Mon. Not. R. Astron. Soc. **000**, 1–12 (2016) Printed 6 April 2016 (MN LATEX style file v2.2)

The Gaia-ESO survey: Confronting Chemodynamical Simulations with Synthetic Observations

B. B. Thompson,[1,2,5] M. Bergemann,[3] C. G. Few,[2,4], A. Serenelli, B. Macfarlane[1] and B. K. Gibs
[1] *Jeremiah Horrocks Institute, University of Central Lancashire, Preston, Lancashire, PR1 2HE, UK*
[2] *E.A. Milne Centre for Astrophysics, University of Hull, Hull, HU6 7RX, UK*
[3] *Max-Planck Institute for Astronomy, Knigstuhl 17, 69117 Heidelberg, Germany*
[4] *Department of Physics & Astronomy, University of Exeter, Exeter, EX4 4QL, UK*
[5] *Institute for Computational Astrophysics, Dept of Astronomy & Physics, Saint Marys University, Halifax, NS, BH3 3C3, Canada*

6 April 2016

ABSTRACT
We use RAMSES-CH, a cosmological adaptive mesh refinement code based on RAMSES with a chemodynamical patch capable of tracing the elemental abundance of H, C, N, O, Ne Mg, Si, and Fe in the intergalactic medium. We compare a sample of stellar particles from a Milky Way-like simulation, selecting a region of space comparable with the solar neighbourhood with the Gaia-ESO survey DR2 and other observational surveys. We find that the model is producing comparable results with the iron and magnesium abundances, although overproducing iron rich stars. In addition, we apply the Synthetic Colour-Magnitude Diagram tool to sample the stellar properties within a simulated solar neighbourhood equivalent region in the same way that an observer would do, which narrows the metallicity distribution function by filtering out red and faint stars. TALK ABOUT PURPOSE OF THIS BEING PAPER ONE. THE CLEAR AIM.

Key words: galaxies: evolution – galaxies: formation – methods: numerical – Galaxy: abundances

1 INTRODUCTION

With the progression of time in astronomy, subsequent observational surveys and instrumentation, and the performance improvements of high performance computing, have lead to successive improvements in resolution and abundance of both survey and simulation datasets. This increasing performance in both observational and computational models have lead to an increased understanding of properties of the Milky Way galaxy and the origins of the stars that reside within.

The improvement of resolution and abundance of datasets has been progressive, whereas the methodology mapping simulations to observations has remained unaltered for decades. It is common, and indeed straight forward to simply take the results from the outputs of an observational survey - such as the Gaia-ESO survey (GES) (Gilmore et al. 2012), and compare it like for like with simulation results of a Milky Way-Like galaxy. This is achieved by simply comparing like-for-like a sub-sample of data from a simulated Milky Way-like galaxy, of which the sub-sample is created from within a spatial region which maps approximately one-to-one with the spatial region which is implied from the position data from the observational survey.

However, this approach - which underpins the majority of analysis work of galaxy simulations in the simulation community today - takes the inherent assumption that the typical spatial selection of stellar data is completely analogous to that of galaxies in nature. As such that the physical and chemical properties of massive stellar-like particles, a consequence of resource and computational limitations of computational models, can be compared directly with individual stars in observational surveys.

One could argue that the treatment of sub-grid physics aims to resolve the problems of miss-match of resolution between "composite" stellar particles and stars from real life surveys. However A spatial selection, along with treatment of the sub-grid physics - although both are quantitatively important - do not feature any inclusion of how one "observes" the simulation. Typically in observational surveys, such as the GAIA-ESO survey (Gilmore et al. 2012), one would invoke an observational selection function which would consist of boundary conditions for individual stellar properties of colour, apparent magnitude, chemistry and gravity in addition to a spatial region for individual stars. In simulations, one is limited primarily to the integrated luminosities and chemistry of the open cluster-scale simple stellar populations, which are used to represent "composite" stellar particles. A such (Miranda, Macfarlane & Gibson 2015) concludes that the impact of how one observes a simulation, whether it be observationally motivated (e.g from the point of view of a simulated observer), to even simple spatial cuts filtering out everything but the effective solar neighbourhood is as quantitatively important just as any of the sub-grid physics treatments within the simulations themselves.

As we enter the Gaia era, it is crucial that the approach to confronting simulation data changes in a manner which better repli-

cates the methodology employed by observers. This would allow for more direct comparisons between observational surveys with theoretical models on a self-consistent level. The Synthetic Colour-Magnitude Diagram (hence-force known as SynCMD) Pasetto, Chiosi & Kawata (2012) is a tool that allows us to apply an observationally-motivated selection function upon the inferred age and metallicity distribution of an analogous 'solar neighbourhood' in a simulation. SynCMD enables you to take the role of a survey designer within a simulated model. As such you can apply colour, magnitude and surface gravity boundary conditions and produce synthetic properties of ages and metal abundances from populating a colour magnitude diagram (CMD) for synthetic star particles, of which is commonly done in observational surveys. Therefore allowing stellar samples to be drawn by an observer situated in a simulation using apparent magnitude and gravity criteria from an analogous 'solar neighbourhood'.

In addition, one of the challenges faced by observational surveys is reducing the magnitude of, and number of sources of errors on observational surveys. Examples of such errors range from the measurement of the spectral line width for stellar spectral abundance measurements, errors on the precision of the measurement themselves or as result of the atmospheric physics for the case of a ground base observation. With simulation work, we are capable of measuring to a much finer precisions various field data of stellar particles. As such which the field data for "composite" stellar particles can be precisely measured and are computed as a result of the sub-grid physics, the hydrodynamical and chemodynamical mode employed and the resolution of position and particle number. With the right methodology, the errors on various fields of observational data can be replicated and implemented on fields in simulation data, so long as the errors are known concisely on an observational survey itself.

In this work, we will present in detail our methodology of transforming a simulation model into the observers plane making use of SynCMD for direct comparison with an observational survey. In subsequent papers in this series, we will the series will explore the implications for galaxy formation/evolution in a more quantitative manner, focusing on the results of the Gaia-ESO survey with galaxy models generated from the cosmological hydrodynamical code RAMSES (Teyssier 2002), but specifically the chemical evolution patch RAMSES-CH (Few et al. 2012a, 2014).

Our observational survey of choice in this work is the high resolution results of the Gaia-ESO survey fourth data release (GES-DR4). The Gaia-ESO Survey (Gilmore et al. 2012) is an ongoing, five-year public spectroscopic survey conducted with the FLAMES/GIRAFFE and UVES instruments at the Very Large Telescope, which began at the end of 2011. The results of the Gaia-ESO Survey will deliver high-quality spectra for up to ∼100,000 stars in the MW by 2017. The recently released internal fourth data release (iDR4) contains nearly 20,000 spectra for stars in the MW field as well as globular and open clusters. With the data taken with the Gaia-ESO survey, which forms one of the basis of this work, we are in a good position to address the age and metal relations of our simulated galaxies (Bergemann et al. 2014).

The choice of RAMSES-CH enables us to model the abundance evolution of stellar ages, metallicites and abundances of Fe, Mg and O for a Milky Way-like galaxy in addition to H and He. RAMSES-CH also includes the energetic feedback from both Type Ia and Type II supernovae, In addition it also the chemical enrichment from low- and intermediate-mass asymptotic giant branch stars. Employing the methodology described in this paper and using a chemodynamical model enables us to compare like-for-like the Gaia-ESO survey with a Milky Way-like galaxy simulated in RAMSES-CH, both using observational selection functions. Our comparison work will focus on the comparison of the results of the Gaia-ESO survey with the chemodynamical properties of a Milky Way-like galaxy - henceforth known as Selene-CH (the chemodynamical simulation of Selene) Few et al. (2012b) - which using the methodology described in this paper, will transform the "Solar Neighbourhood" in Selene-CH into the observers plane with the aid of the convolution of errors of the Gaia-ESO survey fields and the SynCMD tool. The resulting observers plane data will henceforth be known as Selene-SYN.

In the standard and well tested ΛCDM cosmological model, galaxies are found to build up their mass through the successive mergers of smaller galaxies via a process known as hierarchical structure formation (e.g White & Rees 1978; Springel et al. 2005). The result of the merging of smaller galaxies with the Milky Way (MW) will result in the deposit stars, gas and dark matter into the MW galaxy and the destruction of the smaller galaxy within the process (e.g Walker, Mihos & Hernquist 1996; Abadi et al. 2003; Read et al. 2008).

The MW is a unique tool for understanding the formation and evolutionary processes of disc galaxies in the Universe in the cosmic field-like environments. We can study the stellar populations which carry imprints of the history of the MW in greater detail than in any other galaxy.

One of the primary diagnostic tools of galaxy formation is the determination of elemental abundance patterns. We use numerous comparisons of trends between stellar ages, metal abundance and spatial position relative to the galaxy. These trends are able to guide and develop an understanding of the physical processes which undergo within galaxy formation models. Observations of various metal abundance ratios aid our understanding of the nuclear physics involved with α-element production, which are produced on shorter time-scales than iron-peak elements (e.g Carbon et al. 1987; Edvardsson et al. 1993; Reddy, Lambert & Allende Prieto 2006; Ramírez, Allende Prieto & Lambert 2007). As a consequence of this, the mass-dependent nuclear burning processes acting within the relevant progenitor stars.

