A&A 658, A91 (2022)
https://doi.org/10.1051/0004-6361/202142369
© ESO 2022

Astronomy & Astrophysics

Photo-astrometric distances, extinctions, and astrophysical parameters for *Gaia* EDR3 stars brighter than $G = 18.5$★

F. Anders[1,2], A. Khalatyan[2], A. B. A. Queiroz[2,3,4], C. Chiappini[2,3], J. Ardèvol[1], L. Casamiquela[5], F. Figueras[1], Ó. Jiménez-Arranz[1], C. Jordi[1], M. Monguió[1], M. Romero-Gómez[1], D. Altamirano[1], T. Antoja[1], R. Assaad[6], T. Cantat-Gaudin[7], A. Castro-Ginard[8], H. Enke[2], L. Girardi[9], G. Guiglion[2], S. Khan[10], X. Luri[1], A. Miglio[11,12,13], I. Minchev[2], P. Ramos[14], B. X. Santiago[15,3], and M. Steinmetz[2]

[1] Institut de Ciències del Cosmos, Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
e-mail: fanders@icc.ub.edu
[2] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
e-mail: akhalatyan@aip.de
[3] Laboratório Interinstitucional de e-Astronomia – LIneA, Rua Gal. José Cristino 77, Rio de Janeiro 20921-400, Brazil
[4] Institut für Physik und Astronomie, Universität Potsdam, Haus 28 Karl-Liebknecht-Str. 24/25, 14476 Golm, Germany
[5] Laboratoire d'Astrophysique de Bordeaux, Univ. Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[6] Department of Physics, University of Surrey, Guildford GU2 7XH, UK
[7] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[8] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
[9] Osservatorio Astronomico di Padova, INAF, Vicolo dell'Osservatorio 5, 35122 Padova, Italy
[10] Institute of Physics, Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
[11] Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
[12] INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, 40129 Bologna, Italy
[13] School of Physics and Astronomy, University of Birmingham, Edgbaston B15 2TT, UK
[14] Observatoire astronomique de Strasbourg, Université de Strasbourg, CNRS, 11 rue de l'Université, 67000 Strasbourg, France
[15] Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, Porto Alegre 91501-970, Brazil

Received 4 October 2021 / Accepted 5 November 2021

ABSTRACT

We present a catalogue of 362 million stellar parameters, distances, and extinctions derived from *Gaia*'s Early Data Release (EDR3) cross-matched with the photometric catalogues of Pan-STARRS1, SkyMapper, 2MASS, and AllWISE. The higher precision of the *Gaia* EDR3 data, combined with the broad wavelength coverage of the additional photometric surveys and the new stellar-density priors of the `StarHorse` code, allows us to substantially improve the accuracy and precision over previous photo-astrometric stellar-parameter estimates. At magnitude $G = 14$ (17), our typical precisions amount to 3% (15%) in distance, 0.13 mag (0.15 mag) in V-band extinction, and 140 K (180 K) in effective temperature. Our results are validated by comparisons with open clusters, as well as with asteroseismic and spectroscopic measurements, indicating systematic errors smaller than the nominal uncertainties for the vast majority of objects. We also provide distance- and extinction-corrected colour-magnitude diagrams, extinction maps, and extensive stellar density maps that reveal detailed substructures in the Milky Way and beyond. The new density maps now probe a much greater volume, extending to regions beyond the Galactic bar and to Local Group galaxies, with a larger total number density. We publish our results through an ADQL query interface (`gaia.aip.de`) as well as via tables containing approximations of the full posterior distributions. Our multi-wavelength approach and the deep magnitude limit render our results useful also beyond the next *Gaia* release, DR3.

Key words. stars: distances – stars: fundamental parameters – Galaxy: general – Galaxy: stellar content – Galaxy: structure

1. Introduction

Since its launch in 2013 the European Space Agency's flagship mission *Gaia* (Gaia Collaboration 2016) has revolutionised Galactic astronomy and its neighbouring fields (Brown 2021). The precision and accuracy of our knowledge of the Solar System (e.g. Gaia Collaboration 2018a; Bailer-Jones et al. 2018a; Portegies Zwart 2021), stellar astrophysics (e.g. Jao et al. 2018; Lanzafame et al. 2019; Mowlavi et al. 2021), the immediate solar vicinity (e.g. Gaia Collaboration 2021a,b; Reylé et al. 2021), open star clusters (e.g. Cantat-Gaudin et al. 2018; Cantat-Gaudin & Anders 2020; Castro-Ginard et al. 2020), distant regions of the Milky Way (e.g. Ramos et al. 2021; Gaia Collaboration 2021c; Zari et al. 2021) and the Local Group (e.g. Gaia Collaboration 2018b, 2021d; Antoja et al. 2020), the Galactic potential (e.g. Crosta et al. 2020; Cunningham et al. 2020; Hattori et al. 2021), and even the Hubble constant (e.g. Breuval et al. 2020; Riess et al. 2021; Baumgardt & Vasiliev 2021) are constantly increasing thanks to *Gaia*.

In the context of Galactic archaeology, *Gaia* has also enabled a completely new line of precision studies tracing

★ The catalog is available at the CDS via anonymous ftp to `cdsarc.u-strasbg.fr` (`130.79.128.5`) or via `http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/658/A91`

the past accretion events of the Milky Way (Helmi 2020; Kruijssen et al. 2020; Pfeffer et al. 2021). Often this is achieved by combining the complete phase-space information with detailed chemistry from spectroscopic surveys (e.g. Aguado et al. 2021; Gudin et al. 2021; Limberg et al. 2021a,b; Montalbán et al. 2021; Naidu et al. 2021; Shank et al. 2021).

The latest *Gaia* data release, Early Data Release 3 (*Gaia* EDR3; Gaia Collaboration 2021e), covers the first 34 months of observations with positions and photometry for 1.8×10^9 sources (Riello et al. 2021), proper motions and parallaxes for 1.5×10^9 sources (Lindegren et al. 2021a), and radial velocities for 7×10^6 sources (Seabroke et al. 2021; Gaia Collaboration 2021e). With respect to Data Release 2 (*Gaia* DR2; Gaia Collaboration 2018c), the proper motions are by a factor of 2 more precise, and parallax uncertainties are reduced by 20% (see Fabricius et al. 2021 for details).

In a previous work (Anders et al. 2019, hereafter A19) based on *Gaia* DR2, our group derived Bayesian stellar parameters, distances, and extinctions for 265 million stars brighter than $G = 18$ with the `StarHorse` code (Santiago et al. 2016; Queiroz et al. 2018). The combination of precise *Gaia* DR2 parallaxes and optical photometry with the multi-wavelength photometry of Pan-STARRS1 (Chambers et al. 2016), 2MASS (Cutri et al. 2003), and AllWISE (Cutri et al. 2013) substantially improved the accuracy of the extinction and effective temperature estimates provided with only *Gaia* DR2 (Andrae et al. 2018). A selection of the most reliable in- and output data, a sample of 137 million stars, allowed A19 to detect the imprint of the Galactic bar both in the stellar density distribution and in proper motion maps (further studied with APOGEE spectroscopy in Queiroz et al. 2021).

The results of our *Gaia* DR2 `StarHorse` run presented in A19 have been used in a wide variety of science cases, including exoplanetary research (Sozzetti et al. 2021), interstellar extinction (Leike et al. 2020), runaway stars from supernova remnants (Lux et al. 2021), X-ray transients (Lamer et al. 2021), γ-ray astronomy (Steppa & Egberts 2020), ,the Galactic escape speed curve (Monari et al. 2018), the three-dimensionl phase-space structure of the Milky Way disc (Carrillo et al. 2019), and spectroscopic survey simulations (Chiappini et al. 2019).

Anticipating a significant improvement thanks to the new *Gaia* data, we update our analysis using the new EDR3 data in this paper, addressing some of the known caveats of our previous data release and reducing the uncertainties of the main output parameters by a factor of 2. In a parallel effort we will be publishing `StarHorse` results for spectroscopic surveys combined with *Gaia* (≈6 million stars) in Queiroz et al. (in prep.).

This paper is structured as follows: Section 2 presents the input data and Sect. 3 our method. In particular, Sect. 3.1 describes the updates to our code with respect to previous applications, and Sect. 3.3 explains how we flagged the new `StarHorse` results for *Gaia* EDR3. We then present some first astrophysical results in Sect. 4, mainly focussing on colour-magnitude diagrams (CMDs), extinction maps, and stellar density maps. The stellar density maps demonstrate the emergence of substructure beyond the detection of the Galactic bar, for example when focussing on metal-poor stars, the Magellanic Clouds, or the outer Milky Way halo. The precision and accuracy of the `StarHorse` EDR3 parameters are discussed in Sect. 5, providing comparisons to Galactic open clusters (OCs Cantat-Gaudin et al. 2020; Dias et al. 2021), asteroseismically derived parameters for giant stars (Miglio et al. 2021), and spectroscopic stellar parameters from the GALAH survey (Buder et al. 2021). We also make comparisons to previous results obtained from *Gaia* DR2 and EDR3 in Sect. 6. Finally, we conclude the paper with a summary and a brief outlook to the near future in Sect. 7.

2. Data

As input for `StarHorse`, we use the *Gaia* EDR3 data cross-matched with 2MASS, AllWISE, Pan-STARRS1, and SkyMapper (Onken et al. 2019), in the sense that all available good photometric measurements are used in the inference. The calibrations used in this paper are summarised in Table 1.

From *Gaia* EDR3 we use the parallaxes and three-band photometry, together with their associated uncertainties. We recalibrate the parallaxes following the recommendations of Lindegren et al. (2021b) who assessed the variations in the parallax zero point as a function of sky position, magnitude, and colour[1]. Furthermore, we inflate the corresponding parallax uncertainties by a magnitude-dependent factor, following Fabricius et al. (2021, see also El-Badry et al. 2021). In particular, we fit the inflation factor to their worst-case scenario of Fig. 19 in Fabricius et al. (2021) (a crowded Large Magellanic Cloud field) to make sure that our parallax uncertainties are not underestimated. A further discussion of the fidelity of the *Gaia* parallaxes in our context is available in Rybizki et al. (2022) and in our Sect. 3.3.1.

Regarding the *Gaia* photometry, we use the precise EDR3 magnitudes (Riello et al. 2021) without any posterior correction other than the G-band correction advertised in Appendix A of Gaia Collaboration (2021e)[2], since they show much lower systematics ($\lesssim$0.01 mag; e.g. Fabricius et al. 2021; Niu et al. 2021) than the previous DR2 photometry (see e.g. Maíz Apellániz & Weiler 2018). For the BP/RP photometry, we follow the recommendation of Fabricius et al. (2021) and do not use magnitudes `phot_bp_mean_mag` > 20.5 or `phot_rp_mean_mag` > 20 in the inference.

The *Gaia* EDR3 cross-match to the large-area photometric surveys 2MASS, AllWISE, Pan-STARRS1 DR1, and SkyMapper DR2 is documented in Marrese et al. (2021). The main novelty with respect to A19 is the inclusion of SkyMapper data. From the SkyMapper DR2 data we only use the *griz* bands (with zero points recalibrated following Huang et al. 2021a) and refrain from using the u and v bands, because our default extinction law (Schlafly et al. 2016) should not be extrapolated to the ultraviolet.

For Pan-STARRS1, we apply the zero-point corrections recommended by Scolnic et al. (2015), and do not use magnitudes brighter than the saturation limit. With respect to A19, we apply more restrictive filters to the 2MASS and AllWISE photometry: only magnitudes with corresponding photometric quality flags 'A' or 'B' are accepted. The minimum photometric uncertainties used in the inference (reflecting also the systematic uncertainties of the passbands and the bolometric corrections) are given in Table 1. We also note that for $\lesssim$0.5% of the *Gaia* EDR3 sources, the *Gaia* cross-match with 2MASS returns in multiple matches. In these cases (pointing towards possible confusion), we do not use any 2MASS photometry.

[1] Code available at https://gitlab.com/icc-ub/public/gaiadr3_zeropoint/-/tree/master/

[2] https://github.com/agabrown/gaiaedr3-6p-gband-correction/blob/main/GCorrectionCode.ipynb

Table 1. Summary of the calibrations and data curation applied to the astrometric and photometric data for this work.

Parameter	Parameter regime	Calibration choice	Reference
$\varpi^{\rm cal}$		Use `zpt.py` calibration	Lindegren et al. (2021b)
$\sigma_\varpi^{\rm cal}$		$\texttt{uwu}(G) \cdot \texttt{parallax_error}$	Fit to Fabricius et al. (2021) Fig. 19a&b
G	`astrometric_params_solved` = 95	Colour-dependent correction	Gaia Collaboration (2021e) Appendix A
$g_{\rm PS1}$	`g_mean_psf_mag` > 14	`g_mean_psf_mag` − 0.020	Scolnic et al. (2015)
$r_{\rm PS1}$	`r_mean_psf_mag` > 15	`r_mean_psf_mag` − 0.033	
$i_{\rm PS1}$	`i_mean_psf_mag` > 15	`i_mean_psf_mag` − 0.024	
$z_{\rm PS1}$	`z_mean_psf_mag` > 14	`z_mean_psf_mag` − 0.028	
$y_{\rm PS1}$	`y_mean_psf_mag` > 13	`y_mean_psf_mag` − 0.011	
$g_{\rm SM}, r_{\rm SM}$		$E(B-V)$ dependent zero-point shifts	Huang et al. (2021a)
$u_{\rm SM}, v_{\rm SM}$		Not used in inference	Schlafly et al. (2016)
$\sigma_{\rm mag}$	*Gaia* EDR3	$\max\{\sigma_{\rm mag,source}, 0.02{\rm mag}\}$	
	2MASS, AllWISE	$\max\{\sigma_{\rm mag,source}, 0.03\ {\rm mag}\}$	
	Pan-STARRS1 DR1	$\max\{\sigma_{\rm mag,source}, 0.04\ {\rm mag}\}$	
	SkyMapper DR2	$\max\{\sigma_{\rm mag,source}, 0.05\ {\rm mag}\}$	

3. Method: The StarHorse code

StarHorse (Queiroz et al. 2018) is an isochrone-fitting code tailored to derive distances d, extinctions (at $\lambda = 542$ nm) A_V, ages τ, masses m_*, effective temperatures $T_{\rm eff}$, metallicities [M/H], and surface gravities $\log g$ for field stars. In the absence of spectroscopic input data, it takes as input only the measured parallax ϖ and a set of observed magnitudes m_λ to estimate how close a stellar model is to the observed data.

StarHorse also includes priors about the geometry, metallicity and age characteristics of the main Galactic components. The priors adopted here are very similar to those in Queiroz et al. (2018), A19, and Queiroz et al. (2020): a Chabrier (2003) initial mass function; exponential spatial density profiles for thin and thick discs; a spherical halo and a triaxial bulge/bar component, as well as broad Gaussian distributions for the age and metallicity distribution priors. The normalisation of each Galactic component, as well as the solar position, were taken from Bland-Hawthorn & Gerhard (2016).

3.1. Code updates and improvements

With respect to A19 and Queiroz et al. (2020), we have implemented some changes that help to improve the performance of StarHorse in the context of *Gaia* EDR3.

3.1.1. A more informative interstellar extinction prior

One of the drawbacks of the A19 *Gaia* DR2 run was the a priori limit in interstellar extinction to $A_V < 4$ mag for sources with low signal-to-noise parallax measurements (`parallax_over_error` < 5). This resulted in poor convergence or biased results for distant obscured objects in the Galactic plane. For the present EDR3 run we therefore update our previously uninformative top-hat A_V prior to a prior that takes into account our knowledge of Galactic interstellar extinction.

For the region of the sky covered by Pan-STARRS1, we use the large-scale three-dimensional extinction map of Green et al. (2019). For the missing part (1/4) of the sky, we use the 2MASS-derived three-dimensional extinction model by Drimmel et al. (2003). To get the range of possible distances for each star (needed to query the extinction maps), we invert the EDR3 zero-point-corrected parallax measurements to estimate a prior extinction value range for each star (using a maximum of $A_{V_{\rm prior}} =$ 10 mag, considering that our sample is limited by $G < 18.5$). The extinction prior is then defined as a very broad Gaussian distribution around the central value (with $\sigma_{AV,{\rm prior}} = \max\{0.2, 0.33 \cdot A_{V_{\rm prior}}\}$).

3.1.2. Extragalactic priors

In A19 we saw that the stellar populations of the Magellanic Clouds, the Sagittarius (Sgr dSph) galaxy, and a number of relatively nearby globular clusters, whose stellar densities were not accounted for by our priors until now, left a spurious imprint on the posterior Galactic density distribution inferred with StarHorse.

For the new runs we therefore included new extragalactic and globular cluster priors in the calculation of the global prior. For the extragalactic resolved stellar population we use the updated list (October 2019) of Local Group Galaxies from McConnachie (2012), which comprises sky positions, distances, foreground extinctions, apparent dimensions, central densities, metallicities, masses, and other basic quantities for the Local Group. We manually added M31 to this list, and curated the list for the most prominent objects on the sky: the Magellanic Clouds and the Sgr dSph galaxy. For all but these objects we estimate the mass of the external galaxy by inverting the mass-metallicity relation of Panter et al. (2008) (linearly extrapolated below [Fe/H] = −1). For the Galactic globular clusters we used the recent catalogue of Hilker et al. (2020).

The sky distribution of all considered objects is shown in Fig. 1. If a star's celestial coordinates coincide with those of an external galaxy or globular cluster within five half-light radii, we add to the Milky Way foreground prior an additional population corresponding to the characteristics of that object (stellar density, distance, metallicity). For the sake of simplicity, the density profile priors for all objects are assumed to be three-dimensional Gaussians.

3.1.3. Update of the bar angle in the priors

Our knowledge about the large-scale parameters of the Milky Way is constantly improving. In light of the growing evidence for a bar angle around 27 ± 2 deg (see discussion in Bland-Hawthorn & Gerhard 2016, reinforced also by Queiroz et al. 2021), we have updated the angle of the Galactic bar in our prior to that value.

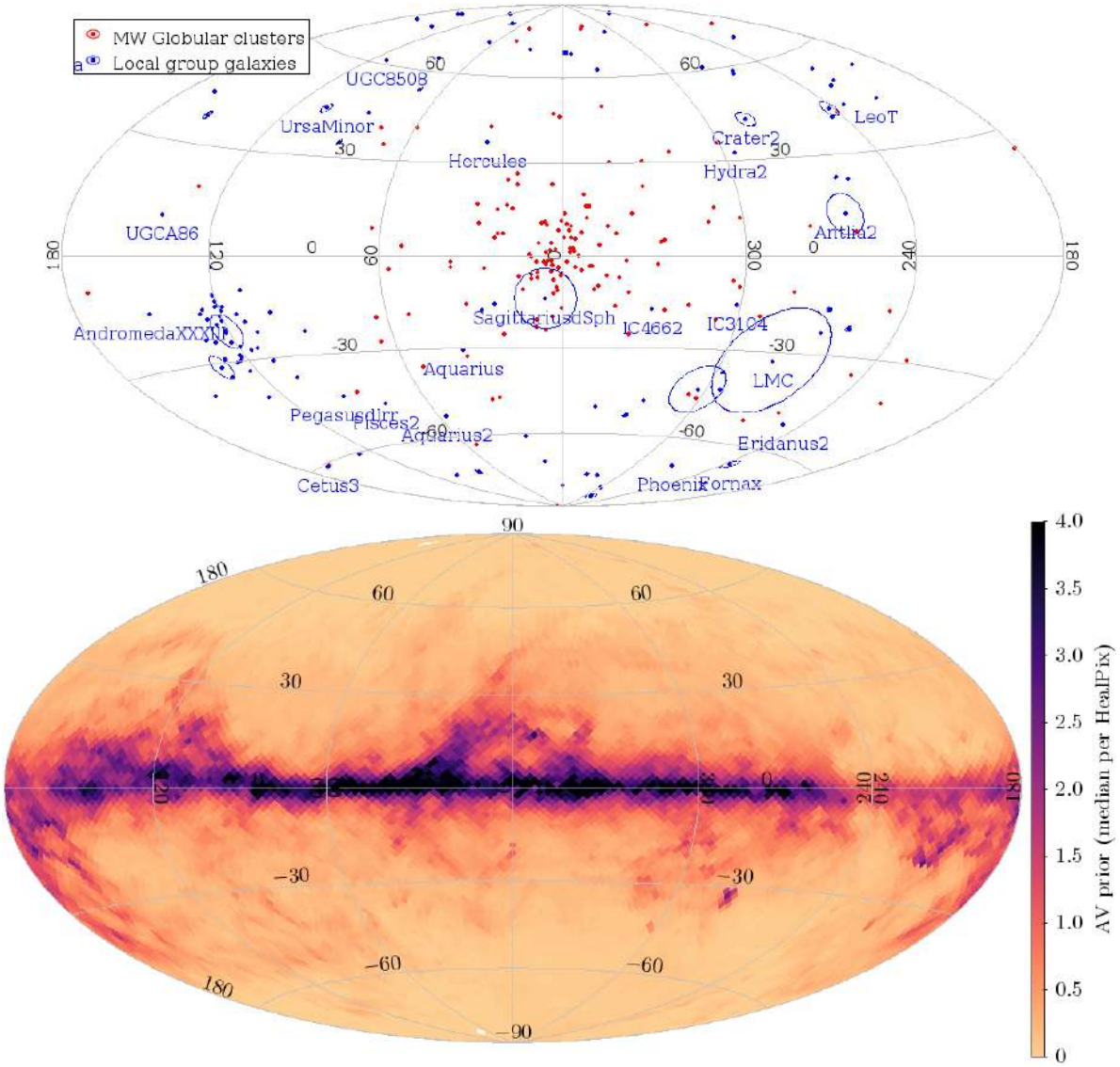

Fig. 1. Newly implemented priors in the `StarHorse` code. *Top panel*: sky distribution (in Galactic coordinates) of extragalactic and globular-cluster priors added in the new `StarHorse` version. The angular extents (5 effective radii) of each of the Local Group priors are shown as circles, highlighting the most prominent objects: the Magellanic Clouds, the Sgr dSph, and Andromeda. *Bottom panel*: median prior V-band extinction per HealPix. The extinction prior is calculated individually for each star from the three-dimensional extinction maps of either Green et al. (2019) or Drimmel et al. (2003).

