Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **487,** 3946–3957 (2019)
Advance Access publication 2019 May 6

doi:10.1093/mnras/stz1239

Yule–Simpson's paradox in Galactic Archaeology

I. Minchev [1]★ G. Matijevic,[1] D. W. Hogg,[2,3,4,5] G. Guiglion,[1] M. Steinmetz,[1] F. Anders,[1] C. Chiappini,[1] M. Martig,[6] A. Queiroz[1] and C. Scannapieco [7]

[1]*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[2]*Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA*
[3]*Center for Data Science, New York University, 60 Fifth Ave, New York, NY 10011, USA*
[4]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[5]*Flatiron Institute, Simons Foundation, 162 Fifth Ave, New York, NY 10010, USA*
[6]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[7]*Facultad de Ciencias Exactas y Naturales, Departamento de Fśica, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina*

Accepted 2019 April 11. in original form 2019 February 7

ABSTRACT

Simpson's paradox, or Yule–Simpson effect, arises when a trend appears in different subsets of data but disappears or reverses when these subsets are combined. We describe here seven cases of this phenomenon for chemo-kinematical relations believed to constrain the Milky Way disc formation and evolution. We show that interpreting trends in relations, such as the radial and vertical chemical abundance gradients, the age–metallicity relation, and the metallicity–rotational velocity relation (MVR), can lead to conflicting conclusions about the Galaxy past if analyses marginalize over stellar age and/or birth radius. It is demonstrated that the MVR in RAVE giants is consistent with being always strongly negative, when narrow bins of [Mg/Fe] are considered. This is directly related to the negative radial metallicity gradients of stars grouped by common age (mono-age populations) due to the inside-out disc formation. The effect of the asymmetric drift can then give rise to a positive MVR trend in high-[α/Fe] stars, with a slope dependent on a given survey's selection function and observational uncertainties. We also study the variation of lithium abundance, A(Li), with [Fe/H] of AMBRE:HARPS dwarfs. A strong reversal in the positive A(Li)–[Fe/H] trend of the total sample is found for mono-age populations, flattening for younger groups of stars. Dissecting by birth radius shows strengthening in the positive A(Li)–[Fe/H] trend, shifting to higher [Fe/H] with decreasing birth radius; these observational results suggest new constraints on chemical evolution models. This work highlights the necessity for precise age estimates for large stellar samples covering wide spatial regions.

Key words: Galaxy: abundances – Galaxy: disc – Galaxy: evolution – Galaxy: kinematics and dynamics – Galaxy: stellar content.

1 INTRODUCTION

Galactic Archaeologists use the observed stellar kinematics, chemical abundances, and derived ages as fossil records to recover the chemo-dynamical history of the Milky Way (Freeman & Bland-Hawthorn 2002; Bland-Hawthorn, Krumholz & Freeman 2010; Binney 2013; Rix & Bovy 2013; Hogg et al. 2016; Minchev 2016). Age information is by far the hardest one to obtain, yet extremely important to break degeneracies among different chemical evolution models (Miglio et al. 2017).

To infer the Galactic disc history, we rely on observed correlations, such as the mean stellar metallicity,[1] [Fe/H], as a function of Galactic disc radius. The negative metallicity gradient observed today suggests that the disc has formed inside out (Matteucci & Francois 1989). The observational situation is improving rapidly: with the expansion of the disc coverage around the Sun due to a number of Galactic surveys, it has become possible in recent years to identify a number of new chemo-kinematical trends.

For example, an inversion of the radial metallicity gradient from negative to positive (and in [α/Fe] gradient from positive to negative)

★ E-mail: iminchev@aip.de

[1]As commonly done in the literature, throughout this work we use 'metallicity' to mean the iron fraction [Fe/H], not [M/H] – the overall metallicity of a star, defined as the total amount of elements heavier than Helium.

for samples at increasingly larger distance from the disc mid-plane, $|z|$, has been observed in most large Milky Way surveys (e.g. SEGUE – Cheng et al. 2012; RAVE – Boeche et al. 2013b; APOGEE – Anders et al. 2014; Gaia-ESO – Recio-Blanco et al. 2014; LAMOST – Wang et al. 2019), explained by Minchev, Chiappini & Martig (2014a) (see also Rahimi, Carrell & Kawata 2014; Miranda et al. 2016; Ma et al. 2017; Schönrich & McMillan 2017) as the effect of inside-out disc formation and disc flaring (increasing disc thickness with Galactocentric radius) of mono-age populations (fig. 10 in Minchev et al. 2014a). It was shown that gradient inversion can result from the variation in the stellar age distribution with radius: at large $|z|$ stars that are old, metal-poor, and [α/Fe]-rich dominate in the inner disc (due to the inside-out formation), while younger, more metal-rich, and more [α/Fe]-poor stars are more abundant in the outer disc (present at large $|z|$ due to disc flaring). This naturally results in a positive metallicity (and a negative [α/Fe]) gradient at large $|z|$, although for groups of stars of common age (mono-age populations) the gradient is always negative.

Another example of a trend that can be linked to the Galaxy's past is disc flaring (e.g. Kazantzidis et al. 2008; Bland-Hawthorn & Gerhard 2016; Bovy et al. 2016; Kawata et al. 2017; Wang et al. 2018). Disc flaring can be related to the dynamical effect of merging satellites, e.g. the Sagittarius dwarf (Laporte et al. 2018; Thomas et al. 2019). Using two suites of simulations of galactic disc formation in the cosmological context, Minchev et al. (2015) showed that flaring is always present for mono-age populations. When considering the total stellar population, however, it was found that the disc flaring was lost or at least significantly decreased. This was explained by a combination of the nested flares of mono-age populations, where younger groups of stars flare and dominate in density at progressively larger radii as a result of the inside-out disc formation.

As a final example, Minchev, Chiappini & Martig (2013, hereafter MCM13) found that in the solar vicinity resulting from their chemo-dynamical model the variation of [Fe/H] with vertical distance from the disc mid-plane, $|z|$, showed a strong negative gradient (~ -0.3 dex kpc^{-1}) but almost flat relations for mono-age populations (see their fig. 9, top). This phenomenon could be understood by taking into account that the fraction of metal-poor old stars increases with $|z|$, thus giving rise to a negative gradient for the total population (see results with SEGUE G-dwarfs – Schlesinger et al. 2012; RAVE – Boeche et al. 2013b; LAMOST data – Wang et al. 2019).

What all of the above examples have in common is that a relation resulting from the total sample is weakened, erased, or even inverted, compared to the trends in subsets grouped by considering a third variable (in this case, age).

This phenomenon is known as Simpson's, or Yule–Simpson's, paradox (hereafter, YSP), in classical statistics (e.g. Yule 1902; Simpson 1951; Thompson 2006; Smith & Goltz 2012). It has mostly been described in social sciences, medicine, biology, and baseball batting averages, but also in quantum mechanics (Li et al. 2013; Selvitella 2017). YSP arises when both the independent and dependent variables of a given relation depend on yet a third variable, often referred to as a 'lurking' or 'confounding' variable (in the above cases – age). A lurking variable is one for which data are unavailable, but it nevertheless has influence on other variables in the study. See for example Brase & Brase (2016) for more information on the topic.

Five cases of the YSP, although not identified as such, occurring in the field of Galactic Archaeology have been described in several works in the last 5 yr (Minchev et al. 2013, 2014a,b, 2015, 2018). We here elaborate on those and discuss two new cases. We will refer to a 'weak' case of Simpson's paradox when a relation is simply flattened or weakened and a 'strong' case when it is completely inverted.

In Sections 2.2–2.5 we make use of the Milky Way chemo-dynamical model by MCM13, who used a hybrid technique to combine a classical semi-analytical chemical evolution model (Chiappini 2009) with a simulation in the cosmological context (Martig et al. 2012). The model has been shown to comply with a range of observational constraints, as well as predict new ones (for a summary, see Minchev 2016). In Sections 2.2 and 2.1 we present YSP cases for global Galactic disc chemo-kinematical trends. In Sections 2.3–2.7 we will be concerned with relations of samples confined to the solar neighbourhood.

2 EXAMPLES OF SIMPSON'S PARADOX

2.1 Variation of disc thickness with Galactocentric radius

As a first illustration of YSP, we consider a cosmological zoom-in hydrodynamical simulation, using initial conditions from one of the Aquarius Project haloes (Springel et al. 2008; Scannapieco et al. 2009) and a TREE-PM SPH code. The spatial and mass resolutions are 300 pc and 4.4×10^5 M$_\odot$, respectively. The stellar mass is 5.5×10^{10} M$_\odot$ and the disc scale length is $h_\mathrm{d} = 4$ kpc. More details about this simulation can be found in Aumer et al. (2013; model Aq-D-5).

In the top panel of Fig. 1 colour curves show the variation of disc scale height, h_z, with galactic radius, r, for mono-age populations of width 1 Gyr and median values as seen in the colour bar. Single exponential models were fitted to get h_z in radial bins of width $0.5\,h_\mathrm{d}$ and distance from the disc mid-plane h_d (see details in Minchev et al. 2015). Strong disc flaring (an increase of scale height with radius) can be seen for the oldest mono-age populations, decreasing for younger subsamples. When the total stellar population is considered, the vertical disc density at all radii requires a sum of two exponentials for a proper fit (Gilmore & Reid 1983). Intriguingly, the resulting thin (short-dash curve) and thick (long-dash curve) discs do not show much flaring, even though they are composed of the strongly flared mono-age subsamples. This is an example of geometrical (also referred to as structural or morphological) definition of galactic thin and thick discs, as opposed to separation by age or by chemistry in the [α/Fe]–[Fe/H] plane for which 'low-' and 'high-[α/Fe] sequences', respectively, are more appropriate names (see discussion in Martig et al. 2016).

