Edinburgh Research Explorer

Information content of BP/RP spectra in Gaia DR3

Information content of BP/RP spectra in Gaia DR3

Callum E. C. Witten,[1]* David S. Aguado,[2,3] Jason L. Sanders,[4] Vasily Belokurov,[1,5] N. Wyn Evans,[1]
Sergey E. Koposov,[6,1] Carlos Allende Prieto,[7,8] Francesca De Angeli [1] and Mike J. Irwin[1]

[1]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, United Kingdom*
[2]*Dipartimento di Fisica e Astronomia, Universit á degli Studi di Firenze, Via G. Sansone 1, I-50019 Sesto Fiorentino, Italy*
[3]*INAF/Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy*
[4]*University College London, Gower St., London WC1E 6BT, UK*
[5]*Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*
[6]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[7]*Instituto de Astrofísica de Canarias, Vía Láctea, 38205 La Laguna, Tenerife, Spain*
[8]*Universidad de La Laguna, Departamento de Astrofísica, 38206 La Laguna, Tenerife, Spain*

ABSTRACT

Gaia Data Release 3 has provided the astronomical community with the largest stellar spectroscopic survey to date (> 220 million sources). The low resolution (R~50) blue photometer (BP) and red photometer (RP) spectra will allow for the estimation of stellar atmospheric parameters such as effective temperature, surface gravity and metallicity. We create mock Gaia BP/RP spectra and use Fisher information matrices to probe the resolution limit of stellar parameter measurements using BP/RP spectra. The best-case scenario uncertainties that this analysis provides are then used to produce a mock-observed stellar population in order to evaluate the false positive rate (FPR) of identifying extremely metal-poor (EMP) stars. We conclude that the community will be able to confidently identify metal-poor stars at magnitudes brighter than $G = 16$ using BP/RP spectra. At fainter magnitudes true detections will start to be overwhelmed by false positives. When adopting the commonly-used $G < 14$ limit for metal-poor star searches, we find a FPR for the low-metallicity regimes [Fe/H] < -2, -2.5 and -3 of just 14%, 33% and 56% respectively, offering the potential for significant improvements on previous targeting campaigns. Additionally, we explore the chemical sensitivity obtainable directly from BP/RP spectra for Carbon and α-elements. We find an absolute Carbon abundance uncertainty of $\sigma_{A(C)} < 1$ dex for Carbon-enriched metal-poor (CEMP) stars, indicating the potential to identify a CEMP stellar population for follow-up confirmation with higher resolution spectroscopy. Finally, we find that large uncertainties in α-element abundance measurements using BP/RP spectra means that efficiently obtaining these abundances will be challenging.

Key words: stars: abundances – stars: fundamental parameters – stars: chemically peculiar

1 INTRODUCTION

The *Gaia* satellite was launched by the European Space Agency in 2013 and is expected to continue data-taking until 2026 (Gaia Collaboration et al. 2016). The main objective of *Gaia* is micro-arcsecond astrometry, but it also collects stellar spectra using the red photometer (RP), blue photometer (BP), and radial velocity spectrometer (RVS). The RVS spectra are only publicly available for stars brighter than $G_{RVS} = 12$, which is about 7.2 million stars in *Gaia* Data Release 2 (Katz et al. 2019).

However, on 13 June 2022, Data Release 3 (DR3) will provide for the first time low-resolution BP and RP spectra for hundreds of millions of stars. The combined BP/RP spectra cover the wavelength range 3300-10500 Å with a resolution between 13 and 85 (Carrasco et al. 2021). These spectra have resolutions that are too poor to allow us to measure individual spectral lines. Nonetheless, the combined BP/RP spectra may be used to estimate some stellar properties such as effective temperature $T_{\rm eff}$, surface gravity $\log g$, and metallicity

[M/H], and hence extract individual objects for further study. Given the huge size of the BP/RP dataset, this will likely contain representatives of extreme objects. We are therefore interested in devising algorithms to extract such objects.

Low metallicity stars in the Milky Way have been the subject of many recent searches (Venn et al. 2004; Beers & Christlieb 2005; Frebel & Norris 2015). They are primitive objects, probably the descendants of Population III stars – the very first, almost metal-free, objects that seeded chemical evolution in the Universe. Beers & Christlieb (2005) introduced the nomenclature of extremely metal poor (EMP) stars to describe those stars with iron abundances below 1/1000 of the Solar, or [Fe/H]<-3. Targeting strategies have been devised to hunt down bright candidates from public surveys like 2MASS near-infrared and WISE mid-infrared photometry for subsequent spectroscopic follow-up (e.g., Schlaufman & Casey 2014; Limberg et al. 2021). Dedicated surveys using narrow band imaging based around the metallicity sensitive near-infrared Ca H&K lines are underway (the PRISTINE survey, Starkenburg et al. 2017) with the goal of uncovering hundreds of candidates (e.g., Starkenburg et al. 2018; Venn et al. 2020). At such low metallicities, the fraction

of the carbon-enhanced metal-poor (CEMP) stars dramatically increases (Carollo et al. 2012; Salvadori et al. 2015; Yoon et al. 2016). Metal-poor CEMP stars have [C/Fe] > +1, so that compared to the Sun they are carbon enhanced at least ten times more than iron. CEMP stars are often further subdivided as to whether r-process or s-process elements are enhanced, while CEMP-no stars have no enhancement.

An earlier study (Allende Prieto 2016) hinted that Gaia's low-dispersion spectra are quite useful to constrain atmospheric parameters down to very low metallicity[1]. The aim of this paper is therefore to examine whether EMP and CEMP stars can be extracted in an efficient manner from the BP/RP spectral database. The paper is arranged as follows. Section 2 describes our methodology for creation of mock BP/RP spectra for Galactic populations. Then, in Section 3 we exploit the Fisher matrix to understand the degeneracies between extracted parameters in these low resolution spectra. Finally, we outline the resolution of Gaia BP/RP in stellar parameter measurements and additionally its ability to efficiently extract EMP candidates.

2   MOCK GAIA BP/RP SPECTRA

In order to probe the potential resolution of stellar parameter measurements using Gaia spectra, we must first produce a mock Gaia BP/RP spectra catalogue. The form of these spectra are dictated by the instrumentation of the Gaia spectrophotometers, such as their transmission curves, spectral resolution, original sampling, and observed line spread function (LSF). According to these parameters we built simulated Gaia BP/RP spectra from existing spectral libraries from the literature.

2.1   Stellar models

The selected stellar models came from two different sources, the PHOENIX (Husser et al. 2013) and Aguado et al. (2017) libraries. These two sets of synthetic stellar spectra cover different aspects needed for this study. The PHOENIX library uses 1D radiative hydrodynamic models (spherically-symetric), leading to a more realistic modelling of stellar atmospheres in low-gravity giant populations, similar to the well known MARCS library (Gustafsson et al. 2008)[2]. However, our choice is further supported by the fact the PHOENIX library covers a wide range in α-elements abundance. On the other hand, the library by Aguado et al. (2017) using plane-parallel symmetry (1D) will allow us to include synthetic models for carbon-enhanced metal-poor (CEMP) stars. This set of stellar models share the main properties of a broader synthetic library from Allende Prieto et al. (2018).

2.1.1   The PHOENIX library

This library, made publicly available by Husser et al. (2013), employs a grid of 1D stellar atmosphere models computed by PHOENIX code. The fact that these models use spherical symmetry make them

very suitable for evolved phases of stellar evolution, especially for cool giants. The synthesis is also performed by a particular mode available in the PHOENIX code. Every model is synthesised by assuming a unique microturbulence derived from convection velocities (Ludwig et al. 1999) within the atmospheric model and it is directly related with the large scale motion in the stellar atmosphere (macroturbulence). We consider this a priori approach an advantage of the PHOENIX library. Remarkably, the authors included some line profiles for neutral atoms of lighter elements in a non thermodynamical equilibrium (NLTE). A more compact version of the library, is available from the PHOENIX webpage[3], smoothed to the X-SHOOTER resolution and covering the following parameters range:

- 2300 K < T_{eff} < 8000 K, ΔT_{eff} = 100 K
- −0.5 < log g < 7.0, Δ log g = 0.5
- −4.0 < [Fe/H] < +1.0, Δ[Fe/H] = 0.5
- −0.4 < [α/Fe] < +1.0, Δ[α/Fe] = 0.2

We download the data and build a super 4D cube of models in a suitable manner for our purposes (see Sec. 2.1.3)

2.1.2   The Aguado et al. (2017) library

For this work we use a grid of synthetic models that is an extension in metallicity and temperature to those publicly available from Aguado et al. (2017). This library is based on a grid of Kurucz (1979) atmospheric models computed with ATLAS12 (Sbordone et al. 2007) assuming a 1D-LTE approach. The synthesis is performed with the radiative transfer code ASSET (Koesterke et al. 2008). These spectra have an assumed [α/Fe] = 0.4 - a typical value for metal-poor halo stars (see e.g., Sneden et al. 1991; Tomkin et al. 1992) and microturbulance of ξ = 2.0 km/s. This library covers a wider range of temperatures than those available in the PHOENIX or MARCS collections and focuses on FGK-type. Additionally, it includes very recent and accurate continuum opacities. Although for the parametric study we present in this work we mostly focus on the PHOENIX library, the chance to extend this library and incorporate carbon abundance in the synthesis of the spectra and also in their models of stellar atmospheres makes the the Aguado et al. (2017) library of great use for this analysis. The parameters covered by the grid are summarised as follows:

- 4500 K < T_{eff} < 7000 K, ΔT_{eff} = 250 K
- 1.0 < log g < 5.0, Δ log g = 0.5
- −4.0 < [Fe/H] < +1.0, Δ[Fe/H] = 0.5
- −1.0 < [C/Fe] < +3.0, Δ[C/Fe] = 1.0

This carbon-grid was designed to account for metal-poor stars with carbon enrichment and has been tested in recent works (see, e.g. Aguado et al. 2021a,b).