3.1.4. Taking into account evolution of surface metallicity

We adopt here the latest version of the PARSEC1.2S + COLIBRI S37 stellar evolutionary model tracks (Bressan et al. 2012; Marigo et al. 2017; Pastorelli et al. 2019). Using these tracks in conjunction with the new CMD web interface allows us to take into account changes in the surface metallicity of stars during stellar evolution. While the effect is typically very small, element diffusion does introduce some small but appreciable decrease in the surface metal content for solar-mass stars before and around the turn-off (Fig. 2; see e.g. Bertelli Motta et al. 2018; Souto et al. 2019 for observational evidence). The effect is much stronger for low-metallicity stars. The opposite (i.e. a strong increase in the surface metallicity) happens for a fraction of the more evolved stars, such as those in the thermally pulsing asymptotic giant branch and Wolf-Rayet phases; this latter effect, however, is much less relevant to our results given that these evolutionary phases are much shorter-lived (and hence rarer) than main sequence stars, especially in nearby volume-limited samples.

3.2. `StarHorse` setup

For the EDR3 run we used a grid of PARSEC 1.2S stellar models (Marigo et al. 2017) in the 2MASS, Pan-STARRS1, SkyMapper, *Gaia* EDR3, and WISE photometric systems available on the PARSEC web page[3]. The model grid was equally spaced by 0.1 dex in log age as well as in initial metallicity [M/H]. The code explores distances within $\{1/(\varpi^{\rm cal} + 3 \cdot \sigma_{\varpi}^{\rm cal}), 1/(\varpi^{\rm cal} - 3 \cdot \sigma_{\varpi}^{\rm cal})\}$.

For the present *Gaia* EDR3 run ($G < 18.5$ mag, 400M stars), the code took on average 0.3 s per star to run (depending slightly on the position in the CMD and the number of photometric measurements available). In total, the computational cost for this

[3] http://stev.oapd.inaf.it/cgi-bin/cmd_3.4; see also http://stev.oapd.inaf.it/cmd_3.4/faq.html

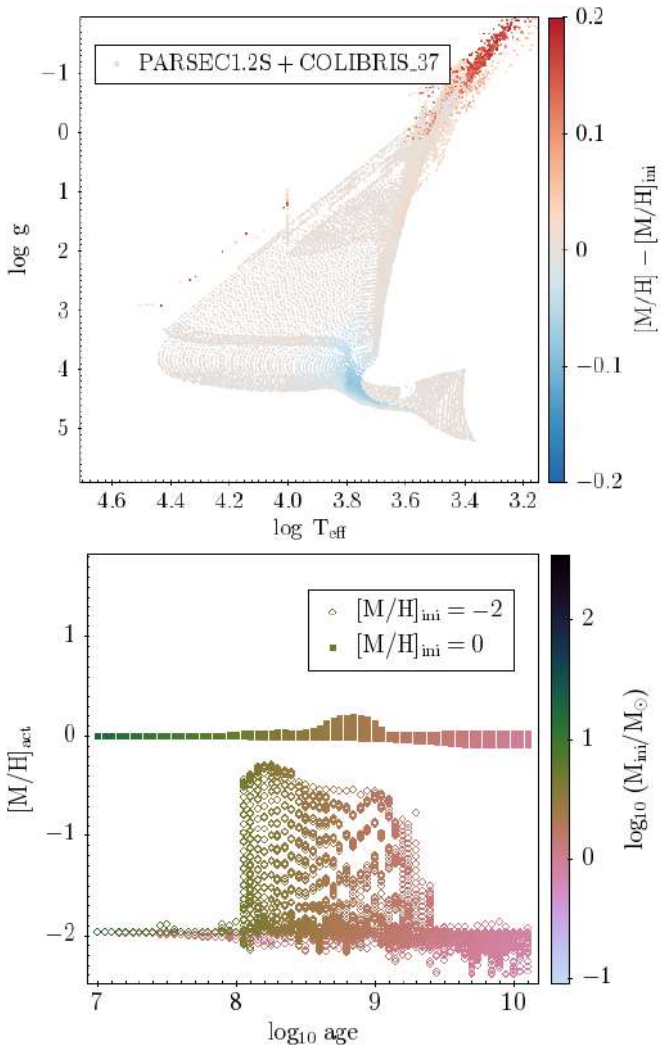

Fig. 2. Stellar evolution effects on surface metallicity in the PARSEC 1.2S + COLIBRI S37 stellar models. *Top*: Kiel diagram colour-coded by the difference between the surface metallicity and the initial metallicity. The evolution effects of diffusion and dredge-up are clearly visible. *Bottom*: age dependence of the surface metallicity, for two initial metallicities, colour-coded by stellar mass.

StarHorse run thus was ~50 000 CPU hours, reducing the CO_2 footprint of StarHorse by a factor of 3 with respect to the *Gaia* DR2 run presented in A19, while increasing the number of stars with reliable output parameters by more than a factor of 2. The global statistics for our output results are summarised in Table 2 and discussed in detail in Sect. 4.

3.3. Input and output flags

Along with the output of our code (median statistics of the marginal posterior in distance, extinction, and stellar parameters), we provide a set of flags to help the user decide which subset of the data to use for their particular science case. These flags correspond to the columns defined in the next few subsections.

3.3.1. *Gaia* EDR3 quality criteria used in this work

In the previous StarHorse *Gaia* DR2 run, we defined a set of input flags (summarised in the column SH_GAIAFLAG) based on the DR2 recommendations by the Gaia Collaboration (e.g. Lindegren et al. 2018). It contained three digits corresponding to astrometric fidelity (in particular the renormalised unit-weight error, ruwe; Lindegren 2018), the photometric fidelity (indicated by the phot_bp_rp_colour_excess), as well as the DR2-native variability_flag.

In this work we make use of the quality criteria established by Rybizki et al. (2022) and Riello et al. (2021) who have addressed these questions in detail and provide recipes to select high-quality EDR3 measurements. We thus follow their recommendations and use the following cuts:

Astrometric fidelity: We cross-matched our catalogue with the astrometric fidelity flag defined by Rybizki et al. (2022), based on a neural-network classifier for EDR3 objects. The classifier uses the twelve EDR3 astrometric columns identified by Gaia Collaboration (2021b) as containing most information about the fidelity of the EDR3 parallaxes and proper motions (and their uncertainties). It was trained on a set of bona fide trustworthy and bona fide bad EDR3 results. Bad astrometric results can be culled by requiring, for example, fidelity > 0.5.

Colour excess factor: The corrected version of the EDR3 phot_bp_rp_colour_excess column, C_* or bp_rp_excess_corr (Riello et al. 2021, see also Appendix B of Gaia Collaboration 2021e), indicates whether the BP/RP photometry of a *Gaia* source may be affected by background flux from neighbouring objects. When cleaning the StarHorse results for potentially affected BP/RP photometry, we recommend using a cut of $|C_*|/\sigma_{C^*} < 5$, where σ_{C^*} is a simple function of the G magnitude, computed according to Eq. (18) in Riello et al. (2021).

3.3.2. sh_photoflag

As in A19, we define the human-readable sh_photoflag that contains the information about the combination of photometric input data used for each object (*Gaia* EDR3, PS1, SkyMapper, 2MASS, AllWISE). For example, if only *Gaia* EDR3 G, G_{RP} and 2MASS HK_s magnitudes were available, the flag reads GRPHKs. PS1 and SkyMapper photometry are separated by a slash (/) in the sh_photoflag: for example, the flag Gg/riW1W2 means that the object in question has good *Gaia* G, PS1 g, SkyMapper ri, and AllWISE $W1W2$ measurements, while G/g means that the object has only *Gaia* G and SkyMapper g.

We note that with respect to A19 we improved the quality filters especially for the input AllWISE and 2MASS data, as well as for the *Gaia* BP/RP photometry (see Sect. 2).

3.3.3. sh_outflag

In A19, we defined a StarHorse output flag, consisting of five digits that informed about the fidelity of the StarHorse output parameters. The first digit served as the main quality indicator and filtered out stars with inconsistent median output parameters. Although the main caveats of the A19 results have been rendered obsolete by EDR3, we still define an output flag for convenience. It contains the following four digits:

The first digit flags low number of consistent models. For some targets, the number of stellar models in our model grid found to be 3σ-consistent with the data is low, indicating either very precise results or (more likely) some tension in the input data. We consider a results unproblematic if the number of models is greater than 30, and apply a (strong) warning flag if this number is between 10 and 30 (below 10): IF nummodels > 30 THEN 0 ELIF nummodels > 10 THEN 1 ELSE 2.

Table 2. Global statistics of some of the currently available astrometric and astro-photometric results based on *Gaia* DR2 and EDR3 data, in comparison to this work.

Reference	Data used	Mag limit	# objects	$\sigma_d/d^{G=17}$	$\sigma_{A_V}^{G=17}$	$\sigma_{T_{\rm eff}}^{G=17}$
Bailer-Jones et al. (2018b)	DR2 parallaxes	$G \lesssim 21$	1330M	24%	–	–
Andrae et al. (2018)	DR2 photo-astrometry	$G \leq 17$	161M	–	–	324 K
			88M	–	0.46 mag	–
Anders et al. (2019)	DR2 + 2MASS+AllWISE	$G < 18$	266M	40 %	0.25 mag	350 K
flag-cleaned sample	+Pan-STARRS1		137M	18 %	0.23 mag	230 K
Green et al. (2019)	DR2 + 2MASS+Pan-STARRS1	$z_{PS1} < 20.9$	799M	20 %	0.15 mag	–
Bai et al. (2019, 2020)	DR2 photo-astrometry	$G \lesssim 17$	133M	–	0.16 mag	350 K
Bailer-Jones et al. (2021)	EDR3 parallaxes	$G \lesssim 21$	1470M	20 %	–	–
	EDR3 photo-astrometry		1310M	16 %	–	–
This work	EDR3 photo-astrometry	$G < 18.5$	402,431,354			
`StarHorse` converged	+2MASS		362,392,321	15 %	0.15 mag	183 K
& `fidelity` > 0.5	+AllWISE		329,646,544			
& $\lvert C^*\rvert/\sigma_{C^*} < 5$	+Pan-STARRS1		321,131,855			
& `sh_outflag`=="0000"	+SkyMapper		281,501,963			

Notes. [G=17]For comparability, we report here the median precision for stars at magnitude $G \approx 17$.

The second digit flags negative extinction. Significantly negative extinctions should be treated with care: IF `AV95` > 0 THEN 0 ELSE 1.

The third digit warns about very large uncertainties. Large uncertainties are not problematic per se, but the corresponding median values are not usually very informative, which is why we provide this flag to be able to filter out very uncertain results quickly. The definition is as follows: IF $0.5 * (\texttt{dist84} - \texttt{dist16})/\texttt{dist50} > 1$ OR $0.5 * (\texttt{AV84} - \texttt{AV16}) > 1$ OR $0.5 * (\texttt{teff84} - \texttt{teff16}) > 1000$ OR $0.5 * (\texttt{logg84} - \texttt{logg16}) > 1$ OR $0.5 * (\texttt{met84} - \texttt{met16}) > 1$ OR $0.5 * (\texttt{mass84} - \texttt{mass16})/\texttt{mass50} > 1$ THEN 1 ELSE 0.

The fourth digit flags very small uncertainties. Very small posterior uncertainties are most likely underestimated and probably indicate poor convergence. These results should also be used with care. The definition is as follows: IF $0.5 * (\texttt{dist84} - \texttt{dist16})/\texttt{dist50} < 0.001$ OR $0.5 * (\texttt{av84} - \texttt{av16}) < 0.01$ OR $0.5 * (\texttt{teff84} - \texttt{teff16}) < 20.$ OR $0.5 * (\texttt{logg84} - \texttt{logg16}) < 0.01$ OR $0.5 * (\texttt{met84} - \texttt{met16}) < 0.01$ OR $0.5 * (\texttt{mass84} - \texttt{mass16})/\texttt{mass50} < 0.01$ THEN 1 ELSE 0.

Unproblematic results from the point of view of `StarHorse` can thus be filtered by requiring `sh_outflag`=="0000".

4. `StarHorse` *Gaia* EDR3 results

4.1. Summary

Table 2 summarises the results of the `StarHorse` run for *Gaia* EDR3 as well as previous results available from the recent literature. We observe that our new `StarHorse` results compare favourably in terms of both sample size and parameter precision. For example, the results have notably improved in precision (typically shrinking the formal uncertainties by a factor of 2) with respect to A19 (see Sect. 6 for a more detailed comparison).

Figure 3 shows the distribution of the `StarHorse` median posterior output values $T_{\rm eff}, \log g$, [M/H], M_*, d, and A_V and their corresponding uncertainties, demonstrating the complexity of the dataset as well as the typical precision (discussed in more detail in Sect. 5.1). Even the median output parameters are highly correlated, either intrinsically (enforced by the stellar models, e.g. $T_{\rm eff}$ vs. $\log g$), due to selection effects (e.g. d vs. M_*), or because of degeneracies related to our method ($\sigma_{T_{\rm eff}}$ vs. σ_{A_V}). The gridding effect in the metallicity panels of Fig. 3 is due to the finite resolution of the model grid.

Figure 4 shows the sky distribution of the input sample (400 M stars with $G < 18.5$), as well as sky maps of the percentage of converged sample and the cleaned sample. The figure shows that the code convergence is lowest in the densest areas of the sky (the innermost bulge and Galactic plane as well as the centre of the Large Magellanic Cloud), and that cleaning the *Gaia* data enhances this effect. For example, the X shape in the inner bulge visible in the bottom panel of Fig. 4 is mainly produced by the quality cut in the *Gaia* EDR3 colour excess factor (compare to Fig. 21 of Riello et al. 2021).

In the following subsections, we present some immediate results that can be obtained from our catalogue, focussing on CMDs (Sect. 4.2), Kiel diagrams (Sect. 4.3), stellar density maps (Sect. 4.4), and extinction maps (Sect. 4.5).

4.2. Extinction-corrected colour-magnitude diagrams

Since the stellar models used in our Bayesian inference have not changed much with respect to A19, the `StarHorse` extinction-corrected CMDs are also similar. The top row of Fig. 5 shows the CMD of the total sample and two interesting subsamples (the *Gaia*-cleaned sample and the fully flag-cleaned sample). When comparing these panels to Fig. 5 in A19, we note that some of the previously noted unphysical features have disappeared (most notably, the 'nose' between the main sequence and the lower red-giant branch). On the other hand, new structure in the top parts of the full CMDs emerges from the explicit inclusion of the Magellanic Clouds in the priors. For illustration, the bottom row of Fig. 5 shows the populations of the Milky Way disc, the Large Magellanic Cloud (LMC), and the Small Magellanic Cloud (SMC).

The second row of Fig. 5 shows the CMDs for three bins in apparent magnitude. The overall appearance of the magnitude-binned CMDs in Fig. 5 resembles those of Fig. 4 in A19, with a few notable differences. For example, in addition to the sharp features of single-star evolution in the $G < 14$ panel, we now also appreciate the unresolved binary sequence right above the low-mass main sequence. We also see the impact of the LMC and SMC populations on the CMD, already in the magnitude bin $16 < G < 17$. The rightmost middle panel, corresponding to $18 < G < 18.5$, shows already significant broadening in the

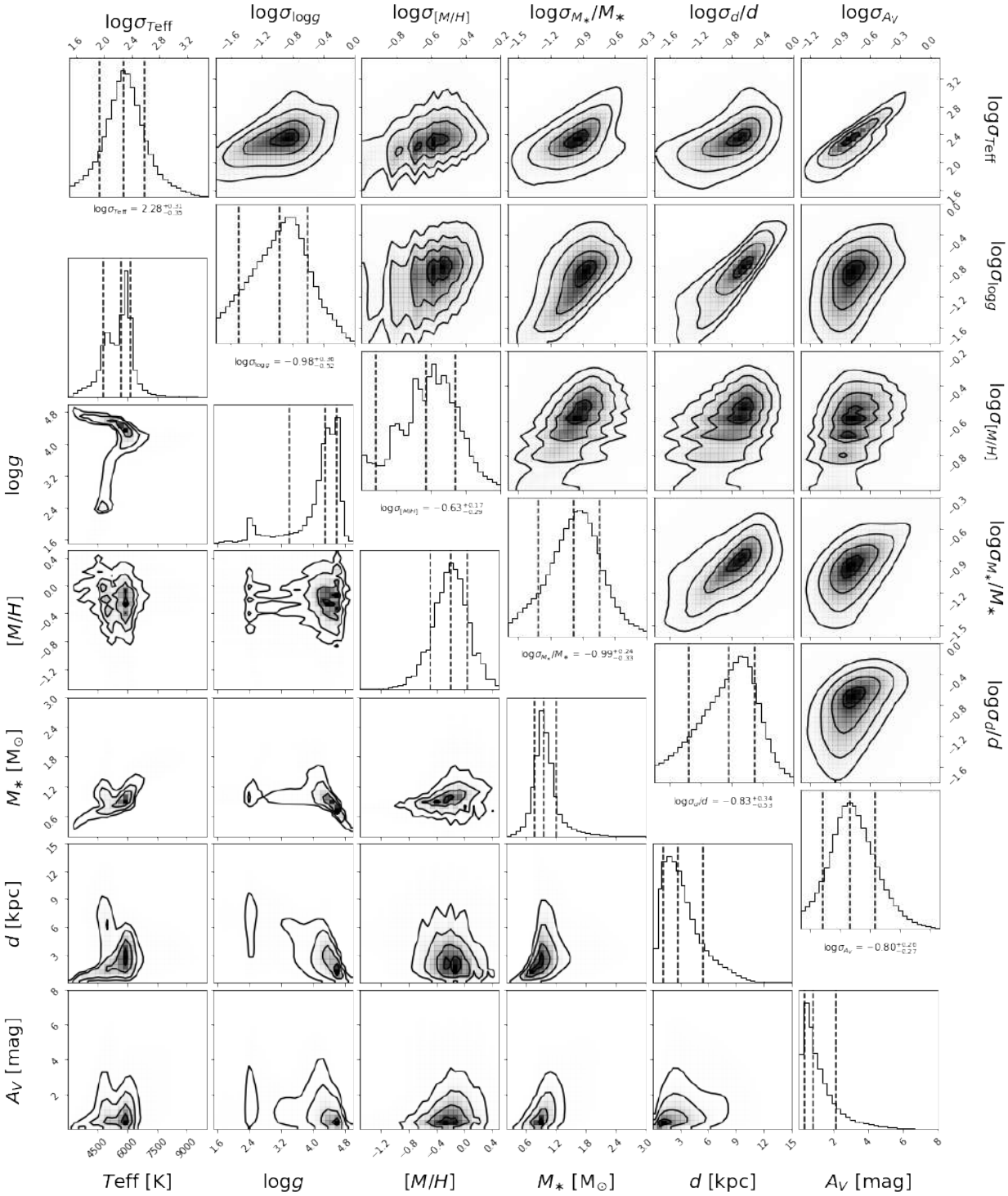

Fig. 3. corner plots showing the correlations and distributions of StarHorse median posterior output values $T_{\rm eff}$, $\log g$, [M/H], M_*, d, and A_V (*lower-left panels*), and their corresponding uncertainties (in logarithmic scale; *top-right panels*) for all stars in our catalogue. The dashed vertical lines in the diagonal panels show the 16th, 50th, and 84th percentiles of each parameter.

CMD features. As in A19, this is a result of the growing uncertainty in the input parameters, especially the parallax. We also recall that the absolute magnitudes and de-reddened colours displayed in Fig. 5 are not a direct output of StarHorse, but were computed from the observed magnitudes and the StarHorse median distance, extinction, and effective temperatures[4].

[4] https://github.com/fjaellet/gaia_edr3_photutils

4.3. Kiel diagrams

Figure 6 shows Kiel diagrams ($T_{\rm eff}$ vs. $\log g$) for the full *Gaia* EDR3 StarHorse sample. The density plot (left panel) shows that most of the sample is classified as FGK stars, as expected. Also clearly visible (both in the left and the middle panel) are the stripe-like overdensities corresponding to the metallicity resolution of the stellar model grid already noted in Sect. 4.1.