This figure highlights a simple example of a 'weak' case of YSP, where the strong variation in h_z with radius for mono-age subsamples is lost when the total population is considered. As discussed by Minchev et al. (2015), this results because the flaring of younger samples typically dominate at larger radii due to the inside-out disc formation. These 'nested flares' constitute a geometric thick disc which does not flare, or in which the flaring is strongly reduced, consistent with the lack of disc flaring found in observations of external galaxies (van der Kruit & Searle 1982; de Grijs 1998; Comerón et al. 2011). It can be seen in the bottom panel of Fig. 1 that, indeed, the density of older mono-age groups drops much faster with radius and it flattens for younger populations.

This model of thick disc formation predicts a negative radial age gradient in geometrically defined thick discs, which was found to be indeed the case for the Milky Way using APOGEE data (Martig et al. 2016). Flaring in Milky Way stellar populations of limited age range has been found by Kalberla et al. (2014; Red Clump), Feast et al. (2014; Cepheids), Carraro et al. (2015; young

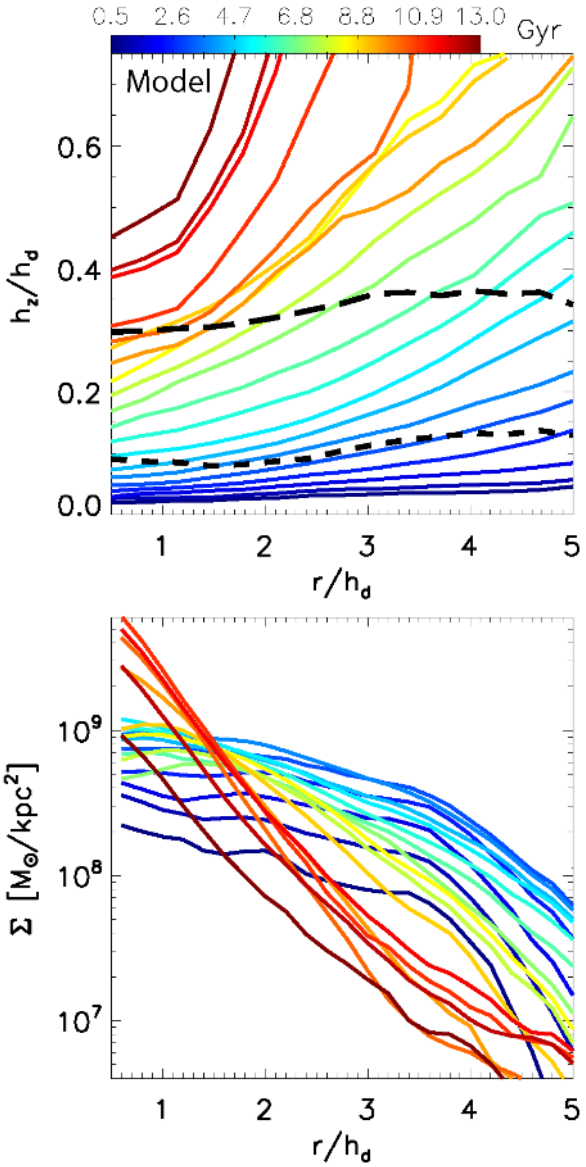

Figure 1. *Top:* Variation of disc scale height, h_z, with galactic radius for a cosmological simulation (see text). Colour lines show mono-age populations, as indicated. Bins of width 1 Gyr are used and exponential models at different radii are fitted out to $|z| = h_d$, where h_d is the disc scale length. Overlaid also are the thin (short-dash curve) and thick (long-dash curve) discs obtained by fitting a sum of two exponentials to stars of all ages. No significant flaring is found for the thin and thick discs. *Bottom:* Disc surface density for mono-age populations showing that older samples are confined to the inner disc as a result of the inside-out disc formation. Flaring is lost in the total population since younger discs dominate in density at progressively larger radii.

clusters), and in LAMOST (Wan, Liu & Deng 2017; Wang et al. 2018; Xiang et al. 2018) and APOGEE (Mackereth et al. 2017) mono-age populations. A proper Milky Way disc mass model is needed to assess the interplay among the flares of different mono-age populations on the overall scale height variation with radius (that of the total disc mass).

Flaring in mono-age populations from a theoretical point of view appears to be unavoidable, shown to result from merger perturbations (Kazantzidis et al. 2008; Villalobos & Helmi 2008; Martig, Minchev & Flynn 2014), misaligned gas infall (Scannapieco et al. 2009; Roškar et al. 2010), and reorientation of the disc rotation axis (Aumer & White 2013). A thick disc model incorporating disc flaring is attractive because flaring becomes a necessity rather than a nuisance (as often seen in numerical works in the past).

As will be shown next, inside-out disc formation combined with disc flaring can also explain the inversion of chemical abundance gradients with distance from the disc mid-plane.

Possible erroneous interpretation in the absence of age measurements: The Milky Way disc presents no flaring and thus dynamical mechanisms known to flare discs were not at work during its evolution. This is not true if you slice by age – a better interpretation is that mono-age discs always flare but flaring is lost in the total (or wide age range) sample due to the inside-out forming disc. This is a weak YSP case.

2.2 Inversion in radial abundance gradients

As discussed in Section 1, inversion in the [α/Fe] (e.g. [Mg/Fe]) gradient with distance from the disc mid-plane has been found in a number of Galactic surveys. We investigate the origin of this in the MCM13 model.

The top left-hand panel of Fig. 2 shows the radial [Mg/Fe] gradient for the MCM13 model (similar to fig. 10 by Minchev et al. 2014a). The black-dashed curve results from the total stellar population with a maximum distance from the disc mid-plane of $|z| = 0.3$ kpc. Different colours correspond to different mono-age populations, as indicated. Although the younger mono-age groups have well-defined positive slopes, that of the total population is strongly flattened inside $r \lesssim 8$ kpc. This can be understood if we consider the relative stellar surface density as a function of radius for each mono-age population, which is indicated by the height of the coloured rectangles (conveying the same information as the disc surface density variations with radius shown in the bottom panel of Fig. 1). The old stars – red and orange bins – dominate in the inner disc, causing the mean [Mg/Fe] at $r \lesssim 8$ kpc to go up. On the other hand, the younger populations dominate in the outer disc, as evident from the total mean closely following the age = 5 Gyr line.

The top right-hand panel of Fig. 2 is similar to the left one, but for a sample at $0.8 < |z| < 1.5$ kpc. For this significantly higher distance from the disc mid-plane, the density of old stars is found to strongly dominate at small radii, causing an increase in mean [Mg/Fe] by about 0.2 dex compared to the sample confined to the disc plane and, thus, a negative gradient.

The bottom row of Fig. 2 presents the effect of uncertainties, as expected from high-resolution spectroscopy (e.g. APOGEE – Majewski et al. 2017; WEAVE – Dalton et al. 2012; 4MOST – de Jong et al. 2012) δ[Mg/Fe] = 0.05 dex and asteroseismic ages (e.g. CoRoT, K2 – Howell et al. 2014; TESS – Ricker et al. 2015; PLATO – Rauer et al. 2014) δage = 20 per cent, drawn from a Gaussian distribution. Interestingly, the gradient of the total sample is only minimally affected, while a decrease of ∼0.05 dex is found in the oldest mono-age group.

The above illustrates a 'strong' YSP case, resulting from the effect of inside-out disc formation combined with disc flaring (see Section 2.1), giving rise to a different mixture of ages as a function of Galactic radius at different $|z|$ – older, [α/Fe]-rich stars are confined to the inner disc and large $|z|$, while younger, [Mg/Fe]-poor stars dominate above the disc plane in the outer disc. Similar explanation holds for the inversion of [Fe/H] gradient from negative close to the disc plane, to positive above it.

Possible erroneous interpretation in the absence of age measurements: (1) The sign shift in the [α/Fe] and [Fe/H] radial