From measuring the chemistry, ages and dynamics of the individual stars of the MW, we can trace their origins. This is known as Galactic Archaeology (Eggen, Lynden-Bell & Sandage 1962; Freeman & Bland-Hawthorn 2002) in which the underlying principle is that surveys provide a fossil record of the formation and evolution of the MW. Despite significant progress due to the advent of large stellar surveys like the Sloan Digital Sky Survey (SDSS) York et al. (2000); Eisenstein et al. (2011), we still do not know the precise merger history of the MW.

Galactic chemical evolution (CE) models use predictions based upon the coupling of both the elemental production sites and time-scales with phenomenological (but, empirically constrained) parametrisations of star formation and gas outflows/inflows. These resulting predicted abundance patterns can be compared directly with observation, as demonstrated in this work, with observations allowing us to shed light on the formation and evolution of the system under study, such as for example our home galaxy, the Milky Way.

Studies of the role that CE plays with the physics of the interstellar medium (ISM) have shown that metallicity-dependent radiative cooling rates of plasmas have an impact on the efficiency of metal transportation through the galactic disc and thus its impact on stellar chemodynamics (Scannapieco et al. 2005). RAMSES-CH uses an Eulearian Adaptive Mesh Refinement (AMR) solver, which

provides diffusion of metals via hydrodynamical mixing of metal rich and metal poor gases. In addition, AMR codes are suited with dealing with physical processes that involved regions of rapidly changing density resolving instabilities.

Even though the recipes for implementing the physics of the formation of galaxies are still widely discussed in detail, the majority of N-body codes are currently able to evolve with time the baryonic component, i.e gas an stars, in their mutual interaction. It is thus possible and mandatory to develop the right tools to compare the results of dynamical N-body simulations with observational data for the stellar content of galaxies.

Building up from the work of (Tantalo et al. 2010) which describes a technique to derive the integrated spectra, magnitudes and colours of the stellar content of simulated galaxies, we use the Synthetic Colour Magnitude Diagram (SynCMD) Pasetto, Chiosi & Kawata (2012) to generate colour-magnitude diagrams of our simulated galaxies, and thus also derive synthetic properties of our simulated galaxies such as the colours, observed magnitudes (from a simulated observer) and log(g). In effect, this allows us to sample our simulations in the same way that observers do, which fits in very well comparison with Gaia-ESO survey, and thus are able to employ a similar selection criteria.

2 METHODOLOGY

use as basis of overview to methodology We apply the code SynCMD (Synthetic Colour Magnitude Diagram), synthetic stellar populations generation tool (Pasetto, Chiosi & Kawata 2012), to our simulated galaxy. SynCMD replaces composite simulation star particles with individual stars spanning the IMF mass range, and post-process the simulation as an observer would do, selecting stars based upon apparent magnitude, colour, and surface gravity criteria similar to that employed by observers. This allows us to adopt a truly observationally-motivated approach to study our simulated galaxy as an additional sampling scheme to compare it with spatial selection which is frequently used in the simulational community.

2.1 RAMSES-CH

Our cosmological simulations are performed with RAMSES (Teyssier 2002). To trace the chemical evolution of the simulated galaxy we employ a chemodynamical patch called RAMSES-CH (Few et al. 2012a, 2014). RAMSES-CH is able to perform N-body and hydrodynamical simulations including star, dark matter and gas. The AMR method used in RAMSES allows for refinement on a cell-by-cell basis increasing the resolution in dense regions of the volume. This refinement allows for a reduction in computing time while maintaining a high resolution around the galaxy and capturing large-scale cosmological phenomena. RAMSES-CH includes treatments of self-gravity, hydrodynamics, star formation, supernova feedback, gas cooling and chemical enrichment.

Dense gas cells form stellar population particles if the density surpasses a number density threshold of n_0=0.1 cm^{-3}. The rate at which star particles are produced is $\dot{\rho}_* = \epsilon_* \rho_g / t_{ff}$, where $t_{ff} = (3\pi/32 G\rho_g)^{1/2}$ is the local gas free-fall time, ρ_g is the gas mass density and star formation efficiency, ϵ_*=0.01. The particles used to trace the stellar mass phase are commonly and colloquially referred to as "star particles" by the community however they do not represent single stars but aggregate coeval stellar populations, in the simulation present here they have a birth mass of 3.3×10^4 M$_\odot$. To avoid confusion we make a a distinction between the actual simulation particles (hereafter "stellar population particles") and the synthetic particles representing individual stars described in §2.3 (hereafter "star particles"). **GARETH: Worth using SSP and Synthetic Star Particle? BEN: We have a definition clash?, Synthetic Star Particle and Simple Stellar Populations could easily be mixed up (or at least I worry this may be the case)** We inject $10^5$1 erg of energy per SN into the local ISM grid cell as thermal energy.

Radiative gas cooling in the simulation is metallicity- and density- dependent. Cooling rates are calculated assuming photoionisation equilibrium with a redshift dependent uniform UV background (Haardt & Madau 1996). Cooling rates due to the presence of metals is calculated from the total metallicity of the gas which is interpolated to the CLOUDY (Ferland et al. 1998) cooling rates at zero and solar metallicity at temperatures exceeding 10^4 K; colder gas takes its metal cooling rates from Rosen & Bregman (1995). We also employ the delayed cooling feedback mechanism from Teyssier et al. (2013) to account for the unresolved multiphase nature of the gas and avoid the spurious loss of thermal energy following SN feedback. In addition to these prescriptions for gas heating/cooling we also impose a polytropic equation of state as a temperature floor. This temperature floor prevents the gas from reaching low enough temperatures that the Jeans' length of the gas is unresolved, which can lead to unphysical fragmentation. The gas is therefore prevented from falling below $T_{min}=T_{th}(n_g/n_0)^{\gamma-1}$ where $\gamma = 2$, T_{th}=188 K, and n_0=0.1 cm^{-3}.

RAMSES-CH allows us to track the elements H, He, C, N, O, Ne, Mg, Si, and Fe from their dominant production sites (SNII, SNIa and AGB stellar winds) into the ISM where they are advected with the gas flow and become imprinted on the stellar population particles. The details of RAMSES-CH are described fully in Few et al. (2014) but we briefly summarise the main components here. Energetic feedback from both type-Ia and type-II supernovae (SNIa and SNII respectively) is included, while AGB stars eject their mass passively into the enclosing grid cell. We use the model B SNII yields of Woosley & Weaver (1995) with a correction applied to the yields after Timmes, Woosley & Weaver (1995) which halves the Fe produced by massive stars and AGB yields are taken from van den Hoek & Groenewegen (1997). We consider stars in the mass range 0.1–8 M$_\odot$ to evolve along the AGB while stars with masses 8–100 M$_\odot$ eject mass and energy as SNII. The mass distribution of stars in each stellar population particle is determined by the choice of Initial Mass Function (IMF), in this work we use the IMF of (Salpeter 1955). The number of SNIa per unit initial stellar mass is also determined by the IMF via the number of stars with masses 3–8 M$_\odot$ in binary systems with either a red giant or main sequence star; the lifetime of these systems is taken of the main sequence lifetime of the secondary star (Kodama & Arimoto 1997). This combination of chemical evolution model parameters is described in Few et al. (2014) as model *S55-uM100-IaK* where the impact of the choice of IMF and SNIa model on simulations similar to that presented here is also discussed.

There are now numerous hydrodynamical simulation code to choose from; grid and particle based codes and the more recently emerging moving-mesh and meshless approaches (Springel 2010; Hopkins 2015). The strengths and weaknesses of these codes are explored in Agertz et al. (2007)**BEN: what other papers should be referenced?**. A key property of our code is its ability to capture metal mixing, this is extremely pertinent to this work as it directly affects the dispersion in the abundance ratios of the gas which becomes imprinted on the stars. Adopting a grid-, rather than particle-

based approach, in general allows for a better treatment diffusion of metals (Pilkington et al. 2012)**BEN: What is ,Yves Revaz's paper**.

2.2 Galaxy Initial Conditions: Selene-CH

We employ a cosmological 'zoom-in' simulation technique using RAMSES-CH to simulate the galaxy, '*Selene-CH*'. This galaxy exists in a box 20 Mpc h^{-1} in size created with cosmological parameters (H_0, Ω_m, Ω_Λ, Ω_b, σ_8) = (70 km s^{-1}, 0.28, 0.72, 0.045, 0.8) and we run the simulation to $z = 0$. The adaptive grid can refine up to 17 levels corresponding to a maximum resolution of 218 pc with a dark matter particle mass resolution of 5.64×10^6 M$_\odot$ and a stellar population particle birth mass of 3.3×10^4 M$_\odot$.