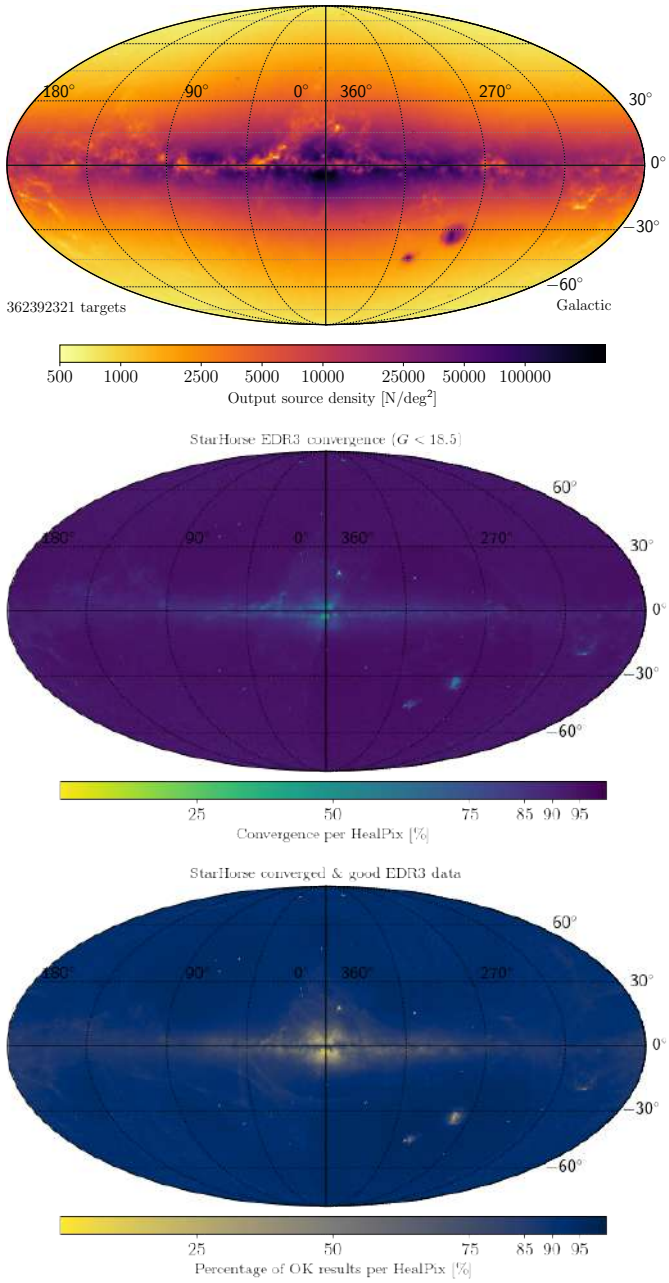

Fig. 4. Sky density map of all converged targets ($G < 18.5$ mag; *top panel*). *Middle and bottom panels*: relative fraction of converged stars and flag-cleaned results with respect to the input data.

We note a much more defined horizontal branch with respect to A19, which is at least in part due to the metallicity prior for globular clusters. We also note a more populated pre-main sequence (region above the lower main sequence), since we now applied a slightly less restrictive age cut ($\log t > 7$, as compared to $\log t > 7.5$ in A19).

The middle panel of Fig. 6 (Kiel diagram colour-coded by metallicity) shows that the posterior metallicity information is consistent with the stellar model grid through most of the parameter space. The only few outliers from the space spanned by the stellar models are stars whose median output parameters lie in-between the main sequence and the giant branch (due to a significantly bimodal posterior). The number of those stars (for which the median `StarHorse` are unreliable) has diminished enormously with respect to A19.

Finally, the right panel of Fig. 6 shows the typical distance range sampled for different regions of the Kiel diagram (also visible in Fig. 3), showing the expected behaviour of large typical distances (even >100 kpc) for the most luminous stars and very small distances for the coolest and least massive dwarf stars (<100 pc; see e.g. Gaia Collaboration 2021b).

4.4. Stellar density maps

4.4.1. Overall density distribution

One of the main motivations for the `StarHorse` project is Galactic cartography, and some of the newly implemented changes in the code (see Sect. 3.1) result in a visible improvement of the stellar density maps. To illustrate this, Fig. 7 shows two-dimensional projections of the stellar density distribution for the full `StarHorse` sample in Cartesian galactocentric coordinates. The left column of the plot focuses on larger structures: the Galactic volume probed by *Gaia* and the neighbouring dwarf galaxies, as indicated in each panel. These populations are now clearly visible as overdensities in the maps, although a considerable amount of stars still has median distances that fall in between the Magellanic Clouds and the Milky Way – a result of the multimodal posterior distance distributions (see e.g. Anders et al. 2019).

The right panels of Fig. 7 zoom into a 20 kpc wide cube centred on the Sun. When we compare these maps to the ones presented in Fig. 7 of A19, we notice: 1. the increase in total stellar number density (from 137 million to 360 million stars), and 2. the greater volume probed by the *Gaia* EDR3 $G < 18.5$ sample.

Direct consequences of the maps shown in Fig. 7 for Galactic cartography, however, are not obvious, since these maps are the result of a complex convolution of the true stellar density distribution, interstellar extinction, the applied magnitude limit, the selection function, and the priors. In the following subsections, we discuss the density maps of some specific stellar populations that are arguably easier to interpret.

4.4.2. Red-clump stars

Core helium burning red-clump stars (for a review see Girardi 2016) are often used as standard candles for mapping Galactic populations. They are numerous, relatively bright, and span a wide range of ages and metallicities.

Figure 8 shows the distribution of disc red-clump stars in the `StarHorse` *Gaia* EDR3 catalogue. The stars have been selected using the Kiel diagram as in Sect. 4.4 of A19: $4500\,\mathrm{K} < T_{\mathrm{eff}} < 5000\,\mathrm{K}$, $2.35 < \log g < 2.55$, $-0.6 <$[M/H]$< +0.4$, $|Z| < 3$ kpc. Figure 8 can thus be directly compared to Fig. 8 in A19.

In Fig. 8 of A19 we observed a very clear overdensity of red-clump stars tracing the Galactic bar. This result was all the more convincing since the bar angle used in the prior was significantly different from the one observed in the posterior distribution. However, Fig. 8 of A19 also displayed some minor artefacts, such as underdensities of red-clump stars both in front of and behind the near side of the bar, or an underdense ring-like structure that arose from the quality cuts necessary to clean the DR2 `StarHorse` data.

The EDR3 version of that figure, shown in Fig. 8, shows that the result of A19 (the detection of the Galactic bar in stellar density) is clearly maintained. The number of red-clump stars has become greater (13.8M vs. 10.8M), the underdensity artefacts of

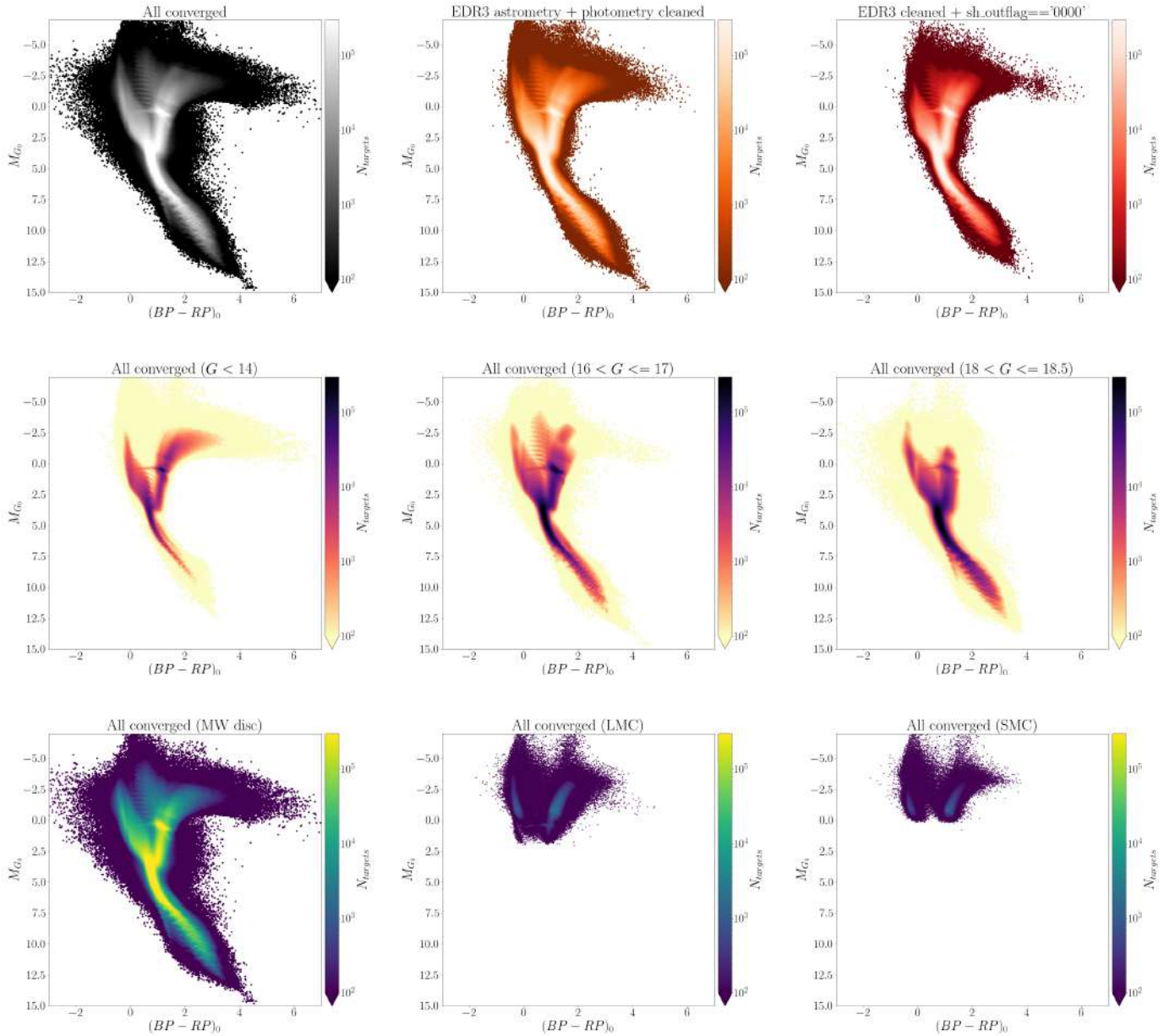

Fig. 5. StarHorse posterior *Gaia* EDR3 CMDs. *Top row, from left to right*: all converged objects (362M), *Gaia* EDR3 cleaned sample (321M), EDR3- and flag-cleaned sample (282M). *Middle row*: CMDs for three broad magnitude bins, showing both the increasing mix of stellar populations (e.g. the giant-star populations of the Magellanic Cloud starting to appear around $M_G \sim -3$ in the $16 < G < 17$ panel) and the decreasing astrometric quality with increasing magnitude. *Bottom row*: separate CMDs for the Milky Way disc (*left*; 339M stars), the LMC (*middle*; 1.09M stars), and the SMC (*right*; 94k stars). The abrupt absolute magnitude cut in the last two panels is caused by the $G < 18.5$ mag cut.

the map are greatly reduced, and the probed Galactic area now extends to regions beyond the (near side of the) bar.

The apparent bar angle is similar to the one in Fig. 8 of A19 and thus still appears to be a few degrees higher than the one assumed in the prior (27 deg; see Sect. 3.1.3). The main overdensity of the bar also appears relatively short compared to recent estimates of $\gtrsim$5 kpc. A quantitative analysis of the bar's structural parameters is, however, beyond the scope of this paper, as this requires careful modelling (e.g. Wegg et al. 2015; Portail et al. 2017) and taking into account selection effects.

Another feature in Fig. 8 is an overdensity appearing around $R_{\rm Gal} \sim 6$ kpc that might correspond to the Sagittarius spiral arm (see e.g. Reid et al. 2019). This feature is much less clear in the red-clump stars than in maps of young stellar populations (e.g. Castro-Ginard et al. 2021; Zari et al. 2021; Poggio et al. 2021), and the map in Fig. 8 shows the underlying density distribution convolved with dust extinction and other selection effects. The clear overdensity in the (logarithmic) red-clump star count map is, however, a strong feature that deserves further investigation, since the strength of the spiral density signature in an intermediate-age population has implications on the modelling of the Milky Way's spiral arms.

Recently, Nogueras-Lara et al. (2021) have used the high angular-resolution infrared photometric survey GALACTICNUCLEUS (Nogueras-Lara et al. 2019) to determine the distances, extinctions, and stellar populations of the inner spiral arms in a small region of the sky containing the Galactic centre. While their data are of clearly superior quality, we suggest that

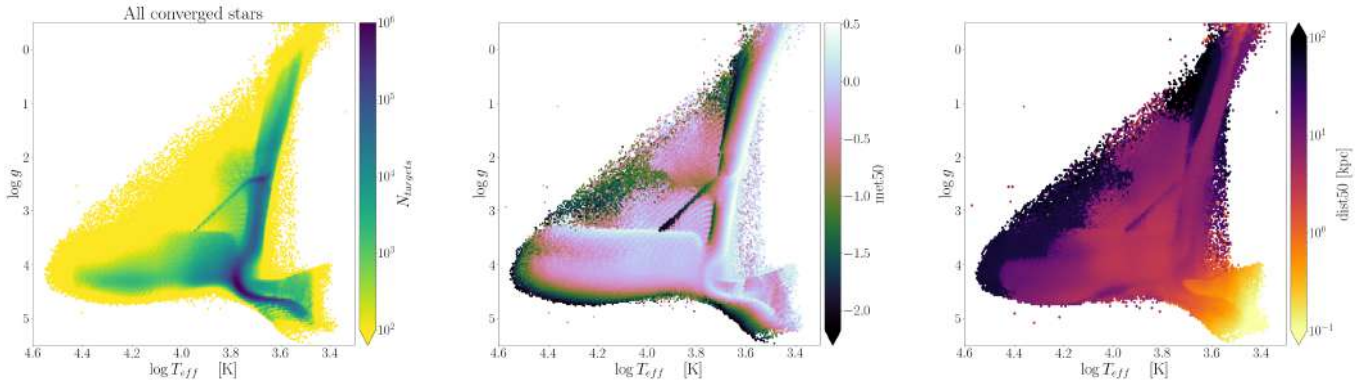

Fig. 6. `StarHorse`-derived Kiel diagrams (before applying any quality cuts). *Left*: density plot. *Middle*: colour-coded by median metallicity. *Right*: colour-coded by median distance.

similar mapping studies could be carried out using *Gaia* and multi-wavelength photometry (and possibly our `StarHorse` catalogue) for the portions of the disc less affected by interstellar extinction.

4.4.3. Magellanic Clouds

The Magellanic Clouds as our immediate galactic neighbours represent a key laboratory to study gravitational interactions and their effects on the structure and kinematics of satellite galaxies. In this section, we analyse our results for the region of the Magellanic Clouds and compare them to the Gaia Collaboration (2021d) results.

In Fig. 9 we show from top to bottom the sky density map, 2D distance distribution, metallicity and extinction maps for the sources around the Large Magellanic Cloud (LMC, left) and Small Magellanic Cloud (SMC, right), respectively, in equatorial coordinates.

For the LMC (left column of Fig. 9), the sky density distribution highlights the main components of the galaxy. The innermost contour encloses the elongated bar, while the second contour highlights the spiral arm. We notice a small region with low star density between the bar and the spiral arm, in agreement with the star counts shown in Gaia Collaboration (2021d, e.g.), but much less smooth, because of the relatively low convergence rate of `StarHorse` in that region (due to crowding issues in the input data; see Fig. 4).

The distance map (second row of Fig. 9) indicates a median heliocentric distance of 49.4 kpc (for comparison, the distance used in the prior is $d_{\rm prior}$ = 50.58 kpc; McConnachie 2012), for the sources inside the outermost contour level, in agreement with previous estimations (e.g. Pietrzyński et al. 2019). It also shows the expected distance gradient from the fact that the LMC is inclined about 34°, being the closer side the one towards larger declinations (Gaia Collaboration 2021d, and references therein).

The LMC metallicity map (third row left panel of Fig. 9) highlights a problematic result: In the inner parts of the LMC, we see a positive metallicity gradient from the bar region towards the outer disc, opposite to the trend observed with red-giant branch (RGB) stars from Magellanic Cloud Photometric Survey (MCPS) and OGLE-III (Choudhury et al. 2016), RR Lyrae stars from OGLE-IV (Skowron et al. 2016), or RGB stars from *Gaia* DR2 (Grady et al. 2021). The median metallicity in the bar region (inside the innermost contour level) is of −0.77 dex, while at the outer disc (between the innermost and outermost levels) is of −0.68 dex. This suggests that the little metallicity information contained the broad-band colours we use in this work is affected by significant systematics, at least for the very dense and complex regions of the Magellanic Clouds. The declining influence of the LMC prior biases the resulting median metallicities and inverts the expected trend (this can possibly be remedied when using the full posterior; see Appendix B).

Analogously, the right column of Fig. 9 shows the corresponding plots for the SMC sample. The sky distribution (top-right panel of Fig. 9) highlights the irregular structure of the SMC and the beginning of the bridge towards the direction of growing right ascension (and decreasing declination). The distance map (second row right panel of Fig. 9) provides a median distance to the SMC of 63.2 kpc (prior: $d_{\rm prior}$ = 63.97 kpc), for the sources inside the outermost level, in agreement with previous estimations (e.g. Cioni et al. 2000). The outer ring with closer distances may be partly an artefact due to the vanishing of the prior contribution towards the outer regions. No clear distance gradient is visible in the SMC.

Two small blobs with slightly smaller distance are visible in the central parts of the SMC, which are also correlated with the metallicity. Again, a small positive metallicity gradient from the inner towards the outer parts of the galaxy is visible, opposite to the expected behaviour observed with the Red Giant Branch sources from MCPS and OGLE-III (Choudhury et al. 2018) or *Gaia* DR2 (Grady et al. 2021). As in the LMCANDE0890215004, the metallicity and extinction appear to be correlated, being the extinction higher towards the central more crowded region of the galaxy (see bottom-right panel of Fig. 9).

4.4.4. Candidate metal-poor stars

The study of metal-poor stars provides a unique window into the formation and accretion history of our Galaxy, since the bulk of those stars were formed at high redshift and conserve abundance patterns unique to their site of formation (Beers & Christlieb 2005).

Although the broad-/intermediate-band photometry used in this work is only marginally sensitive to metallicity (in fact, only when including optical *griz* photometry can we expect to detect some metallicity information; see Sect. 5), low metallicities may manifest themselves in the broad-band colours (especially in the ultraviolet; e.g. Norris et al. 1999). We therefore venture to look at candidate metal-poor stars as determined by `StarHorse`, by defining a candidate metal-poor sample as `met84` < -1,

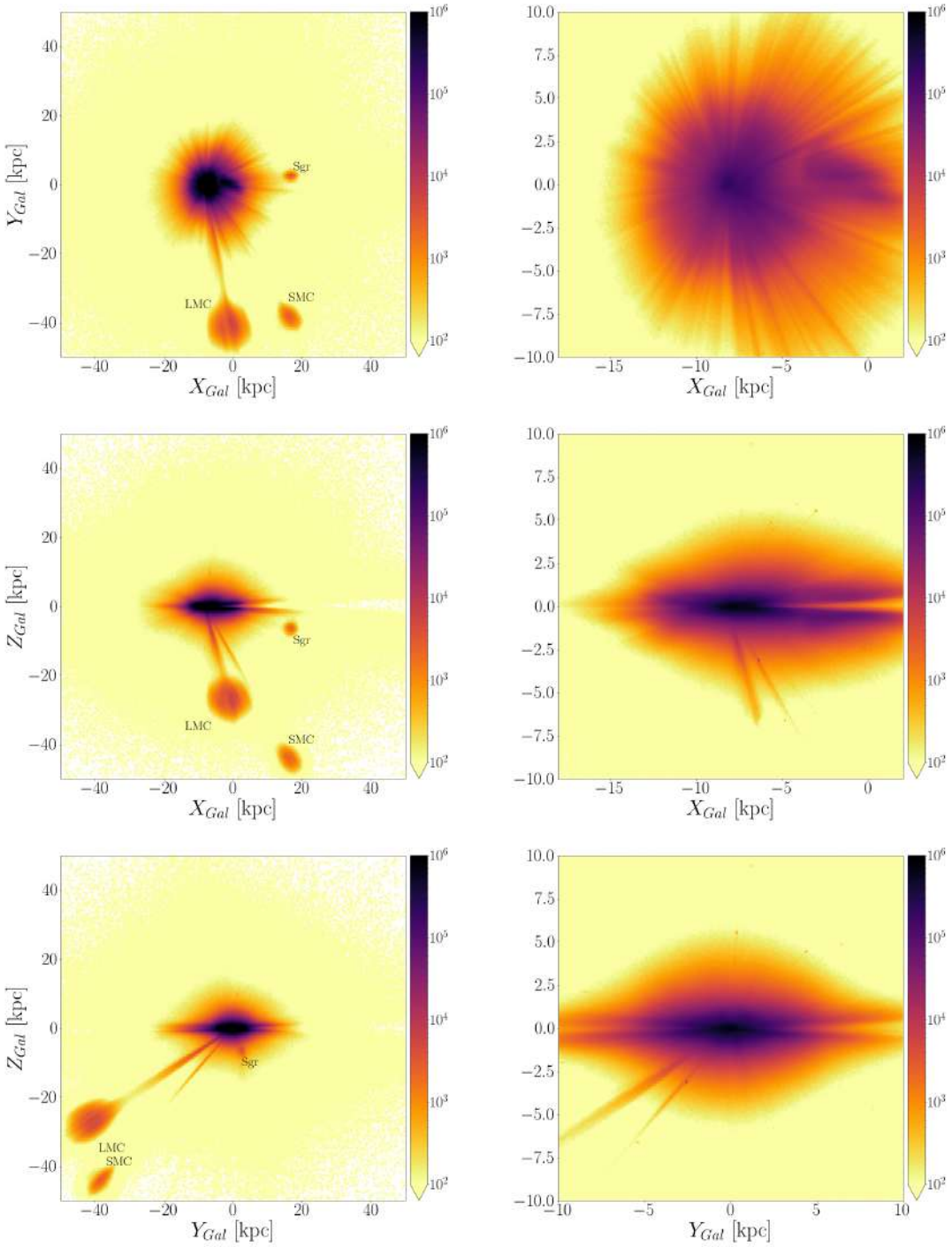

Fig. 7. StarHorse density maps (*from top to bottom*: *XY*, *XZ*, and *YZ*) in galactocentric coordinates. *Left column*: 100 kpc wide cube centred on the Galactic centre, while *right column*: zooms into a 20 kpc wide cube centred on the Sun.

corresponding to a 1σ confidence-level cut. This selection yields 1.58 million objects (without applying any further quality cuts).

