00	–	€8.81
Download Worldwide Destinations (excluding Asia & Australia)	Cloud Storage	22EB-AAE8-FBCD	132.97 gibibyte	€2.69	€0.00	–	€2.69

FIG. 8: Daily costs of BGM FASt Cloud Computing. N1 refers to the first cluster, and N2 includes the second and third. The columns indicate the name and type of service, the resource usage and their cost as well as the applied discounts.

dynamical parameters of the model and we consider a non-parametric SFH for the MS. In both figures, it is marked with rectangles the part of the code where the parameters are set (blue), where it reads the catalogue data (green), and where it generates the PS and loops the process with ABC (grey).
Finally, in the new version of the package (Fig. 10) it is also shown at the bottom-right part of the scheme a box containing the additional scripts developed in the current work:

- `Gaia_instrument_model_filter.py`, that includes the error models applied to the MS (see Sect. IV).
- `set_inputs_for_bgmfast.py`, which performs the cut in G and V magnitudes and displays the columns of the data as expected by `master_bgmfast.py`.
- `analysis_tools.py`, comprising the statistical tools developed in Sect. III as well as the possibility of quickly building plots for the analysis of the results.

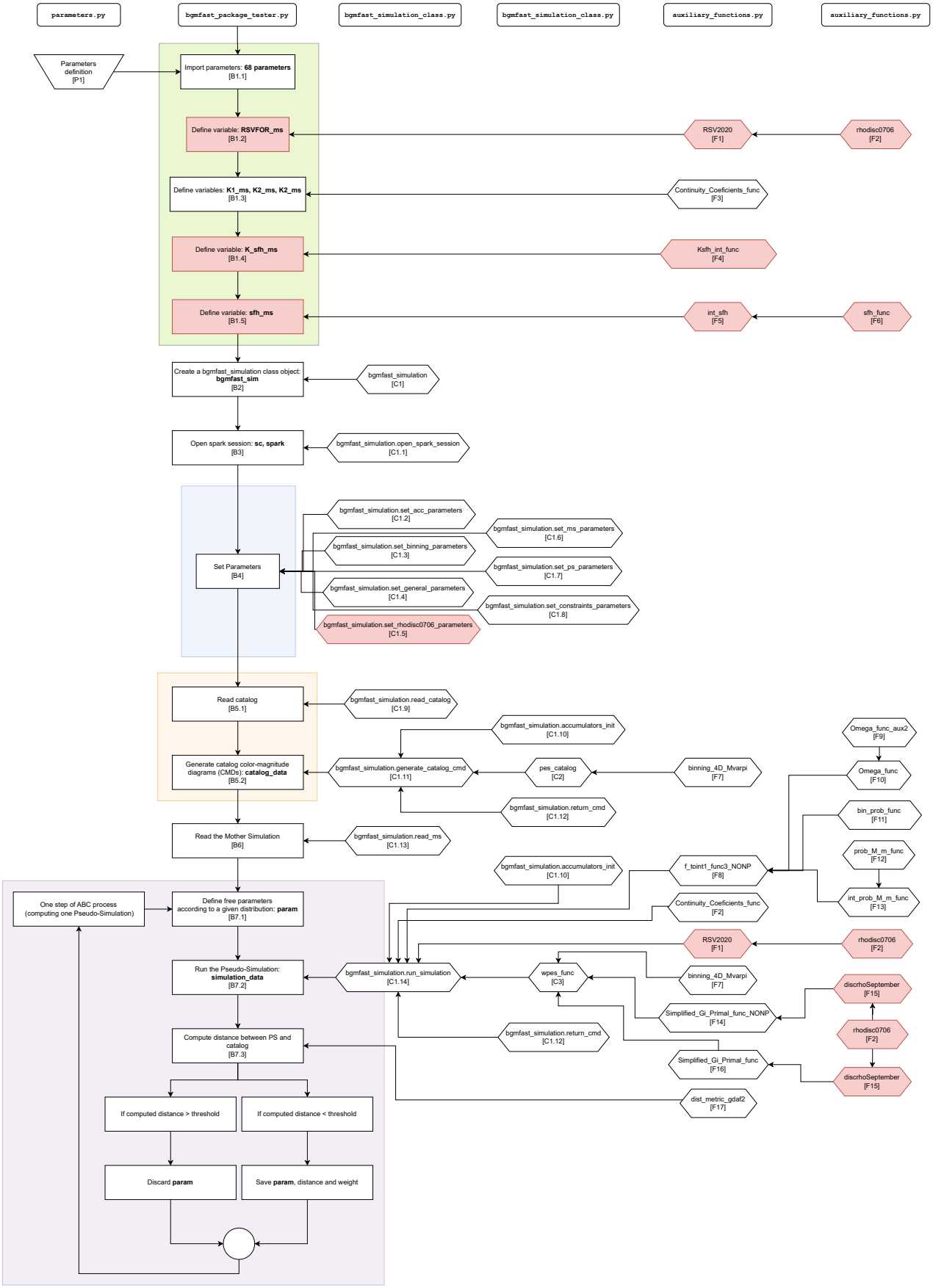


FIG. 9: Flux diagram of the old version of the BGM FASt Python package.