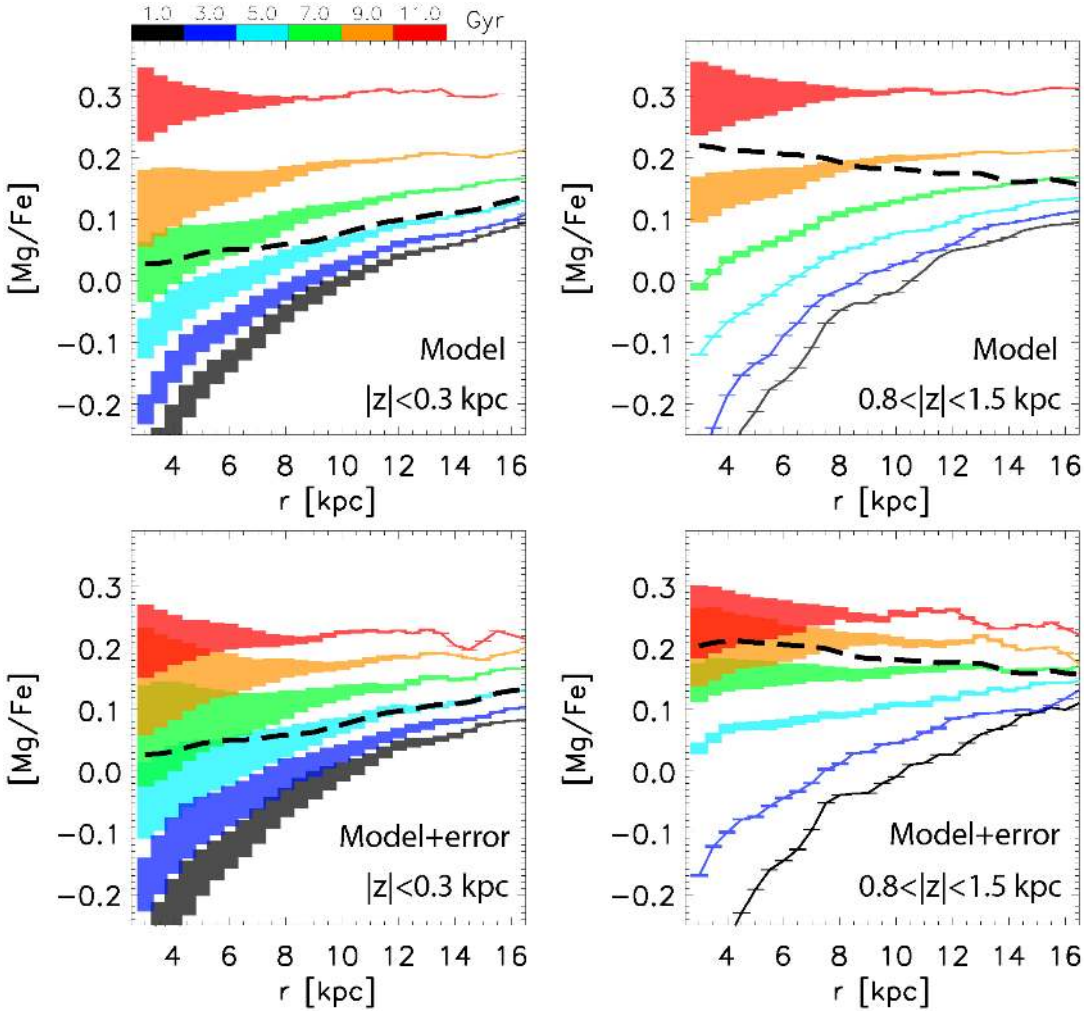

Figure 2. *Top:* Variation of azimuthally averaged [Mg/Fe] gradient with distance from the disc mid-plane for the MCM13 model. Dashed black curves show the [Mg/Fe] variation with Galactic radius for stellar samples at two different distances from the disc mid-plane, $|z|$, as marked in each panel. Different colours correspond to different age groups, as indicated in the colour bar. The height of the coloured rectangles reflects the stellar surfaces density of each radial bin. The positive gradient seen in the total population close to the plane (left-hand panel) is reversed at high $|z|$ (right-hand panel). The gradient of each individual age bin, however, is positive or flat (for the oldest age group). *Bottom:* Same as above, but with convolved uncertainties as expected from precise observations: δ[Mg/Fe] = 0.05 dex and δage = 20 per cent. This figure illustrates both a weak (left column) and a strong case (right column) of Simpson's paradox.

gradients found in observations at larger distance form the disc mid-plane resulted from an outside-in (thick) disc formation or a dominating accreted population at large $|z|$. A better interpretation, according to our model, may be that [Mg/Fe] ([Fe/H]) depends negatively (positively) on r only when one marginalizes over age. This is a strong YSP case. (2) The flattened gradient close to the disc plane is in disagreement with classical chemical evolution models. Steeper gradients are expected in mono-age populations. This is a weak YSP case.

2.3 Birth radius versus vertical velocity dispersion

We now demonstrate that disc flaring and heating as a function of radius can be seen in the local velocity distribution due to the effect of stellar radial migration.

In the left-hand panel of Fig. 3 we show the variation of the vertical velocity dispersion, σ_z, with birth radius, $r_{\rm birth}$, for stars found in the simulated solar neighbourhood of the MCM13 model. σ_z is computed as the standard deviation of the vertical velocity in each radial bin. The black-dashed curve shows the total population, giving the impression that the hottest stars in the solar vicinity today have arrived from the inner disc. Looking at this relation it may be tempting to conclude that it is radial migration of hot inner disc stars that give rise to a thick disc. However, dissecting by age (colour-coded curves) an opposite relation is revealed for the oldest stellar populations showing that stars born at 2–3 kpc have half the velocity dispersion of coeval populations born at the solar radius. This predicts that the kinematically hottest stars near the Sun today were born locally or in the outer disc – a natural result of the flaring in mono-age discs (see Section 2.1), which is unavoidable in simulations for different reasons, the most likely possible being perturbations from orbiting satellites.

The strong positive slope seen in old stars flattens towards redshift zero (present time), becoming slightly negative for age $\lesssim$4 Gyr, in agreement with the expected effect of migration in quiescent discs (Minchev et al. 2012; Roškar, Debattista & Loebman 2013). Note that the flaring induced by outward migration of stars with larger vertical actions and inward migration of outer disc born stars

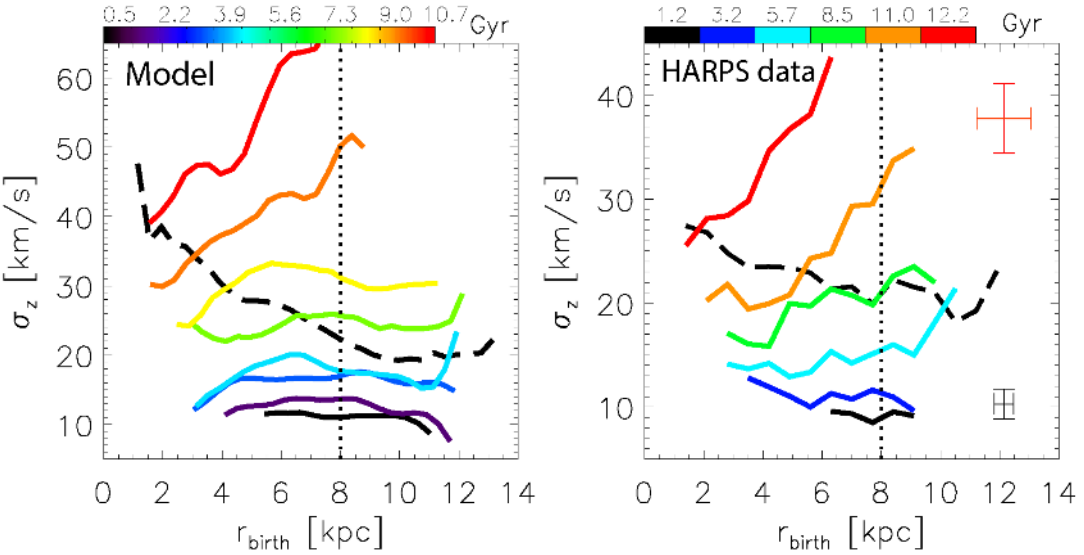

Figure 3. *Left:* Predicted vertical velocity dispersion, σ_z, as a function of mean birth radius, $r_{\rm birth}$, for the simulated solar neighbourhood stellar sample of the MCM13 model ($d < 0.5$ kpc). The black-dashed line shows the total population, suggesting that the hottest stars were born in the inner disc. Curves of different colours correspond to different age groups, with medians shown in the colour bar. For the oldest samples stars arrive at the solar vicinity much cooler than stars born *in situ*, due to the stronger effect of mergers on the outer disc and the decreasing probability of migration with increasing velocity dispersion. The large positive gradient found for old populations (red colours) turns negative for the youngest samples (blue and black), indicating a quiescent regime, where stars arriving from the inner disc heat slightly the local velocity distribution. *Right:* Same as top but resulting from HARPS-GTO dwarf stars (ages from Anders et al. 2018, birth radii from Minchev et al. 2018). This plot, resulting from observations, is remarkably similar to the model shown in the left-hand panel.

with small vertical actions (∼50 per cent increase in h_z in ∼4 disc scale lengths, see Minchev et al. 2012) is trumped by the impact of mergers and the effect of a cosmological context in general, in which case flaring results from locally born stars, i.e. non-migrators.

The $r_{\rm birth}$–σ_z relation was first presented in order to understand why the velocity dispersion increase with [Mg/Fe] in RAVE Giants and SEGUE G-dwarfs showed an abrupt drop at the highest [Mg/Fe] end (Minchev et al. 2014b). It was found that the observations could be explained by kinematically cool, old (thus low metallicity and high [Mg/Fe]) stars arriving to the solar neighbourhood from low Galactic radii (∼2–4 kpc). Indeed, larger changes in guiding radii (or angular momenta) are expected for the coldest subpopulations in a given age group (e.g. Vera-Ciro et al. 2014; Daniel & Wyse 2018).

Minchev et al. (2018) recently presented a largely model-independent method for estimating stellar birth radii based on age and [Fe/H] measurements, using AMBRE:HARPS (de Laverny et al. 2013; De Pascale et al. 2014; Hayden et al. 2017) and HARPS-GTO data (Adibekyan et al. 2012; Delgado Mena et al. 2017) combined with stellar ages computed with the STARHORSE code (Santiago et al. 2016; Queiroz et al. 2018), as described in Anders et al. (2018). This opened the possibility to test the $r_{\rm birth}$–σ_z relation in observations.

In the right-hand panel of Fig. 3 we show σ_z versus $r_{\rm birth}$ using HARPS-GTO data for six mono-age populations with median values indicated in the colour bar and a bin width Δage = 2 Gyr. Also shown are typical error bars, corresponding to two standard deviations of 1000 realizations in a bootstrapping calculation. As in Minchev et al. (2018), we imposed the following quality criteria on ages and abundances: δ[Mg/Fe] < 0.07 dex, δage/age < 0.25, or δage < 1 Gyr, resulting in a sample of 603 stars.