Figure 10 shows the distribution of the metal-poor candidates in galactocentric cylindrical coordinates ($R_{\rm Gal}$ vs. $Z_{\rm Gal}$). We clearly see the imprint of the globular-cluster priors in this figure: All noticeable point-like overdensities correspond to prominent globular clusters, as annotated in the plot. We also note the overdensities in the direction of the Magellanic Clouds, corresponding to stars with bimodal distance probability density function (PDFs), resulting in a median distance in-between the inner halo and the Magellanic Clouds (see Sect. 4.4.3). A similar, less obvious structure, is also visible in the direction of the core of the Sgr dSph galaxy (located towards $(l, b) \sim (5, -14)$), resulting in an elongated overdensity around $(R_{\rm Gal}, Z_{\rm Gal}) \sim (0-3, -1)$.

Apart from these expected features, we also note a very prominent overdensity of local dwarf stars, many of them also following a disc-like density profile, and a diffuse overdensity in the nearby Galactic halo. The disc-like overdensity is likely mostly due to sample contamination, although even very metal-poor stars have been found on disc-like orbits recently in the Milky Way (Sestito et al. 2020) as well as in simulations (Sestito et al. 2021). The diffuse overdensity at larger

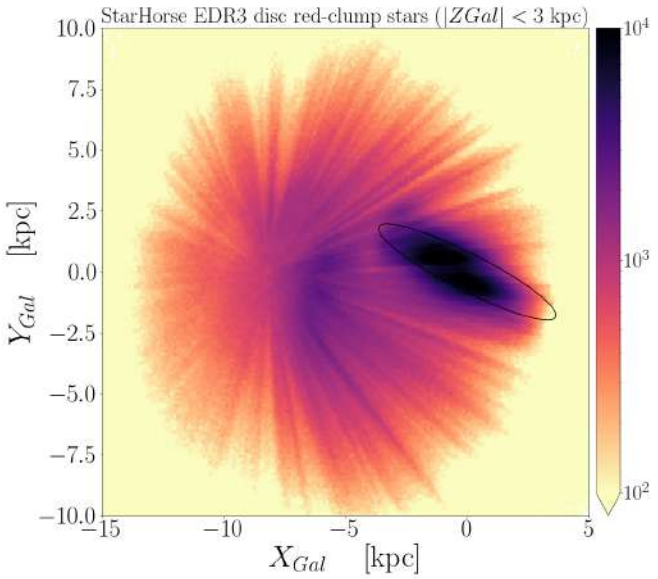

Fig. 8. XY density map, selecting all (13.8M) red-clump stars less than 3 kpc away from the Galactic midplane. The ellipse shows the orientation (27 deg with respect to the Sun-Galactic centre line) and approximate extent (semi-major axes $a = 4.07$ kpc and $b = 0.76$ kpc) of the Galactic bar assumed in the prior.

heliocentric distances is produced by more distant giant stars of the inner halo, expected from the combination of our selection function ($G < 18.5$) and our halo prior. Its members can be regarded as potential targets for future/ongoing spectroscopic surveys. Another possible overdensity is seen in the central parts of the Galaxy, where indeed many of the Milky Way's oldest stars are expected to reside (e.g. Tumlinson 2010; Koch et al. 2016; Starkenburg et al. 2017; Horta et al. 2021; Queiroz et al. 2021).

Although methods explicitly tailored to detect metal-poor star candidates from combined broad- and narrow-band colours can be expected to perform much better (e.g. Beers et al. 1985; Youakim et al. 2017; Da Costa et al. 2019; Thomas et al. 2019; Arentsen et al. 2020; Chiti et al. 2021; Huang et al. 2021b), our approach yields a large number of metal-poor star candidates for possible follow-up observations with multi-object spectroscopic surveys such as 4MOST (de Jong et al. 2019; Chiappini et al. 2019; Helmi et al. 2019).

4.4.5. Outer halo and Local Group

Figure 11 focuses on the density distribution of distant stars in the Galactic halo (defined by $|b| > 15$ deg, `dist50` > 10 kpc). The two top panels (showing Aitoff projections of the sky in ecliptic coordinates) highlight the long tidal tails of the Sgr dSph galaxy, also called the Sgr stream (e.g. Law et al. 2016). This feature, although not included in our priors, appears clearly both in the density map (top panel of Fig. 11) and the median distance map (middle panel), superseding the extent of the previous membership maps of the Sgr stream, for example the one produced by Antoja et al. (2020) based on *Gaia* DR2 proper motions.

The lower panel of Fig. 11 shows that the $G < 18.5$ sample encompasses also a significant amount of individual stars in dwarf galaxies of the Local Group other than the Magellanic Clouds and the Sgr dSph. For many of them (e.g. the Draco dSph, Bootes I, the Carina dSph, or the Ursa Minor dSph), the more informative extragalactic priors of the new `StarHorse` results can help to improve membership probabilities. For others (e.g. Sculptor dSph, Fornax dSph), the prominent pencil-beams between the halo and the expected location of the respective dwarf galaxy hint a problematic prior (e.g. imprecise central coordinates or too low galaxy masses in the Local Group tables used) that results in typically bimodal distance posterior PDFs.

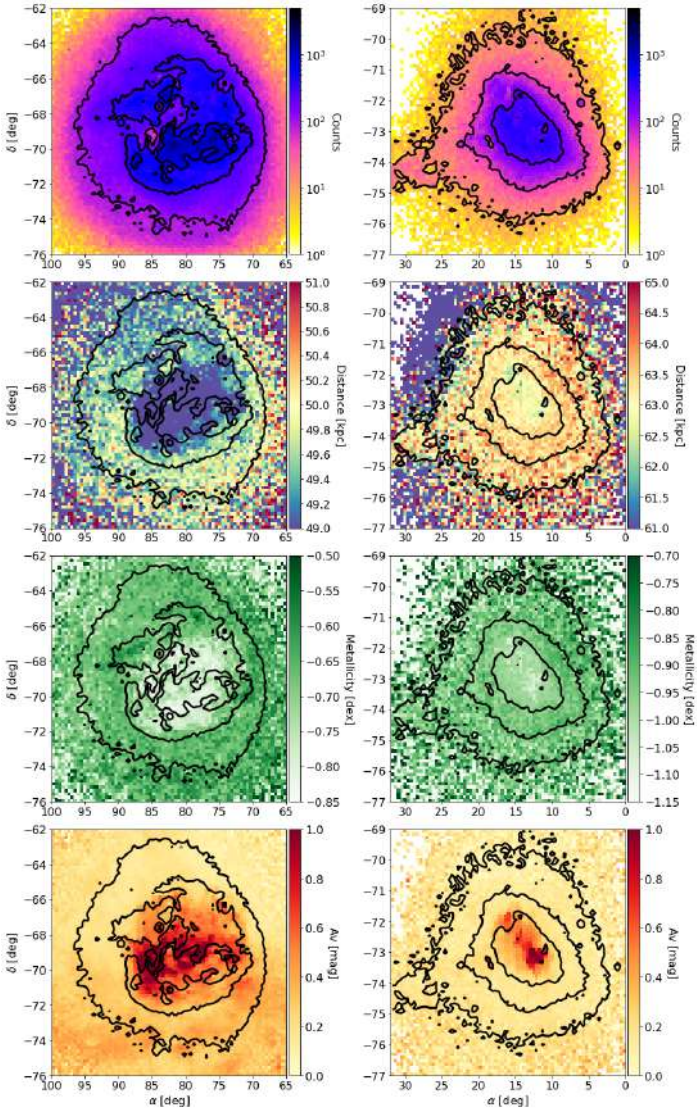

Fig. 9. Median sky density, distance, metallicity, and extinction maps (*from top to bottom*) of the Magellanic Clouds as seen by `StarHorse` (in equatorial coordinates and only including objects with `dist50` > 25 kpc). *Left panels*: centred on the LMC, *right panels*: on the SMC. The contour lines in each of the panels are derived from the sky density plots in the *top panels*. For the LMC, the contours are drawn at stellar densities of [100, 300, 700] per pixel (from outside inwards), with 905 205 sources within the outermost contour. For the SMC, the contour lines correspond to levels [10, 50, 200], with 195 634 sources contained inside the outermost contour.

4.5. Extinction maps

Figure 12 shows the median `StarHorse`-derived line-of-sight extinction per HealPix cell in four consecutive distance bins out to 2.5 kpc (from top to bottom), illustrating the gradual increase

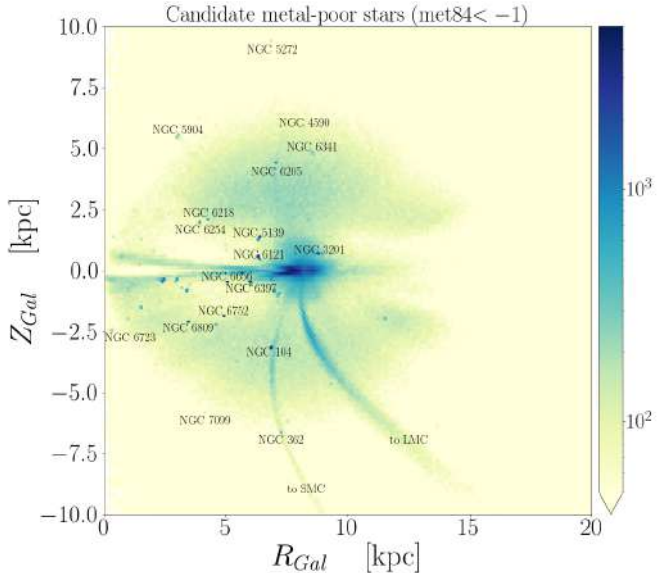

Fig. 10. Density map for bona fide candidate metal-poor stars (`met84` < −1; 1.58M stars) in galactocentric coordinates. Some prominent overdensities corresponding to Galactic globular clusters and the direction towards the Magellanic System are annotated.

in interstellar extinction as a function of distance and sky position. As expected, these maps are similar to the large-scale integrated dust extinction maps of, for example, Green et al. (2019). Since we have used the three-dimensional extinction maps of Green et al. (2019) and Drimmel et al. (2003) (albeit convolved with quite broad Gaussians) in our prior, this is not too surprising.

In principle, our extinction results can be used to infer precise distances to individual dust clouds (e.g. Wolf 1923; Zucker et al. 2020) and to infer the three-dimensional distribution of dust (e.g. Lallement et al. 2019; Leike et al. 2020). The top panel of Fig. 12 shows the presence of high-latitude dust within the 500 pc sphere around the Sun, confirming that the so-called North Polar Spur (the dust filament reaching up to $b \sim 45$ deg at $l \sim 0$ deg) is a local structure and not related to the Fermi bubbles produced by the Galactic centre (see Das et al. 2020 for a comprehensive discussion).

5. Precision and accuracy

5.1. Internal precision

Along with the median statistics of each output parameter, `StarHorse` also delivers the corresponding confidence intervals (defined as the 16th and 84th percentile of the marginal posterior). The overall distribution of the output uncertainties (defined as, for example, $\sigma_{\rm Teff} = 0.5 \cdot$ (`teff84` − `teff16`), etc.) and the correlations between the output uncertainties are shown in the top-right corner plot of Fig. 3. This plot shows the complete sample of converged stars and demonstrates that the output uncertainties are typically highly correlated (we note the logarithmic scaling of the plot axes). The highest correlations are seen, as expected, between effective temperature and extinction, and between distance and surface gravity.

The precision of the results, however, depends first and foremost on the quality of the *Gaia* EDR3 parallaxes and the availability of multi-band photometry for each source. Both these criteria are, to first approximation, functions of the *Gaia G* magnitude. In Fig. 13 we therefore show the formal uncertainties as a function of G magnitude for a random sample of 1 million stars. The orange two-dimensional histogram in the background shows the uncertainty distribution of all objects, while the red line shows the smoothed median trend. We can appreciate that the distance uncertainties for stars with $G < 14$ are typically around 2%, growing to about 8% around $G \approx 16$, and reaching 20% at $G \approx 18$. The improvement in precision with respect to our DR2 run (A19, black line in Fig. 13) is mainly due to the improvement in both precision and accuracy brought by the *Gaia* EDR3 parallaxes.

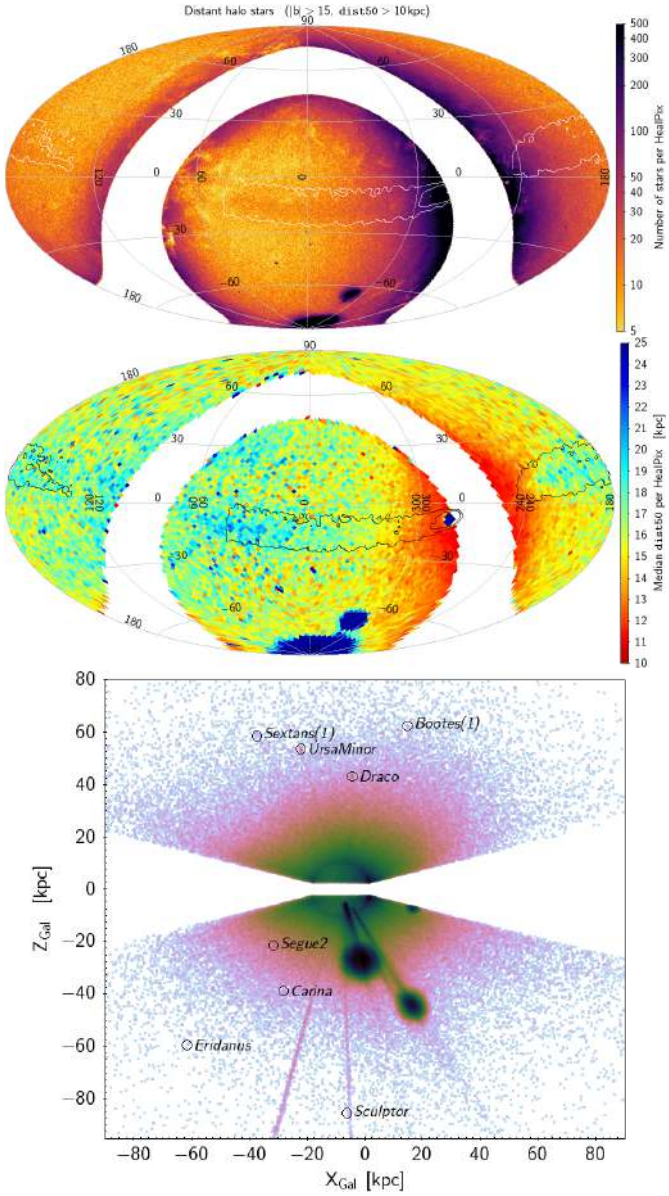

Fig. 11. Distribution of distant halo stars, selected by excluding the Galactic plane and a cut in median distance ($|b| > 15$ deg, `dist50` > 10 kpc; 2.55M stars). *Top panel*: sky distribution in ecliptic coordinates, highlighting the presence of the Sagittarius stream close to the ecliptic plane. *Middle panel*: same projection, colour-coded by the median distance per HealPix. In both panels the contour overlay shows the location of the Sagittarius stream candidates from Antoja et al. (2020). The *bottom panel* shows a Cartesian projection ($X_{\rm Gal}$ vs. $Z_{\rm Gal}$), highlighting some of the less prominent Local Group objects included in the priors.

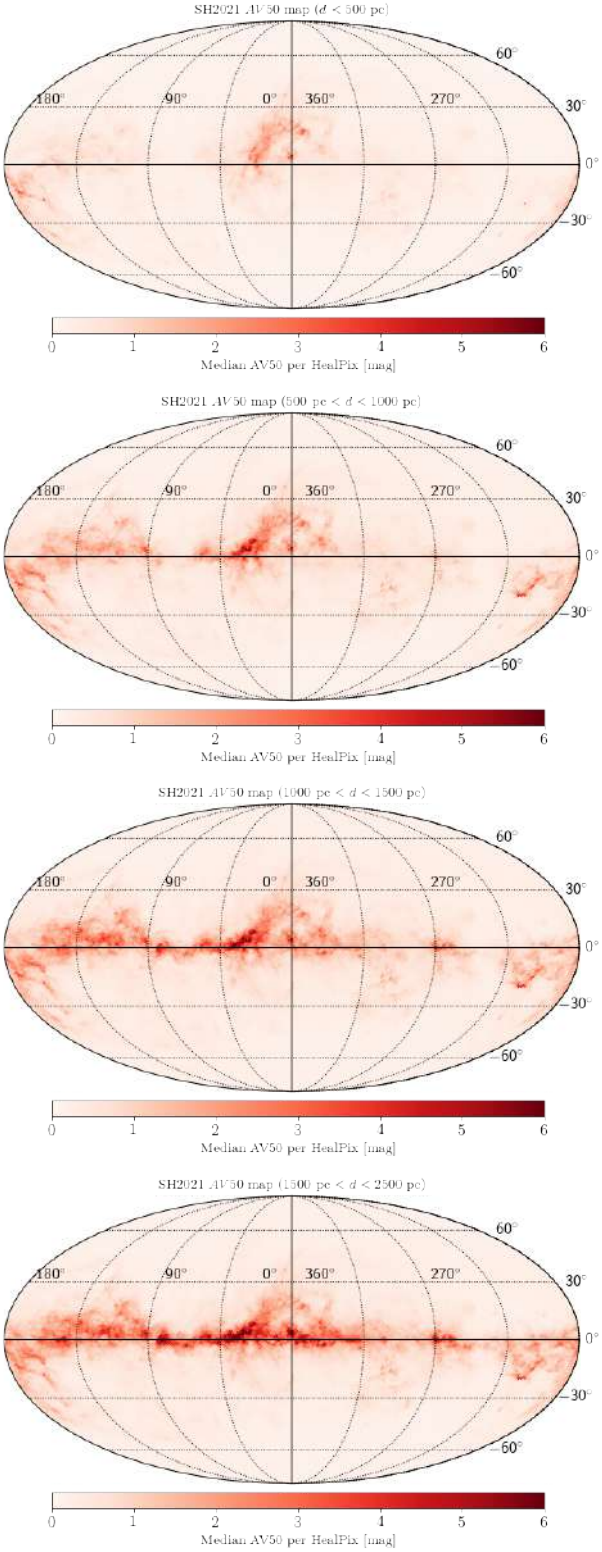

Fig. 12. All-sky median `StarHorse` extinction map for four wide distance bins up to 2.5 kpc, as indicated in each of the subplots.

The bottom row of Fig. 13 show the median formal uncertainties as a function of position in the Galaxy, again for a random set of 1 million stars. Many of the features in these uncertainty maps can already be appreciated (although at a different absolute scale) in Fig. 13 of A19. Apart from the overall precision improvement (by a factor of ∼2) the major changes are: 1. a slight increase of the 'parallax sphere' (the region for which parallaxes are determined with a precision of $\lesssim$20%), 2. the disappearance of the bulk of stars with very high distance uncertainties that had to be flagged because they were compatible with both dwarf- and giant-star solutions, and 3. a slightly lower impact of the missing PS1 photometry on the $T_{\rm eff}$ and A_V precisions below a declination of −20 deg ($Y_{\rm Gal} < 0, X_{\rm Gal} \gtrsim -10$ kpc) thanks to the use of SkyMapper data (and, in fact, a higher precision in the region where both catalogues overlap).

The precision of the secondary output parameters ($\log g$, [M/H], and M_*), not shown in Table 2 and Fig. 13, behave similarly as a function of G, although the improvement in precision with respect to the DR2 results is slightly less pronounced (by a factor of 1.5). At magnitude $G \approx 17$, the median uncertainties for the secondary output parameters amount to $\sigma^{G=17}_{\log g} = 0.23$ dex, $\sigma^{G=17}_{\rm [M/H]} = 0.10$ dex, and $\sigma_{M_*}/M_*^{G=17} = 9.5\%$.