The galaxy presented here is a chemodynamical resimulation of the '*Selene*' initial conditions first presented in Few et al. (2012b). The feedback scheme used in simulating *Selene-CH* is different to the original version and so, while the galaxy has roughly the same environment and assembly history, some differences are to be expected. The galaxy inhabits a dark matter halo with a mass of 5.245×10^{11} M$_\odot$ that is more than 3 Mpc distant from any other haloes more massive than 3×10^{11} M$_\odot$. The halo and its properties are identified using the AMIGA halo finder (AHF) (Knollmann & Knebe 2009; Gill, Knebe & Gibson 2004). The assembly history of this halo is relatively quiescent (it was selected as such) with no major mergers after redshift $z = 1.0$. The assembly history of the original version of *Selene* is described in Few et al. (2012b) and more extensively with relation to the effect of its assembly on the metallicity and age distribution in (Ruiz-Lara et al. 2016).

A gas surface density projection of *Selene-CH* is shown in Figure 1 demonstrating the presence and shape of the spiral arms. The cross at $x = 4.0$ kpc and $y = 6.93$ kpc is the region where we place our simulated observer as described in Section 2.6.1, 8 kpc from the galactic centre in a spiral arm. We use stars from a spherical region 2 kpc in radius around this point which is treated as our simulated solar neighbourhood. The size of this region is discussed in Section 2.6.1.

: GARETH: Why are we showing the rotation curve? : BEN: I agree, no real need We include plots of the rotational velocity of this galaxy which are shown in Figure **??**.

: GARETH: If including this here we need to say more about it : BEN: I think it should be included, but barely discussed. The star formation and supernova rates for our simulated galaxy are shown in Figure **??**.

2.3 SynCMD

BEN: Since this is the first "published" paper using SynCMD as a tool, I would like emphasise the application detail more so The SynCMD synthetic stellar populations generation tool (Pasetto, Chiosi & Kawata 2012) is a toolkit designed to examine simulation data in a similar manner to how an observational survey would sample real life stellar populations by applying observationally-motivated selection functions to simulated stellar population particles. The first step in this process is to split the particles into individual stars by stochastically populating a colour-magnitude diagram based on the population's age. Due to resolution limits in simulations of this kind the particles represent 'averaged' stellar populations rather than individual stars. Stellar population particles typically have a mass of $\sim10^4$–10^6 M$_\odot$ and therefore a stochastic approach to dividing that mass into single stars is an appropriate way of populating the colour-magnitude diagram. This allows the calculation of properties of stars that are not natively part of the simulated stellar population particles, i.e. the effective temperature, surface gravity and magnitudes in different colour bands.

The details of the SynCMD code are given in Pasetto, Chiosi & Kawata (2012) but we summarise the process here. A database of simple stellar populations (SSP) is initially calculated based on the stellar models as described in Bertelli et al. (2008, 2009) which covers a wide grid of helium and metal abundances (Y and Z respectively) and an enrichment ratio $\Delta Y/\Delta Z$. This database includes the stellar isochrones generated from the chosen initial mass function which includes the effect of mass loss by stellar wind and the thermally pulsing AGB phase as described by the models within Marigo & Girardi (2007). The code used to generate the database of SSP 'under-the-hood' is a modified version of YZVAR which has been used in many studies (for instance Chiosi, Bertelli & Bressan 1986; Chiosi et al. 1986, 1989; Chiosi & Greggio 1981; Ng et al. 1995; Aparicio et al. 1996; Bertelli et al. 1995; Bertelli & Nasi 2001; Bertelli et al. 2003) and has recently been extrapolated to compute isochrones and SSPs in an extended region of the Z-Y plane. The details of the interpolation scheme at a given $\Delta Y/\Delta Z$ are **GARETH: what is the deal with the interpolation/extrapolation BEN: Interpolation occurs inbetween the isochrone tracks. Extrapolation *I think* is needed to extend beyond the 0.15 and 20 M$_\odot$ range YZVAR has** given in Bertelli et al. (2008, 2009), but for SynCMD, this is extended this so we can interpolate between the range of 0.08 M$_\odot$ (the hydrogen burning limit) and 100.0 M$_\odot$ (the upper limit of most initial mass functions), rather than the range of 0.15 and 20.0 M$_\odot$ as described in Bertelli et al. (2008, 2009). This extrapolation occurs by TODO FILL THIS IN.

GARETH: I've edited this section but I have NOT fact checked it, I'm trusting the meaning of the text I was supplied with. BEN: This is one of the topics I would like to discuss in the meeting. I feel that there should be some strong emphasis on how SynCMD is used to from from your inputs to your outputs. I.E from the IMF and your star particles to your output. Using this database, and simulated stellar particle masses, ages, metallicities and distances from the simulated observer - with the use of the isochrones and simple stellar populations (SSPs) - a catalog of artificial stars is created tracing a synthetic colour magnitude diagram (CMD). The next step is to remove the stars that do not fulfill the selection functions of the observations to which the simulation is to be compared. The stars retain the age and elemental abundances of the simulation particle they are created from.

GARETH: If the below section is described in the original paper then cut it from here BEN: technically it is, but in my opinion not very well, hence why I would like to summerise the application of SynCMD here. (i.e it's inputs and outputs). e.g being able to filter through any colour band of UBVRIJHK - unless maybe that idea is more for the discussion/conclusion? The CMD is binned into a grid of cells, from each cell and the database, we can compute the number of stars per elemental cell of the CMD, and thus the properties of each synthetic star. This is done instead of computing the number of synthetic star particles per stellar particle because once the selection function is include, the majority of these synthetic star particles would have been discarded and CPU resources would have been unnecessarily wasted. Thus obtaining the synthetic stars from the CMD bins is much more practical.

BEN: Guessing you feel that we should discuss this pa-

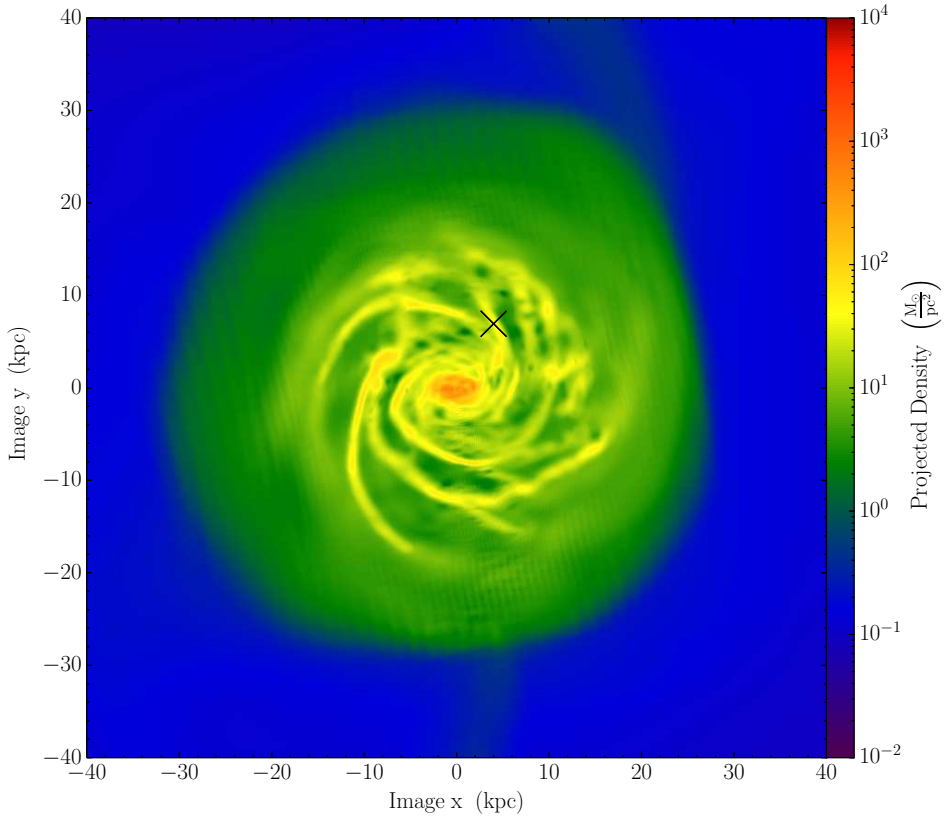


Figure 1. A face-on gas surface density projection plot of the galaxy used in this work Selene-CH, a Milky Way-like galaxy simulated within a cosmological context using the RAMSES-CH code. The galaxy was visualised using the YT visualisation toolkit (Turk et al. 2011) with a projection depth of 20 kpc. The black cross in this figure at $x = 4.0$ kpc and $y = 6.93$ kpc is the region where we place our simulated observer as described in Section 2.6.1, in a star spiral arm 8 kpc from the galactic centre.

per positivly? briefly described in Miranda, Macfarlane & Gibson (2015). **BEN: More specifically, it should be mentioned somewhere that unlike that paper, we use the GAIA-ESO selection function rather than the RAVE. The results of this will be discussed in the results (since RAVE changes things more dramatically than GAIA-ESO?**

2.4 The GAIA-ESO Survey

Since we reside in a spiral galaxy within a field/loose group environment, we can study in high spatial and spectral resolution its stellar and gas-phase constituents. Missions such as ESA's Gaia (Soubiran et al. 2013) provide exquisite positional and kinematic data for billions of stars, while ground-based complements such as the Gaia-ESO Survey (Gilmore et al. 2012; Bergemann et al. 2014) yield powerful age and chemical information for significant subsets of these stars.