2.1.3   The FERRE interpolation

To prepare both libraries to be converted to Gaia BP/RP shape we firstly interpolate to produce a PHOENIX grid with finer steps in both log g and [Fe/H] of 0.25 dex and 0.1 dex respectively. Accordingly, we also interpolated the carbon-grid in [Fe/H] and [C/Fe] to steps of 0.1 and 0.25 dex respectively. This interpolation produces no additional information but allows us to produce smoother variations across changing stellar parameters. This interpolation was

---

[1] Allende Prieto (2016) indicates that photometric colors can be used, in some instances, to constrain stellar parameters, one could therefore use the publicly avalaible GaiaXPy code (https://gaia-dpci.github.io/GaiaXPy-website/) to simulate Gaia photometry in order to confirm this, however, it cannot yet be used to simulate Gaia BP/RP spectra.

[2] It is important to note the MARCS library does not extend the spherically-symetric models to dwarf populations while our selected PHOENIX library actually does.

[3] http://phoenix.astro.physik.uni-goettingen.de/

completed using FERRE[4] from Allende Prieto et al. (2006). We use a cubic Bézier interpolation, where FERRE operates over the nodes of the closest stellar parameter values and is able to interpolate simultaneously across the entire parameter space. This arrives at a global solution that is considerably smoother than a linear interpolation. We then restricted the wavelength coverage to between 300-1100 nm and smoothed the models to a resolving power of R = 1000.

2.2 BP/RP simulation

A large set of calibrators were used to produce passbands released alongside the Early Data Release 3 (EDR3) (see Riello et al. 2021). We apply these photon transmission curves, $T_{phot}(\lambda)$, to our raw synthetic spectra, $f(\lambda)$ (erg cm^{-2} s^{-1} Å^{-1}), to replicate the transmitted BP/RP flux in each pixel, f_{trans}(photons s^{-1}),

$$f_{trans} = \int T_{phot}(\lambda) \, \lambda \, f(\lambda) \, d\lambda, \tag{1}$$

where the integration range is defined by the resolution element of the spectrograph. This resolution element is both a function of wavelength and the object's position in the Gaia focal plane (Carrasco et al. 2021). The variation with position in the focal plane is however minimal and hence we fit a single curve to the data from Carrasco et al. (2021) and use this to estimate the width of the wavelength bins. We start at the lowest wavelengths for which BP and RP have non-zero photon transmission, 325 nm and 600 nm respectively, taking the corresponding bin width for that minimum wavelength, building the following bins from the previous bin's upper bound wavelength.

These steps replicate the bandwidths and wavelength resolutions of the Gaia BP and RP spectra. However we must also take into account the smearing effects of the LSF. The result of the LSF is that light of a given wavelength is registered in not just one wavelength bin/pixel but instead also in the neighbouring pixels. In applying this LSF to our transmitted spectrum, f_{trans}, we create the final mock BP/RP spectrum, f_λ.

Carrasco et al. (2021) quantifies the full-width half-maximum (FWHM) of the LSF to be ~ 1 BP/RP pixel, but this once again varies as a function of wavelength and focal plane position as well as the field of view and elapsed time throughout the mission. These many parameters are not possible to incorporate within our model and hence we make a tentative assumption, based on figures in Carrasco et al. (2021), to model the line spread function with a Gaussian with FWHM of 1.1 BP/RP pixels. We then convolve our transmitted flux by the LSF in order to produce a complete mock observed Gaia DR3 BP/RP stellar spectra. An example of the effect of each step in the process of creating the mock BP/RP spectra can be seen in Fig. 1. The final step taken to produce the BP/RP spectra is to normalize the flux by dividing through by the total *G*-band flux in the spectrum to ensure any change in total flux in the raw spectrum does not artificially create Fisher information.

We note that the employed grids of synthetic spectra do not model for dust reddening. This is something we neglect and therefore recommend anyone using our analysis to dictate their observing strategy to focus on stars least affected by reddening. We have also considered the effects of a component of radial velocity with respect to Gaia on the observed BP/RP spectra. This could potentially act to shift spectroscopic structures into different wavelength bins than those we may ordinarily expect for a given set of stellar parameters. We ultimately

find that for a star of high radial velocity (V_{Rad} = 500 km/s) we see no significant change in its spectra.

3 ANALYSIS OF BP/RP SPECTRA

3.1 Fisher information matrix

Fisher information is the expectation value of the Hessian of the likelihood and has been shown to be useful in gauging the amount of information that an observable variable carries about the model parameters (see Fisher 1935). Fisher information matrix is computed at the point of maximum likelihood and is the inverse of the estimate of the covariance matrix (see e.g. Tegmark et al. 1997). In our case the observable is the BP/RP spectrum and the parameters of interest are the numbers controlling the properties of the stellar atmosphere (T_{eff}, log(g) and [Fe/H]). We define our position in the stellar parameter space as the vector $\vec{x} = [T_{eff}, \log(g), [Fe/H]]$. We can use Fisher information matrices to indicate the information available at the position $\vec{x}$ as a function of wavelength, or alternatively, by summing the Fisher information matrices across the wavelength range we can obtain the total Fisher information for the spectrum. For example, assuming that the likelihood is Gaussian and unimodal, following Bonaca & Hogg (2018), the total Fisher information matrix at position $\vec{x}$ is defined as:

$$\mathbf{F}(\vec{x}) = \mathbf{C}(\vec{x})^{-1} = \left[\left(\frac{\partial \vec{f}}{\partial \vec{x}} \right)^T \cdot \mathbf{C}_f^{-1} \cdot \left(\frac{\partial \vec{f}}{\partial \vec{x}} \right) \right] + \mathbf{V}^{-1}$$

$$= \sum_\lambda [\mathbf{F}(\vec{x}, \lambda)] + \mathbf{V}^{-1} = \sum_\lambda \left[\left(\frac{\partial \vec{f_\lambda}}{\partial \vec{x}} \right)^T \mathbf{C}_{f_\lambda}^{-1} \left(\frac{\partial \vec{f_\lambda}}{\partial \vec{x}} \right) \right] + \mathbf{V}^{-1}, \tag{2}$$

where $\mathbf{C}$ is the covariance matrix of the parameters we are interested in (T_{eff}, log(g) and [Fe/H]) at a fixed position in the stellar parameter space, $\vec{x}$. $\vec{f}$ denotes the vector flux across the wavelength space, $\mathbf{C}_f$ is the covariance matrix of the flux across the wavelength space, and finally, $\mathbf{V}$ is the covariance matrix of any priors on the measurement of the stellar parameters T_{eff}, log(g) and [Fe/H]. In the second line of equation (2) we express the dot product in summation notation, indicating how we can extract the Fisher information at a given wavelength, $\mathbf{F}(\vec{x}, \lambda)$.

We calculate the differential of flux with respect to the position in the stellar parameter space, $\vec{x}$, by varying each stellar parameter by the minimum grid step and observing the change in flux. In order to obtain the diagonal matrix, $\mathbf{C}_f$, we estimate the uncertainty in flux across the wavelength range by taking the publicly available[5] ESA/Gaia/DPAC signal-to-noise ratio as a function of wavelength for Gaia DR3 mean spectra. This provides us with the uncertainty in flux as a function of wavelength and G-band magnitude. We have additionally considered the impact of varying temperature on the signal-to-noise ratio (SNR), however the impact is not significant. The maximum SNR difference between temperatures is ~ 20% and thus we see that magnitude is the most important parameter in defining the SNR as a function of wavelength. Equation (2) returns the maximum information available in each stellar parameter at a given position, $\vec{x}$, in the stellar parameter space (T_{eff}, log$g$, [Fe/H]).

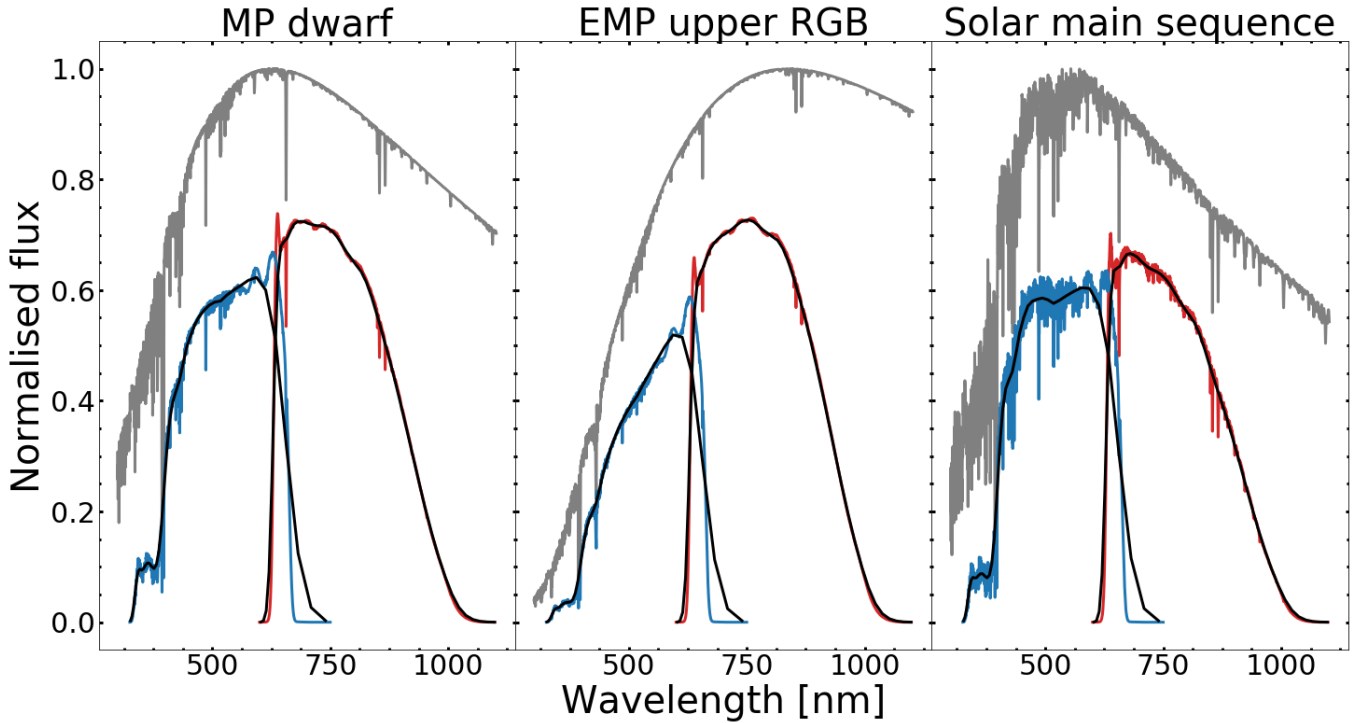[4] FERRE is available from http://github.com/callendeprieto/ferre