5.2. Comparison to open clusters

Member stars of an OC are expected to have, to first order, the same age, metallicity, distance, and interstellar extinction. They thus constitute excellent samples fro evaluating the precision and accuracy of our astrophysical parameters.

Figure 14 shows comparisons of the `StarHorse` distance, extinction, and metallicity scales for the five most populated and well-studied OCs (NGC 6791, NGC 7789, Collinder 261, NGC 3532, and NGC 188) in the *Gaia* DR2 OC catalogue of Cantat-Gaudin et al. (2020). These OCs are those with the most identified members, mainly by virtue of being relatively massive and nearby (but not so nearby that they are extended in the sky and in proper-motion space, like the Hyades). Each panel of Fig. 14 shows a `StarHorse` output parameter as a function of effective temperature in comparison to the literature values for the particular cluster. The five clusters are diverse enough in their physical characteristics to appreciate some first trends as a function of effective temperature, surface gravity, metallicity, and cluster age.

For example, for the old metal-rich cluster NGC 6791 the `StarHorse`-derived metallicities are clearly underestimated (with respect to the spectroscopically derived cluster metallicity [Fe/H]; Casamiquela et al. 2017) and show a quite large scatter (which is, however, both expected and reflected in the quoted uncertainties). Similarly, the code finds a slightly lower extinction and distance than derived by Cantat-Gaudin et al. (2020).

A more quantitative comparison for the bulk of the known Galactic OC population (the 1867 OCs with astrophysical parameters from the Cantat-Gaudin et al. 2020 catalogue) is shown in Figs. 15 and 16. Cantat-Gaudin et al. (2020) determined the distance, extinction, and age of each cluster homogeneously with an artificial neural network trained on a set of high-quality measurements (mostly relying on Bossini et al. 2019). In Fig. 15 we plot the `StarHorse` median values per cluster (for FGK-type stars, 3800 K < $T_{\rm eff}$ < 6000 K) compared with the Cantat-Gaudin et al. 2020 determinations of the distance and extinction. The colour represents the median absolute deviation (MAD) obtained for the cluster members in SH. Figure 15 shows that the OCs cover a broad range of physical parameters: 90% of the clusters are nearer than 4.4 kpc and have less than 2.5 magnitudes of extinction, and the age range (10–90th percentile) covers $\log \tau$ from 7.2 to 9.1.

A complementary catalogue of astrophysical parameters for OCs was recently presented by Dias et al. (2021). It contains parameters of 1743 OCs (in their vast majority also contained in Cantat-Gaudin et al. 2020) determined by isochrone fitting of

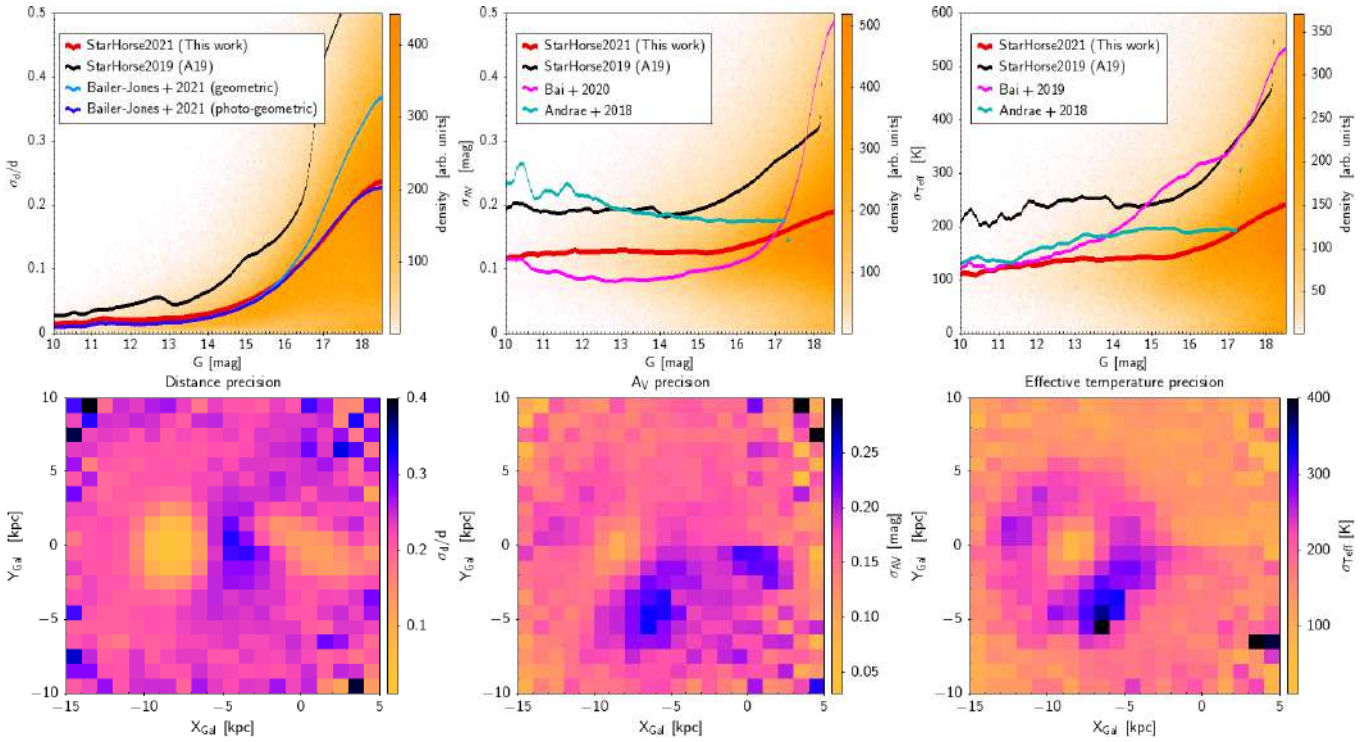

Fig. 13. StarHorse formal output uncertainties. *Top row*: uncertainties in distance (relative distance uncertainty; *left*), extinction (*middle*), and effective temperature (*right*) as a function of G magnitude. In each top panel we show two-dimensional histograms of a random sample of 1 million *Gaia* EDR3 stars in orange, along with the running median smoothed by an Epanechnikov kernel (width = 0.2; thick red line). For comparison we also show the corresponding values obtained from the (unfiltered) *Gaia* DR2 run A19 in black, as well as the results from Bailer-Jones et al. (2021) for distances, from Andrae et al. (2018) and Bai et al. (2020) for extinctions, and from Andrae et al. (2018) and Bai et al. (2019) for effective temperatures. *Bottom row*: median formal output uncertainties as a function of Galactic position for the same random sample.

Gaia DR2 photometry. We used this catalogue as an additional reference to test if the discrepancies between StarHorse and the Cantat-Gaudin et al. (2020) catalogue can partly be attributed to systematics in the OC catalogues as well. We only show the comparison with Cantat-Gaudin et al. (2020); the comparison with the Dias et al. (2021) catalogue leads to the very similar general conclusions.

The top row of Fig. 15 shows that the concordance with the OC distance scale is reasonable. The majority of both the clusters and the member stars present less than 20% deviation. The deviating clusters are mostly distant objects with very few member stars, partly uncertain membership, and thus a large internal dispersion of StarHorse parameters (red dots). We see a trend of negative differences with respect to the OC catalogue: on average, our EDR3 distances are shorter by −3.5%. An opposite trend of similar magnitude, however, is seen in the comparison with the Dias et al. (2021) catalogue: our distances are larger than theirs by +3.8% on average. No significant trends of distance difference with neither extinction nor age are found.

For extinction, on the other hand, some systematics similar to those seen in A19 can be appreciated: in particular, a slight systematic overestimation for nearby, low-extinction objects. This may in part be due to the fact that StarHorse treats every object as a single star and tries to adjust its parameters to a PARSEC isochrone. For similar-mass unresolved binaries on the main sequence this typically leads to an overestimated effective temperature, an underestimated $\log g$ (moving the object towards the sub-giant or lower red-giant branch), and an overestimated extinction to compensate for the extra brightness (compared to a single star). We also refer to Appendix D.

In Fig. 16 we further investigate possible systematic biases depending on sky position and spectral type. We find a rather uniform sky distribution of relative distance differences that is consistent with Fig. 15. The distance systematics are typically very small and lightly negative ($\lesssim$5%; see also the last row of Fig. 15). A sky pattern is hardly discernable, but may be related to the parallax bias present in *Gaia* DR2 (and thus also in the Cantat-Gaudin et al. 2020 and Dias et al. 2021 catalogues), which has been largely accounted for in EDR3 (using the corrections proposed by Lindegren et al. 2021b).

Both the parallax improvement with respect to *Gaia* DR2 and the inclusion of a dust map in the new priors allow a slightly smoother distribution of extinction differences than in A19. However, we see that extinction is generally 0.1–0.2 mag higher than the one estimated by Cantat-Gaudin et al. (2020). Our extinction estimates are, on the other hand, slightly lower than the ones in the catalogue of Dias et al. 2021, so that the absolute scale is far from well defined.

Furthermore, the top-right panel of Fig. 16 shows significant systematic trends of extinction with position the Kiel diagram (being most severe in sparsely populated areas). For example, it seems that StarHorse tends to slightly underestimate extinctions for metal-rich (redder) giant stars, while it overestimates extinctions for metal-poor giants. For dwarf stars, extinction biases are generally low, except for (probable) binary stars close to the turn-off phase (see Appendix D.1).

StarHorse also tends to severely overestimate the extinction of the stars hotter than 7000 K (Pantaleoni González et al. 2021). Due to the initial-mass-function prior, stellar models with $T_{\rm eff} \gtrsim 10^4$ K are highly suppressed in the posterior

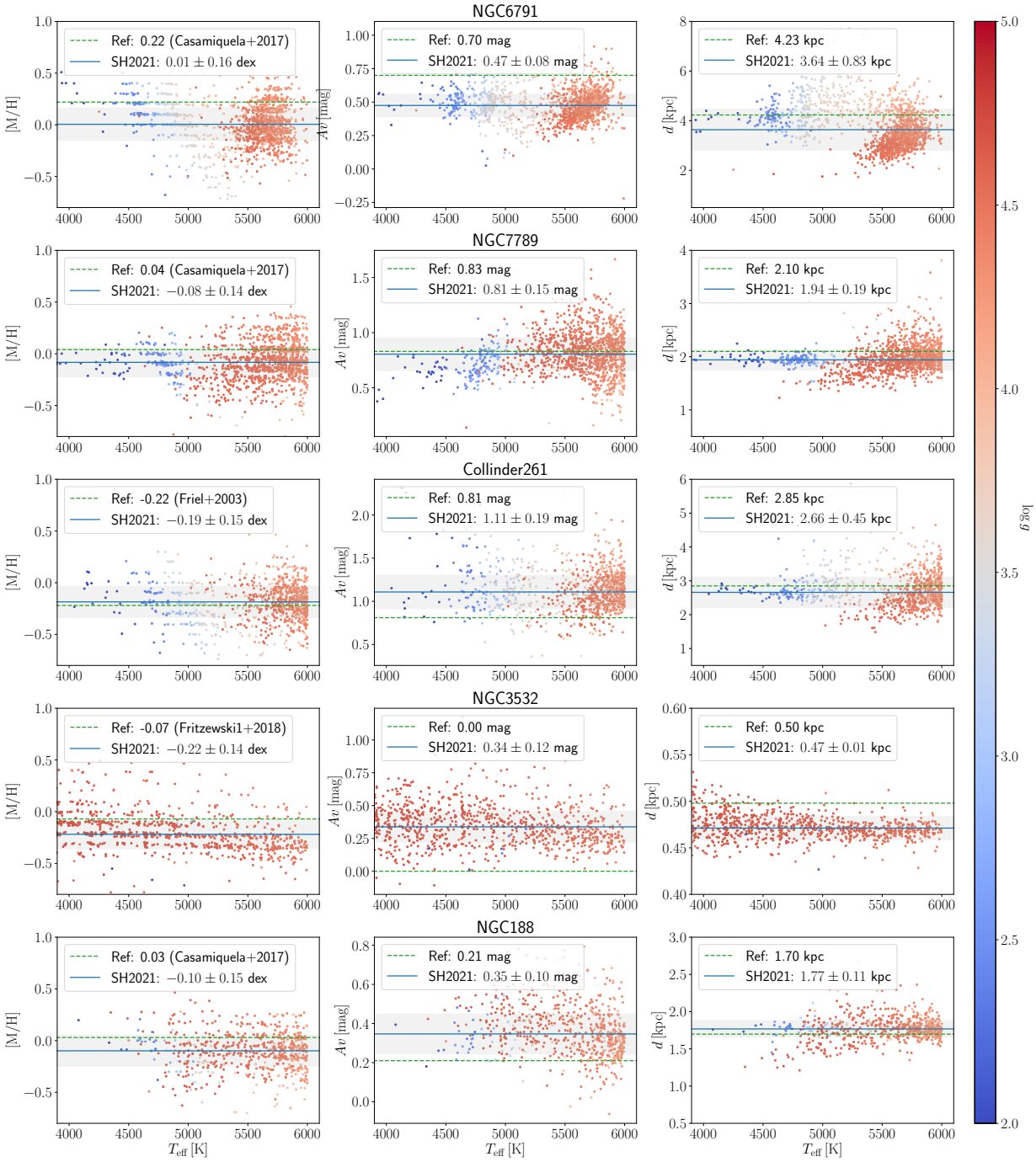

Fig. 14. Metallicity, extinction, and distance results for FGK star members of the five most populated Galactic OCs in the Cantat-Gaudin et al. (2020) catalogue, reflecting the typical precision of the `StarHorse` results as well as some systematic trends with effective temperature and surface gravity. The blue lines refer to the cluster median and the blue-shaded area to the median absolute deviation, while reference values are plotted as dashed orange lines. We note that the reference cluster metallicities are in fact iron abundances [Fe/H], which are only approximately equal to the total metallicity [M/H] determined by `StarHorse`.

– which leads to significantly biased results for massive stars (see Appendix D.5 for details).

5.3. Comparison to asteroseismology

Asteroseismology provides a unique way to peer into the interior of stars (for a recent review see Aerts 2021) and thus also to test the accuracy of our derived stellar parameters. In particular, asteroseismology of solar-type and red-giant stars (Chaplin & Miglio 2013) can provide very precise stellar surface gravities and masses (Ulrich 1986; Kjeldsen & Bedding 1995), evolutionary stages (Bedding et al. 2011; Mosser et al. 2011), and thereby also distances and extinctions (Rodrigues et al. 2014).

In Fig. 17 we compare our photo-astrometric results with the most precise and accurate parameters obtained for red-giant field stars (outside the immediate solar vicinity) to date. Miglio et al. (2021) combined asteroseismic observations

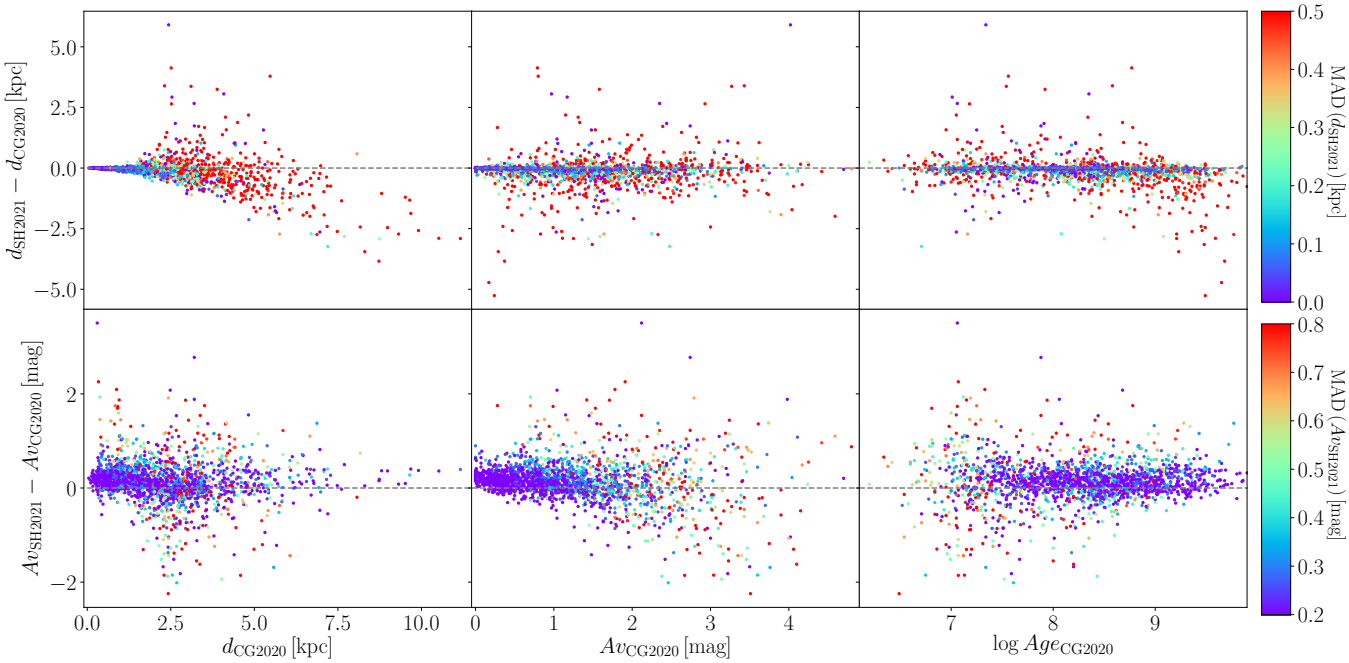

Fig. 15. Distance (*top row*) and extinction (*bottom row*) comparison with the OC parameter catalogue of Cantat-Gaudin et al. (2020). Each point represents one star cluster. We show the systematic difference between StarHorse (calculated as the median of all FGK member stars) and the reference value as a function of distance, A_V, and log age. The colour denotes the intrinsic dispersion (MAD) within each cluster.

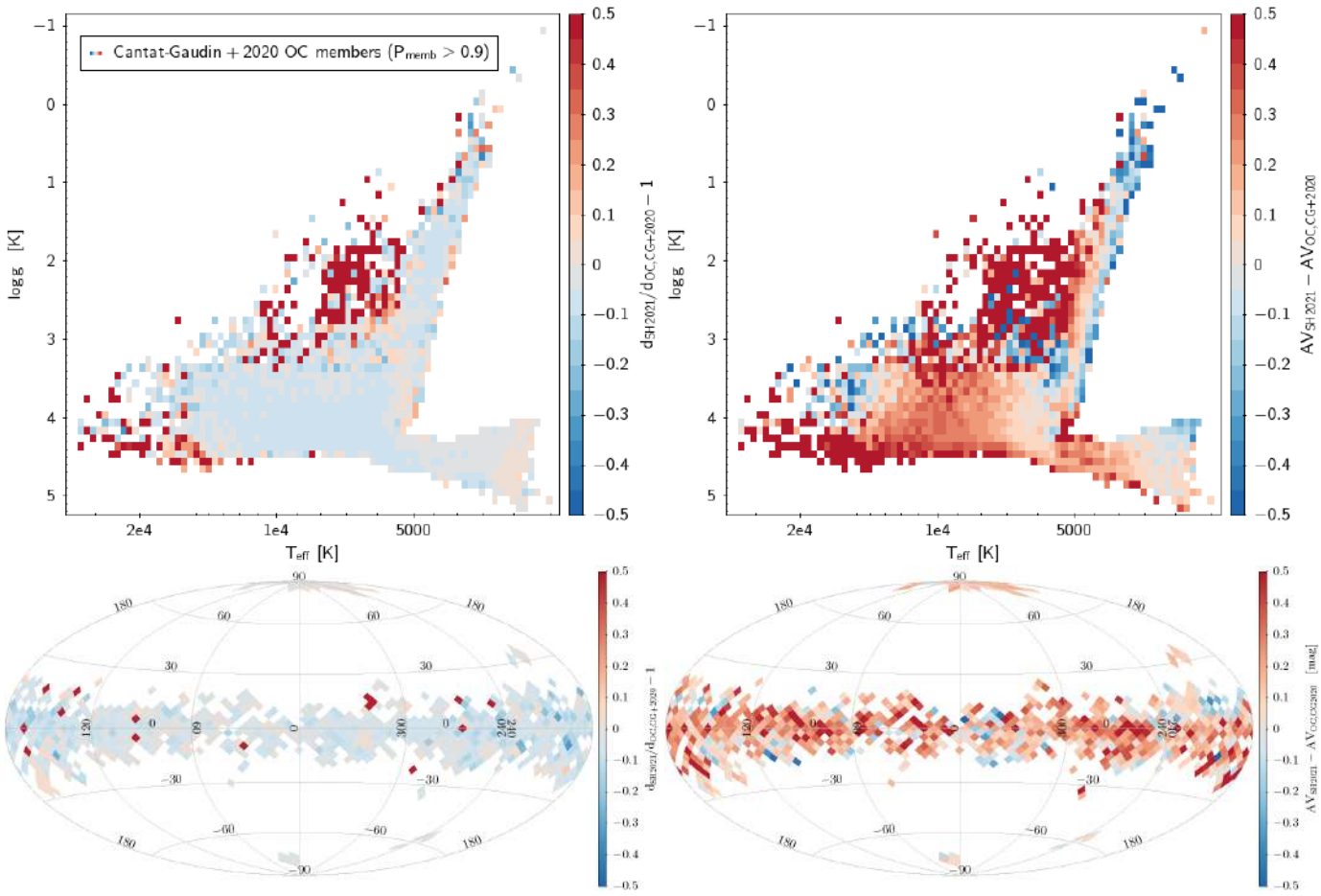

Fig. 16. StarHorse results for OC members: comparison to the distance and extinction scale of Cantat-Gaudin et al. (2020). *Top panels*: Kiel diagrams colour-coded by median relative distance difference (*left*) and absolute extinction difference (*right*) per pixel. *Bottom panels*: sky distribution colour-coded by median differences.