BEN: Maria should look at this. The Gaia-ESO survey (GES-DR4). is the largest high-resolution spectroscopic survey of stars in the Milky Way. The Gaia-ESO Survey has been awarded 300 nights on the Very Large Telescope in Chile to perform high- and medium-resolution spectroscopy of 100,000 stars in the Milky Way. In the high-resolution ($R \sim 47\,000$) mode, the goal is to acquire spectra for about 5 000 field stars, probing distances up to 2 kpc from the Sun; 100 000 spectra will be acquired in medium-resolution mode with Giraffe spectrograph. Here we use the results from the fourth data release of the survey, which includes all Giraffe and UVES spectra for the first 18 months of the survey. We focus on the high-resolution UVES data, for which accurate Teff, log(g), and metallicities are available.

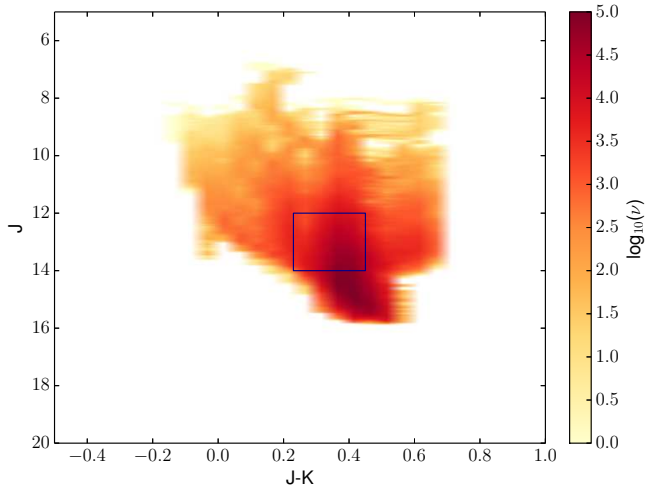


Figure 2. A plot of the synthetic colour-magnitude diagram (CMD) of J vs J-K in apparent magnitude space with $3.5 \geq \log(g) \geq 4.5$. The blue rectangle highlights the J and J-K region probed in this study within Selene-SYN.

BEN: Is the footnote here still True? The UVES solar neighbourhood targets were chosen according to their colours to maximise the fraction of un-evolved FG stars within 2 kpc in the solar neighbourhood. The selection box was defined using the 2MASS photometry: $12 < J < 14$ and $0.23 < J - K < 0.45 + 0.5E(B - V)$ (The $+0.5E(B - V)$ term is in order to correct for dust extinction); if there were not enough targets the red edge was extended[1]. According to these selection criteria, the majority of stars are FG stars with magnitudes down to $V = 16.5$

For the analysis of the spectra, several state-of-the-art spectrum analysis codes are used (Gilmore et al. 2012; Smiljanic et al. 2014; Randich, Gilmore & Gaia-ESO Consortium 2013). The observed spectra were processed by 13 research groups within the Gaia-ESO survey collaboration with the same model atmospheres and line lists (Heiter et al. 2014), but different analysis methods: full spectrum template matching, line formation on-the-fly, and the equivalent width method. The model atmospheres are 1D LTE spherically-symmetric ($log(g) \leq 3.5$) and plane-parallel ($log(g) \geq 3.5$) MARCS (Gustafsson et al. 2008). The final parameter homogenisation involves a multi-stage process, in which both internal and systematic errors of different datasets are carefully investigated. Various consistency tests, including the analysis of stellar clusters, benchmark stars with interferometric and astroseismic data, have been used to assess each group's performance. The final parameters are median of the multiple determinations, and the uncertainties of stellar parameters are median absolute deviations, which reflect the method-to-method dispersion.

BEN: The new Mg data does not have NLTE corrections. Is other aspects of this paragraph still true The mean uncertainty of stellar parameters are 100 K in Teff, 0.15 dex in surface gravity, and 0.1 dex in metallicity. This accuracy could be achieved because of very careful selection of diagnostics features, very broad spectral coverage and good S/N ratio of the observed spectra, and the application of NLTE spectral models in the spectroscopic analysis.

GARETH: Do we need these two? It's an honest question, I don't truly appreciate them so can somebody with an observers head give a good reason for these staying? BEN: Not Sure, Over to Maria

The ages and masses were determined consistently using the Bayesian method of Serenelli et al. (2013). The input stellar evolution models are GARSTEC. The code was validated on asteroseismiology measurements of ages in Bergemann et al. (2014). Distances were computed as described in Ruchti et al. (2013)), using the 2MASS photometry.

The important new parameter in our work is chemical abundances. Moreover, a part of the GES-DR4 abundances suffer from the problem of spectral data reduction and erroneous transition probabilities of the two Mg lines provided to the survey in that analysis run. The problem of spectral data reduction and erroneous transition probabilities of the two Mg lines has a strong impact on the Mg and O abundance, because for each of these elements there are only a handful of spectral lines in an observed spectrum and they are extremely sensitive to the input atomic physics. We have therefore recomputed chemical abundances for the 3 observed spectral datasets using exactly the same procedure and atomic data, i.e oscillator strengths and damping. We used the method of spectral synthesis, fitting simultaneously for macro-turbulence and element abundance.

[1] The targets selected before April (2012) had a brightest cut on J of 11 instead of 12. If the number of objects in the field within the box was less than the number of UVES fibres, then the red-edge of the colour-box was shifted to have enough targets to fill the fibres.

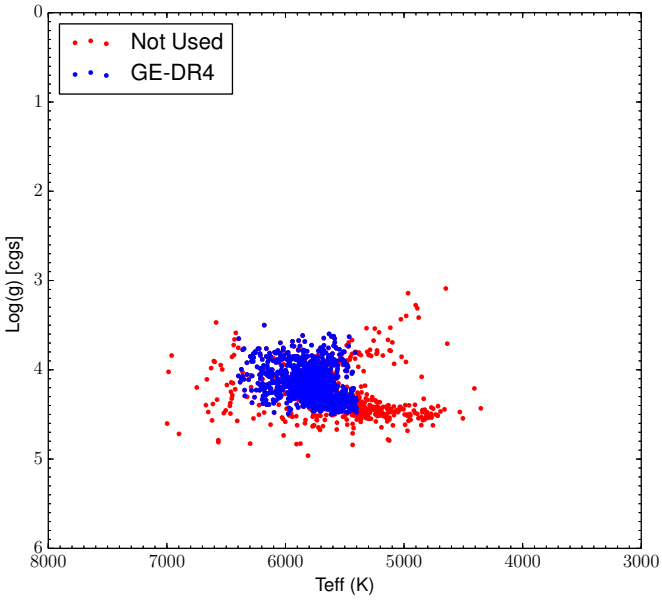


Figure 3. Teff vs Log_{10}(g) for GES-DR4 (blue) plotted alongside the stars filtered out (red), of which the filters are described in section 2.5

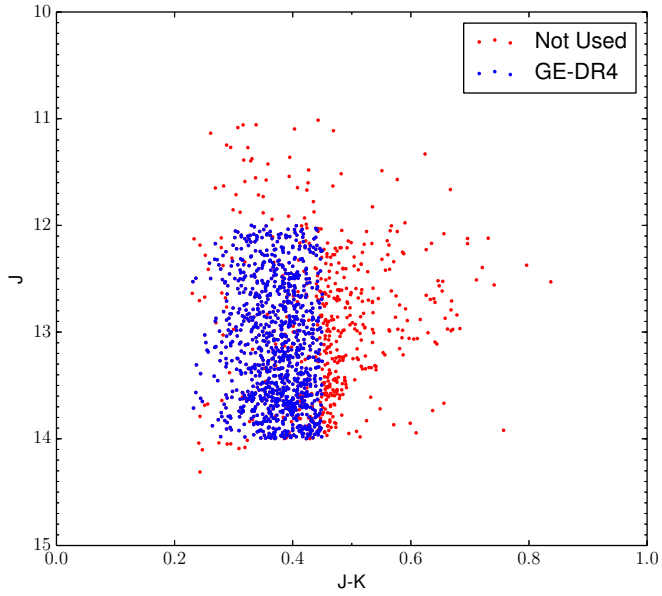


Figure 4. Hertzprung-Russel diagram showing J-K against J plot for GES-DR4 (blue) plotted alongside the stars filtered out (red), of which the filters are described in section 2.5. Plot this over the synthetic CMD somewhere?

BEN: to be removed? The resulting Herzsprung-Russel diagram of the Gaia-ESO sample is shown in Figure 3 The distribution of the observed sample in the colour-magnitude diagram is shown in Figure 4.

2.5 Gaia-ESO Survey: The Observational Survey Dataset

GARETH: I have moved this section above the part on the different simulations sets because I think it depends upon this,

it could be combined in some way with the previous? We use the fourth data release of the high resolution Gaia-ESO survey (Gilmore et al. 2012). Before applying the observational selection function, we further post-process the GES-DR4 data for the analyses presented here.

The oxygen abundances have been re-determined since the initial data release values were being over-estimated by the inclusion of telluric lines contaminating the 6300Å forbidden line of OI employed by the Gaia-ESO Survey. Such contamination is estimated to affect ~50% of the initial data in the Gaia-ESO Survey.