Figure 1. The normalised photon count (counts s^{-1} nm^{-1}) as a function of wavelength, showing the different steps taken in the production of the mock BP/RP spectra. The grey line shows the initial model spectra from Allende Prieto et al. (2018). The blue and red lines reveal the effect of applying the respective transmission curves of the BP and RP spectra. The black line shows the final mock observed BP/RP spectra after also reducing the resolution and applying the LSF. These are for three model stars, (left) a metal-poor dwarf (T_{eff} = 5500 K, log g = 4.4, [Fe/H] = -2), (middle) an extremely-metal-poor red-giant branch star (T_{eff} = 4500 K, log g = 1, [Fe/H] = -3) and (right) a typical Solar main sequence star (T_{eff} = 6000 K, log g = 4.5, [Fe/H] = 0). These spectra are in units of counts s^{-1} nm^{-1} for consistency, however for the final analysis of the BP/RP spectra we work in units of counts s^{-1}.

$\mathbf{F}(\vec{x}, \lambda)$, in Equation (2), allows us to analyse the Fisher information in each stellar parameters at the position $\vec{x}$ as a function of wavelength as shown in Figure 2. We can see clear peaks in metallicity information in Figure 2 that correlate to large or multiple absorption lines in Figure 1 such as the Calcium triplet and Balmer series. We do see in some panels a notably higher information at the overlap between BP and RP spectra. This is due to a combination of factors: the largest signal-to-noise ratio is observed for many stellar types at $\sim$ 630 nm, with a significant decrease at larger wavelengths, we additionally see the largest transmission across BP and RP spectra at $\sim$ 630 nm in RP spectra. These all combine to amplify any information in this region. While the information available in surface gravity varies as a function of wavelength for different stellar types, the temperature information is largely confined to the continuum of the spectrum.

Although we are able to assess the Fisher information as a function of wavelength, it is most useful to assess the total information that we can extract from BP/RP spectra about a given star's stellar parameters. We can see from equation (2) that the total Fisher information matrix is the sum of the Fisher matrices across the wavelength range and this is the best-case co-variance matrix for our stellar parameters at a given position in the stellar parameter space that is achievable given our model BP/RP grids. In summary, our application of the Fisher equation provides uncertainties such as $\sigma(T_{\text{eff}})$, $\sigma(\log(g))$ and $\sigma([\text{Fe/H}])$.

In addition, given the co-variance matrix, we can also calculate the correlation between parameters, $\rho_{x,y} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y}$, in order to assess how external constraints on stellar parameters can help to reduce the uncertainties determined using our BP/RP spectra.

3.2 Effective temperature and surface gravity priors

The methods used in Section 3.1 give us both the best-case uncertainty on each stellar parameter across the parameter space and also the correlation between these parameters. We can first use these uncertainties to assess the feasibility of stellar parameter measurements, but we can additionally use the correlations between parameters to allow us to identify in what parameter space priors become most advantageous.

Although BP/RP spectra will prove invaluable for the determination of T_{eff} and $\log(g)$, we note that Gaia's precise broadband photometry and parallaxes will allow us to obtain additional information on their probability distribution (Anders et al. 2022). These prior uncertainties will allow us to add extra information into equation (2), reducing the resultant uncertainty in metallicity.

Anders et al. (2022) found using Gaia EDR3 they were able to constrain T_{eff} to $\sim$ 140 K and $\log(g)$ to $\sim$ 0.2 dex. The ability to constrain these parameters without analysing BP/RP spectra, not only allows us to study false positive rates for specific effective temperatures and surface gravities as these can be independently identified, we can also use these to reduce the metallicity uncertainties and hence false positive rates.

4 FISHER MATRIX RESULTS

4.1 Parameter uncertainties

In this subsection we analyse the results of our Fisher analysis while neglecting any priors that may be available. This allows us to un-
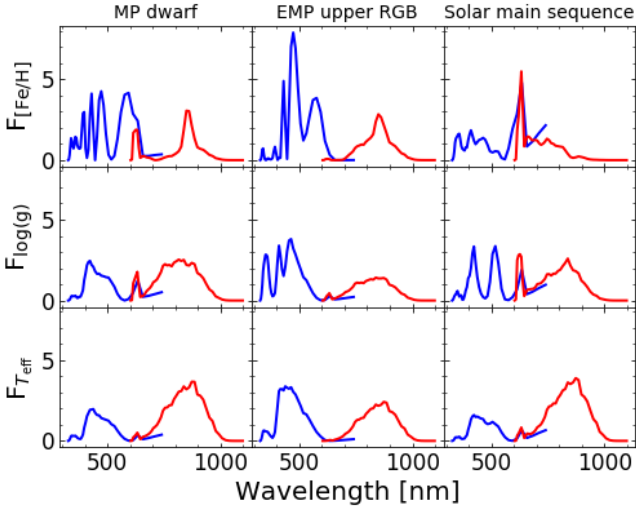
Figure 2. The Fisher information available at each wavelength of the BP and RP spectra normalised by the mean Fisher information value. The top row indicates metallicity information, the middle row shows surface gravity information and the bottom row shows effective temperature information. The blue line indicates the BP spectra information, while the red line shows the RP spectra information. These are shown for the same three example stars as shown and described in Figure 1 at a fixed magnitude $G = 16$
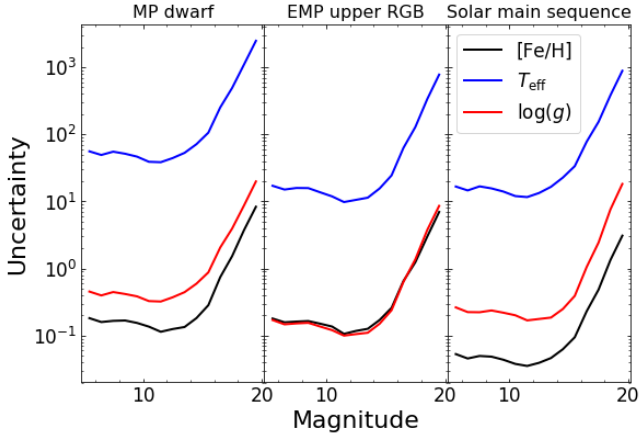
.

Figure 3. Parameter uncertainty as a function of G-band magnitude for the three example stars in Figure 1 and Figure 2. The blue, red and black lines represent $\sigma(T_{\rm eff}\,[{\rm K}])$, $\sigma(\log(g)\,[{\rm dex}])$ and $\sigma([{\rm Fe/H}]\,[{\rm dex}])$ respectively.

derstand the information that is implicit to BP/RP spectra. We firstly assess the change in parameter uncertainty as a function of magnitude in Figure 3 for our example stellar types. Figure 3 shows stellar parameter uncertainties of the same order in the range of magnitudes $5 \lesssim G \lesssim 16$. We see a slight increase in uncertainty at magnitudes brighter than $G = 11$, due to a gradual increase in the signal-to-noise ratio at the brightest magnitudes. For $G > 16$ we see an exponential growth in all parameter uncertainties. Given this pivot point, for the following analysis we use a magnitude of $G = 16$ as it gives us the upper-bound of stable parameter uncertainties before they start to exponentially grow with magnitude.

Figure 4 shows the best-case uncertainties we find on each stellar parameter for a range of input stellar parameters. We see little change in effective temperature and metallicity uncertainties across
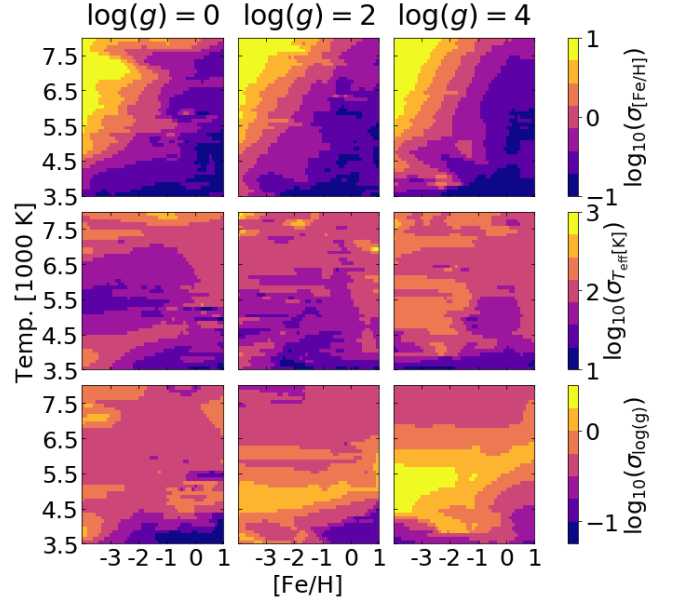

Figure 4. A grid of plots indicating the uncertainty of each parameter for different stellar parameters at magnitude $G = 16$. The x and y axes show increasing metallicity and effective temperature respectively, while each column shows a different surface gravity. Each row shares a common color-bar, the upper row's color-bar indicates the metallicity uncertainty, the middle row color-bar indicates effective temperature uncertainty and the lower row shows surface gravity uncertainty.

changing surface gravity. We clearly see that metallicity uncertainty is a function of both metallicity and effective temperature. Increasing temperature acts to decrease the relative depth of absorption lines, which provide information on the metallicity of a star, thus increasing the metallicity uncertainty. Decreasing metallicity has the same effect as increasing temperature, and hence we see a gradient in metallicity uncertainty across the effective temperature and metallicity parameter space, as seen in the top row of grids in Figure 4.