The trends in the data seen in the right-hand panel of Fig. 3 are remarkably similar to the MCM13 model expectation shown in the left-hand panel. The result that stars born at larger radii are kinematically hotter than those born at smaller radii is remarkable as it provides, for the first time, observational evidence that outward migration cools the local disc, rather than heating it, as has been proposed in the past (e.g. Schönrich & Binney 2009b; Roškar et al. 2013). This finding supports work showing that migration does not contribute to thick disc formation (e.g. Minchev et al. 2012; Minchev et al. 2014a; Vera-Ciro et al. 2014; Grand et al. 2016, see also discussion in Section 1), except in redistributing stars heated by external mechanisms (Quinn, Hernquist & Fullagar 1993; Villalobos & Helmi 2008) or turbulent gas clouds at high redshift (Bournaud, Elmegreen & Martig 2009; Forbes, Krumholz & Burkert 2012).

Possible erroneous interpretation in the absence of age measurements: The trend in the total population of the $r_{\rm birth}$–σ_z relation suggests that migration from the inner disc created the (local) Milky Way thick disc. This is not true if we slice by age, finding exactly the opposite trend for old and intermediate ages in both model and data. A better interpretation is that hotter stars arrived from outer radii due to disc flaring. The outward migrators kinematically cool the local disc. This is a strong YSP case.

2.4 The vertical metallicity gradient

The left-hand panel of Fig. 4 presents the mean [Fe/H] as a function of distance from the disc mid-plane, $|z|$, for stars in the simulated solar vicinity of the MCM13 model. A negative trend in the total population (black-dashed curve) is seen, consistent with SEGUE G-dwarf (Schlesinger et al. 2012; Rix & Bovy 2013) and RAVE data (Boeche et al. 2013b). In contrast, mono-age populations (colour-coded curves) show much weaker variations – some decrease for old and an increase for the two youngest age bins.

In the right-hand panel of Fig. 4 we show the effect of observational uncertainties, considering δage = 20 per cent and δ[Fe/H] = 0.1 dex and a Gaussian probability distributions (as in Fig. 2). Similarly to the radial [Mg/Fe] gradients (Section 2.2), the

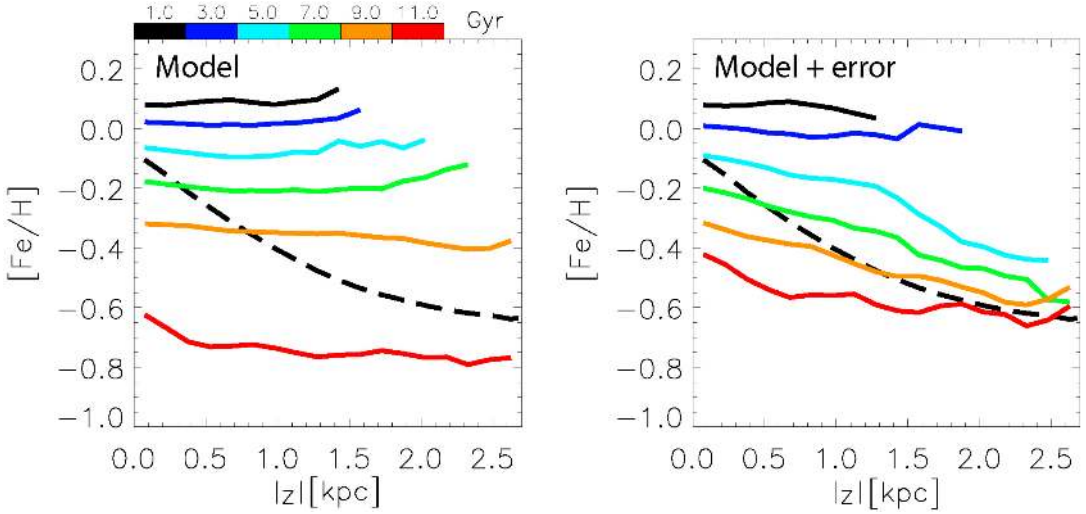

Figure 4. *Left:* Mean [Fe/H] as a function of distance from the disc mid-plane, $|z|$, for stars in the simulated solar vicinity of the MCM13 model. The strong negative trend seen in the total population (black-dashed curve) is consistent with SEGUE G-dwarf data (see fig. 14 in MCM13). In contrast, mono-age populations show much weaker variations. *Right:* Same as left, but including uncertainties of δage = 20 per cent and δ[Fe/H] = 0.1 dex, considering a Gaussian probability distribution. While the mean of the total sample is not much affected, a notable negative gradient has appeared for mono-age population older than ∼4 Gyr (cyan through red curves).

observational error does not seem to affect much the total sample, however, notable negative trends appear for mono-age populations older than ∼4 Gyr (cyan through red curves). A remarkable drop of ≈0.3 and ≈0.4 dex at high $|z|$ is seen for the green and cyan curves, respectively. Recently, Ciucă et al. (2018) found similar negative trends in the older populations in RAVE DR5 (Kunder et al. 2017) data, which, given the results of this section, may be caused by observational uncertainties.

The explanation for the negative vertical metallicity gradient in the total population is again related to the interplay among stellar ages, chemical enrichment, and stellar density, similarly to the radial [Mg/Fe] gradient inversion with $|z|$ (Section 2.2). Younger metal-rich mono-age groups dominate closer to the disc mid-plane, while older metal-poor ones are more abundant at larger $|z|$. The fraction of younger to older stars varies with Galactic radius. This both causes a variation in the vertical metallicity gradient (see e.g. Hayden et al. 2014 using APOGEE data) and an inversion of [Fe/H] and [Mg/Fe] gradients with increasing $|z|$ discussed in Section 2.2. Depending on the observational biases of different data set, different slopes in [Fe/H] (or [M/H]) with $|z|$ would be measured in the total population as summarized by Schlesinger et al. (2014), which may result in different interpretations.

Ciucă et al. (2018) reasoned that the mild negative gradient seen in the oldest population suggested mono-age discs must have been born flared, which after radial mixing would bring hotter, metal-poor stars from the outer disc to the solar neighbourhood, for a given mono-age population. The effect will be the same, however, even if mono-age discs were heated subsequently by external perturbations, i.e. the same scenario producing the nested flares in Fig. 1. The slightly positive gradient in the youngest populations then signifies the opposite effect – migrators from the inner disc weakly heat the local disc, while those arriving from the outer disc cool it, mostly cancelling the overall contribution (as seen in fig. 5 by Minchev et al. 2012 and in the negative slope of the youngest stars in Fig. 3). The latter effect results only in disc evolution lacking external perturbations.

This section presents a weak YSP case.

2.5 The metallicity–velocity relation (MVR)

A correlation between the observed total metallicity, [M/H], and rotational velocity in the Galactic plane, V_ϕ, of stars in the solar neighbourhood was already reported by Carney, Latham & Laird (1990). Haywood (2008) suggested that this MVR, can serve as a test for the amount of stars migrated into the solar neighbourhood. He showed that metal-poor solar neighbourhood stars at low [α/Fe] possess large angular momenta. This meant that they were visitors from the outer disc, currently near the Sun on their pericenters. Schönrich & Binney (2009a) argued that the presence of a negative correlation in the MVR for stars in the chemically defined thin disc (or more correctly, the low-[α/Fe] sequence) is an indication of migration. Using a compilation of data for stars in the solar neighbourhood, Navarro et al. (2011) found a flat MVR for low-[α/Fe] stars, concluding that migration in the solar neighbourhood was unimportant. Lee et al. (2011) found a negative relation for low-[α/Fe] stars, in contrast to Navarro et al. (2011). A positive slope in the MVR for high-[α/Fe] stars was described by Spagna et al. (2010), Lee et al. (2011), Kordopatis et al. (2011), and Adibekyan et al. (2013). More recently, Allende Prieto, Kawata & Cropper (2016) used the APOGEE-TGAS sample which utilizes the *Gaia* DR1 data (Gaia Collaboration 2016), confirmed the positive and negative slopes of the high- and low-[α/Fe] stellar populations in the local disc, respectively.

It will be shown below that the positive slope in the MVR of high-[α/Fe] stars spanning a wide range in age is just another strong YSP case resulting from the combination of MVRs of mono-age (or mono-[Mg/Fe][2]) populations with well-defined negative slopes.[3]

We use a subsample from RAVE DR5 (Kunder et al. 2017) with high-quality chemical and kinematical information. We selected a

[2][α/Fe] is a good proxy for age for samples confined to a small radial bin (Haywood et al. 2013; Bergemann et al. 2014), but not for data sets spanning large Galactic radii (Minchev et al. 2017).

[3]These results were first presented at the RAVE collaboration meeting in Ljubliana, Slovenia in 2014, using RAVE-DR4 (Kordopatis et al. 2013).