We also compute the [Mg/Fe] abundances using the Non-local thermodynamical equilibrium (NLTE) calculation (Bergemann et al. 2015), the model atmospheres and linelists in the GES-DR4 datasets are all 1D LTE (Local Thermodynamical Equilibrium) in which departures from LTE if neglected at the low densities of RSG atmospheres could introduce systematic errors. NLTE calculations are shown to produce stronger Mg i absorption lines in the J-band than in LTE (Bergemann et al. 2015).

We only include galactic field stars (could be halo or disk, there is no information about the stellar orbits) and remove bulge stars and stars that belong to an open or globular cluster.

Furthermore, we apply the following photometric selection function to the GES-DR4 dataset, of which:

(i) Include stars that have a galacticentoric radius of $6.0 \leq r \leq 10.0$ kpc;
(ii) Include stars that are within $z \leq |2.0|$ kpc of the plane of the disk;
(iii) Include stars with a surface gravity of $log(g) \geq 3.5$ and $T_{eff} \geq 5400$ inclusive;
(iv) Include stars with a surface gravity of $log(g) \leq 4.5$;
(v) Include stars with an effective temperature of $T_{eff} \leq 6400$;
(vi) Include stars with a J-band magnitude $12.0 \leq J \leq 14.0$;
(vii) Include stars with a J-K colour of $0.23 \leq J - K \leq 0.45 + 0.5E(B - V)$;

Discussing in detail the selection function, the $6.0 \leq r \leq 10.0$ kpc and $z \leq |2.0|$ kpc filters are specifically chosen to best replicate the analogous 2 kpc sphere selection region that is used with the simulation datasets. These distances are computed using the dust extinction model as described in Bergemann et al. (2014) which also accounts for the $+0.5E(B - V)$ term in the J-K colour component of the selection function.

The T_{eff} and the $log(g)$ fields of the selection function is specifically chosen ($log(g) \geq 3.5$ and $T_{eff} \geq 5400$ inclusive and $log(g) \leq 4.5$ and $T_{eff} \leq 6400$) since for stars with $T_{eff} \lesssim 5300 - 5500$ are not well determinable. This is clear from the HRD in which the main-sequence dwarfs show a strange up-turn pattern which does not exist according to stellar evolution. Likewise stellar ages are not very well determined for very hot stars with $T_{eff} \gtrsim 6500$. This selection would thus include a selection of red sub-giants $log(g) < 4.0$ and main sequence dwarves $log(g) > 4.0$.

In addition, we also remove stars with bad ages and metal abundances which have either non-sensible values for the respected age or [Fe/H] field or their errors . We further remove additional stars which have none sensible [Mg/Fe] values and errors for comparisons with simulation related [Mg/Fe] data and similarly for [O/Fe] values. This is because [Fe/H] is better quantifiable for these observations and it would be unnecessary to remove stars from the dataset if [Fe/H] values are fine but [O/Fe] is unknown for comparisons with [Fe/H]. After applying the selection function as described, we return 265 stars with correct age and [Fe/H] data; 208 stars with correct age, [Fe/H] and [Mg/Fe] data; 100 stars with correct age, [Fe/H], [Mg/Fe] and [O/Fe] data.

2.6 Different Ways of Analysing Simulated Solar Neighbourhood

The aim of this work is to best examine how different ways of processing the same simulation data give variations in the results obtained for the distribution of element abundances in the solar neighbourhood. A first-order approach that is commonly used is to simply take a spatial region within a simulation that matches the region of interest in a galaxy and compare that with observational data. This approach samples the entire stellar population (something that no observational survey does) and neglects the fact that simulations are not subject to observational errors meaning that the distributions are not directly comparable.

In this work, we compare the Gaia-ESO survey results, with the following variants of the simulated galaxy Selene-CH in order to demonstrate the influence of each component of the process used to mimic observational limits;

- *Selene-CH* is the the unaltered and unmodified galaxy. We select 1000 SPPs at random from within a 2 kpc sphere around the simulated observer. The SPPs are directly compared with the GES-DR4 results. This kind of direct comparison demonstrates the methodology employed in the 'traditional sense', i.e with spatial cuts alone. We have elected to select 1000 SPPs to give a similar number of points as our observational dataset. The SFH (Star Formation History) is unlikely to be a 1:1 match to the Milky Way's as illustrated in TODO CITATION. For example if the simulation possessed a massive star formation burst at early times, and 1000 composite star particles are selected, then we would overpopulate the distribution at early times. Since our own Milky Way's disk is more or less consistent with star formation over the past 10 Gyrs TODO citation, the spread of the composite particles that are selected is recommended to be equal in age. **GARETH: I still think the motivation for selection 1000 stars is pretty bogus GARETH: What is all this stuff about sampling the SFH, we don't make a selection based on this do we? BEN: It's literally just a random sample of the stars (which strictly speaking is a sample of the SFH specifically within the 2kpc sphere, although this is implied BEN: 1000 stars is easier on the eye plot than ~ 8000 in an age-met plot and reduces computational time slightly of SynCMD. Overall I agree, the motivation does not stand up on it's own per se since the whole advantage of having a simulation dataset is that we can sample more stars easily (and would reduce the noise *sightly*). Overall it won't make a difference with the results though**
- *Selene-GES* is a modification of the previous comparison. In this case we also apply stochastic scattering to the ages and abundance ratios of the simulated SPPs to emulate observational uncertainties. This scattering is purely based on the error bounds taken from the GAIA-ESO dataset to mimic the effect for the simulated stellar particles.
- *Selene-SYN* is the result of applying the SynCMD toolkit, as described in §2.3, to the scattered SPPs ages and metallicities in *Selene-GES* (i.e the statistically scattered results of Selene-CH). This dataset includes the application of any selection functions for $\log(g)$, $T_{\rm eff}$, J-band magnitude, and J-K colour. This dataset illustrates the effect of a more rigorous attempt to mimic the select effects inherent in the observations.

GARETH: A referee may ask why the scattering isn't applied

to the stars from synCMD which might be the more sensible order to apply the steps in. BEN: The scattering is applied to the stars which are used as the input for SynCMD. The stars that you see in Selene-GES are the inputs for Selene-SYN. More specifically, you transform like this. Selene-CH -¿ (scatter) Selene-GES -¿ (SynCMD) -¿ Selene-SYN, in that linear order. However what would be interesting to see is what happens if you run SynCMD on unscattered data (have I done this before? I think you get a very similar narrow distribution still),

In short, *Selene-CH* represents a first-order comparison between simulations of observations similar to that found in the majority of the literature, *Selene-GES* shows the effect of applying observational scatter, and *Selene-SYN* demonstrates the influence of selection effects. We now describe the post-processing used to create each of these samples in detail.

2.6.1 Selene-CH: The 'Standard' Simulation Approach

We identify a solar neighbourhood within the simulated galaxy Selene-CH, a position 8 kpc from the galactic centre on a spiral arm. This position is shown by a cross in Figure 1 at (x,y) = (4.0, 6.93) kpc; the solar neighbourhood is centred on this point. We have repeated the analysis that follows with stars from different positions at a galactocentric radius of 8 kpc and find that our results are robust to changes in the position of the simulated observer. This is due to the azimuthal homogeneity in the age and elemental abundances; the mean [Fe/H], [Mg/Fe] and [O/Fe] peak-to-peak variations are 0.01 dex, 0.005 dex and 0.02 dex. and the age variation is only 0.5 Gyr. **GARETH: What is a peak-to-peak variation? BEN: difference between maximum and minimum peaks**

We apply a kinematic disc selection to eliminate halo SPPs from our sample in the same manner as (Knebe et al. 2013), stars are SPPs are selected if they fulfill the following criteria, $0.7 \leq \frac{J_z}{J_{\rm circ}} \leq 1.1$, where $J_{\rm circ} = R \times \sqrt{\frac{GM(<R)}{R}}$ in which $M(<R)$ is the enclosed mass at radius R.

We normalise the element abundances to the Asplund et al. (2009) solar values and then apply a shift to our element abundance ratios such that stars with ages 4–5 Gyr have solar ratios. Previous works have applied different variations of *a posteriori* re-normalisations (e.g Pagel & Tautvaisiene 1995; François et al. 2004; Henry et al. 2010) and/or employed GCE models to infer revised sets of stellar yields (François et al. 2004). This may seem arbitrary, however the amount by which we normalise is not very large compared to the width of the distribution and we are primarily concerned here with the dispersion of the element ratios. Furthermore it is demonstrated in Few et al. (2014) that variations in abundance ratios (particularly those of α elements to Fe) are effectively shifted in the same way depending on the IMF. The need to apply such a shift implies that the sub-grid chemical evolution model is not quite correct which is hardly surprising given the uncertainties in the underlying yields and chemical evolution model. Therefore, while the renormalisation of abundance ratios introduces a slight inconsistency to the model it by no means negates our results.

Finally, we select 1000 SPPs from a 2 kpc sphere centred on the simulated observer. We select 1000 stars so that we have a similar number of points as the GES-DR4. A 2 kpc sphere is effectively the size of the region in the Gaia-ESO dataset due to magnitude limits so that we approximate the sample with a simple spatial selection.