We can see that although there are mild structures and trends in effective temperature uncertainty across the parameter space, the clear result is that using Gaia BP/RP spectra we have an effective temperature resolution of order 100 K.

The clear result for surface gravity measurements is an increasing uncertainty for increasing surface gravity. This however appears to occur most significantly in a effective temperature range of 4500 K $< T_{\rm eff} < 6500$ K. This creates a change from a resolution of ~ 0.3 dex for $\log(g) = 0$, up to a resolution of ~ 3 dex for $\log(g) = 4$. Outside of this region of significant variation with surface gravity, we see a resolution of order 0.3 dex for $T_{\rm eff} > 6500$ K, down to a best resolution of 0.1 dex for some regions of the parameter space with $T_{\rm eff} < 4500$ K. We can therefore clearly see that for many intrinsic stellar parameters we shall be able to discriminate between dwarfs and giants from surface gravity measurements using Gaia BP/RP spectra.

We can see from Figure 4 that for low temperature stars (< 4500 K) and for stars with metallicity [Fe/H] > −1 we are able to constrain metallicity uncertainties to of order $\sigma_{[{\rm Fe/H}]} \sim 0.1$ dex. This uncertainty rapidly grows as we move to higher temperatures and to metallicities below [Fe/H] < −1, with uncertainties of order 1 dex, reaching a maximum uncertainty larger than 10 dex as we reach temperatures of 7500 K.
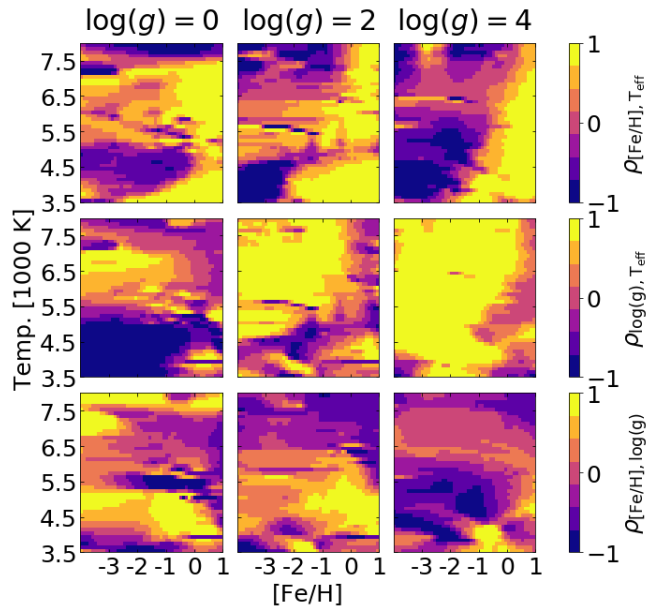
Figure 5. A grid of plots indicating the correlations between each of the stellar parameters we investigate. The x and y axes show increasing metallicity and effective temperature respectively, while each column shows a different surface gravity. The upper row has a common color-bar indicating the correlation between metallicity and effective temperature, the middle line color-bar indicates the correlation between surface gravity and effective temperature and the lower row shows surface the correlation between metallicity and surface gravity.

4.2 Parameter correlations

The uncertainties shown in Figure 4 could potentially be reduced by including temperature and surface gravity priors, evaluated using Gaia's precise photometry and parallaxes, in equation (2). We must first however assess whether there are correlations between parameters that would allow for the exploitation of such priors. This is done using methods discussed in section 3.1. Figure 5 shows the results of this analysis, indicating the correlations between all free stellar parameters in our BP/RP mock spectra grid.

We can clearly see in Figure 5 that metallicity is largely correlated with both temperature and surface gravity. We expect that in regions of Figure 4 where temperature uncertainty is greater than 140 K, and where Figure 5 shows temperature and metallicity are strongly correlated, that the metallicity uncertainty will be further constrained with the use of a temperature prior. This is similarly true for surface gravity priors.

We can see from Figure 4 that we have effective temperature resolution $\sigma_{T_{\mathrm{eff}}} \sim 100$ K across most of the stellar parameter space. There are however small regions, for instance at $\log(g) = 4$, for metallicities $[\mathrm{Fe/H}] < -2$ and temperatures 4500 K $< T_{\mathrm{eff}} < 6000$ K where we see temperature uncertainties greater than the 140 K prior. Within this region we see a strong anti-correlation between temperature and metallicity uncertainties, and therefore we can expect a small reduction in metallicity in this region when we account for temperature priors. We expect no contribution from a temperature prior in the remainder of the stellar parameter space.

In addition, we can see from Figure 4 that the uncertainty in surface gravity is larger than the prior of 0.2 dex across almost the entire parameter space. Figure 5 indicates largely strong correlations between surface gravity and metallicity across the parameter space. We therefore expect any improvement in metallicity uncertainty acquired due

to the inclusion of priors to be driven by surface gravity priors. We see the lowest surface gravity uncertainties around $T_{\mathrm{eff}} \sim 3500$ K and we therefore expect minimal metallicity uncertainty improvement with a surface gravity prior at this temperature. We additionally see at $\log(g) = 2$ and 4, for temperatures above 6000 K respectively, there is minimal correlation between metallicity and surface gravity. Further to this, when regions move from positive to negative correlation, we see no correlation at these boundaries. At all of these regions of low or no correlation between metallicity and surface gravity, we expect to see minimal or no improvement in metallicity uncertainty when we include a surface gravity prior.

Due to the surprising nature of some of these changes in correlations between parameters from positive to negative, we have additionally used a basic minimisation of the reduced chi-squared statistic model to assess the correlation between parameters. We simulate observed spectra by adding random noise to our model spectra based on the SNR data used in section 3.1 and by iterating over all model spectra to minimise the reduced chi-squared statistic we find the best-fit stellar parameters. We repeat this multiple times to find the spread in best-fit stellar parameters and hence estimate the uncertainty and correlation between parameters. We ultimately find correlations and uncertainties consistent with those seen in Figure 5 and 4 using Fisher information matrices.

4.3 Metallicity uncertainty with priors

Figure 6 shows the metallicity uncertainty after we include temperature and surface gravity priors and additionally the fractional improvement in the metallicity uncertainty provided by including these priors. Figure 6 confirms the hypotheses made in Section 4.2 – we see a notable reduction in metallicity uncertainty that mostly traces strong correlations in the surface gravity correlation plots in Figure 5. We see little fractional improvement in the upper regions of the $\log(g) = 2$ and 4 and additionally, along the boundaries of regions of positive and negative correlations, where we have minimal correlation between metallicity and surface gravity.

We can see in the lower row of Figure 6 that the inclusion of surface gravity and effective temperature priors can improve metallicity uncertainties by up to 60%. We now find that across all surface gravities, for the half of the plot that has either low effective temperature, high metallicity or both has a metallicity uncertainty of less than 0.3 dex.

It is clear from these plots that for large regions of the stellar parameter space, we are able to reduce the metallicity uncertainties to low enough values as to confidently constrain the metallicity of a star using BP/RP spectra.