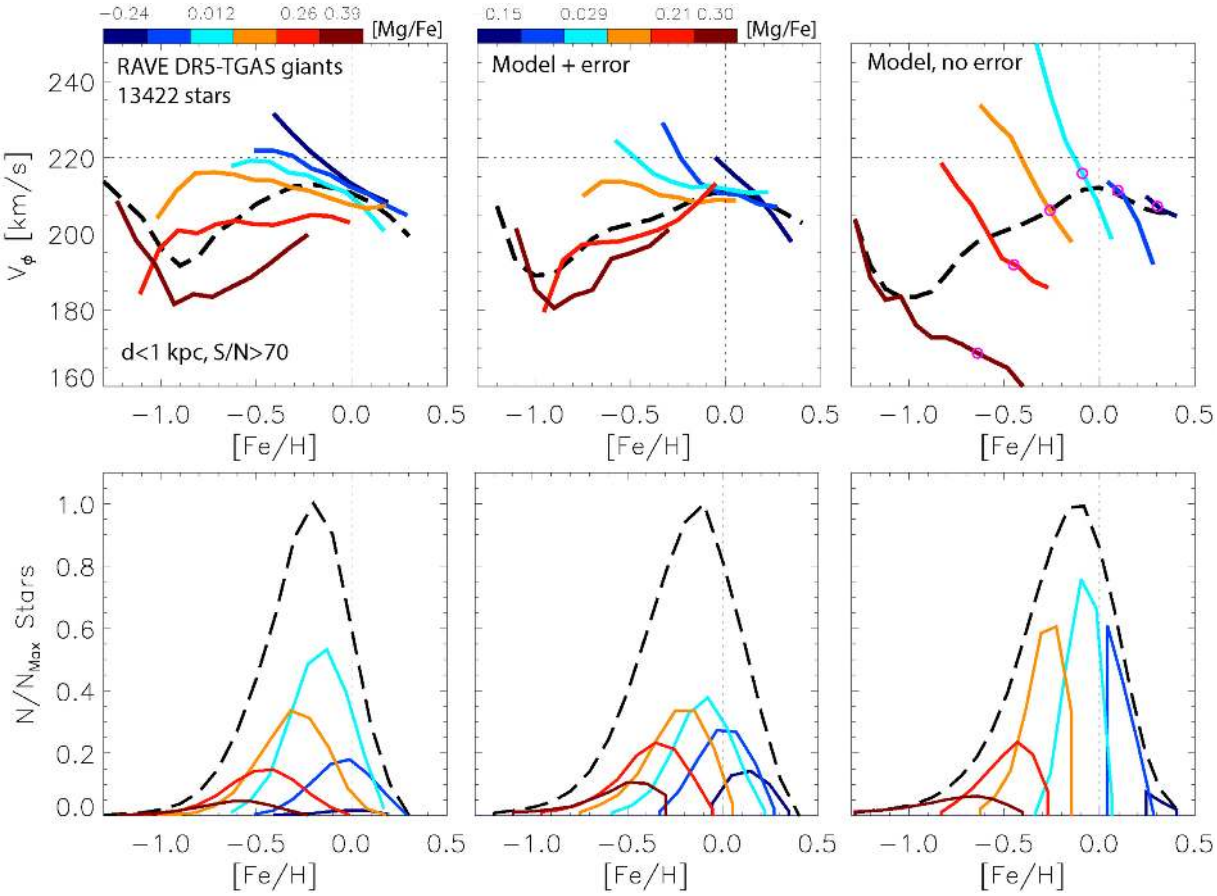

Figure 5. Comparison between the metallicity–rotational velocity relation (MVR) in RAVE-TGAS giants and the MCM13 model. *Left column:* RAVE-TGAS data with S/N > 70 and $d < 1$ kpc. The black-dashed curve in the top panel shows the rotational velocity as a function of metallicity for the total population, while colour-coded solid curves correspond to different median [Mg/Fe] bins, as indicated in the colour bar. The bottom panel shows the distribution of the total (black dashed) and [Mg/Fe] subpopulations. Stars with $V_\phi < 100$ km s^{-1} have been discarded in both data and model. *Middle column:* Same as left, but for the MCM13 model simulated neighbourhood, including velocity and abundance uncertainties (see text). For both the data and model an S-shape in the total population MVR is seen, as well as a remarkably good match among different subpopulations. *Right column:* Same as middle, but *excluding* uncertainties. It can be seen that for each [Mg/Fe] subpopulation the slope is negative. The MVR can become positive at [Fe/H] < 0 if (1) a sample spans a large range of [Mg/Fe] or (2) uncertainties are not sufficiently small.

sample of giant stars close to a random magnitude-limited sample as done in Minchev et al. (2014b). We excluded giants with log(g) < 0.5 to avoid possible effects due to the boundaries of the learning grid used for the automated parametrization, and considered stars in the temperature range 4000 < T_{eff} < 5500 K (thus avoiding horizontal branch stars; see Boeche et al. 2013a). Our final sample consists of stars with signal to noise S/N > 70 and a distance from the Sun $d < 1$ kpc.

The top left-hand panel of Fig. 5 presents the MVR for our RAVE-TGAS sample. The total population (black-dashed curve) shows a distinct S-shape rotated by ∼90°. A positive trend is found in the range −1 < [Fe/H] < 0, while a negative one can be seen at [Fe/H] > 0 and [Fe/H] < −1 dex. When we consider narrow bins in [Mg/Fe] (colour-coded curves), however, the MVRs of low-[Mg/Fe] groups show strong negative slopes, flattening and inverting as [Mg/Fe] increases. A V-shape is seen in the highest [Mg/Fe] bin.

The top middle panel of Fig. 5 is similar to the left one, but showing the MCM13 model, including uncertainties similar to what we expect for the data: δ[Mg/Fe] = 0.12 dex, δ[Fe/H] = 0.1 dex, and $\delta V_\phi = 2$ km s^{-1}. A remarkable match is seen to the RAVE-TGAS data for both the total population and the subsamples, notably, the gradual inversion from negative to positive with increasing [Mg/Fe] and the V-shape of the highest [Mg/Fe] bin.

Finally, in the top right-hand panel of Fig. 5 we present the model *excluding* the uncertainties, which shows a drastic change in shape for individual subpopulations. The MVR trend is always negative, slightly decreasing in steepness for the highest [Mg/Fe] bin. The slope measures $dV_\phi/d[Fe/H] = -127$ km s^{-1} dex^{-1} for stars with median [Mg/Fe] of 0.13 dex (cyan curve), flattening to $dV_\phi/d[Fe/H] = -50$ km s^{-1} dex^{-1} for the highest [Mg/Fe] bin with median value of 0.34 dex (maroon curve).

2.5.1 *Effect of observational uncertainties*

To understand the effect of observational uncertainties seen in the model, in the bottom row of Fig. 5 we show the [Fe/H] distributions of the total samples and the corresponding mono-[Mg/Fe] populations. In both data and model we see similar shifts in the peaks to lower [Fe/H] with increasing [Mg/Fe], as expected.

The small pink circles in the top right-hand panel of Fig. 5 indicate the approximate [Fe/H] values at which the density peaks for each mono-[Mg/Fe] population. We see that those act as pivot points

around which the MVR flattens when neighbouring bins are allowed to communicate, i.e. when uncertainties are convolved with the model (middle column). The only exception is the highest [Mg/Fe] bin, because it is significantly lower in density than the intermediate bins, and it lacks a clear peak.

In a range of [Fe/H] over which the densities of two adjacent mono-[Mg/Fe] populations are comparable, both MVRs are deflected similarly. For example, the two highest [Mg/Fe] bins in the range $-1.0 < \mathrm{[Fe/H]} < -0.5$ dex show similar densities, consequently, the maroon curve shifts upwards and red curve downwards as error is implemented. More generally, the MVRs of density tails are affected by the neighbouring bin's peaks in the corresponding direction: since the slope is always negative, metal-poor tails are shifted towards larger V_ϕ and vice versa. It should be noted that relatively similar densities in the observations are not to be taken at face value due to the selection function of a particular survey. For example, the RAVE survey is known to favour low-[α/Fe] stars (Boeche et al. 2013b), which is evident in the distributions of mono-[Mg/Fe] populations in the bottom left-hand panel of Fig. 5.

The negative MVR slope at $\mathrm{[Fe/H]} < -0.9$ dex is minimally affected, as these are the stars with the highest [Mg/Fe] values and the uncertainties are not large enough to allow other [Mg/Fe] bins to reach that end of the metallicity distribution. This distinct shape is present in both data and model, as well as seen in other works, although not acknowledged. For example, such an inversion can be found (1) in the bottom panel of fig. 7 by Lee et al. (2011), if one focused on the mean values, rather than the line fit, and (2) in the top panel of fig. 4 by Adibekyan et al. (2013) at $\mathrm{[Fe/H]} \approx -1.1$ dex.

Similarly to the effect of errors, using wider [Mg/Fe] ranges would result in overlapping between the neighbouring bins (all of which show clearly negative MVRs), approaching the black-dashed line in the right-hand panel of Fig. 5. If we split into the high- and low-[Mg/Fe] sequences, as often done in the literature, we would end up with a positive and a negative MVR, respectively.

In summary, the genuinely negative MVR can become positive if (1) a sample spans a large range of [Mg/Fe] or (2) uncertainties are not sufficiently small.

2.5.2 *What causes the invariantly negative MVR?*

The negative trends for all mono-[Mg/Fe] populations (right-hand panel of Fig. 5) are easy to understand. While a star at the solar radius, r_0, moving on a perfectly circular orbit rotates with the value of the circular velocity (here $V_0 = 220\ \mathrm{km\,s^{-1}}$), stars with guiding radii inside/outside r_0 will rotate slower/faster than that. Due to the exponential stellar surface density decrease with radius, however, more low-angular momentum stars are present in the solar neighbourhood (this is the asymmetric drift effect, see Binney & Tremaine 2008). Consequently, the mean tangential velocity, V_ϕ, of the total population is always below V_0. Since our sample is selected from a narrow radial range, the angular momentum is given by $L \approx V_\phi r_0$ and in terms of the stellar guiding radius, r_g, as $L \approx V_0 r_\mathrm{g}$. Therefore, $V_\phi \approx (V_0/r_0) r_\mathrm{g}$, i.e. V_ϕ can be thought of as a proxy for r_g. We will not find the youngest stars with guiding radii far away from r_0, as those will be on orbits too close to circular to be able to reach our sample. The total population shows a peak at solar [Fe/H] because these are the youngest stars near the solar radius, thus having the most circular orbits.