2.6.2 Selene-GES

This dataset extends the methodology described above to generate *Selene-CH* by applying a stochastic scattering based on the GES-DR4 error bars for age abundance ratios to mimic the effect of the unavoidable uncertainties found in observations on the precisely known (but not necessary accurate) simulated values.

We degrade the precision of our abundance ratios on a particle-by-particle basis using a Gaussian distribution, centred on the original simulated value, that has a standard deviation equal to the mean error found in the GES-DR4 dataset: $\sigma_{\rm [Fe/H]}$=0.05, $\sigma_{\rm [O/Fe]}$=0.16, and $\sigma_{\rm [MgFe]}$=0.10. A new value for each SPP is chosen randomly from this distribution.

The age value for each SPP is also scattered this way except that the distribution from which the new value is chosen at random is not symmetric. The GES-DR4 ages have a mean lower age error of $\sigma_{age,l}$=0.72 Gyr and a mean upper age error of $\sigma_{age,u}$=1.49 Gyr. We construct a piecewise function from two half-Gaussians with these standard deviations respectively to scatter the simulated ages. This process not only broadens the distributions but also shifts the ages on average to be slightly older. **Any objections to using σ here instead of e for the std? BEN: Nope, that is better.**

2.6.3 Selene-SYN

Our final version of *Selene* takes the scattered SPP properties from *Selene-GES* and inputs those particles to SynCMD creating a third dataset referred to here as *Selene-SYN*. The mechanics of SynCMD are described in §2.3 but the key here is to split the SPPs in individual synthetic stars with a realistic distribution of star properties so that we can apply photometric, surface gravity and effective temperature cuts to exactly mimic the GES-DR4 method. The selection criteria are stated in §2.5 however we do not apply the dust extinction correction to the J-K upper limit because our synthetic stars are unaffected by dust. The synthetic stars that remain after this are used as our sample of stars analogous to the GES-DR4 dataset so that we can compare the simulations in a more like-for-like manner. **GARETH: We need to get this straight - syncmd doesn't actually output a number of stars so we are unable to select 1000 of them, but really the full number of stars from selene should be fed in to give a smoother distribution GARETH: Need to add the fourth line to the distribution function plots. GARETH: We shouldn't mass-weight the observations, then we shouldn't mass-weight the selene-syn data, but we should mass-weight the original simulation because it makes it more like a number of stars... now I have no idea whether syncmd does something like mass weight implicitely by ignoring remnants in the age distribution but it doesn't make a big difference really! GARETH: BEN: Yep, SynCMD does not *explicitly* output the number of stars, it outputs the gridded data (CMD), of which FeH abundance ratios are outputted from this, of which *I THINK* is simply normalised to the total number of stars that get through the selection function (see the Z axis in the synthetic plots). *Strictly speaking* you can grab each individual synthetic star particle *if you really want to*, since this is the result of using the IMF distribution by number... I will explain this in more detail tomorrow (or today) BEN: Stefano said in an email to me that SynCMD does mass weight the output, although really I *think* he meant numerical weighting**

3 RESULTS AND DISCUSSION

TODO: SOMEWHERE/SOMEHOW I need to cite this (Ness et al. 2015)

We discuss the results and implications of the comparison between GES-DR4, Selene-CH, Selene-GES and Selene-SYN. The results of this demonstrate the work of this preliminary study into the impact different 'observational methods' of simulation datasets compare with real observational datasets.

Included are the circular velocities (V_{circ}) of the gas, stellar populations, dark matter and the overall abundance. Also included is the comparison with the rotational velocity (V_{rot}) of the gas, cold gas, stars, young stars and stars within disk-like orbits (Fig 1). Alongside a face-on gas density plot (Figure 1) and a plot of the star formation and supernova rates (Figure **??**). The SFH of Selene-CH does not vary much more by a factor of 3, so the weighting in this case by 1000 stars is less necessary in this instance than in other datasets, for example Selene in Few et al. (2012b) which varies more than a factor of 3. Overall the SFH variation is small, and thus comparable to that of the Milky Way (TODO citation)

Figure 5 shows the age-metallicty relation, comparing the GES of Selene-CH, the post processing results from SynCMD Selene-SYN(the colour-map) and GES-DR4. The similarities between Selene-CH and GES-DR4 exists, in regards to the distribution. But the shape of the SynCMD results (henceforth known as Selene-SYN) shown in the colour map is narrower both in distribution and intensity in comparison to the simulated and observational datasets. The SynCMD results in particular do not map one-to-one with the original Selene-CH results, which is more particularly noticeable in the older stellar populations.

Figure 6 shows the magnesium abundance and age relation for Selene-CH, Selene-SYN and GES-DR4. Unlike the age-metallicity distribution for the Selene-SYN results, the age-magnesium relation shows more dispersion akin to the results of the Selene-CH simulation and the observational datasets.

Figure 7 shows the magnesium abundance and iron abundance relation for Selene-CH, Selene-SYN and GES-DR4. The narrower distribution in the Selene-SYN in the iron abundance plane is similar, but the shape of the distribution between the datasets is apparent.

The properties of the mass-weighted metallicity distribution function (MDF) is sensitive to the properties of the sub-grid physics (Pilkington et al. 2012), but is additionally claimed to be sensitive to how we observe it (Miranda, Macfarlane & Gibson 2015). Figure 8 shows the metallicity distribution function for Selene-CH, GES-DR4 and the results from SynCMD (Selene-SYN). We have chosen to compare with GES-DR4 since the photometric filers are consistent with what we have used within SynCMD. The mode/peak value of the three distributions appear at the same point ($\sim$ 0.0 dex), but other statistical properties of the MDF's vary. Namely the observational MDF is less skewed with a value of -0.984 in comparison to the simulation and synthetic results, with values of -1.48 and -0.65 respectively. The observational results is more positively skewed since our simulation model is not producing stars with metallicities comparably high with the observational datasets. The simulated (and thus synthetic) MDF has a narrow range of stars with [Fe/H] $>$ 0.0, whereas the observational datasets have a larger spread of stars from the statistical mean trend (including the more likely observation of stars that are more metal rich than CE models imply exist).

The kurtosis of the MDF results for Selene-CH, Selene-SYN and the GES-DR4 results is 3.24, 0.0814 and 2.28 respectively. The synthetic results are more peaked at 0.0 dex in comparison to the simulation results. The photometric boundary conditions in SynCMD result in removing redder and fainter stars (stars with higher J-K and higher apparent magnitude values). More noticeably, the knee in the MDF distribution at -0.2 dex in Figure 8 for Selene-SYN illustrates the sharp reduction in metal poor stellar population in comparison to the simulation results, in which is accounted for by the loss of fainter and older stars. Appendix A illustrates this point in more detail. In summary, analysing out simulation data in this way yields different results than what a spatial filter alone would produce, although to match the similar MDF's as found in observation, work is still needed to be done in modelling the formation rates and metal abundance of stars.

4 CONCLUSIONS

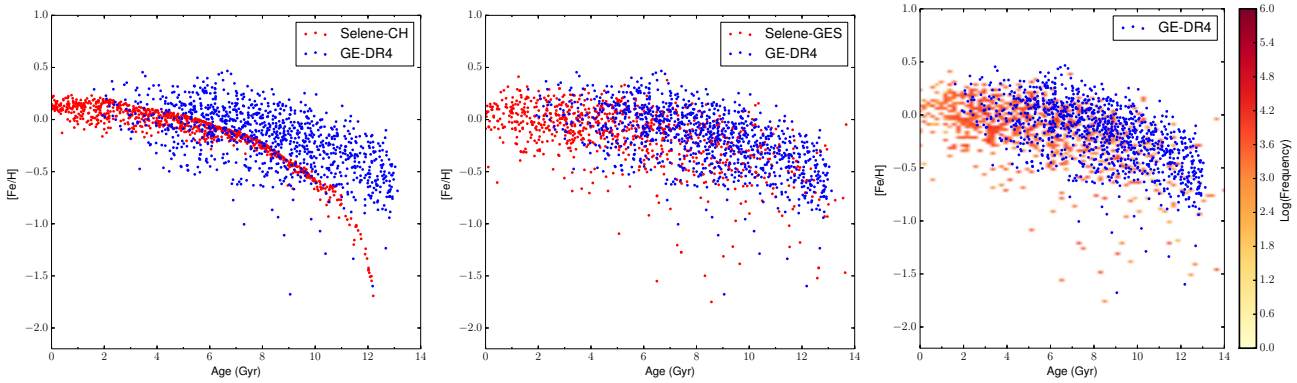
We compared the results from the fourth data release of the Gaia-ESO survey (GES-DR4), with a Milky Way-like galaxy simulated with the chemodynamical patch to RAMSES, RAMSES-CH. After running this simulation, we renormalise the results to match the solar metal abundances (0.0 dex) at the solar age (~5 Gyr) and convolve this with the errors of the gaia-eso survey. In addition, we follow up the convolved simulation results with SynCMD, to produce simulated observation results under the title of Selene-SYN.