5 IDENTIFYING EMP STARS WITH GAIA BP/RP

5.1 Mock observed population

In order to assess the effects these uncertainties may have on the observed metallicity distribution one may observe using Gaia BP/RP, we must start wish some initial distribution of stellar parameters. We chose to use Gaia Object Generator (GOG) to produce our intrinsic distribution (Antiche et al. 2014). GOG is a simulation tool based on the Gaia Universe model to produce a catalogue of the potential observable stellar population (with magnitudes brighter than $G = 20$). The assumed metallicity distribution function of the model is composed of individual metallicity distributions for each stellar population - the disc, halo and bar. Each population is represented
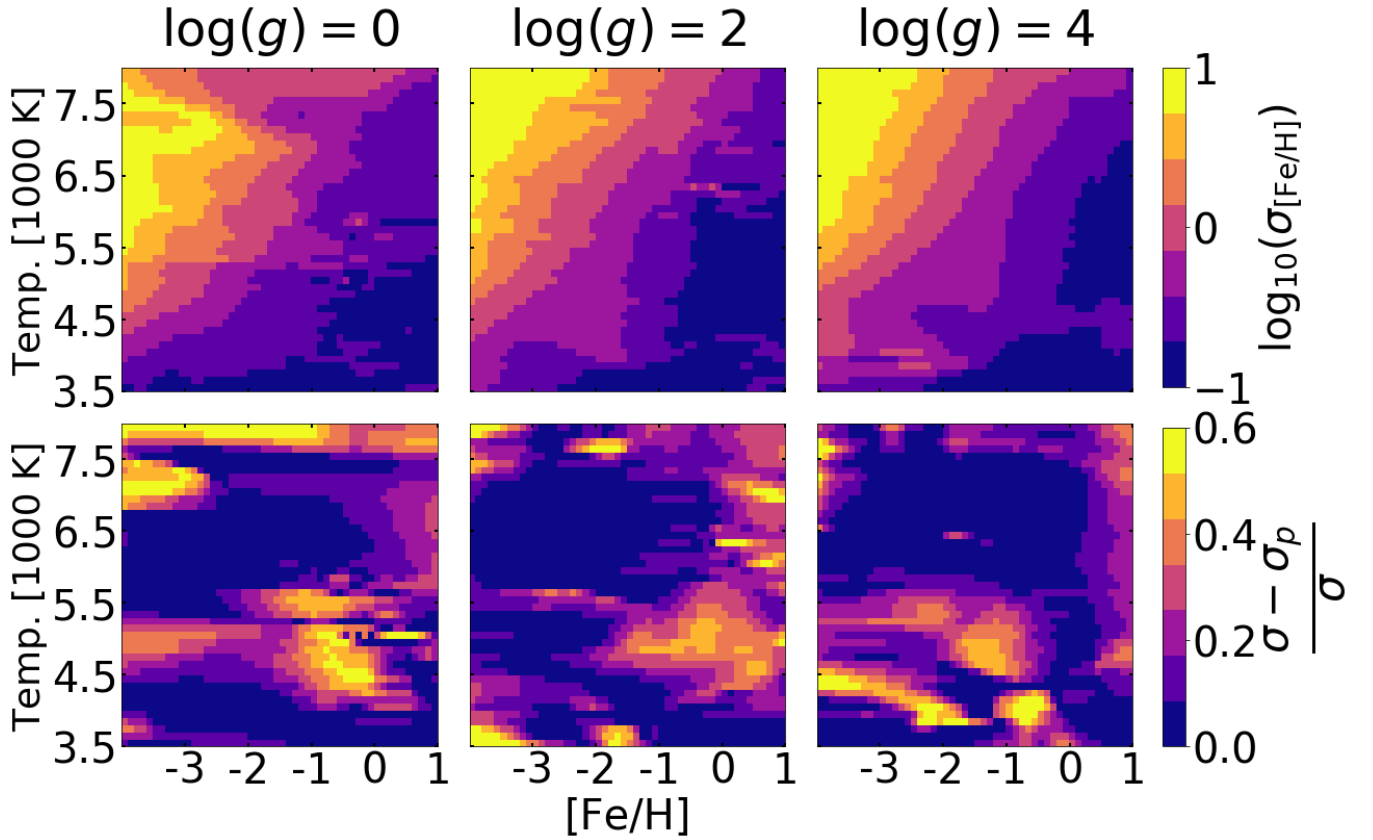
Figure 6. The top row shows metallicity uncertainty plots, for a star of magnitude $G = 16$, when we include effective temperature (140 K) and surface gravity (0.2 dex) priors. The bottom row indicates the fractional improvement from the original metallicity uncertainty, σ, when we include priors to produce σ_P. The x and y axes show increasing metallicity and effective temperature respectively, while each column shows a different surface gravity. Each row shares a common color-bar indicating the metallicity uncertainty and fractional improvement in metallicity uncertainty for the upper and lower rows respectively.

by a normal distribution centred on a mean value and truncated at 5 sigma, as such the minimum metallicity in the mock population is $[Fe/H] = -4$ [6].

Some of the data included within this catalogue are the stellar parameters that we are interested in: magnitude, effective temperature, surface gravity and metallicity. This therefore gives us an ideal intrinsic stellar population for which we are able to estimate the best-case uncertainties of individual stars.

We take this intrinsic population and focus primarily on their metallicites. We focus on three "low-metallicity" regimes, $[Fe/H] < -2, -2.5$ and -3 (0.5%, 0.07% and 0.004% of the total GOG population respectively). We take this intrinsic metallicity population and randomly re-draw each metallicity from a Gaussian distribution centred on its intrinsic metallicity with a standard deviation equal to the uncertainty derived from Section 3.1. Doing so produces a mock observed metallicity distribution. We use this to test the numbers of "high metallicity" stars that are shifted into the low-metallicity regime due to the uncertainty in their measurement using Gaia BP/RP spectra. This finally allows us to assess the false positive rates (FPR) for detections of stars with low metallicities for the three low-metallicity regimes. We produce this analysis using the uncertainties from Sec-

tion 4.3 including the assumed effective temperature and surface gravity priors. Ignoring these priors has a minimal impact on our results.

We note that our assumption that the metallicity distribution is a Gaussian centred on the intrinsic metallicity becomes poor for large uncertainties. Through the use of our reduced chi-squared minimisation method (discussed in section 4.2) we analysed the biases in the predicted metallicity distribution. We find that for Fisher information uncertainties greater than roughly 1 dex, the assumption of a Gaussian distribution breaks down. The distribution is instead biased towards higher metallicities, therefore, stars with non-Gaussian distributions are more likely to be scattered away from "low-metallicity" regions, and hence are less likely to pollute our low-metallicity detections. We therefore conclude that our inability to incorporate these distributions into our predictions will act to increase our FPR, meaning the results from this analysis may be slightly over-estimating the FPR.

5.2 Simulating EMP detections

Although GOG gives us detailed information on a sample of 523 million simulated Gaia observed sources, due to data processing constraints, we choose to reduce to a sub-sample of 215 million stars with simulated effective temperatures, surface gravities and metallicities.

Following the analysis described in Section 5.1 we first find the

---

[6] The values of the mean and sigma for each metallicity distribution can be found at https://gea.esac.esa.int/archive/documentation/GEDR3/Data_processing/chap_simulated/sec_cu2UM/ssec_cu2starsgal.html

Table 1. Table indicating the false-positive rate of EMP star detections for a range of low metallicity regimes and at varying stellar magnitudes. The true-positive rate can be found by taking (1− FPR) reported in the table.

Mag.	False Positive Rate		
	[Fe/H] < -2	[Fe/H] < -2.5	[Fe/H] < -3
$20 \geq G > 19$	0.97	1.00	1.00
$19 \geq G > 18$	0.95	0.99	1.00
$18 \geq G > 17$	0.87	0.96	1.00
$17 \geq G > 16$	0.66	0.88	0.98
$16 \geq G > 15$	0.31	0.61	0.90
$15 \geq G > 14$	0.20	0.42	0.73
$14 \geq G > 13$	0.14	0.35	0.69

FPR for a range of magnitude bins in Table 1. Given Figure 3 indicates that metallicity uncertainty exponentially grows for magnitudes $G > 16$, we would expect the FPR to have a similar trend, which can be observed in Table 1 where we see the largest growth in FPR at $G \sim 16$. These results indicate that for stars of magnitude $G > 16$, for the metallicity regimes considered, many intrinsically metal-poor stars are lost within crowds of false positive detections. However, we note that for stars of magnitude brighter than $G = 16$, for the metallicity regime [Fe/H] < -2, at least 3 in every 4 low-metallicity detections are true positives. This rate is at least 3 in every 5 for the metallicity regime [Fe/H] < -2.5 at magnitude $G < 15$, and roughly 1 in every 2 for [Fe/H] < -3 at magnitudes $G < 14$. We would expect the FPR for [Fe/H] < -3 to further improve up to $G = 11$, however, small number statistics mean we cannot confidently comment at such bright magnitudes.

In order to best understand the driving mechanisms behind false positive detections, we analysed the stellar parameter distribution for "low-metallicity" stars. Figure 7a and 7b show the FPR within a set of bins for stars with observed metallicity [Fe/H] < -2. They additionally indicate the distribution of false and true positive detections across each dimension of the parameter space in the form of histograms and the FPR is overlaid onto these plots. These allow us to best understand the driving processes behind false-positive detections of metal-poor stars using Gaia BP/RP spectra.

Figure 7a shows the population of stars with magnitude brighter than $G = 16$. It clearly indicates that the source of true positive detections fall mostly along the isochrones where we would expect to find low-metallicity stars, while at cooler temperatures away from these isochrones, where we know MP stars are rarer, we find a near 100% FPR. Further analysis of the metallicity distribution finds that the source of false positive detections at such bright magnitudes are in fact stars that have metallicities near to our "low-metallicity" boundary. We find bins that are largely false-positive detections in Figure 7a have a median intrinsic metallicity of [Fe/H] $\lesssim -1.5$.

Therefore our harsh metallicity cut-off makes our FPR appear more pessimistic than reality, as a false positive detection for a [Fe/H] < -2 star has a high probability of still being a "low-metallicity" star, just not at the extreme of [Fe/H] < -2. This is emphasised in Table 1, where we see large increases in FPR between the three "low-metallicity" regimes at bright magnitudes, indicating that when we observe these large increases, the increase in FPR is driven by stars with metallicities between such metallicity regimes. This can therefore further act to support a [Fe/H] < -3 detection using Gaia BP/RP spectra, as it appears even if any such detection is a false positive, there is a reasonably good change the intrinsic metallicity will still fall between -2 < [Fe/H] < -3.

Figure 7a also indicates that there is an increase in the FPR at tem-

peratures between 3500 K < $T_{\rm eff}$ < 4000 K and additionally a peak at $T_{\rm eff}$ = 5550 K. These features are repeatedly observed at a variety of magnitudes, and although the small number of stars in this population means that any such local peak in FPR doesn't significantly effect the global FPR, we advise that any "low-metallicity" detections within these temperature ranges be treated with caution. We additionally see a mild trend, outside of the offending temperature range, of increasing FPR with increasing temperature. This is expected given the increasing metallicity uncertainty with temperature discussed in Section 4.1, however the minimal overall increase within the temperature ranges present in the data shown in Figure 7a means there is no need to place any selection criteria on low-metallicity stellar detections based on this trend. The FPR appears to increase with increasing surface gravity, however any such increase is minimal enough that we still have a relatively low FPR across our surface gravity range and hence readers do not need to target specific surface gravity ranges in their "low-metallicity" identification strategy. Our results ultimately find that if exterior measurements on effective temperature and surface gravity allow a star to be confined to one of the bins in Figure 7a, there are regions that allow for a near 100% success-rate for metal-poor detections. Constraining stars to such regions, and consequently finding stars in regions with large FPRs, will allow the community to hand-pick from their population of metal-poor detections, to create a population of high-confidence metal-poor detections for follow-up high-resolution observations.