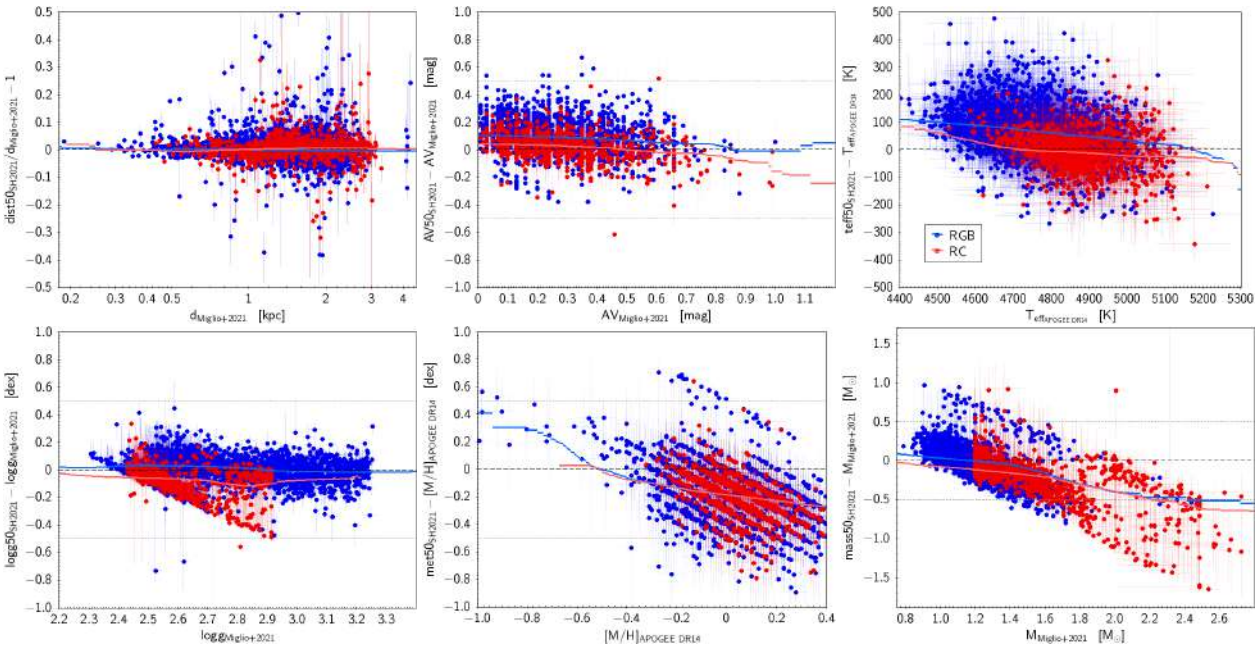

Fig. 17. Comparison of `StarHorse` EDR3 distances, extinctions, and effective temperatures (*top row*), as well as surface gravities, metallicities, and masses (*bottom row*) with the high-precision asteroseismic+spectroscopic red-giant catalogue of Miglio et al. (2021). In each panel we show the parameter difference as a function of the parameter itself, where the blue dots refer to RGB stars, while the red dots refer to core He-burning red-clump stars.

by *Kepler* (Gilliland et al. 2010) with APOGEE DR14 spectroscopy (Majewski et al. 2017; Abolfathi et al. 2018) and used the PARAM tool (da Silva et al. 2006; Rodrigues et al. 2017) to determine precise stellar parameters, distances, and extinctions. The authors also tested the influence of different stellar modelling assumptions (atomic diffusion, initial He abundance, [α/Fe]-enhancement, etc.)

The sample comprises 3195 stars in the *Kepler* field, and thus the systematic trends seen in Fig. 17 do not necessarily apply to the full sky, but the plots give a fair impression of the typical precision and accuracy that can be expected for giant stars. Similar to the comparison shown in A19 for the *Kepler* field, we do not see any significant trend in terms of distances, indicating again the improvement of the parallax zero-point calibration achieved by *Gaia* EDR3 and the corrections proposed by Lindegren et al. (2021b). For extinctions, we detect a slight overestimation (~0.1 mag) with respect to Miglio et al. (2021) for the RGB stars, while no significant offset is seen for the red-clump stars. As expected, a similar behaviour is seen in the effective temperature differences: the RGB $T_{\rm eff}$ scale of APOGEE DR14 is typically 100 K cooler than our inferred effective temperatures. We suggest that these slightly different trends for RGB stars and red-clump stars can probably be generalised to the full sky. We caution, however, that the absolute $T_{\rm eff}$ scales of both spectroscopy and stellar models are uncertain to within similar levels (e.g. Rodrigues et al. 2017; Miglio et al. 2021).

The second row of Fig. 17 shows the comparison for the secondary (naturally less reliable) output parameters $\log g$, [M/H], and mass. In each of these panels we see negative offsets and trends for red-clump stars, and similar, but typically milder ones for the RGB stars.

5.4. Comparison to GALAH DR3

Large-area spectroscopic stellar surveys like RAVE (Steinmetz et al. 2006, 2020), LAMOST (Deng et al. 2012; Zhao et al. 2012), or APOGEE (Majewski et al. 2017) are ideal to detect possible stellar parameter trends for astrophotometrically derived results. In A19 we showed a comparison with APOGEE DR14 (see also previous subsection); here we choose another example: the stellar parameters from the third data release (Buder et al. 2021) of the GALAH survey (De Silva et al. 2015).

Figure 18 shows the parameter comparison to GALAH DR3 for effective temperature, surface gravity, and metallicity. In line with Fig. 17, we see some slight trends (typically an overestimation by 100–200 K) for effective temperatures in the range of FGK (both dwarfs and giants) stars, where most of the common stars are located and for which the GALAH pipeline works best (Buder et al. 2018). Perhaps with the exception of the large $\log g$ spread generated by the red clump (indicating some impurity of the `StarHorse` red-clump sample), a similar pattern as for $T_{\rm eff}$ is observed for $\log g$, with the difference that both the median offset and the dispersion decrease with $\log g$, because for dwarf stars the `StarHorse` surface gravity is typically well constrained by the *Gaia* EDR3 parallax.

The least constrained parameters, as expected for our technique (combining parallaxes and broad-band photometry), are certainly mass and metallicity. The right panel of Fig. 18 shows clearly how our metallicity estimates are dominated by the (broad) Galactic metallicity priors. They clump around zero and exhibit a tail towards negative metallicity, but show little concordance trend (even for metal-poor stars) with metallicities determined from the high-resolution GALAH survey. We therefore remind the reader to use these estimates with caution.

5.5. Caveats

In A19 we enumerated a list of caveats that applied for the `StarHorse` *Gaia* DR2 results. While many of them have been addressed by the improvements presented in this work, some important caveats remain and should be taken into account when

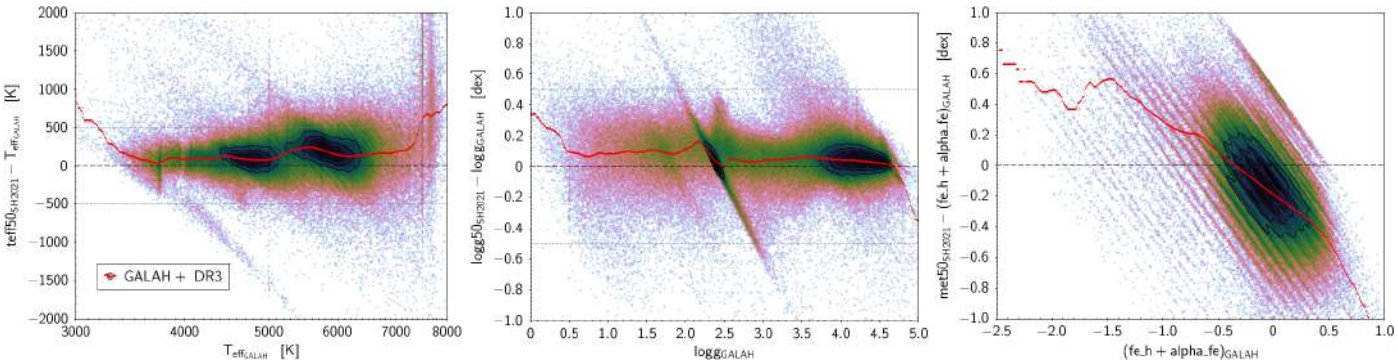

Fig. 18. Comparison of StarHorse EDR3 effective temperatures (*left*), surface gravities (*middle*), and metallicities (*right*) with the spectroscopically derived labels from GALAH+ DR3 (Buder et al. 2021). In each panel the red line corresponds to the running median.

using the results presented here. We discuss these in some detail in Appendix D.

6. Comparison to previous results

In this section we compare our results to some previous attempts to determine astrophysical parameters for massive amounts of *Gaia* stars. A comprehensive comparison to all such datasets is beyond the scope of this paper, so we choose some illustrative examples. In particular we compare to the *Gaia* EDR3 distances of Bailer-Jones et al. (2021), the astrophysical parameters of A19, and the effective temperatures of Bai et al. (2019) and the extinctions of Bai et al. (2020), both obtained by machine-learning algorithms.

6.1. Comparison to the Bailer-Jones et al. (2021) EDR3 distances

Shortly after *Gaia* EDR3, Bailer-Jones et al. (2021) published two sets of distance estimates for 1.47 billion objects based on EDR3 data. The first set, dubbed geometric distances, used solely *Gaia* parallaxes and a sophisticated prior for the stellar density distribution in the Milky Way (Rybizki et al. 2020), analogous to the *Gaia* DR2 catalogue published by the same group (Bailer-Jones et al. 2018b). The second set, dubbed photo-geometric distances, also used the *Gaia* EDR3 photometry to refine the distance prior for each star, thus providing more precise (and arguably also more accurate) results.

Figure 19 shows a comparison of our distance estimates with the two sets of distances obtained by Bailer-Jones et al. (2021) for random sample of 1 million stars. We see a remarkable concordance over almost the entire sky, especially with the set of photo-geometric distances (excluding only the Magellanic Clouds and the centremost Galactic regions; see the second panel in Fig. 19). Slightly larger differences (in the sense that StarHorse typically delivers smaller distances) are present in the comparison to the purely geometric distances, for the region around the Galactic plane. The high differences in the Magellanic Clouds are expected, since the priors of Bailer-Jones et al. (2021) do not include any extragalactic stellar populations, while we explicitly included these in our prior.

This is reassuring, since the method of Bailer-Jones et al. (2021) is quite different in both its prior assumptions (distance scale lengths are derived from a synthetic Milky Way model) and in the implementation of the posterior calculation (Markov chain Monte Carlo sampling). Also, the derived stellar density maps (bottom panels of Fig. 19) are very similar for Bailer-Jones et al. (2021) and StarHorse: both show almost the same density contours, perhaps with the exception that the Galactic bar appears slightly more prominent in our map, and that the StarHorse halo priors seem to allow for slightly more distant objects to be present. In addition, Fig. 13 (left panel) shows that the internal uncertainties obtained by Bailer-Jones et al. (2021) for their distance estimates are very similar to our StarHorse results.

6.2. Comparison to DR2-derived parameters

6.2.1. *Gaia* DR2 StarHorse

We have described the methodological differences with respect to our previous StarHorse results derived from *Gaia* DR2 in Sect. 3.1. The most important difference is, however, the clearly superior quality of the *Gaia* EDR3 catalogue (see Fabricius et al. 2021 for numerous examples showing the increased precision and accuracy of EDR3 compared to DR2). The direct comparison to the A19 catalogue (shown in Fig. 20) is therefore interesting, but of limited value as a true benchmark test; it is highlighting mostly the shortcomings of the previous catalogue.

As a first example of the improvements made since A19, Fig. 8 in this paper shows the spatial distribution of the RC stars selected as in Fig. 8 of A19. In the A19 plot the astrometric quality flag to select 'good' astrometric sources was applied, while for EDR3 (our Fig. 8) no quality cut was applied. The final EDR3 sample of red-clump stars contains 13 640 423 sources. Even without applying any quality cuts, improvements from DR2 to EDR3 are clearly evident in this figure. For example, the ring-like feature between 2 and 3 kpc reported in A19 and Rybizki et al. (2022, their Fig. 8), mostly provoked by bad astrometric solution, has almost disappeared. Also the unphysical paucity of red-clump stars in front of the Galactic bar visible in Fig. 8 of A19 has vanished.

Furthermore, Fig. 13 (uncertainties as a function of G) shows that the precision of the EDR3 results is significantly improved: typically the internal EDR3 StarHorse uncertainties in all output parameters (red lines in Fig. 13) are smaller than the corresponding A19 ones (black lines) by a factor of 2 at any given magnitude.

Figure 20 shows a direct comparison with the A19 results. In each panel we show a HealPix map of the median differences as a function of sky position. The top panel (distance comparison) shows the relative difference ($\mathtt{dist50}/\mathtt{dist50}_{A19} - 1$), while the other panels show absolute differences (this work − A19).

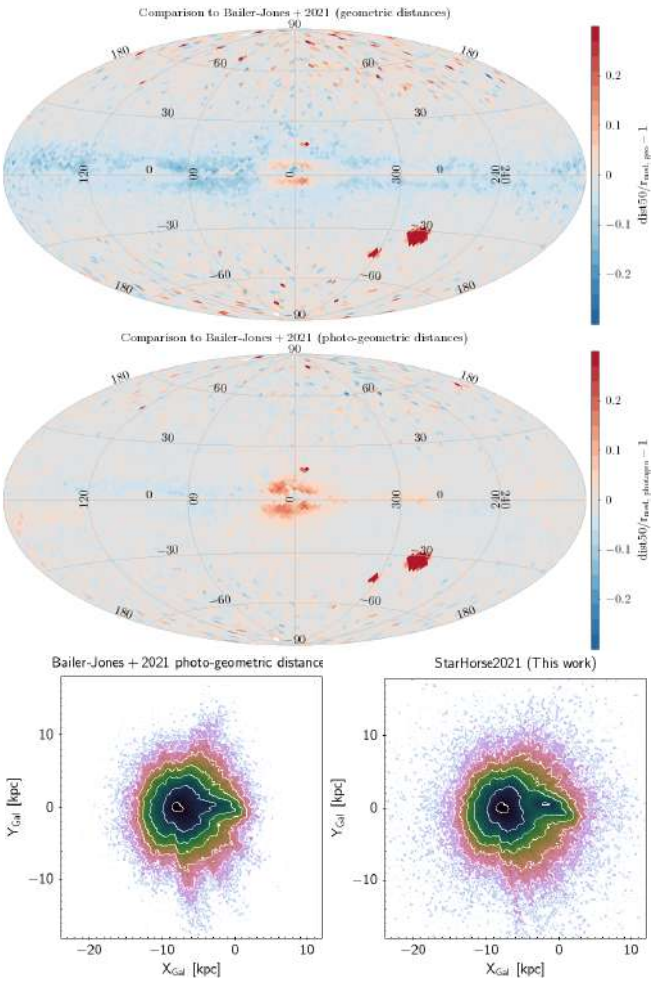

Fig. 19. Comparison of `StarHorse` EDR3 distances with the EDR3 distances from Bailer-Jones et al. (2021) using 1 million random stars. *Top panel*: sky map showing the relative distance difference with respect to the geometric distances (computed using only the *Gaia* EDR3 parallaxes). *Middle panel*: same for the photo-geometric distances (using also the EDR3 photometry in the distance inference). *Bottom panels*: visual appearance of the Cartesian Galactic maps derived from the photo-geometric distances of Bailer-Jones et al. (2021) (*left*) and `StarHorse` (*right*) for the same random sample. In both *bottom panels* the contour lines are logarithmically spaced.

6.2.2. Bai et al. (2019, 2020) effective temperatures and extinctions

Shortly after the release of *Gaia* DR2, Bai et al. (2019) produced a catalogue of stellar effective temperatures for 133 million *Gaia* DR2 stars, using a random-forest regressor trained on spectroscopically measured temperatures from a variety of stellar surveys, achieving precise ($\sigma_{\rm Teff}$ = 191 K) results for stars in the test and control samples. In a subsequent paper (Bai et al. 2020) the authors used a similar regression technique to determine precise $E(B-V)$ reddening values for the same stars (also using their previously derived effective temperatures).

In Fig. 21 we compare our new `StarHorse` results to the values of Bai et al. For effective temperature (lower-left panel of Fig. 21), we find relatively good mean concordance for FGK stars (that comprise the training set of Bai et al. 2019), while for hotter and cooler stars (according to `StarHorse`) the machine-learning pipeline of Bai et al. (2019) seems to force the $T_{\rm eff}$ values into the range of the training set. We also find significant systematics with sky position for both effective temperature and reddening (top and middle panels of Fig. 21) that seem to correlate partly with Galactic extinction and partly with the sky coverage of the training set used by Bai et al. (2019).

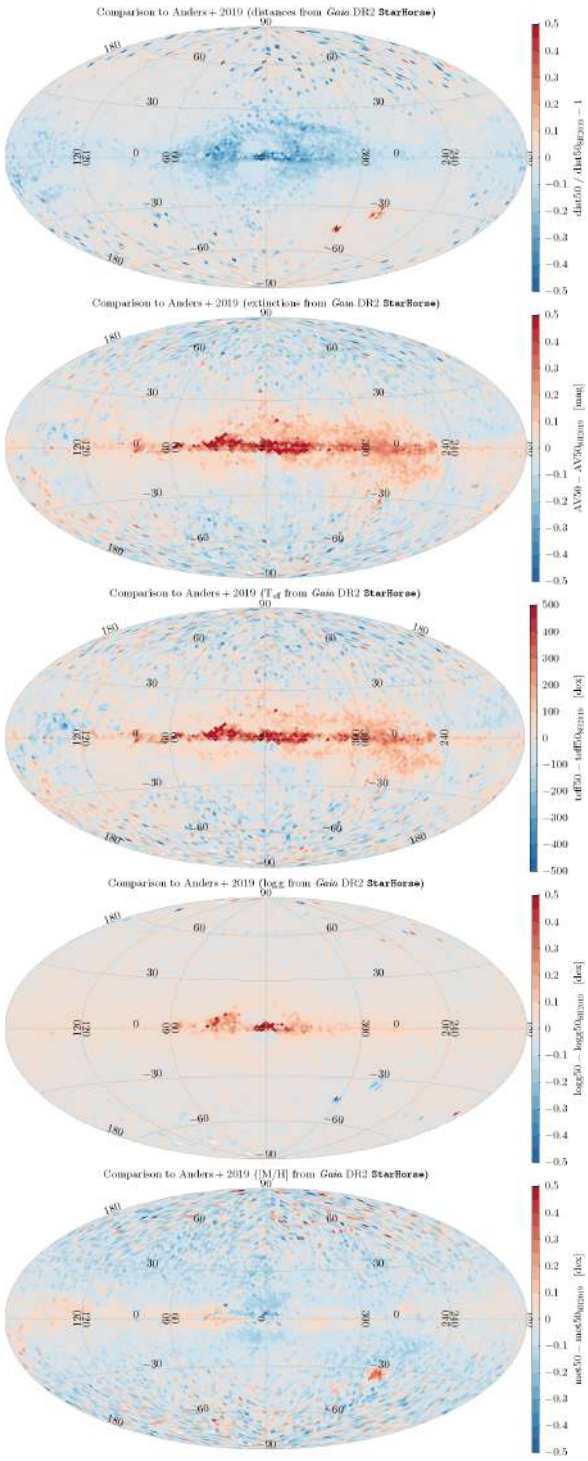

Fig. 20. Comparison of the `StarHorse` EDR3 results with the `StarHorse` DR2 results from Anders et al. (2019), for a random sample of 1 million stars. *From top to bottom*: sky distribution of median distance, extinction, effective temperatures, surface gravity, and metallicity differences.