The agreement between the realisation of a milky way-like galaxy with RAMSES-CH - after convolving observational errors with the simulated stellar particles - and the GES-DR4 along with the other observational datasets is remarkably good. Although observational datasets show a larger number of stars which deviate from the mean trend, the chemical evolution model in RAMSES-CH maps very closely the age-metal and metal-metal abundance relations. This shows that the chemical evolution model within RAMSES-CH is comparable with modern day galactic archaeology surveys such as the Gaia-ESO survey and reproduces observational trends.

Including observational selection effects and photometric limits within our simulated model removes the contribution of some red and faint (older) stars from the sample, and thus the metallicity distribution function. This approach generates a chemical evolution model which favours even more of an over abundance of stars with solar metallicity whilst effectively smoothing the distribution of stars below the solar metallicity and increasing the contribution favouring metal rich stars and bright (younger) stars.

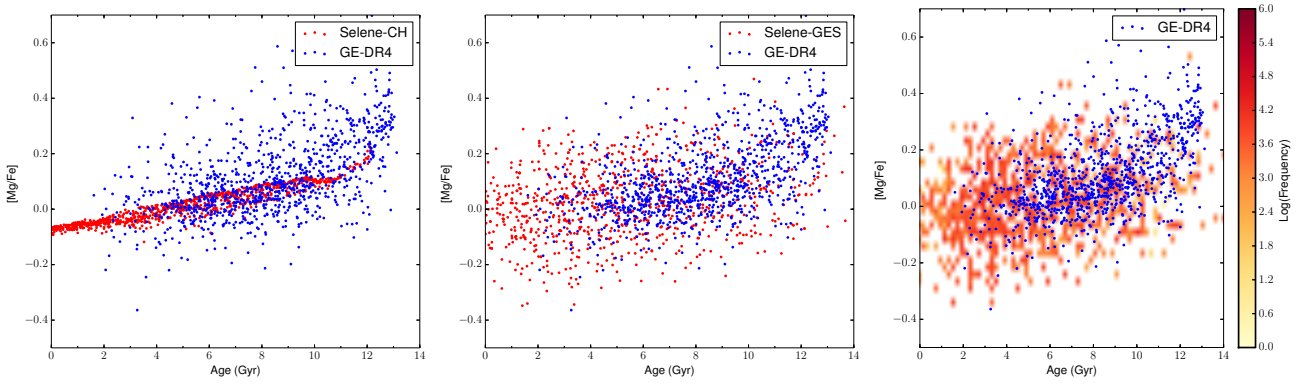
Future chemical evolution models should aim to produce a higher abundance of fainter, metal poor and older stars. Some of these will be filtered out from applying observational survey masks similar to the GAIA-ESO survey, as such the simulated MDF would ideally appear to be more skewed in appearance, as such applying observational survey masks would remove a sample of red and faint stars to match the observational sample (after convolving observational survey errors onto the simulation data).

The distribution of the age-metal and metal-metal plots in comparison between the observational, simulation and synthetic datasets are very remarkable, and demonstrates that chemical evolution models today are successful at producing the abundance distributions. Future work for chemical evolution models should aim to fit better for the lower and higher metallicity distribution of stars.



(a) (b) (c)

Figure 5. Plots of the present day (z = 0) age metallicty relation for Selene-CH, the fourth data release of the Gaia-ESO survey (GES-DR4) and the synthetic observational analysis approach Selene-SYN. Plot a) shows the comparison between GES-DR4 and Selene-CH, represented by blue and red circles respectively. Plot b) shows the comparison between GES-DR4 and Selene-GES. Plot c) shows the comparison of GES-DR4 wth Selene-SYN, of which Selene-SYN is represented by the heat map in which the transition from white to yellow to red represents an increase in the number of synthetic particles which occupying bins of 0.05 dex and 0.5 Gyr logarithmically.



(a) (b) (c)

Figure 6. Plots of the present day (z = 0) [Mg/Fe]-age relation for Selene-CH, the fourth data release of the Gaia-ESO survey (GES-DR4) and the synthetic observational analysis approach Selene-SYN. Plot a) shows the comparison between Selene-CH and GES-DR4, represented by red and blue circles respectively. Plot b) shows the comparison between GES-DR4 and Selene-GES. Plot c) shows the comparison of GES-DR4 wth Selene-SYN, of which Selene-SYN is represented by the heat map in which the transition from white to yellow to red represents an increase in the number of synthetic particles which occupying bins of 0.05 dex and 0.5 Gyr logarithmically.

ACKNOWLEDGEMENTS

We acknowledge the insightful comments and support provided by our colleagues, Stefano Pasetto and Ben Macfarlane. BBT acknowledges the support of STFC through its PhD Studentship programme (ST/F007701/1). The generous allocation of resources from STFCs DiRAC Facility (COSMOS: Galactic Archaeology), the DEISA consortium, co-funded through EU FP6 project RI-031513 and the FP7 project RI-222919 (through the DEISA Extreme Computing Initiative), and the PRACE-2IP Project (FP7 RI-283493), are gratefully acknowledged.

ACKNOWLEDGEMENTS

We acknowledge the insightful comments and support provided by our colleagues, Stefano Pasetto and Ben Macfarlane. BBT acknowledges the support of STFC through its PhD Studentship programme (ST/F007701/1). The generous allocation of resources from STFCs DiRAC Facility (COSMOS: Galactic Archaeology), the DEISA consortium, co-funded through EU FP6 project RI-031513 and the FP7 project RI-222919 (through the DEISA Extreme Computing Initiative), and the PRACE-2IP Project (FP7 RI-283493), are gratefully acknowledged.

REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 597, 21
Agertz O. et al., 2007, MNRAS, 380, 963
Aparicio A., Gallart C., Chiosi C., Bertelli G., 1996, ApJL, 469, L97
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Bergemann M., Kudritzki R.-P., Gazak Z., Davies B., Plez B., 2015, ApJ, 804, 113
Bergemann M. et al., 2014, A&A, 565, A89

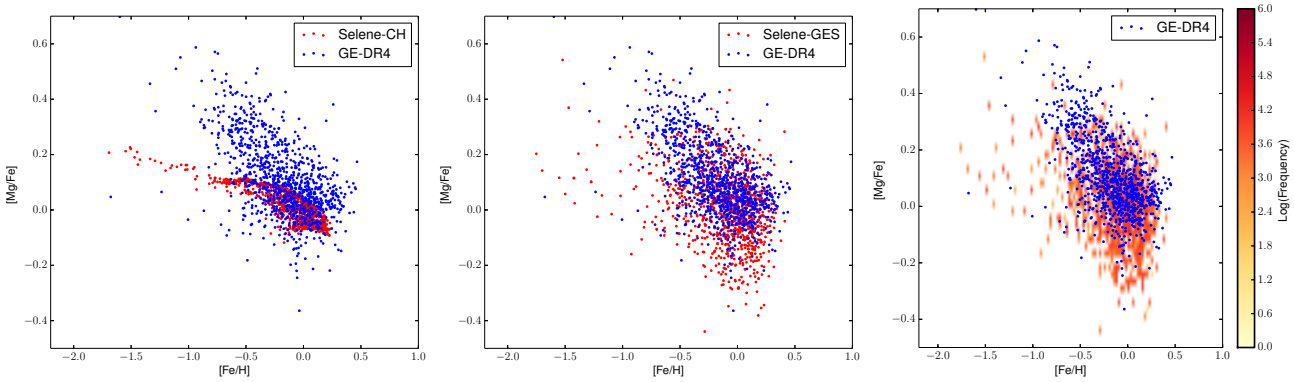


Figure 7. Plots of the present day (z = 0) [Mg/Fe]-[Fe/H] relation for Selene-CH, the fourth data release of the Gaia-ESO survey (GES-DR4) and the synthetic observational analysis approach Selene-SYN. Plot a) shows the comparison between Selene-CH and GES-DR4, represented by red and blue circles respectively. Plot b) shows the comparison between GES-DR4 and Selene-GES. Plot c) shows the comparison of GES-DR4 wth Selene-SYN, of which Selene-SYN is represented by the heat map in which the transition from white to yellow to red represents an increase in the number of synthetic particles which occupying bins of 0.05 dex logarithmically.