What is clear from this analysis is that [Fe/H] < -2 detections for magnitudes brighter than $G = 16$ are not polluted by stars of high metallicity that have exceedingly large uncertainties. This situation becomes the driver for false positive detections at dimmer magnitudes, $G > 16$. The exponential growth of uncertainties at larger magnitudes means that we collect false positive detections from a larger range of stellar parameters, due to their large uncertainties, as seen in Figure 7b. We investigate whether we have stellar parameters for which we can accurately identify metal-poor stars ([Fe/H] < -2) when considering the magnitude range dimmer than $G = 16$ in Figure 7b.

We initially apply a cut to the hottest stars and white dwarfs, as modelling of their isochrones are notoriously challenging and hence intrinsic metallicity estimates are spurious. All such stars had a 100% FPR, and therefore are not of interest. We encourage anyone using Gaia BP/RP spectra to identify metal-poor stars to adopt a similar strategy as these stellar parameters are not observed for magnitudes brighter than $G = 16$ in GOG, and are exclusively false-positive detections for magnitudes $G > 16$. We can see that when observing magnitude $G > 16$ stars, their large uncertainties mean that we have are heavily polluted with false positive detections across the entire temperature and surface gravity range considered. The FPR is ∼ 100% for the vast majority of stellar parameters, however, along the isochrones where we ordinarily expect to find metal-poor stars, the FPR can fall to values around 60%. This may allow for metal-poor detections of reasonable confidence, especially when either temperature or surface gravity has a more precise prior than we consider or additionally for magnitudes that lie close to $G = 16$.

5.3 Comparison with observational metal-poor surveys

There are several photometric surveys aiming to successfully identify metal-poor stars as tracers of formation and evolution of the Galactic halo. One of the pioneering examples was the Ca II H&K objective-prism survey by Beers et al. (1985) that reported more than a hundred bright very metal-poor (VMP) candidates ([Fe/H]<-2). This represented a leap in Galactic archaeology and showed a

Figure 7. A Kiel diagram indicating the stellar parameter distribution of GOG stars with simulated BP/RP observed metallicities [Fe/H] < -2. The population of "low-metallicity" stars is binned into effective temperature and surface gravity bins and the FPR within each bin is indicated by the colorbar. Histograms on a logarithmic scale, located on the x and y axes, indicate the distribution of true positive (hatched blue) and false positive (filled grey) detections. The FPR as a function of each stellar parameter is additionally over-plot in red.
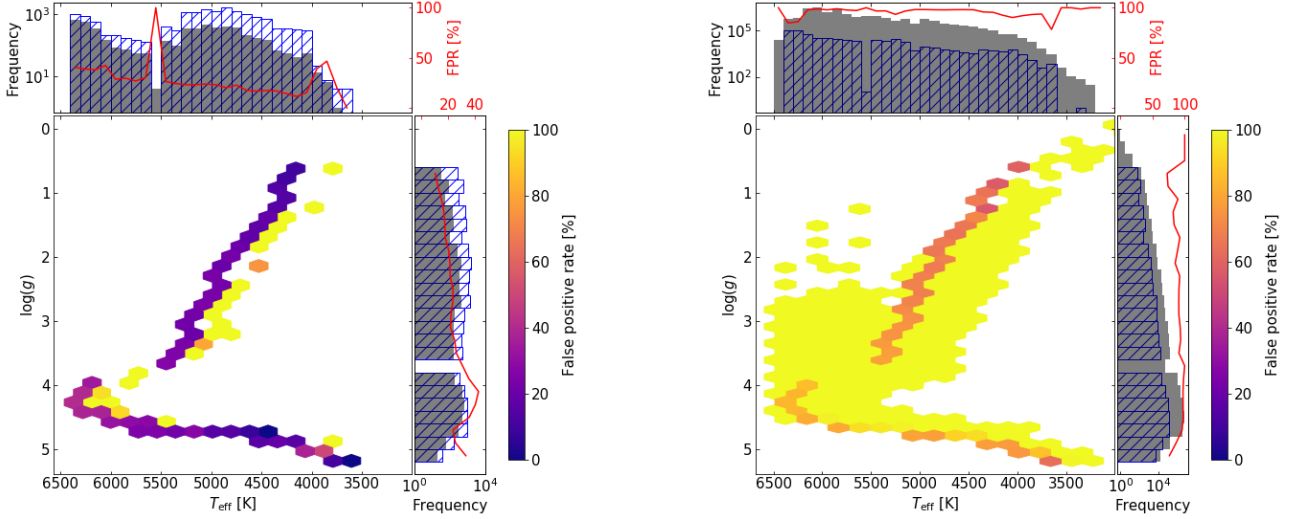
(a) This Kiel diagram only includes stars with intrinsic magnitudes brighter than $G = 16$.

(b) This Kiel diagram only includes stars with intrinsic magnitudes dimmer than $G = 16$.

promising methodology to identify stars from this rare class. Then, a very important contribution was provided by the Hamburg/ESO survey (Wisotzki et al. 2000; Christlieb et al. 2001) also based on the Ca II resonance lines around 395 nm. Thanks to this objective-prism survey, Frebel et al. (2006) identified 1718 metal-poor candidates with B_{mag} < 13, and confirmed 174, 98, 23 to be more metal-poor than [Fe/H]=-2.0, -2.5, and -3.0 respectively. Some of the most metal-poor stars ever known were primarily identified by the Hamburg/ESO survey (Christlieb et al. 2002; Frebel et al. 2005). A few years later, a dedicated survey, the Skymapper (Keller et al. 2007), used narrow-band filter photometry to identify metal-poor candidates. With this technique a larger number of fainter objects could be targeted and the amount of metal-poor candidates increased drastically. More recently, thanks to the Skymapper narrow-band filter, Da Costa et al. (2019) reported from more than 2600 metal-poor candidates followed-up with medium-resolution spectroscopy, 93%, 41% and 18% success ratios for stars with [Fe/H]<-2.0, -2.75, and -3.0 respectively. Some other relevant searches were completed as part of the Sloan Digital Sky Survey (SDSS, York et al. 2000) such as the TOPoS survey (Caffau et al. 2013) or follow-up on EMP stars by Aguado et al. (2016). Also by means of the large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST, Deng et al. 2012) like the 10,000 VMP stars reported by Li et al. (2018). Moreover, in the northern hemisphere, the Pristine survey (Starkenburg et al. 2017) used even narrower photometry around the H&K area and provided high-quality photometric metallicities for around 2 million metal-poor stars. Aguado et al. (2019) reported after a medium-resolution follow-up of more than a thousand metal-poor candidates, success ratios of 88%, 56%, and 23% for stars with [Fe/H]<-2.0, -2.5, and -3.0 respectively. Additionally, a multi-filter photometric survey, J-Plus (Whitten et al. 2019), is also identifying metal-poor stars based on their colours. In a recent paper by Galarza et al. (2022), the authors observed 177 metal-poor candidates selected with J-Plus photometry and confirmed 64% of them to be more metal-poor than [Fe/H]=-2.5. Finally, the Best and the Brightest survey (Schlaufman

& Casey 2014) using a combination of public photometric data in different bands identified efficiently bright (V_{mag} < 14) metal-poor candidates (32%, and 4% for [Fe/H]<-2.0, and -3.0 respectively). According to Limberg et al. (2021) when this multi-band selection is combined with Gaia EDR3 astrometry (Gaia Collaboration et al. 2021) the efficiency grows significantly (76%, 28%, and 4% for [Fe/H]≤ -2.0, ≤ -2.5 and ≤ -3.0, respectively).

We find that when we apply a similar magnitude cut of $G < 14$ to the GOG data, we have a success rate of metal-poor identification of 86%, 67% and 44% for [Fe/H]≤ -2.0, ≤ -2.5 and ≤ -3.0, respectively. These values are a significant improvement on the results from Limberg et al. (2021) that uses only Gaia's astrometry, indicating that the inclusion of BP/RP spectra in such surveys has the potential to drastically improve the success-rate of metal-poor star detections. We additionally see that the success rate for the lowest metallicity regime ([Fe/H] < -3) offers a significant improvement on all other narrow-band photometric surveys indicating that Gaia BP/RP spectra have the potential of more accurately identifying populations of extremely-metal-poor stars than any previous metal-poor survey.

6 CHEMISTRY WITH GAIA BP/RP

6.1 Carbon Sensitivity

Carbon enrichment in halo stars is a crucial issue in Galactic archaeology. The fraction of metal-poor stars that show [C/Fe]>0.7 increases dramatically with decreasing metallicity (see e.g., Cohen et al. 2005; Placco et al. 2014; Bonifacio et al. 2015; Yoon et al. 2018; Arentsen et al. 2021, and references therein). There are a number of explanations invoking both extrinsic (see e.g., Herwig 2005) and intrinsic (see e.g., Umeda & Nomoto 2003; Ishigaki et al. 2014) processes for the creation of carbon-enriched metal-poor (CEMP) stars. It is crucial to be able to pre-select metal-poor stars that are carbon rich.