Splitting the model sample in narrow bins of [Mg/Fe] in the top right-hand panel of Fig. 5 showed that the MVR is always negative. This simply reflects the radial negative metallicity gradient of a mono-age, or in this case, a mono-[Mg/Fe] population. Since $V_\phi \sim r_\mathrm{g}$, for a given [Mg/Fe] bin, a negative metallicity gradient is reflected into a negative MVR gradient.

In summary, the positive MVR gradient in the high-[α/Fe] stars measured in a number of works (Carney et al. 1990; Rocha-Pinto et al. 2006; Spagna et al. 2010; Lee et al. 2011; Adibekyan et al. 2013; Allende Prieto et al. 2016) is simply the result of the asymmetric drift and the negative radial metallicity gradient of mono-age populations.

Possible erroneous interpretation in the absence of age measurements: The MVR slope of a mixed-age population tells us something about the Milky Way migration history. In fact the MVR slope is strongly dependent on the sample selection function and the age range of the sample. When dissected by age (or [α/Fe]), one finds that all groups have negative slopes closely related to the radial metallicity gradients of mono-age populations. This is a strong YSP case.

2.6 The age–metallicity relation (AMR)

The AMR in the solar vicinity is believed to give us a measure of how fast the iron content of stars has increased with cosmic time. It has been found somewhat confusing that the AMR is quite flat, or even reversed, for stars with age $\lesssim 10$ Gyr (e.g. Edvardsson et al. 1993; Lin et al. 2018; Vickers & Smith 2018). This has been taken sometimes as evidence that stars at the solar radius have not enriched in iron during that period, which is in disagreement with predictions from Milky Way chemical evolution models, or thought to result from the target selection of a given sample. While it is widely accepted now that radial migration is responsible for the scatter in [Fe/H] at the same age (Sellwood & Binney 2002; Roškar et al. 2008; Schönrich & Binney 2009a), it has only recently been suggested that stars migrating from different radii present well-defined AMRs but conspire to flatten the relation resulting from the total population (Minchev et al. 2018).

In Fig. 6 we plot the AMR for the local HARPS-GTO data, which was used in the right-hand panel of Fig. 3. The black-dashed curve shows the relation for the total sample. The colour-coded lines represent five mono-r_birth populations, as colour coded, of width 2 kpc and median values given in the colour bar. Birth radii were estimated from the stellar age and [Fe/H], as described by Minchev et al. (2018). This is not an unbiased result since the r_birth estimate is based on age and [Fe/H]. Nevertheless, this figure reflects how the functional form of the interstellar medium (ISM) metallicity evolution with radius and cosmic time (also estimated by Minchev et al. 2018) determines the shape of the mono-r_birth AMRs.

In Fig. 6 we measure a gradient of $0.02\ \mathrm{dex\,Gyr^{-1}}$ for age < 8.5 Gyr in the total population (black-dashed curve). However, the slope in the AMR of stars born in the range $6 < r_\mathrm{birth} < 8$ kpc (blue curve) is $0.04\ \mathrm{dex\,Gyr^{-1}}$ and is similar for the other r_birth bins. This figure provides an explanation for the flatness of the local AMR, when radial migration is not taken into account. The slope of the total population will depend crucially on selection effects of a particular survey and the quality cuts applied. In this particular case we find quite a steep one compared to other data sets. For example, from fig. 7 by Lin et al. (2018) we estimated $\sim 0.013\ \mathrm{dex\,Gyr^{-1}}$ for age < 10 Gyr, while Vickers & Smith (2018) find a double peak at ~ 3 and ~ 8 Gyr. While the shape indeed depends on the particular

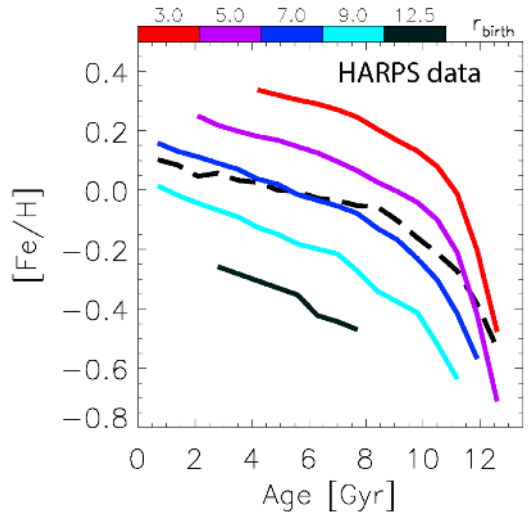

Figure 6. The local AMR using our local HARPS-GTO data. The black-dashed curve shows the mean resulting from the total population. The colour-coded lines represent five mono-$r_{\rm birth}$ populations, with median values indicated in the colour bar and a bin width $\Delta r_{\rm birth} = 2$ kpc. Well-defined AMRs exist for stars with common birth radii. The AMR is shifted downward for outer radii, which is a result of the negative ISM radial metallicity gradient. Note that there is no galactic model involved here.

target selection when the total sample is considered (Casagrande et al. 2016), such effects will be significantly reduced if we split in mono-$r_{\rm birth}$ populations, relying on the assumption that stars with a particular value of [Fe/H] and age were born at roughly the same radius. Therefore, we do not expect the AMRs of mono-$r_{\rm birth}$ populations to deviate much in other data sets from what we show here.

In summary, the chemical enrichment and the radial migration history are so closely entwined in the data that they can only be inferred together, as showed by Minchev et al. (2018) and Frankel et al. (2018).

Possible erroneous interpretation in the absence of $r_{\rm birth}$ measurements: The metallicity at the solar radius has not increased in the last 8–10 Gyr of evolution. This is not true if mono-$r_{\rm birth}$ populations are considered, showing well-defined slopes in the AMR of stars born at a given radial bin. This is a weak YSP case.

2.7 Lithium abundance variation with [Fe/H]

The left-hand panel of Fig. 7 shows the abundance of lithium, A(Li), versus [Fe/H] for our local HARPS data used in Sections 2.3 and 2.6. The difference here is that the lithium is estimated from the AMBRE:HARPS sample (as detailed by Guiglion et al. 2016a,b) and cross-matched with the ages derived from HARPS-GTO stellar parameters (Anders et al. 2018, using the STARHORSE code), resulting in 326 dwarf stars. The black-dashed curve shows the mean resulting from the total population and the colour-coded lines represent six mono-age populations, as in Fig. 3. While the relation from the total sample has a positive slope, mono-age populations show negative trends, except for the two youngest age bins. The variation of mean A(Li) with [Fe/H] for mono-age groups, to our knowledge, has not been shown before in observations, although it could be argued would result from the scatter plots shown in figs 5 and 7 by Ramírez et al. (2012) and should correspond to bins of common stellar mass, shown in several works (e.g. Nissen & Schuster 2012; Ramírez et al. 2012; Delgado Mena et al. 2015; Bensby & Lind 2018). The flattening we find in A(Li)-[Fe/H] for younger groups of stars is very similar to the flattening seen with increasing stellar mass in fig. 6 by Bensby & Lind (2018), as expected from the correspondence between stellar mass and age. This is, once again, a clear strong case of YSP, akin to the $r_{\rm birth}$–σ_z relation of Section 2.3 and the MVR of Section 2.5, which all show a dramatic reversal of the trend in the total local population.

Studies are usually concerned with the *upper envelope* (not shown here) in the A(Li)–[Fe/H] relation of dwarf stars, reasoning that this is closer to the original lithium abundance in the ISM, since depletion over the lifetime of stars takes place (e.g. Rebolo, Molaro & Beckman 1988; Romano et al. 1999; Lambert & Reddy 2004; Guiglion et al. 2016b; Cescutti & Molaro 2019; Guiglion et al. 2019). This *upper envelope*, which will be traced by extending the metal-poor tails of the mono-age populations in Fig. 7, is found to delineate stars from old to young when moving from low to high [Fe/H]. Further work is needed to understand if this is simply due to the inability of young stars to reach the solar neighbourhood on apo/pericenters as their orbits are close to circular, and the insufficient time for migration.

Splitting the same data by birth radius (right-hand panel of Fig. 7), we find that mono-$r_{\rm birth}$ populations exhibit mostly positive trends as in the total sample, but with significantly steeper slopes (a weak YSP case, similar to the AMR relation, see Fig. 6). This general shift from high to low [Fe/H] for inner to outer $r_{\rm birth}$ bins is also seen in the chemical evolution models by Prantzos et al. (2017) and Chiappini (2009) (shown by Guiglion et al. 2019). In those models, however, the variation of A(Li) with [Fe/H] is shown for the gas, i.e. at the time of stellar birth. Since lithium is depleted throughout stars' lifetime, we need to consider the *upper envelope* for each $r_{\rm birth}$ bin. The above models provide a good match to the trends seen in the right-hand panel of Fig. 7 assuming the scatter around the mean for all mono-$r_{\rm birth}$ groups is similar.

Delgado Mena et al. (2015) described a lithium decrease at [Fe/H] > 0 using HARPS data, a result subsequently confirmed by Guiglion et al. (2016b) using a larger data set. Recently Guiglion et al. (2019) reasoned that the observed decline in lithium abundance at supersolar metallicity can be explained as the effect radial migration, as stars in the metal-rich tail are expected to have been born roughly in the range $2 < r < 5$ kpc (see fig. 8 by Minchev et al. 2018). The right-hand panel of Fig. 7 shows this is indeed the case, using our estimated birth radii for HARPS data.

It should be noted that our results are not dependent on any chemical evolution or dynamical modelling, except for a simple assumption on keeping the mono-$r_{\rm birth}$ distributions physically meaningful (see for details Minchev et al. 2018). It is clear that, depending on the birth radius bin, the *upper envelope* will have different shape, notably, shifting to higher [Fe/H] and having a faster drop at the [Fe/H]-poor tail, as $r_{\rm birth}$ decreases. This opens up the possibility to study the production/destruction of lithium as a function of birth radius, that will provide stronger constraints on chemical evolution models.