Bertelli G., Bressan A., Chiosi C., Ng Y. K., Ortolani S., 1995, A&A, 301, 381

Bertelli G., Girardi L., Marigo P., Nasi E., 2008, A&A, 484, 815

Bertelli G., Nasi E., 2001, AJ, 121, 1013

Bertelli G., Nasi E., Girardi L., Chiosi C., Zoccali M., Gallart C., 2003, AJ, 125, 770

Bertelli G., Nasi E., Girardi L., Marigo P., 2009, A&A, 508, 355

Carbon D. F., Barbuy B., Kraft R. P., Friel E. D., Suntzeff N. B., 1987, PASP, 99, 335

Chiosi C., Bertelli G., Bressan A., 1986, Mem. Societa Astronomica Italiana, 57, 507

Chiosi C., Bertelli G., Bressan A., Nasi E., Pigatto L., 1986, in Star-forming Dwarf Galaxies and Related Objects, Kunth D., Thuan T. X., Tran Thanh Van J., Lequeux J., Audouze J., eds., pp. 449–463

Chiosi C., Bertelli G., Meylan G., Ortolani S., 1989, A&A, 219, 167

Chiosi C., Greggio I., 1981, A&A, 98, 336

Edvardsson B., Andersen J., Gustafsson B., Lambert D. L., Nissen P. E., Tomkin J., 1993, A&AS, 102, 603

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748

Eisenstein D. J. et al., 2011, AJ, 142, 72

Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, PASP, 110, 761

Few C. G., Courty S., Gibson B. K., Kawata D., Calura F., Teyssier R., 2012a, MNRAS, 424, L11

Few C. G., Courty S., Gibson B. K., Michel-Dansac L., Calura F., 2014, MNRAS, 444, 3845

Few C. G., Gibson B. K., Courty S., Michel-Dansac L., Brook C. B., Stinson G. S., 2012b, A&A, 547, A63

François P., Matteucci F., Cayrel R., Spite M., Spite F., Chiappini C., 2004, A&A, 421, 613

Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487

Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399

Gilmore G. et al., 2012, The Messenger, 147, 25

Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, A&A, 486, 951

Haardt F., Madau P., 1996, ApJ, 461, 20

Heiter U., Soubiran C., Netopil M., Paunzen E., 2014, A&A, 561, A93

Henry R. B. C., Kwitter K. B., Jaskot A. E., Balick B., Morrison M. A., Milingo J. B., 2010, ApJ, 724, 748

Hopkins P. F., 2015, MNRAS, 450, 53

Knebe A. et al., 2013, MNRAS, 428, 2039

Knollmann S. R., Knebe A., 2009, ApJS, 182, 608

Kodama T., Arimoto N., 1997, A&A, 320, 41

Marigo P., Girardi L., 2007, A&A, 469, 239

Miranda M. S., Macfarlane B. A., Gibson B. K., 2015, ArXiv e-prints

Ness M., Hogg D. W., Rix H., Martig M., Pinsonneault M. H., Ho A. Y. Q., 2015, ArXiv e-prints

Ng Y. K., Bertelli G., Bressan A., Chiosi C., Lub J., 1995, A&A, 295, 655

Pagel B. E. J., Tautvaisiene G., 1995, MNRAS, 276, 505

Pasetto S., Chiosi C., Kawata D., 2012, A&A, 545, A14

Pilkington K. et al., 2012, A&A, 540, A56

Ramírez I., Allende Prieto C., Lambert D. L., 2007, A&A, 465, 271

Randich S., Gilmore G., Gaia-ESO Consortium, 2013, The Messenger, 154, 47

Read J. I., Lake G., Agertz O., Debattista V. P., 2008, MNRAS, 389, 1041

Reddy B. E., Lambert D. L., Allende Prieto C., 2006, MNRAS, 367, 1329

Rosen A., Bregman J. N., 1995, ApJ, 440, 634

Ruchti G. R., Bergemann M., Serenelli A., Casagrande L., Lind K., 2013, MNRAS, 429, 126

Ruiz-Lara T., Few C. G., Gibson B. K., Pérez I., Florido E., Minchev I., Sánchez-Blázquez P., 2016, A&A, 586, A112

Salpeter E. E., 1955, ApJ, 121, 161

Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2005, MNRAS, 364, 552

Serenelli A. M., Bergemann M., Ruchti G., Casagrande L., 2013, MNRAS, 429, 3645

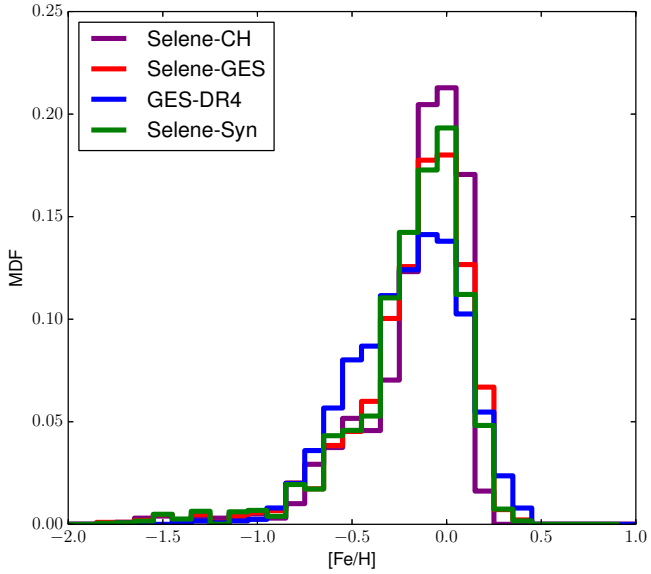
Smiljanic R. et al., 2014, A&A, 570, A122

Soubiran C., Jasniewicz G., Chemin L., Crifo F., Udry S., Hestroffer D., Katz D., 2013, A&A, 552, A64

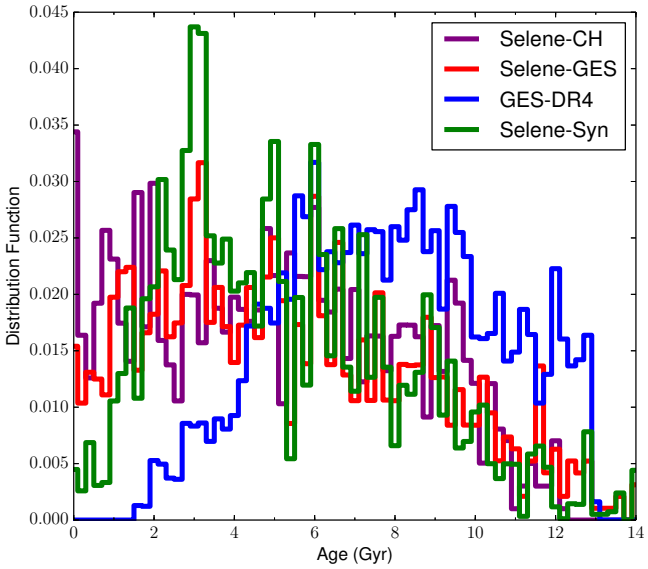
Springel V., 2010, MNRAS, 401, 791

Springel V. et al., 2005, Nature, 435, 629

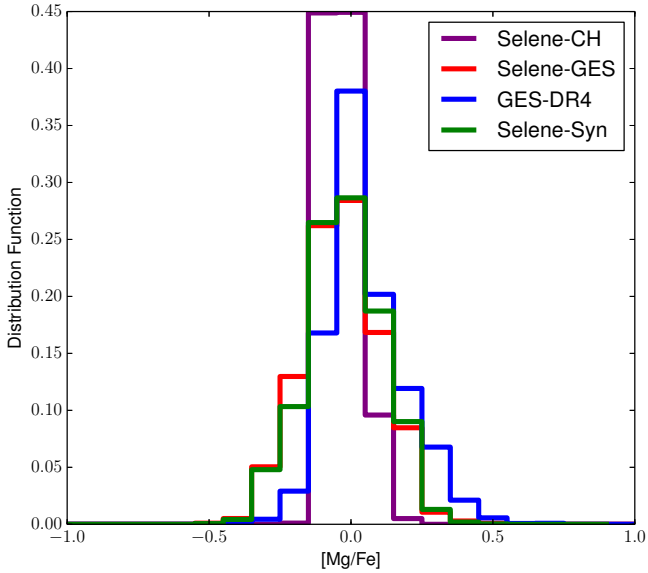
Tantalo R., Chinellato S., Merlin E., Piovan L., Chiosi C., 2010,



(a) skewness : SYN: -0.6504, Comp: -1.4839, Obs: -0.9836,, kurtosis : SYN: -0.0813, Comp: 3.249, Obs: 2.2846,



(b) skewness : SYN: -0.444, Comp: 0.178, Obs: 11.35,,kurtosis : SYN: -0.037, Comp: -0.843, Obs: 160.91,



(c) skewness : SYN: -0.7955, Comp: -0.029, Obs: 1.34,, kurtosis : SYN: -0.840, Comp: -0.033, Obs: 2.83 ,

Figure 8. A representation of the present day (z = 0) mass-weighted metallicity and age distribution functions with GES-DR4, Selene-GES and Selene-Syn represented with blue, red and green lines respectively. Plot a) shows the metallicity distribution function in [Fe/H] space, plot b) shows the age distribution function and plot c) shows the magnesium distribution function. The three metal distribution functions have a bin width 0.1 dex, and the age distribution function have a bin width of 0.5 Gyr.

A&A, 518, A43
Teyssier R., 2002, A&A, 385, 337
Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, MNRAS, 429, 3068
Timmes F. X., Woosley S. E., Weaver T. A., 1995, ApJS, 98, 617
Turk M. J., Smith B. D., Oishi J. S., Skory S., Skillman S. W., Abel T., Norman M. L., 2011, ApJS, 192, 9
van den Hoek L. B., Groenewegen M. A. T., 1997, A&AS, 123
Walker I. R., Mihos J. C., Hernquist L., 1996, ApJ, 460, 121
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181
York D. G. et al., 2000, AJ, 120, 1579