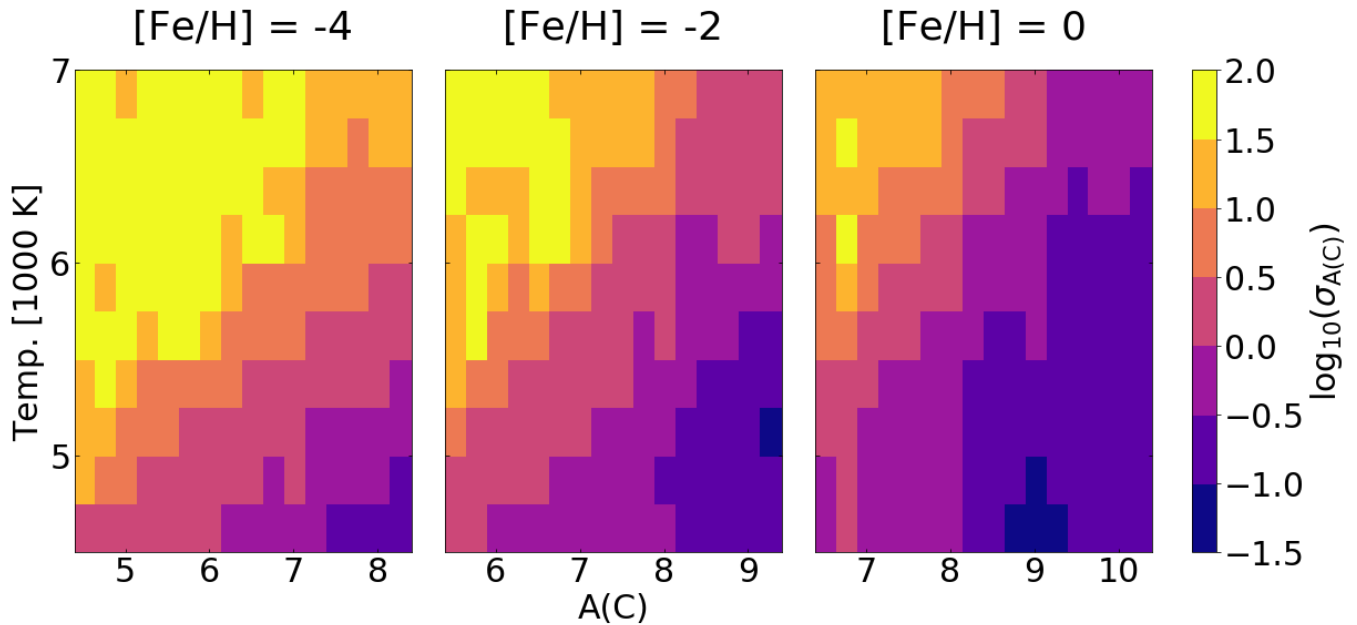We investigate the potential for Gaia BP/RP spectra to resolve

Figure 8. A row of Carbon abundance uncertainty plots, for a star of magnitude $G = 16$. The x and y axes show increasing absolute Carbon abundance and effective temperature respectively, while each column shows a different metallicity. The plots share a common color-bar indicating the absolute Carbon abundance uncertainty.

absolute carbon abundance, A(C), using the synthetic spectra grid, discussed in Section 2.1.2. Although these grids supply spectra that only extend from $355 - 1149$ nm, thus missing the bluest region of the BP spectrum, due to low transmission in this region, we expect this to have a minimal impact on the accuracy of our mock BP/RP spectra analysis. The synthetic spectra library includes Carbon as a free parameter and therefore following the same steps discussed in Sections 2.2 and 3 we can evaluate the potential resolution of Carbon abundance measurements using BP/RP spectra as well as the resolution on all other stellar parameters discussed in this paper.

We initially fix the Carbon abundance at [C/Fe] = 0, in order to compare our results from the Carbon grid to the results from the grid used throughout the rest of this work (discussed in Section 2.1.1). For the stellar parameters for which the two grids overlap, we find consistent results regarding their uncertainties and correlation, confirming our results are invariant with the spectral library we use.

Following this we focus on the absolute Carbon abundance, A(C) = [C/Fe] + [Fe/H] + 8.39[7], where the uncertainty in absolute Carbon abundance can be found by combining $\sigma_{[C/Fe]}$ and $\sigma_{[Fe/H]}$ in quadrature. Figure 8 shows the results of the absolute Carbon abundance analysis indicating the absolute Carbon abundance uncertainty as a function of metallicity and effective temperature on the x and y axes respectively, with each column showing a different metallicity. The analysis was completed for a star with fixed surface gravity $\log(g) = 2$ as we found that absolute Carbon abundance uncertainty was not a function of the stellar parameter surface gravity. The results are for a fixed magnitude of $G = 16$.

Figure 8 shows an increasing absolute Carbon uncertainty as a function of increasing temperature and decreasing absolute Carbon abundance. Therefore, in order to place a reasonable constraint on Carbon abundance of $\sigma_{A(C)} < 0.5$ the observed star must have an

intrinsic absolute Carbon abundance A(C) > 8 and must have a temperature below 5000 K, 6000 K and 6500 K for metallicities [Fe/H] = -4, -2 and 0 respectively.

We use a criterion of [C/Fe] > 0.7 for CEMP stars. Given this criterion, we can see that this equates to A(C) > 6.1, 7.1 and 8.1 for the three metallicity regimes shown in Figure 8. If we place a temperature constraint of $T_{\rm eff} < 6000$ K for CEMP stars, we can see that for metallicities [Fe/H] = -2 and 0, we have an absolute Carbon abundance uncertainty $\sigma_{A(C)} < 1$. We obtain a similar uncertainty for metallicity [Fe/H] = -4 when we constrain effective temperatures to below 5000 K. These uncertainties are large enough that we may not be able to confidently identify individual CEMP stars with such stellar parameters, especially when we consider false positive detections. However, these reasonable uncertainties for CEMP stars mean that the false negative rate will be low for CEMP stars. Therefore Gaia BP/RP spectra will be crucial for identifying candidate CEMP stars for high-resolution spectroscopic follow-up. Although the population identified as CEMP stars using Gaia BP/RP spectra will have some level of false positive pollution, it will contain the vast majority of true CEMP stars and therefore this population will be essential for follow-up studies of such stars.

While discovering Carbon-enhanced, hyper metal-poor stars ([Fe/H] ~ -5) would be of huge interest and represent the frontier of the earliest chemical enrichment of the universe, using BP/RP from Gaia DR3 we unfortunately find it to be very unlikely. Our analysis extends down to [Fe/H] = -4.0, but by extrapolating our results in Figure 8 to lower metallicities, we expect that only very cool, extremely carbon-rich (A(C) ~ 9 or 10) stars could possibly show Carbon features in hyper metal-poor stars. Nevertheless, we note that EMP stars ([Fe/H] ~ -2, -3) with high carbon abundances play an important role in galactic archaeology and in that case Gaia BP/RP will help to identify some of them.

---

[7] The solar carbon abundance considered here is the one reported by Asplund et al. (2005), A(C)$_\odot$ = 8.39.
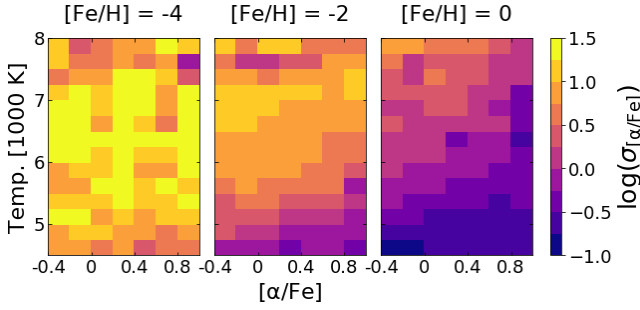
Figure 9. A plot of α-element abundance for a star of magnitude $G = 16$. The horizontal and vertical axes show increasing α-element abundance and effective temperature respectively. The colour-bar shows the α-element abundance resolution.

6.2 α-element abundance Sensitivity

Elements created by fusing helium nuclei (O, Mg, Si, S, Ca, Ti) are so-called α-elements and play an important role in Galactic Archaeology (see e.g., Shetrone et al. 2003). Pioneering works reporting the α-excess in halo metal-poor stars (see e.g., Aller & Greenstein 1960; Wallerstein 1962) helped to understand the relation between the relative abundances of α-elements with respect to iron and the star formation history of a system such as the Milky Way or its satellites. Since then, hundreds of authors have studied the large scale properties related to the chemistry of α-elements (see e.g., Edvardsson et al. 1993; Nissen & Schuster 1997, and references therein). Moreover, significant efforts are still underway to try to understand the important role of α-production in different cosmic environments.

Throughout the analysis discussed in Section 3 we focused on the stellar parameters: effective temperature, surface gravity and metallicity. The stellar library discussed in Section 2.1.1 however additionally includes α abundance as a free parameter. Similarly to Section 6.1 we repeat our original analysis while including α-element abundance as a parameter.

The results of this analysis can be seen in Figure 9 where we see that α-element abundance uncertainty increases as a function of increasing effective temperature and additionally as a function of decreasing abundance, a feature that is common across all abundance uncertainty plots shown in this paper (Figures 4, 6 and 8). We note that we only see an α-element abundance uncertainty of $\sigma_{[\alpha/\text{Fe}]} < 0.3$ dex for high metallicity, [Fe/H] = 0, and low temperature, $T_{\text{eff}} \lesssim 5000$ K. Assuming α-abundance varies between $-0.4 < [\alpha/\text{Fe}] < 1$ dex, a resolution that is of order 1 dex does not allow for any reasonable constraint on the α-element abundance of the star. Therefore, for many stellar parameters we will be unable to use BP/RP spectra to constrain α-element abundance, as suggested by Gavel, Alvin et al. (2021). Our results suggest that when temperature and metallicity measurements identify a star as being high metallicity ([Fe/H] = 0) and low temperature ($T_{\text{eff}} \lesssim 5000$ K), we will likely have a low α-element abundance uncertainty. This low uncertainty may allow us to identify populations of low or high α-element abundance stars in this stellar parameter range, however, the large increase in uncertainty for any other stellar parameters means studying α-element abundance variations across the stellar parameter space will potentially be problematic.