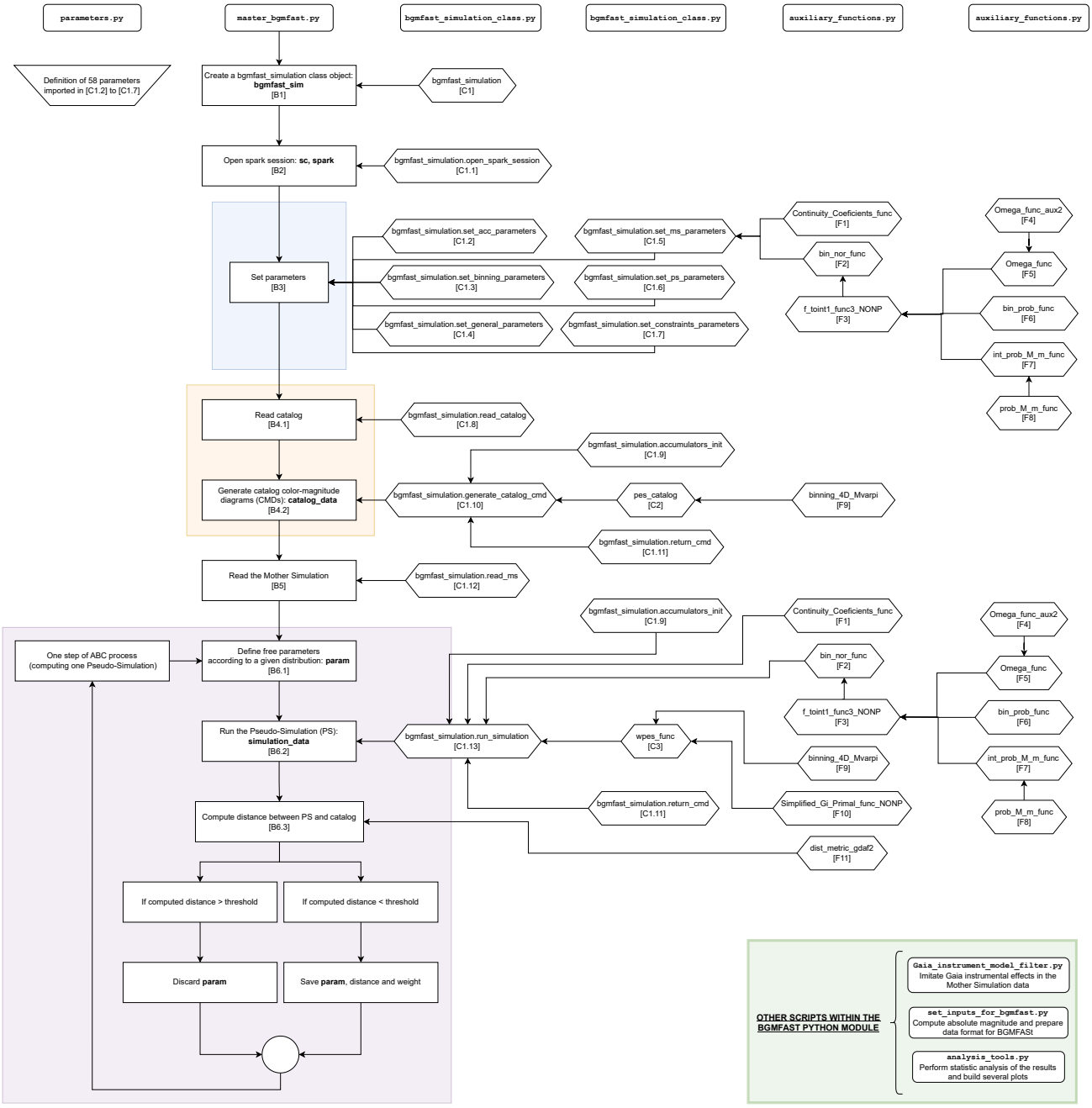


FIG. 10: Flux diagram of the new version of the BGM FASt Python package.