Based on the previous comparisons, we suggest that systematics in the results of Bai et al. are likely causing the bulk of the differences seen in Fig. 21. This comparison also reminds

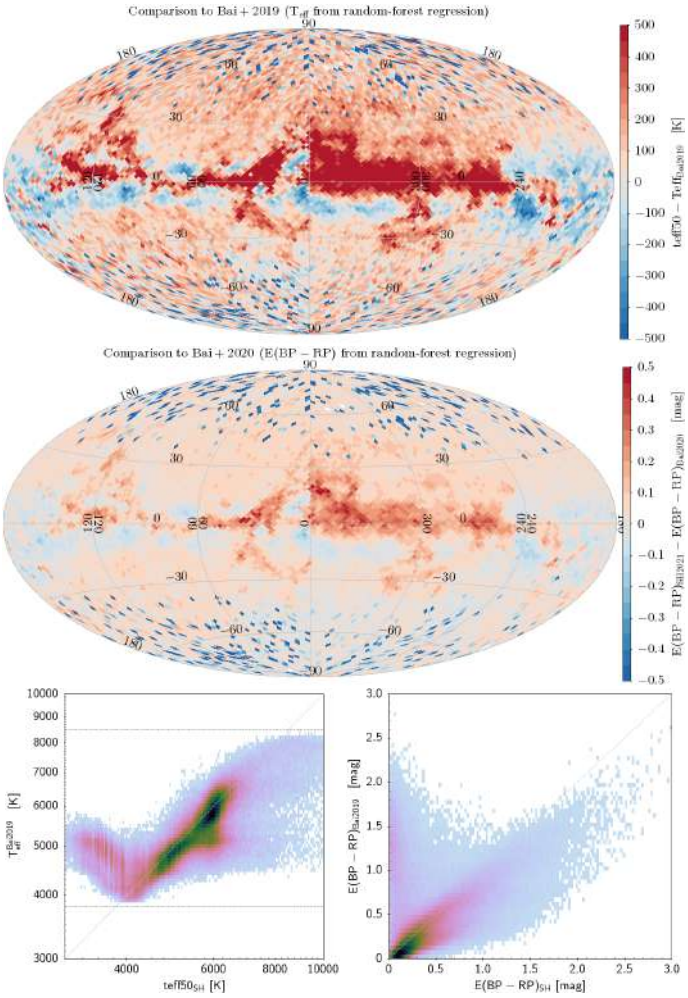

Fig. 21. Comparison of the StarHorse EDR3 results with the effective temperatures from Bai et al. (2019, *top panel*), and the reddenings from Bai et al. (2020, *middle panel*). The colour scale is the same as in Fig. 20. *Bottom panels*: one-to-one comparisons for both parameters.

us that extreme caution is due when interpreting the results of a machine-learning algorithm outside the (multi-dimensional) range of training data.

7. Conclusions

We present a catalogue of 362 million stellar parameters, distances, and extinctions based on *Gaia* EDR3, Pan-STARRS1, SkyMapper, 2MASS, and AllWISE. The new data and computational updates in our code serve to substantially improve the accuracy and precision over previous photo-astrometric stellar-parameter estimates (typically by a factor of 2 compared to A19).

The typical precisions, at magnitude $G = 14$ (17), amount to 3% (15%) in distance, 0.13 mag (0.15 mag) in V-band extinction, and 140 K (180 K) in effective temperature. Our results are validated by comparisons with OCs, as well as with asteroseismic and spectroscopic measurements, indicating systematic errors smaller than the nominal uncertainties for the vast majority of objects. We also provide distance- and extinction-corrected CMDs, extinction maps, and extensive stellar density maps that reveal detailed substructures in the Milky Way and beyond.

The new density maps now probe a much greater volume, extending to regions beyond the Galactic bar and to Local Group galaxies, with a larger total number density. The Galactic bar remains a very prominent feature in the density maps, especially when focussing on red-clump stars. Other subtler features, such as spiral arms or the Sagittarius stream, also start to appear in the density maps.

Our *Gaia* EDR3 StarHorse catalogue can be queried through CDS or the *Gaia* mirror archive[5] hosted by the Leibniz-Institut für Astrophysik Potsdam (AIP). In addition, we also provide approximations to the full posterior PDFs for download in HDF5 format and instructions for bulk data download[6] (see Appendix B for details).

In the near future, *Gaia* DR3 (planned for Q2 2022)[7] will provide new *Gaia* astrophysical parameters for ~500M stars, in part determined using also the *Gaia* BP/RP and RVS spectra, allowing for a further increase in precision for many millions of stars that might possibly supersede some parts of this catalogue. However, we expect that our Bayesian multi-wavelength approach will continue to be relevant and useful for fainter sources without BP/RP spectra ($G \gtrsim 17$), including after the release of the *Gaia* DR3 stellar parameters.

Acknowledgements. We warmly thank Anthony Brown (Leiden) and the referee for comments on the manuscript. During the analysis, we have made extensive use of the astronomical java software TOPCAT and STILTS (Taylor 2005), Aladin Lite (Bonnarel et al. 2000; Boch et al. 2014), as well as the python packages numpy and scipy (Oliphant 2007), astropy (Astropy Collaboration 2013), healpy (Górski et al. 2005; Zonca et al. 2019), dustmaps (Green 2018), pomegranate (Schreiber 2017), dask (Dask Development Team 2016), HoloViews (http://holoviews.org), matplotlib (Hunter 2007), and corner (Foreman-Mackey 2016). This research has made use of the SVO Filter Profile Service (http://svo2.cab.inta-csic.es/theory/fps/; Rodrigo & Solano 2020) supported from the Spanish MINECO through grant AYA2017-84089. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (http://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, http://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. F.A. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 800502 H2020-MSCA-IF-EF-2017 and from MICINN (Spain) through the Juan de la Cierva-Incorporación program under contract IJC2019-04862-I. This work was partially funded by the Spanish MICIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe" by the "European Union" through grant RTI2018-095076-B-C21, and the Institute of Cosmos Sciences University of Barcelona (ICCUB, Unidad de Excelencia 'María de Maeztu') through grant CEX2019-000918-M. T.A. acknowledges the grant RYC2018-025968-I funded by MCIN/AEI/10.13039/501100011033 and by "ESF Investing in your future". L.C. acknowledges support from "programme national de physique stellaire" (PNPS) and the "programme national cosmologie et galaxies" (PNCG) of CNRS/INSU. A.M. acknowledges support from the European Research Council Consolidator Grant funding scheme (project ASTEROCHRONOMETRY, G.A. n. 772293, http://www.asterochronometry.eu). P.R. acknowledges the support of the Agence Nationale de la Recherche (ANR project SEGAL ANR-19-CE31-0017) and the European Research Council (ERC grant agreement No. 834148).

References

Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42
Aerts, C. 2021, Rev. Mod. Phys., 93, 015001
Aguado, D. S., Myeong, G. C., Belokurov, V., et al. 2021, MNRAS, 500, 889
Anders, F., Khalatyan, A., Chiappini, C., et al. 2019, A&A, 628, A94
Anderson, L., Hogg, D. W., Leistedt, B., Price-Whelan, A. M., & Bovy, J. 2018, AJ, 156, 145
Andrae, R., Fouesneau, M., Creevey, O., et al. 2018, A&A, 616, A8
Antoja, T., Ramos, P., Mateu, C., et al. 2020, A&A, 635, L3
Arenou, F., Luri, X., Babusiaux, C., et al. 2018, A&A, 616, A17
Arentsen, A., Starkenburg, E., Martin, N. F., et al. 2020, MNRAS, 491, L11

[5] gaia.aip.de
[6] data.aip.de/starhorse21
[7] https://www.cosmos.esa.int/web/gaia/dr3

Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Bai, Y., Liu, J., Bai, Z., Wang, S., & Fan, D. 2019, AJ, 158, 93
Bai, Y., Liu, J., Wang, Y., & Wang, S. 2020, AJ, 159, 84
Bailer-Jones, C. A. L., Farnocchia, D., Meech, K. J., et al. 2018a, AJ, 156, 205
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018b, AJ, 156, 58
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Baumgardt, H., & Vasiliev, E. 2021, MNRAS, 505, 5957
Bedding, T. R., Mosser, B., Huber, D., et al. 2011, Nature, 471, 608
Beers, T. C., & Christlieb, N. 2005, ARA&A, 43, 531
Beers, T. C., Preston, G. W., & Shectman, S. A. 1985, AJ, 90, 2089
Bertelli Motta, C., Pasquali, A., Richer, J., et al. 2018, MNRAS, 478, 425
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Boch, T., & Fernique, P. 2014, in Astronomical Data Analysis Software and Systems XXIII, eds. N. Manset, & P. Forshay, ASP Conf. Ser., 485, 277
Bonnarel, F., Fernique, P., Bienaymé, O., et al. 2000, A&AS, 143, 33
Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019, A&A, 623, A108
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Breuval, L., Kervella, P., Anderson, R. I., et al. 2020, A&A, 643, A115
Brown, A. G. A. 2021, ARA&A, 59, 59
Buder, S., Asplund, M., Duong, L., et al. 2018, MNRAS, 478, 4513
Buder, S., Sharma, S., Kos, J., et al. 2021, MNRAS, 506, 150
Cantat-Gaudin, T., & Anders, F. 2020, A&A, 633, A99
Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, A&A, 618, A93
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1
Carrillo, I., Minchev, I., Steinmetz, M., et al. 2019, MNRAS, 490, 797
Casamiquela, L., Carrera, R., Blanco-Cuaresma, S., et al. 2017, MNRAS, 470, 4363
Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, A&A, 635, A45
Castro-Ginard, A., McMillan, P. J., Luri, X., et al. 2021, A&A, 652, A162
Chabrier, G. 2003, PASP, 115, 763
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints [arXiv:1612.05560]
Chaplin, W. J., & Miglio, A. 2013, ARA&A, 51, 353
Chiappini, C., Minchev, I., Starkenburg, E., et al. 2019, The Messenger, 175, 30
Chiti, A., Frebel, A., Mardini, M. K., et al. 2021, ApJS, 254, 31
Choudhury, S., Subramaniam, A., & Cole, A. A. 2016, MNRAS, 455, 1855
Choudhury, S., Subramaniam, A., Cole, A. A., & Sohn, Y. J. 2018, MNRAS, 475, 4279
Cioni, M. R. L., van der Marel, R. P., Loup, C., & Habing, H. J. 2000, A&A, 359, 601
Crosta, M., Giammaria, M., Lattanzi, M. G., & Poggio, E. 2020, MNRAS, 496, 2107
Cunningham, E. C., Garavito-Camargo, N., Deason, A. J., et al. 2020, ApJ, 898, 4
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of Point Sources
Cutri, R. M., Wright, E. L., Conrow, T., et al. 2013, Explanatory Supplement to the AllWISE Data Release Products, Tech. rep
Da Costa, G. S., Bessell, M. S., Mackey, A. D., et al. 2019, MNRAS, 489, 5900
da Silva, L., Girardi, L., Pasquini, L., et al. 2006, A&A, 458, 609
Das, K. K., Zucker, C., Speagle, J. S., et al. 2020, MNRAS, 498, 5863
Dask Development Team 2016, Dask: Library for Dynamic Task Scheduling
de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, The Messenger, 175, 3
Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, Res. Astron. Astrophys., 12, 735
De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604
Dias, W. S., Monteiro, H., Moitinho, A., et al. 2021, MNRAS, 504, 356
Drew, J. E., Gonzales-Solares, E., Greimel, R., et al. 2016, VizieR Online Data Catalog: II/341
Drimmel, R., Cabrera-Lavers, A., & López-Corredoira, M. 2003, A&A, 409, 205
Duquennoy, A., & Mayor, M. 1991, A&A, 500, 337
El-Badry, K., Rix, H.-W., & Heintz, T. M. 2021, MNRAS, 506, 2269
Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4
Fabricius, C., Luri, X., Arenou, F., et al. 2021, A&A, 649, A5
Foreman-Mackey, D. 2016, J. Open Source Software, 1, 24
Fuhrmann, K., Chini, R., Kaderhandt, L., & Chen, Z. 2017, ApJ, 836, 139
Gaia Collaboration (Brown, A. G. A., et al.) 2016, A&A, 595, A2
Gaia Collaboration (Spoto, F., et al.) 2018a, A&A, 616, A13
Gaia Collaboration (Helmi, A., et al.) 2018b, A&A, 616, A12
Gaia Collaboration (Brown, A. G. A., et al.) 2018c, A&A, 616, A1
Gaia Collaboration (Klioner, S. A., et al.) 2021a, A&A, 649, A9
Gaia Collaboration (Smart, R. L., et al.) 2021b, A&A, 649, A6
Gaia Collaboration (Antoja, T., et al.) 2021c, A&A, 649, A8
Gaia Collaboration (Luri, X., et al.) 2021d, A&A, 649, A7
Gaia Collaboration (Brown, A. G. A., et al.) 2021e, A&A, 649, A1
Gilliland, R. L., Brown, T. M., Christensen-Dalsgaard, J., et al. 2010, PASP, 122, 131
Girardi, L. 2016, ARA&A, 54, 95
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Grady, J., Belokurov, V., & Evans, N. W. 2021, ApJ, 909, 150
Green, G. 2018, Journal Open Source Software, 3, 695
Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, ApJ, 887, 93
Gudin, D., Shank, D., Beers, T. C., et al. 2021, ApJ, 908, 79
Hattori, K., Valluri, M., & Vasiliev, E. 2021, MNRAS, 508, 5468
Helmi, A. 2020, ARA&A, 58, 205
Helmi, A., Irwin, M., Deason, A., et al. 2019, The Messenger, 175, 23
Hilker, M., Baumgardt, H., Sollima, A., & Bellini, A. 2020, in Star Clusters: From the Milky Way to the Early Universe, eds. A. Bragaglia, M. Davies, A. Sills, & E. Vesperini, 351, 451
Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2021, MNRAS, 500, 1385
Huang, Y., Yuan, H., Li, C., et al. 2021a, ApJ, 907, 68
Huang, Y., Beers, T. C., Wolf, C., et al. 2021b, ApJ, submitted [arXiv:2104.14154]
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jao, W.-C., Henry, T. J., Gies, D. R., & Hambly, N. C. 2018, ApJ, 861, L11
Kjeldsen, H., & Bedding, T. R. 1995, A&A, 293, 87
Koch, A., McWilliam, A., Preston, G. W., & Thompson, I. B. 2016, A&A, 587, A124
Kruijssen, J. M. D., Pfeffer, J. L., Chevance, M., et al. 2020, MNRAS, 498, 2472
Lallement, R., Babusiaux, C., Vergely, J. L., et al. 2019, A&A, 625, A135
Lamer, G., Schwope, A. D., Predehl, P., et al. 2021, A&A, 647, A7
Lanzafame, A. C., Distefano, E., Barnes, S. A., & Spada, F. 2019, ApJ, 877, 157
Law, D. R., & Majewski, S. R. 2016, in The Sagittarius Dwarf Tidal Stream(s), eds. H. J. Newberg, & J. L. Carlin, 420, 31
Leike, R. H., Glatzle, M., & Enßlin, T. A. 2020, A&A, 639, A138
Limberg, G., Rossi, S., Beers, T. C., et al. 2021a, ApJ, 907, 10
Limberg, G., Santucci, R. M., Rossi, S., et al. 2021b, ApJ, 913, L28
Lindegren, L. 2018, Re-normalising the Astrometric chi-square in Gaia DR2
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021a, A&A, 649, A2
Lindegren, L., Bastian, U., Biermann, M., et al. 2021b, A&A, 649, A4
Lux, O., Neuhäuser, R., Mugrauer, M., & Bischoff, R. 2021, Astron. Nachr., 342, 553
Maíz Apellániz, J., & Weiler, M. 2018, A&A, 619, A180
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Marigo, P., Girardi, L., Bressan, A., et al. 2017, ApJ, 835, 77
Marrese, P. M., Marinoni, S., Fabrizio, M., & Altavilla, G. 2021, Gaia EDR3 Documentation Chapter 9: Cross-match with External Catalogues, Gaia EDR3 Documentation
McConnachie, A. W. 2012, AJ, 144, 4
Miglio, A., Chiappini, C., Mackereth, J. T., et al. 2021, A&A, 645, A85
Mohr-Smith, M., Drew, J. E., Napiwotzki, R., et al. 2017, MNRAS, 465, 1807
Monari, G., Famaey, B., Carrillo, I., et al. 2018, A&A, 616, L9
Montalbán, J., Mackereth, J. T., Miglio, A., et al. 2021, Nat. Astron., 5, 640
Mosser, B., Barban, C., Montalbán, J., et al. 2011, A&A, 532, A86
Mowlavi, N., Rimoldini, L., Evans, D. W., et al. 2021, A&A, 648, A44
Naidu, R. P., Conroy, C., Bonaca, A., et al. 2021, ApJ, 923, 92
Niu, Z., Yuan, H., & Liu, J. 2021, ApJ, 908, L14
Nogueras-Lara, F., Schödel, R., Gallego-Calvente, A. T., et al. 2019, A&A, 631, A20
Nogueras-Lara, F., Schödel, R., & Neumayer, N. 2021, A&A, 653, A33
Norris, J. E., Ryan, S. G., & Beers, T. C. 1999, ApJS, 123, 639
Oliphant, T. E. 2007, Comput. Sci. Eng., 9, 10
Olivares, J., Sarro, L. M., Bouy, H., et al. 2020, A&A, 644, A7
Onken, C. A., Wolf, C., Bessell, M. S., et al. 2019, PASA, 36, e033
Pantaleoni González, M., Maíz Apellániz, J., Barbá, R. H., & Reed, B. C. 2021, MNRAS, 504, 2968
Panter, B., Jimenez, R., Heavens, A. F., & Charlot, S. 2008, MNRAS, 391, 1117
Pastorelli, G., Marigo, P., Girardi, L., et al. 2019, MNRAS, 485, 5666
Pfeffer, J., Lardo, C., Bastian, N., Saracino, S., & Kamann, S. 2021, MNRAS, 500, 2514
Pietrzyński, G., Graczyk, D., Gallenne, A., et al. 2019, Nature, 567, 200
Poggio, E., Drimmel, R., Cantat-Gaudin, T., et al. 2021, A&A, 651, A104
Portail, M., Gerhard, O., Wegg, C., & Ness, M. 2017, MNRAS, 465, 1621
Portegies Zwart, S. 2021, A&A, 647, A136
Queiroz, A. B. A., Anders, F., Santiago, B. X., et al. 2018, MNRAS, 476, 2556
Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2020, A&A, 638, A76
Queiroz, A. B. A., Chiappini, C., Perez-Villegas, A., et al. 2021, A&A, 656, A156
Ramos, P., Antoja, T., Mateu, C., et al. 2021, A&A, 646, A99
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, ApJ, 885, 131

Reylé, C., Jardine, K., Fouqué, P., et al. 2021, A&A, 650, A201
Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3
Riess, A. G., Casertano, S., Yuan, W., et al. 2021, ApJ, 908, L6
Rodrigo, C., & Solano, E. 2020, in Contributions to the XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society, 182
Rodrigues, T. S., Girardi, L., Miglio, A., et al. 2014, MNRAS, 445, 2758
Rodrigues, T. S., Bossini, D., Miglio, A., et al. 2017, MNRAS, 467, 1433
Rybizki, J., Demleitner, M., Bailer-Jones, C., et al. 2020, PASP, 132
Rybizki, J., Green, G., Rix, H. W., et al. 2022, MNRAS, 510, 2597
Santiago, B. X., Brauer, D. E., Anders, F., et al. 2016, A&A, 585, A42
Schlafly, E. F., Meisner, A. M., Stutz, A. M., et al. 2016, ApJ, 821, 78
Schlafly, E. F., Peek, J. E. G., Finkbeiner, D. P., & Green, G. M. 2017, ApJ, 838, 36
Schreiber, J. 2017, ArXiv e-prints [arXiv:1711.00137]
Scolnic, D., Casertano, S., Riess, A., et al. 2015, ApJ, 815, 117
Seabroke, G. M., Fabricius, C., Teyssier, D., et al. 2021, A&A, 653, A160
Sestito, F., Martin, N. F., Starkenburg, E., et al. 2020, MNRAS, 497, L7
Sestito, F., Buck, T., Starkenburg, E., et al. 2021, MNRAS, 500, 3750
Shank, D., Beers, T. C., Placco, V. M., et al. 2021, ArXiv e-prints [arXiv:2109.08600]
Skowron, D. M., Soszyński, I., Udalski, A., et al. 2016, Acta Asrton., 66, 269
Souto, D., Allende Prieto, C., Cunha, K., et al. 2019, ApJ, 874, 97
Sozzetti, A., Damasso, M., Bonomo, A. S., et al. 2021, A&A, 648, A75
Starkenburg, E., Oman, K. A., Navarro, J. F., et al. 2017, MNRAS, 465, 2212
Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645
Steinmetz, M., Matijevič, G., Enke, H., et al. 2020, AJ, 160, 82
Steppa, C., & Egberts, K. 2020, A&A, 643, A137
Taylor, M. B. 2005, in Astronomical Data Analysis Software and Systems XIV, eds. P. Shopbell, M. Britton, & R. Ebert, ASP Conf. Ser., 347, 29
Thomas, G. F., Annau, N., McConnachie, A., et al. 2019, ApJ, 886, 10
Tumlinson, J. 2010, ApJ, 708, 1398
Ulrich, R. K. 1986, ApJ, 306, L37
Wegg, C., Gerhard, O., & Portail, M. 2015, MNRAS, 450, 4050
Wolf, M. 1923, Astron. Nachr., 219, 109
Youakim, K., Starkenburg, E., Aguado, D. S., et al. 2017, MNRAS, 472, 2963
Zari, E., Rix, H. W., Frankel, N., et al. 2021, A&A, 650, A112
Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, Res. Astron. Astrophys., 12, 723
Zonca, A., Singer, L., Lenz, D., et al. 2019, J. Open Source Software, 4, 1298
Zucker, C., Speagle, J. S., Schlafly, E. F., et al. 2020, A&A, 633, A51

Appendix A: Data model

Table A.1 provides the data model for the provided `StarHorse` output tables.