7 CONCLUSIONS

Gaia DR3 will see the release of over 100 million stellar spectra, producing a wealth of opportunities to study stellar parameters. However, the low resolution ($R \sim 50$) of the on-board BP and RP spectrographs were expected to lead to large uncertainties on measurements of stellar parameters such as metallicity. We produced mock BP/RP spectra using the known instrumentation of the spectrographs in order to study the feasibility of extracting the effective temperature, surface gravity and in particular metallicity of a star observed by Gaia. We produced a mock BP/RP spectra catalogue by transforming a synthetic spectra catalogue using the known transmission curve, spectral resolution and line-spread-function of the spectrographs.

In order to obtain the best-case uncertainties on stellar parameter measurements, we analysed these mock spectra using Fisher information matrices. We found that for most intrinsic stellar parameters, for a magnitude $G = 16$ star, we can constrain the effective temperature to within 100 K, while surface gravity can often be constrained to a resolution that allows us to distinguish between dwarfs and giants. This however shows that, in combination with observed high correlations between stellar parameters, expected effective temperature and surface gravity priors of 140 K and 0.2 dex respectively will help to reduce metallicity uncertainties. Before introducing these priors, we found metallicity uncertainties that increase with decreasing metallicity and increasing temperature, ranging from of order 0.1 dex for low temperature, high metallicity stars, up to of order 10 dex for high temperature, low metallicity stars. Following the inclusion of temperature and surface gravity priors, we saw the same trends, but with improvements in metallicity uncertainty of up to 60% for some stellar parameters.

We took simulated Gaia observed stellar parameters of 215 million stars and used these parameters to derive their associated metallicity uncertainty. Using this uncertainty we re-drew their metallicity to produce a mock-observed metallicity distribution. This then allowed us to probe the false positive rate of identifying a "low-metallicity" star using Gaia BP/RP spectra for a range of "low-metallicity" definitions. We found that for magnitudes $G \leqslant 16$, we had a FPR for the metallcity regime [Fe/H] < -2 of just 1 in 4, while this is 1 in 2 for the metallicity regime [Fe/H] < -3 for stars of magnitude $G \leqslant 14$.

However, given the exterior effective temperature and surface gravity constraints that are available we considered how the FPR varies for known intrinsic stellar parameters. We found that when considering stars with magnitude $G \leqslant 16$, for the metallicity regime [Fe/H] < -2, the FPR was less than 40% when avoiding 3500 K < T_{eff} < 4000 K and $T_{\text{eff}} \sim 5500 K$. We additionally found that with a sufficient effective temperature and surface gravity prior, any star can be placed more specifically into regions of the Kiel diagram which allow for a clearer diagnosis of the FPR given that low-metallicity stars are frequently confined to specific isochrones. Our findings when considering stars with magnitudes $G > 16$ show very large FPRs however, we can identify regions of the Kiel diagram where the FPR is lower, $\sim 60\%$, and hence metal-poor detections of moderate confidence are possible.

We additionally found that Gaia DR3 has the potential to drastically out-perform other low-metallicity stellar spectroscopic surveys regarding the success-rate of metal poor detections. When we adopt a similar magnitude limit to literature (magnitudes dimmer than $G = 14$) we found a success rate of metal-poor star identifications of 86%, 67% and 44% for [Fe/H] $\leqslant$ -2.0, -2.5 and -3.0 respectively.

Our results suggest that Gaia DR3 has the potential to accurately identify extremely-metal-poor stars ([Fe/H] < -3), however, those attempting to identify them must be conscious of the potential for

false positive detections and adjust their strategy for identification appropriately. We advise observing bright ($G \leqslant 16$) stars and preferably treating detections within the effective temperature range 3500 K < T_{eff} < 4000 K and those within 50 K of $T_{\text{eff}} \sim 5550K$ with caution, due to the locally high FPR.

We finally conclude that uncertainties on Carbon abundance should allow for the identification of a CEMP candidate population for follow-up high-resolution spectroscopic observations. However, due to large uncertainties on α-abundance, diagnosing α-abundance with Gaia BP/RP will be very challenging.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data underlying this article may be shared on reasonable request to the corresponding author.

REFERENCES

Aguado D. S., Allende Prieto C., González Hernández J. I., Carrera R., Rebolo R., Shetrone M., Lambert D. L., Fernández-Alvar E., 2016, A&A, 593, A10
Aguado D. S., González Hernández J. I., Allende Prieto C., Rebolo R., 2017, A&A, 605, A40
Aguado D. S., et al., 2019, MNRAS, 490, 2241
Aguado D. S., et al., 2021a, MNRAS, 500, 889
Aguado D. S., et al., 2021b, ApJ, 908, L8
Allende Prieto C., 2016, A&A, 595, A129
Allende Prieto C., Beers T. C., Wilhelm R., Newberg H. J., Rockosi C. M., Yanny B., Lee Y. S., 2006, ApJ, 636, 804
Allende Prieto C., Koesterke L., Hubeny I., Bautista M. A., Barklem P. S., Nahar S. N., 2018, Astronomy & Astrophysics, 618, A25
Aller L. H., Greenstein J. L., 1960, ApJS, 5, 139
Anders F., et al., 2022, Astronomy & Astrophysics, 658, A91
Antiche E., Masana E., Julbe F., Borrachero R., 2014, in EAS Publications Series. pp 355–355, doi:10.1051/eas/1567064
Arentsen A., et al., 2021, MNRAS, 505, 1239
Asplund M., Grevesse N., Sauval A. J., 2005, in Barnes Thomas G. I., Bash F. N., eds, Astronomical Society of the Pacific Conference Series Vol. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis. p. 25
Beers T. C., Christlieb N., 2005, ARA&A, 43, 531
Beers T. C., Preston G. W., Shectman S. A., 1985, AJ, 90, 2089
Bonaca A., Hogg D. W., 2018, The Astrophysical Journal, 867, 101
Bonifacio P., et al., 2015, A&A, 579, A28
Caffau E., et al., 2013, A&A, 560, A71
Carollo D., et al., 2012, ApJ, 744, 195
Carrasco J. M., et al., 2021, Astronomy & Astrophysics, 652, A86
Christlieb N., Green P. J., Wisotzki L., Reimers D., 2001, A&A, 375, 366
Christlieb N., et al., 2002, Nature, 419, 904
Cohen J. G., et al., 2005, The Astrophysical Journal, 633, L109
Da Costa G. S., et al., 2019, MNRAS, 489, 5900

Deng L.-C., et al., 2012, Research in Astronomy and Astrophysics, 12, 735
Edvardsson B., Andersen J., Gustafsson B., Lambert D. L., Nissen P. E., Tomkin J., 1993, A&A, 275, 101
Fisher R. A., 1935, J. Roy. Stat. Soc., 98, 39
Frebel A., Norris J. E., 2015, ARA&A, 53, 631
Frebel A., et al., 2005, Nature, 434, 871
Frebel A., et al., 2006, ApJ, 652, 1585
Gaia Collaboration et al., 2016, A&A, 595, A1
Gaia Collaboration et al., 2021, A&A, 649, A1
Galarza C. A., et al., 2022, A&A, 657, A35
Gavel, Alvin Andrae, René Fouesneau, Morgan Korn, Andreas J. Sordo, Rosanna 2021, A&A, 656, A93
Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, A&A, 486, 951
Herwig F., 2005, ARA&A, 43, 435
Husser T. O., Wende-von Berg S., Dreizler S., Homeier D., Reiners A., Barman T., Hauschildt P. H., 2013, A&A, 553, A6
Ishigaki M. N., Tominaga N., Kobayashi C., Nomoto K., 2014, ApJ, 792, L32
Katz D., et al., 2019, A&A, 622, A205
Keller S. C., et al., 2007, Publ. Astron. Soc. Australia, 24, 1
Koesterke L., Allende Prieto C., Lambert D. L., 2008, ApJ, 680, 764
Kurucz R. L., 1979, ApJS, 40, 1
Li H., Tan K., Zhao G., 2018, ApJS, 238, 16
Limberg G., et al., 2021, ApJ, 913, 11
Ludwig H.-G., Freytag B., Steffen M., 1999, A&A, 346, 111
Nissen P. E., Schuster W. J., 1997, A&A, 326, 751
Placco V. M., Frebel A., Beers T. C., Stancliffe R. J., 2014, ApJ, 797, 21
Riello M., et al., 2021, A&A, 649, A3
Salvadori S., Skúladóttir Á., Tolstoy E., 2015, MNRAS, 454, 1320
Sbordone L., Bonifacio P., Castelli F., 2007, in Kupka F., Roxburgh I., Chan K. L., eds, 2007 International Astronomical Unio Vol. 239, Convection in Astrophysics. pp 71–73, doi:10.1017/S1743921307000142
Schlaufman K. C., Casey A. R., 2014, ApJ, 797, 13
Shetrone M., Venn K. A., Tolstoy E., Primas F., Hill V., Kaufer A., 2003, AJ, 125, 684
Sneden C., Kraft R. P., Prosser C. F., Langer G. E., 1991, AJ, 102, 2001
Starkenburg E., et al., 2017, MNRAS, 471, 2587
Starkenburg E., et al., 2018, MNRAS, 481, 3838
Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22
Tomkin J., Lemke M., Lambert D. L., Sneden C., 1992, AJ, 104, 1568
Umeda H., Nomoto K., 2003, Nature, 422, 871
Venn K. A., Irwin M., Shetrone M. D., Tout C. A., Hill V., Tolstoy E., 2004, AJ, 128, 1177
Venn K. A., et al., 2020, MNRAS, 492, 3241
Wallerstein G., 1962, ApJS, 6, 407
Whitten D. D., et al., 2019, A&A, 622, A182
Wisotzki L., Christlieb N., Bade N., Beckmann V., Köhler T., Vanelle C., Reimers D., 2000, A&A, 358, 77
Yoon J., et al., 2016, ApJ, 833, 20
Yoon J., et al., 2018, ApJ, 861, 146
York D. G., et al., 2000, AJ, 120, 1579

This paper has been typeset from a TeX/LaTeX file prepared by the author.