Possible erroneous interpretation in the absence of age and $r_{\rm birth}$ measurements: In the total sample the *upper envelope* in the A(Li)–[Fe/H] plane will be biased towards older stars at the higher and (especially) lower [Fe/H] tails (see fig. 2 by Minchev et al. 2018), in which case the inferred ISM evolution of lithium will be incorrect. This problem can be resolved with the help of age and $r_{\rm birth}$ estimates. Dissecting the A(Li)–[Fe/H] relation by age presents a strong YSP case and that by $r_{\rm birth}$ – a weak one.

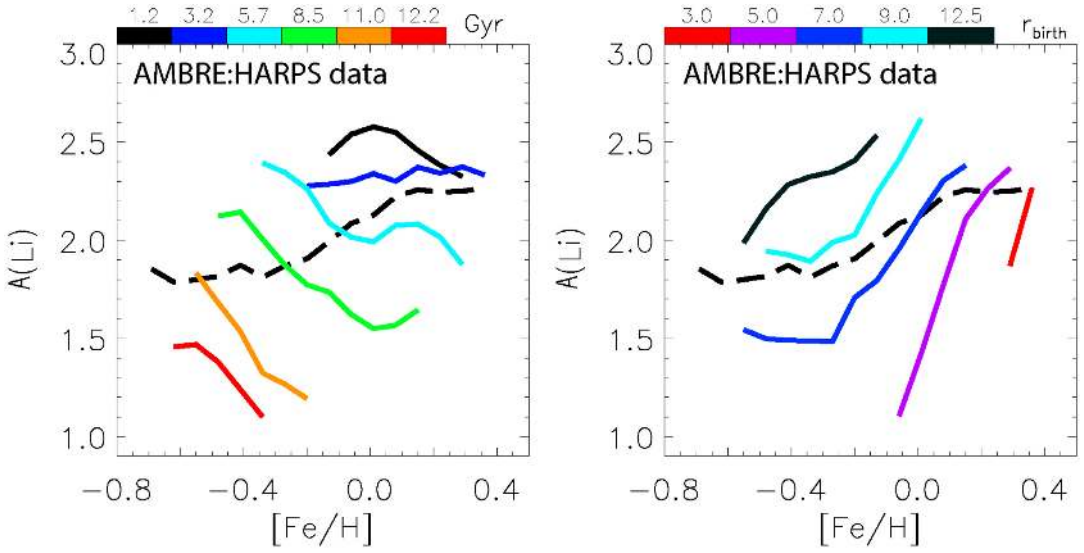

Figure 7. *Left:* Lithium abundance, A(Li), versus [Fe/H] for the local AMBRE:HARPS data (Guiglion et al. 2016b). The black-dashed curve shows the mean resulting from the total population. The colour-coded lines represent six mono-age populations, with median values indicated in the colour bar and a bin width 2.5 Gyr. Mono-age populations show typically negative trends while the total populations have a positive slope. *Right:* Same as left, but colour curves show mono-$r_{\rm birth}$ populations. As in the case of the AMR, steeper slopes are found for stars coming from similar radii, compared to the total sample. This figure presents a strong (left-hand panel) and a weak (right-hand panel) YSP cases.

3 CONCLUSIONS

We showed in this work that YSP is omnipresent in the field of Galactic Archaeology. Chemo-kinematical relations believed to constrain the Milky Way past can look completely different, depending on whether a subset of stars with common age or a mixture of ages is considered (and similarly for birth radius). In other words, a statistical relationship between two variables does not necessarily represent a cause-and-effect relationship. We can summarize our results as follows:

(i) We described seven YSP cases, which we classified as 'weak' if a relation was flattened or weakened, or 'strong' if it was completely reversed. Weak cases were found in the variation of galactic disc thickness with Galactocentric radius (Section 2.1), the vertical metallicity gradient (Section 2.4), the AMR (Section 2.6), and the A(Li)–[Fe/H] relation when split by birth radius (Section 2.7). Strong cases appeared in the inversion of radial abundance gradients with distance from the disc mid-plane (Section 2.2), the $r_{\rm birth}$–σ_z relation (Section 2.3), the MVR (Section 2.5), and the A(Li)–[Fe/H] relation when split by age (Section 2.7).

(ii) YSP is found in both global (Sections 2.1 and 2.2) and local (samples confined to the solar neighbourhood, Sections 2.3–2.7) chemo-kinematical trends.

(iii) The shape of the MVR of stars in the solar vicinity has been suggested as a discriminant for the amount of radial migration suffered by the local disc, where a flat (inverse) relation has been argued to signify the unimportance (importance) of migration. Using RAVE data and the MCM13 chemo-dynamical model, we showed that the slope in the MVR for any narrow [Mg/Fe] subpopulation (used as a proxy for age) is always negative. An inversion from a negative to a positive slope of high-[α/Fe] stars can be caused by the increasingly lower angular momentum of older, kinematically hotter stars at [Fe/H] $\lesssim -0.5$ dex with a slope measurement dependent on the selection function of a particular survey, as well as the observational uncertainties. The persistent MVR negative slope for any [α/Fe] bin simply reflects the Milky Way negative metallicity gradient of mono-age populations due to the inside-out disc formation, combined with the effect of the asymmetric drift.

(iv) Using estimates of stellar birth radii for AMBRE:HARPS data, we showed that the A(Li)–[Fe/H] relation in the solar neighbourhood looks different for different mono-$r_{\rm birth}$ populations. This can provide new constraints on chemical evolution models by requiring to match the *upper envelope* of the relation in different $r_{\rm birth}$ bins, as well as their combination at a given final radius expected from the effect of radial migration.

Depending on the observational biases of different data set, different trends can be measured in the total population of a given relation, thus resulting in various interpretations. Dissecting by age, however, decreases strongly such biases, as coeval stars will be affected by dynamical processes similarly, for a given radius. Moreover, one need not worry what age a particular stellar sample is biased towards, as long as the relation of interest does not involve the stellar density. For example, in the case of the vertical metallicity gradient, assuming stars with the same metallicity and age are born at the same radius (i.e. no significant intrinsic abundance scatter at the time of star formation, cf, Spitoni et al. 2018), the only effect we need to worry about is radial migration.

We demonstrated that splitting a stellar sample by age or birth radius is of utmost importance for recovering the Galaxy evolution. Given that our birth radius estimate depends crucially on precise ages measurements, we conclude that ages are urgently needed to make progress in the field of Galactic Archaeology. We expect to get those from the K2 (Howell et al. 2014), TESS (Ricker et al. 2015), and PLATO (Rauer et al. 2014) asteroseismic missions in the near future, as well as isochrone age determination from spectroscopic surveys, such as GALAH (De Silva et al. 2015), Gaia-ESO (Gilmore et al. 2012), APOGEE, WEAVE, SDSS-V's Milky Way Mapper (Kollmeier et al. 2017), and 4MOST, combined with Gaia astrometry.

Simpson's paradox undoubtedly exists in other Galactic Archaeology relations, as well as other areas of astrophysics. There may be other lurking variables in addition to age and birth radius we discussed here, which must be accounted for in statistical experiments, in order to avoid non-meaningful results. We hope this work brings awareness to different communities and helps advance our understanding of the Milky Way formation and evolution.

ACKNOWLEDGEMENTS

Many thaks to the second referee of the manuscript. IM acknowledges support by the Deutsche Forschungsgemeinschaft under the grant MI 2009/1-1. FA 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.