Table A.1. Data model of the *Gaia* EDR3 `StarHorse` catalogue released via the *Gaia* mirror at `gaia.aip.de`.

ID	Column name	Unit	Description
0	`source_id`		*Gaia* EDR3 unique source identifier
1	`dist05`	kpc	distance, 5th percentile
2	`dist16`	kpc	distance, 16th percentile
3	`dist50`	kpc	distance, 50th percentile
4	`dist84`	kpc	distance, 84th percentile
5	`dist95`	kpc	distance, 95th percentile
6	`av05`	mag	line-of-sight extinction at $\lambda = 5420$ Å, A_V, 5th percentile
7	`av16`	mag	line-of-sight extinction at $\lambda = 5420$ Å, A_V, 16th percentile
8	`av50`	mag	line-of-sight extinction at $\lambda = 5420$ Å, A_V, 50th percentile
9	`av84`	mag	line-of-sight extinction at $\lambda = 5420$ Å, A_V, 84th percentile
10	`av95`	mag	line-of-sight extinction at $\lambda = 5420$ Å, A_V, 95th percentile
11	`teff16`	K	effective temperature, 16th percentile
12	`teff50`	K	effective temperature, 50th percentile
13	`teff84`	K	effective temperature, 84th percentile
14	`logg16`	[cgs]	surface gravity, 16th percentile
15	`logg50`	[cgs]	surface gravity, 50th percentile
16	`logg84`	[cgs]	surface gravity, 84th percentile
17	`met16`		metallicity, 16th percentile
18	`met50`		metallicity, 50th percentile
19	`met84`		metallicity, 84th percentile
20	`mass16`	$M_\odot$	stellar mass, 16th percentile
21	`mass50`	$M_\odot$	stellar mass, 50th percentile
22	`mass84`	$M_\odot$	stellar mass, 84th percentile
23	`ag50`	mag	line-of-sight extinction in the G band, A_G, 50th percentile, derived from `av0` and `teff50`
24	`abp50`	mag	line-of-sight extinction in the G_{BP} band, A_{BP}, 50th percentile, derived from `av50` and `teff50`
25	`arp50`	mag	line-of-sight extinction in the G_{RP} band, A_{RP}, 50th percentile, derived from `av50` and `teff50`
26	`bprp0`	mag	dereddened colour, derived from `phot_bp_mean_mag, phot_rp_mean_mag, abp50, arp50`
27	`mg0`	mag	absolute magnitude, derived from `phot_g_mean_mag` (recalibrated), `dist50`, and `ag50`
28	`xgal`	kpc	Galactocentric Cartesian X coordinate, derived from `dist50` and assuming $X_0 = -8.2$ kpc
29	`ygal`	kpc	Galactocentric Cartesian Y coordinate, derived from `dist50` and assuming $X_0 = -8.2$ kpc
30	`zgal`	kpc	Galactocentric Cartesian Z coordinate, derived from `dist50` and assuming $Z_0 = 0$
31	`rgal`	kpc	Galactocentric planar distance, derived from `xgal` and `ygal`
32	`fidelity`		*Gaia* EDR3 astrometric fidelity flag (Rybizki et al. 2022)
33	`bp_rp_excess_corr`		*Gaia* EDR3 photometric BP/RP flux excess factor (corrected following Riello et al. 2021)
34	`sh_photoflag`		`StarHorse` photometry input flag
35	`sh_outflag`		`StarHorse` output quality flag

Appendix B: Approximation of the full posterior

In our previous catalogues we could not publish the full posterior PDFs, since they are typically very heavy files that are not stored to disc. We since implemented an approximation of the joint posterior of the five main spectroscopic output parameters (mass, age, metallicity, distance, extinction) using the fast weighted multivariate Gaussian Mixture Model included in the `python` module `pomegranate` (Schreiber 2017). We thus provide, along with the previously available output parameter PDF quantiles, a representation of the full posterior, stored in custom HDF5 files.

These HDF5 files can be accessed at `data.aip.de`. They contain, for each converged star, the weights, means, and covariances of the three Gaussian functions used to approximate the posterior. Five examples of the approximated PDFs, showing the varying complexity of the data, can be appreciated in Fig. B.1.

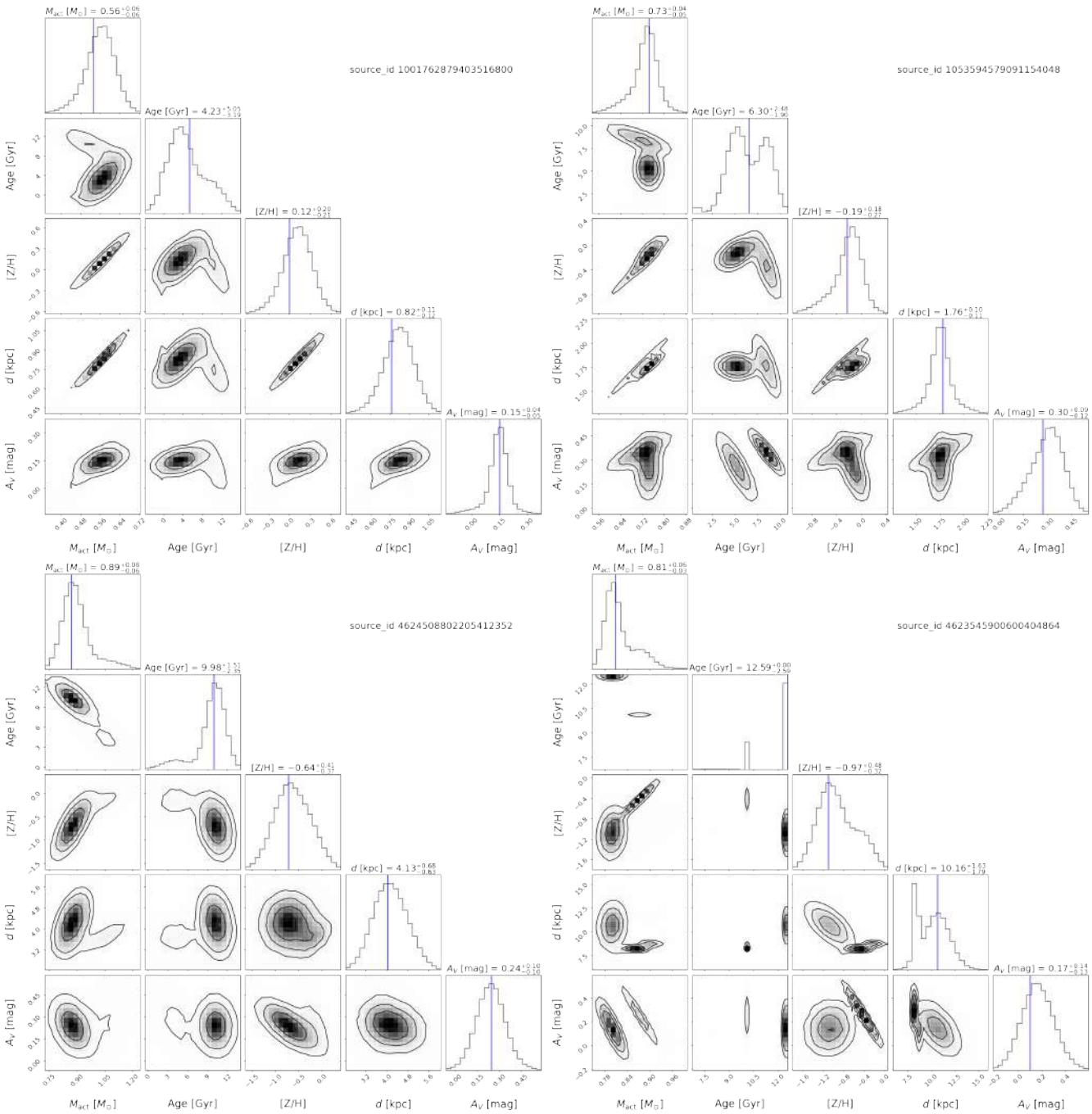

Fig. B.1. StarHorse posterior probability distributions for four example stars. In the off-diagonal sub-panels we show the two-dimensional projections of the five-dimensional posterior PDF (mass, age, metallicity, distance, and extinction) as approximated by a three-component Gaussian mixture model as black contours, while the diagonal panels show the 1D posterior approximations. The blue vertical lines show the median values directly inferred from the full posterior (available in the CDS tables and through ADQL).

Appendix C: Example ADQL queries

In this appendix we show some example ADQL queries that can be used to access the StarHorse *Gaia* EDR3 results via the *Gaia* mirror archive at gaia.aip.de. For example, To inspect the first 50 rows of the dataset, it is sufficient to write:

```
SELECT TOP 50 *
FROM gaiaedr3_contrib.starhorse
```

The second example shows how to access the first 50 rows of our results, cross-matched with the *Gaia* EDR3 catalogue, cleaned only for the first digit of the StarHorse output flag (see Sect. 3.3.3), sh_outflag[0]=="0":

```
SELECT TOP 50 g.ra, g.dec, s.*
FROM gaiaedr3.gaia_source AS g,
     gaiaedr3_contrib.starhorse AS s
WHERE g.source_id = s.source_id
     AND s.sh_outflag LIKE '0%%%'
```

The first 50 rows of the red-clump sample shown in Fig. 8 can be selected using this query:

```
SELECT TOP 50 s.*
FROM gaiaedr3_contrib.starhorse AS s
WHERE s.teff50 < 5000 AND s.teff50 > 4500
AND   s.met50  < .4   AND s.met50 > -.6
AND   s.logg50 < 2.55 AND s.logg50 > 2.35
AND   abs(s.zgal) < 3
```

A de-reddened CMD for a random sample can be obtained with a query like this:

```
SELECT bp_rp_index / 40 AS bp_rp,
       g_abs_index / 10 AS g_abs, n
FROM ( SELECT FLOOR(s.bprp0 * 40) AS bp_rp_index,
              FLOOR(s.mg0 * 10)   AS g_abs_index,
              COUNT(*) AS n
       FROM gaiaedr3.gaia_source AS g,
            gaiaedr3_contrib.starhorse AS s
       WHERE g.source_id = s.source_id
       AND   g.random_index < 1000000
       GROUP BY bp_rp_index, g_abs_index )
AS subquery
```

To retrieve a number of columns from both *Gaia* EDR3 and `StarHorse` for a random sample of 1 000 stars (including stars that are not in the `StarHorse` catalogue and even stars with missing parallaxes), one can use this type of query:

```
SELECT g.source_id, g.ra, g.dec,
       g.phot_g_mean_mag, g.parallax,
       g.parallax_error, s.dist50, s.teff50,
       s.av50, s.sh_outflag, s.sh_photoflag,
       s.fidelity
FROM gaiaedr3.gaia_source AS g
LEFT OUTER JOIN gaiaedr3_contrib.starhorse AS s
ON (g.source_id=s.source_id)
WHERE g.random_index < 1000
```

If one is interested in objects for which `StarHorse` did not converge (e.g. white dwarfs, galaxies, stars with problematic input data), this last example query shows how to retrieve them:

```
SELECT TOP 50
       g.source_id, g.l, g.b, g.parallax,
       g.parallax_error, g.phot_g_mean_mag,
       g.phot_bp_mean_mag, g.phot_rp_mean_mag
FROM gaiaedr3.gaia_source AS g
LEFT OUTER JOIN gaiaedr3_contrib.starhorse AS s
ON (g.source_id=s.source_id)
WHERE g.phot_g_mean_mag <= 18.5
AND   g.astrometric_params_solved > 3
AND   s.source_id IS NULL
```

Appendix D: Caveats

As mentioned in Sect. 5.5, many of the caveats present in our previous catalogue have been addressed in this work, but some important drawbacks remain and are discussed in this appendix.

D.1. Unresolved multiple stars

Many stars, both in the field and in star clusters, come in multiple systems. *Gaia* is able to resolve millions of wide binaries out to significant distances (e.g. El-Badry et al. 2021), but most multiple systems are still either completely or partially unresolved (resolved in the *G* band and in astrometry, unresolved in the BP/RP photometry; see e.g. Sect. 2 in Fabricius et al. 2021).

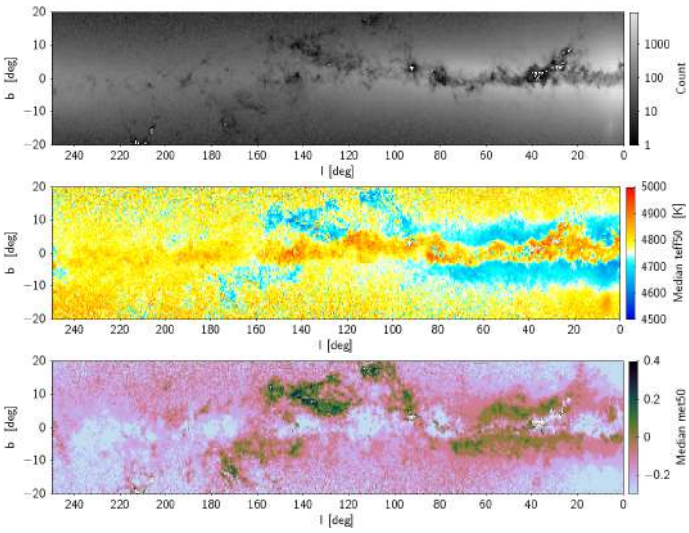

Fig. D.1. Sky map of the red-clump star sample for a region close to the Galactic plane ($0 < l < 250$, $-20 < b < 20$), revealing systematics possibly related to variations in the Galactic extinction law (compare with Fig. 1 of Schlafly et al. 2017). The top panel is colour-coded by number density, the second panel by median effective temperature, and the bottom panel by median metallicity.

A main drawback of `StarHorse` and most similar codes is that they do not take into account unresolved stellar multiplicity. Especially in the case of nearly equal-mass binaries or higher-order systems on the main sequence (which are quite abundant; see e.g. Duquennoy & Mayor 1991; Fuhrmann et al. 2017), we may expect significantly biased results. For example, our code might fit nearby main-sequence binaries (stars that are slightly brighter than predicted by single star models and for which the parallaxes are very well constrained by the *Gaia* EDR3 data) by moving them towards the sub-giant branch (i.e. higher effective temperatures) and to higher extinction values, so that the reddened synthetic magnitudes match the observed ones. This explains, at least in part, why our extinctions tend to be overestimated on average for nearby dwarf stars.

Properly taking into account multiplicity is beyond the scope of this work, but one way to allow for high mass-ratio binaries in the analysis would be to use data-driven stellar models (e.g. Anderson et al. 2018).

D.2. Low `StarHorse` convergence in crowded fields

In the middle panel of Fig. 4 we observed that `StarHorse` tends to converge less for objects located close to the Galactic plane, especially towards the centre of the Galaxy and in the Magellanic Clouds. The main reason for this is that the *Gaia* BP/RP aperture photometry is prone to systematics in crowded regions (e.g. Evans et al. 2018; Arenou et al. 2018; Fabricius et al. 2021; Riello et al. 2021). Filtering the data by the astrometric fidelity and the BP/RP excess factor increases the convergence inhomogeneities on the sky (bottom panel of Fig. 4), thereby rendering direct comparisons to simulations (without taking into account crowding effects in the *Gaia* selection function) even harder.

D.3. Variations in the extinction law induce systematic effective temperature shifts

As mentioned already in A19, our results rely to some degree on the validity of the assumed extinction curve, which we fixed

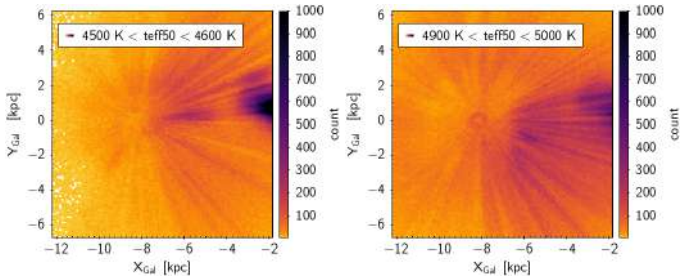

Fig. D.2. Galactic distribution for two thin effective temperature slices of the disc red-clump sample ($|Z_{Gal}| < 3$ kpc). The left panel shows a slight underdensity in the region where Pan-STARRS1 photometry is missing, while the right panel shows an overdensity in the region where SkyMapper photometry is available.

to the one recommended by Schlafly et al. (2016). Figure D.1 shows sky maps for red-clump stars close to the Galactic plane, highlighting some systematic trends in both effective temperature and metallicity. While most of the systematics may possibly be explained by selection effects provoked by the complex three-dimensional dust distribution, some of the trends maight also be correlated with the highly variable total-to-selective extinction ratio (R_V; see the detailed study of Schlafly et al. 2017, for example their Fig. 1).

D.4. Limited sky coverage of input catalogues causes slight inhomogeneities

While the input catalogues *Gaia* EDR3, 2MASS, and AllWISE cover the full sky with considerable homogeneity, the sky coverage of Pan-STARRS1 and SkyMapper photometry is limited by the location of the respective telescopes, which to some degree also affects the homogeneity of the resulting StarHorse catalogue (see e.g. the bottom panels of Fig. 13). The effect is alleviated by the similar filter system of Pan-STARRS1 and SkyMapper (and certainly less prominent than in A19 where no SkyMapper data were used), but should nonetheless be mentioned.

Another example is given in Fig. D.2, where we show the Galactic maps for the red-clump sample shown in Fig. 8, but for two narrow bins in effective temperature. Some of the underdensities in the left panel correspond to overdensities in the right panel (and vice versa), which is in part a plausible population effect, but in part also reflects the sky regions covered by Pan-STARRS1 and SkyMapper (compare to Fig. 8, bottom-right panel), which cautions us about the blind use of narrow bins in effective temperature or metallicity without taking into account their uncertainties.

D.5. Unreliable stellar parameters for massive stars

The OC comparisons discussed in Sect. 5.4 have already shown that for hot stars the present StarHorse stellar parameters are often biased. Here we therefore investigate this statistically small but astrophysically important subset in more detail using a known sample of OB stars.

Mohr-Smith et al. (2017) selected O-B3 stars in the far Carina spiral arm using VPHAS+ data (Drew et al. 2016). Their method to detect OB stars has proven to be quite reliable (confirmed by spectroscopy for some of the targets), thanks in part to the u filter (not used in our work). It also joins in 2MASS information in order to provide $T_{\rm eff}$, A_0 and R_V. We crossmatched this sample with the StarHorse EDR3 catalogue, resulting in 4,658 stars with good ($\chi^2 < 7.82$) parameters from Mohr-Smith et al. 2017, and compared the effective temperatures and extinctions with the ones obtained from VPHAS+.

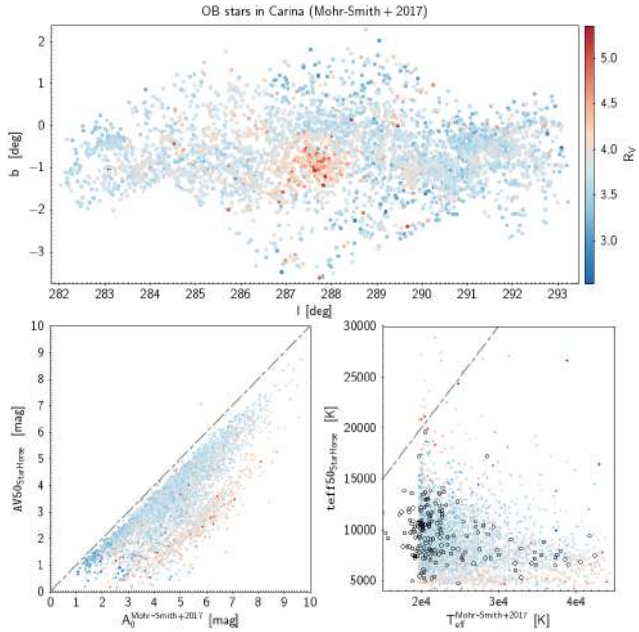

Fig. D.3. Comparison to the Carina OB star sample of Mohr-Smith et al. (2017). The top panel shows the sky distribution, while the bottom panels show one-to-one comparisons for extinction (left) and effective temperature (right). In each panel, the points are colour-coded by R_V, as determined by Mohr-Smith et al. (2017). The black circles in the bottom-right panel show the spectroscopically measured effective temperatures, available for a subset of the OB stars.

Figure D.3 shows that the StarHorse extinctions compare relatively well to the ones of the external catalogue (modulo a small offset that also depends on R_V; see Sect. D.3). The StarHorse effective temperatures for the O-B3 star candidates of Mohr-Smith et al. (2017), however, are in a completely different range than estimated by those authors. While it could be argued that this photometrically selected OB star sample may still be contaminated by lower-mass field stars, the observed $T_{\rm eff}$ differences are too drastic to be explained by contamination only. We therefore caution that our results for very massive stars are very likely to be unreliable in most cases.

Similar conclusions regarding our A19 results for OB stars have been reached by Pantaleoni González et al. (2021). The reason is that for a generic field-star approach such as StarHorse, the initial-mass-function prior strongly suppresses hot-star solutions, since they are a very small minority among the Galactic stellar population. Especially for rare stellar populations (e.g. OB stars: Zari et al. 2021; or open star clusters: Cantat-Gaudin et al. 2020; Olivares et al. 2020), specifically tailored algorithms are therefore expected to outperform our results.