REFERENCES

Adibekyan V. Z. et al., 2013, A&A, 554, A44
Adibekyan V. Z., Sousa S. G., Santos N. C., Delgado Mena E., González Hernández J. I., Israelian G., Mayor M., Khachatryan G., 2012, A&A, 545, A32
Allende Prieto C., Kawata D., Cropper M., 2016, A&A, 596, A98
Anders F. et al., 2014, A&A, 564, A115
Anders F., Chiappini C., Santiago B. X., Matijevič G., Queiroz A. B., Steinmetz M., Guiglion G., 2018, A&A, 619, A125
Aumer M., White S. D. M., 2013, MNRAS, 428, 1055
Aumer M., White S. D. M., Naab T., Scannapieco C., 2013, MNRAS, 434, 3142
Bensby T., Lind K., 2018, A&A, 615, A151
Bergemann M. et al., 2014, A&A, 565, A89
Binney J., 2013, New Astron. Rev., 57, 29
Binney J., Tremaine S., 2008, Galactic Dynamics, 2ndedn. Princeton Univ. Press, Princeton, NJ
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bland-Hawthorn J., Krumholz M. R., Freeman K., 2010, ApJ, 713, 166
Boeche C. et al., 2013a, A&A, 553, A19
Boeche C. et al., 2013b, A&A, 559, A59
Bournaud F., Elmegreen B. G., Martig M., 2009, ApJ, 707, L1
Bovy J., Rix H.-W., Schlafly E. F., Nidever D. L., Holtzman J. A., Shetrone M., Beers T. C., 2016, ApJ, 823, 30
Brase C., Brase C., 2016, Understandable Statistics: Concepts and Methods, Cengage Learning, Boston, USA
Carney B. W., Latham D. W., Laird J. B., 1990, AJ, 99, 572
Carraro G., Vázquez R. A., Costa E., Ahumada J. A., Giorgi E. E., 2015, AJ, 149, 12
Casagrande L. et al., 2016, MNRAS, 455, 987
Cescutti G., Molaro P., 2019, MNRAS, 482, 4372
Cheng J. Y. et al., 2012, ApJ, 746, 149
Chiappini C., 2009, in Andersen J., Bland-Hawthorn J., Nordström B., eds, Proc. IAU Symp. 254, The Galaxy Disk in Cosmological Context. Kluwer, Dordrecht, p. 191
Ciucă I., Kawata D., Lin J., Casagrande L., Seabroke G., Cropper M., 2018, MNRAS, 475, 1203
Comerón S. et al., 2011, ApJ, 741, 28
Dalton G. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460P
Daniel K. J., Wyse R. F. G., 2018, MNRAS, 476, 1561
de Grijs R., 1998, MNRAS, 299, 595
de Jong R. S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460T
de Laverny P., Recio-Blanco A., Worley C. C., De Pascale M., Hill V., Bijaoui A., 2013, Messenger, 153, 18
De Pascale M., Worley C. C., de Laverny P., Recio-Blanco A., Hill V., Bijaoui A., 2014, A&A, 570, A68
De Silva G. M. et al., 2015, MNRAS, 449, 2604
Delgado Mena E. et al., 2015, A&A, 576, A69
Delgado Mena E., Tsantaki M., Adibekyan V. Z., Sousa S. G., Santos N. C., González Hernández J. I., Israelian G., 2017, A&A, 606, A94
Edvardsson B., Andersen J., Gustafsson B., Lambert D. L., Nissen P. E., Tomkin J., 1993, A&A, 275, 101
Feast M. W., Menzies J. W., Matsunaga N., Whitelock P. A., 2014, Nature, 509, 342
Forbes J., Krumholz M., Burkert A., 2012, ApJ, 754, 48
Frankel N., Rix H.-W., Ting Y.-S., Ness M., Hogg D. W., 2018, ApJ, 865, 96
Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487
Gaia Collaboration, 2016, A&A, 595, A2
Gilmore G. et al., 2012, Messenger, 147, 25
Gilmore G., Reid N., 1983, MNRAS, 202, 1025
Grand R. J. J., Springel V., Gómez F. A., Marinacci F., Pakmor R., Campbell D. J. R., Jenkins A., 2016, MNRAS, 459, 199
Guiglion G. et al., 2019, A&A, 623, A99
Guiglion G., de Laverny P., Recio-Blanco A., Worley C. C., de Pascale M., Masseron T., Prantzos N., Mikolaitis S., 2016a, VizieR Online Data Catalog. p. 359
Guiglion G., de Laverny P., Recio-Blanco A., Worley C. C., De Pascale M., Masseron T., Prantzos N., Mikolaitis Š., 2016b, A&A, 595, A18
Hayden M. R. et al., 2014, AJ, 147, 116
Hayden M. R., Recio-Blanco A., de Laverny P., Mikolaitis S., Worley C. C., 2017, A&A, 608, L1
Haywood M., 2008, MNRAS, 388, 1175
Haywood M., Di Matteo P., Lehnert M. D., Katz D., Gómez A., 2013, A&A, 560, A109
Hogg D. W. et al., 2016, ApJ, 833, 262
Howell S. B. et al., 2014, PASP, 126, 398
Kalberla P. M. W., Kerp J., Dedes L., Haud U., 2014, ApJ, 794, 90
Kawata D., Grand R. J. J., Gibson B. K., Casagrande L., Hunt J. A. S., Brook C. B., 2017, MNRAS, 464, 702
Kazantzidis S., Bullock J. S., Zentner A. R., Kravtsov A. V., Moustakas L. A., 2008, ApJ, 688, 254
Kollmeier J. A. et al., 2017, preprint (arXiv:1711.03234)
Kordopatis G. et al., 2011, A&A, 535, A107
Kordopatis G. et al., 2013, AJ, 146, 134
Kunder A. et al., 2017, AJ, 153, 75
Lambert D. L., Reddy B. E., 2004, MNRAS, 349, 757
Laporte C. F. P., Johnston K. V., Gómez F. A., Garavito-Camargo N., Besla G., 2018, MNRAS, 481, 286
Lee Y. S. et al., 2011, AJ, 141, 90
Li Y.-L. et al., 2013, Phys. Rev. A, 88, 015804
Lin J., Dotter A., Ting Y.-S., Asplund M., 2018, MNRAS, 477, 2966
Ma X., Hopkins P. F., Wetzel A. R., Kirby E. N., Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Quataert E., 2017, MNRAS, 467, 2430
Mackereth J. T. et al., 2017, MNRAS, 471, 3057
Majewski S. R. et al., 2017, AJ, 154, 94
Martig M., Bournaud F., Croton D. J., Dekel A., Teyssier R., 2012, ApJ, 756, 26
Martig M., Minchev I., Flynn C., 2014, MNRAS, 442, 2474
Martig M., Minchev I., Ness M., Fouesneau M., Rix H.-W., 2016, ApJ, 831, 139
Matteucci F., Francois P., 1989, MNRAS, 239, 885
Miglio A. et al., 2017, Astron. Nachr., 338, 644
Minchev I. et al., 2014b, ApJ, 781, L20
Minchev I. et al., 2018, MNRAS, 481, 1645
Minchev I., 2016, Astron. Nachr., 337, 703
Minchev I., Famaey B., Quillen A. C., Dehnen W., Martig M., Siebert A., 2012, A&A, 548, A127
Minchev I., Chiappini C., Martig M., 2013, A&A, 558, A9
Minchev I., Chiappini C., Martig M., 2014a, A&A, 572, A92
Minchev I., Martig M., Streich D., Scannapieco C., de Jong R. S., Steinmetz M., 2015, ApJ, 804, L9

Minchev I., Steinmetz M., Chiappini C., Martig M., Anders F., Matijevic G., de Jong R. S., 2017, ApJ, 834, 27
Miranda M. S. et al., 2016, A&A, 587, A10
Navarro J. F., Abadi M. G., Venn K. A., Freeman K. C., Anguiano B., 2011, MNRAS, 412, 1203
Nissen P. E., Schuster W. J., 2012, A&A, 543, A28
Prantzos N., de Laverny P., Guiglion G., Recio-Blanco A., Worley C. C., 2017, A&A, 606, A132
Queiroz A. B. A. et al., 2018, MNRAS, 476, 2556
Quinn P. J., Hernquist L., Fullagar D. P., 1993, ApJ, 403, 74
Rahimi A., Carrell K., Kawata D., 2014, Res. Astron. Astrophys., 14, 1406
Ramírez I., Fish J. R., Lambert D. L., Allende Prieto C., 2012, ApJ, 756, 46
Rauer H. et al., 2014, Exp. Astron., 38, 249
Rebolo R., Molaro P., Beckman J. E., 1988, A&A, 192, 192
Recio-Blanco A. et al., 2014, A&A, 567, A5
Ricker G. R. et al., 2015, J. Astron. Telesc. Instrum. Syst., 1, 014003
Rix H.-W., Bovy J., 2013, A&AR, 21, 61
Rocha-Pinto H. J., Rangel R. H. O., Porto de Mello G. F., Bragança G. A., Maciel W. J., 2006, A&A, 453, L9
Romano D., Matteucci F., Molaro P., Bonifacio P., 1999, A&A, 352, 117
Roškar R., Debattista V. P., Quinn T. R., Stinson G. S., Wadsley J., 2008, ApJ, 684, L79
Roškar R., Debattista V. P., Brooks A. M., Quinn T. R., Brook C. B., Governato F., Dalcanton J. J., Wadsley J., 2010, MNRAS, 408, 783
Roškar R., Debattista V. P., Loebman S. R., 2013, MNRAS, 433, 976
Santiago B. X. et al., 2016, A&A, 585, A42
Scannapieco C., White S. D. M., Springel V., Tissera P. B., 2009, MNRAS, 396, 696
Schlesinger K. J. et al., 2012, ApJ, 761, 160
Schlesinger K. J. et al., 2014, ApJ, 791, 112
Schönrich R., Binney J., 2009a, MNRAS, 396, 203
Schönrich R., Binney J., 2009b, MNRAS, 399, 1145
Schönrich R., McMillan P. J., 2017, MNRAS, 467, 1154
Sellwood J. A., Binney J. J., 2002, MNRAS, 336, 785
Selvitella A., 2017, J. Math. Phys., 58, 032101
Simpson E. H., 1951, J. R. Stat. Soc. B, 13, 238
Smith M. L., Goltz H. H., 2012, Health Promotion Pract., 13, 637
Spagna A., Lattanzi M. G., Re Fiorentin P., Smart R. L., 2010, A&A, 510, L4
Spitoni E., Cescutti G., Minchev I., Matteucci F., Silva Aguirre V., Martig M., Bono G., Chiappini C., 2018, preprint (arXiv:1811.11196)
Springel V. et al., 2008, MNRAS, 391, 1685
Thomas G. F. et al., 2019, MNRAS, 483, 3119
Thompson B., 2006, Foundations of behavioral statistics: An insight-based approach. Guilford, New York
van der Kruit P. C., Searle L., 1982, A&A, 110, 61
Vera-Ciro C., D'Onghia E., Navarro J., Abadi M., 2014, ApJ, 794, 173
Vickers J. J., Smith M. C., 2018, ApJ, 860, 91
Villalobos Á., Helmi A., 2008, MNRAS, 391, 1806
Wang C. et al., 2019, MNRAS, 482, 2189
Wang H.-F., Liu C., Xu Y., Wan J.-C., Deng L., 2018, MNRAS, 478, 3367
Wan J.-C., Liu C., Deng L.-C., 2017, Res. Astron. Astrophys., 17, 079
Xiang M. et al., 2018, ApJS, 237, 33
Yule G. U., 1902, Biometrika, 121, 134

This paper has been typeset from a TEX/LATEX file prepared by the author.