

# Ultra-faint dwarf galaxies: unveiling the minimum mass of the first stars

Martina Rossi,<sup>1,2</sup>★ Stefania Salvadori<sup>1,2</sup> and Ása Skúladóttir<sup>1,2</sup>

<sup>1</sup>Dipartimento di Fisica e Astrofisica, Università degli Studi di Firenze, via G. Sansone 1, Sesto Fiorentino 50019, Italy

<sup>2</sup>INAF/Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, Firenze 50125, Italy

Accepted 2021 March 16. Received 2021 March 16; in original form 2020 November 18

## ABSTRACT

The non-detection of zero-metallicity stars in ultra-faint dwarf galaxies (UFDs) can be used to constrain the initial mass function (IMF) of the first (PopIII) stars by means of a statistical comparison between available data and predictions from chemical evolution models. To this end we develop a model that follows the formation of isolated UFDs, calibrated to best reproduce the available data for the best studied system: Boötes I. Our statistical approach shows that UFDs are the best suitable systems to study the implications of the persisting non-detection of zero-metallicity stars on the PopIII IMF, i.e. its shape, the minimum mass ( $m_{\min}$ ), and the characteristic mass ( $m_{\text{ch}}$ ). We show that accounting for the incomplete sampling of the IMF is essential to compute the expected number of long-lived PopIII stars in inefficiently star-forming UFDs. By simulating the colour–magnitude diagram of Boötes I, and thus take into account the mass-range of the observed stars, we can obtain even tighter constrains on  $m_{\min}$ . By exploiting the 96 stars with measured metallicities ( $i < 19$ ) in the UFDs represented by our model, we demonstrate that  $m_{\text{ch}} > 1 M_{\odot}$  or  $m_{\min} > 0.8 M_{\odot}$  at 99% confidence level. This means that a present-day IMF for PopIII stars is excluded by our model, and a top-heavy PopIII IMF is strongly favoured. We can limit  $m_{\min} > 0.8 M_{\odot}$  independent of the PopIII IMF shape by targeting the four UFDs Boötes I, Hercules, Leo IV, and Eridanus II with future generation instruments, such as ELT/MOSAIC ( $i < 25$ ), which can provide samples of  $> 10,000$  stars.

**Key words:** stars: luminosity function, mass function – stars: Population III – galaxies: dwarf – cosmology: theory – dark ages, reionization, first stars .

## 1 INTRODUCTION

The first stars are expected to have played a crucial role in the evolution of the primordial Universe since they represent the first sources of ionizing photons, dust, and chemical elements heavier than helium, i.e. metals. The amount of ionizing photons, dust, and different chemical species produced by the first stars strongly depends on their mass. Thus, understanding the initial mass function (IMF) of the first stars is a fundamental problem in cosmology. Yet, this function is very difficult to infer, both from an observational and a theoretical perspective.

First stars, also known as Population III (PopIII) stars, are predicted to form at  $z \sim 15$ –30, in so-called *minihaloes* with masses  $M_h \approx 10^6 M_{\odot}$  and virial temperature  $T_{\text{vir}} \leq 10^4$  K (e.g. Bromm, Coppi & Larson 2002; Yoshida et al. 2003). Since they form out of gas of primordial composition they are expected to be completely metal-free. These metal-free stars should produce key emission lines that might allow us to easily identify their host galaxies (e.g. Schaefer & Pelló 2002). However, since PopIII stars predominantly formed in poorly star forming minihaloes, they are likely too faint to be detected, even with new-generation telescopes (e.g. Wise et al. 2012). Therefore, we have to deal with a lack of direct observations of PopIII stars to infer their mass distribution.

Our current knowledge of the mass properties of PopIII stars derives on one hand from computer simulations, and on the other

from Stellar Archaeology. The latter indirectly provides information about the first stars by exploiting the stellar chemical abundances measured in the oldest and most metal-poor components of our Milky Way and its dwarf galaxies satellites (e.g. McWilliam 1998; Freeman & Bland-Hawthorn 2002; Gratton, Sneden & Carretta 2004; Venn et al. 2004; Beers & Christlieb 2005; Tolstoy, Hill & Tosi 2009; Frebel & Norris 2015; Skúladóttir et al. 2015; Ji et al. 2016; Bonifacio et al. 2018; Li, Tan & Zhao 2018; Starkenburg et al. 2018; Aguado et al. 2019; Bonifacio et al. 2019; Sestito et al. 2019; Reichert et al. 2020; Chiti et al. 2021). Various different works (e.g. Hartwig et al. 2015; Salvadori, Skuladottir & Tolstoy 2015; de Bennassuti et al. 2017; Magg et al. 2017; Ishigaki et al. 2018) have interpreted these results as indirect evidence of the existence of massive PopIII stars ( $[10$ –\$60]  $M_{\odot}$ ), as well as the proof of existence of pair-instability supernovae,  $m_{\text{PopIII}} = [140$ –\$260]  $M_{\odot}$  (Aoki et al. 2014; Salvadori et al. 2019; Tarumi, Hartwig & Magg 2020).

In parallel, hydrodynamic simulations of PopIII stars formation predict that the first stars are more massive than stars formed today, with characteristic masses of  $[10$ –\$20]  $M_{\odot}$  and the maximum mass possibly extending up to  $1000 M_{\odot}$  (Omukai & Palla 2001; Abel, Bryan & Norman 2002; Bromm et al. 2002; Tan & McKee 2004; Yoshida et al. 2006; O’Shea 2007; Hosokawa et al. 2011; Bromm 2013; Hirano et al. 2014). However, 3D simulations that study the cooling of primordial gas reveal that proto-stellar gas clouds can experience strong fragmentation and PopIII stars could have masses lower than  $1 M_{\odot}$  (e.g. Machida, Matsumoto & Inutsuka 2008; Turk 2009; Smith et al. 2010; Clark et al. 2011; Greif et al. 2011; Dopcke et al. 2013; Stacy et al. 2013; Stacy & Bromm 2014; Stacy, Bromm

\* E-mail: martina.rossi@unifi.it

& Lee 2016; Susa 2019; Wollenberg et al. 2020). Yet, we have to stress that these high-resolution simulations follow the evolution of the proto-star for a very short time compared to the time-scale of star formation. Hence, we do not know if these fragments will eventually merge into the central object (e.g. Hirano & Bromm 2017) or can be thrown out and form long-lived first stars that may have survived until today.

As a result of the unpredictable and stochastic star formation process, the mass distribution of the first stars is expected to cover a wide range of masses, PopIII  $m_\star = [0.1\text{--}1000]M_\odot$ , possibly reaching down into the substellar regime. However, if PopIII stars with mass below  $0.8M_\odot$  were able to form they should be alive today, in particular in the oldest and least metal-poor systems of the Local Group, where stars can be observed individually.

Despite long searches, zero-metallicity stars have never been found among the ancient stars in our Milky Way and its dwarf galaxy satellites. The apparent absence of long-lived PopIII stars can be attributed to two main reasons: the first is that PopIII stars are so rare and challenging to find that we have not caught them yet; the second is that PopIII stars were more massive than present-day stars and hence did not survive until today, i.e. with minimum mass  $m_{\min} > 0.8 M_\odot$  (e.g. Oey 2003; Salvadori, Schneider & Ferrara 2007; Tumlinson 2007; Salvadori et al. 2010). Non-detection of long-lived Pop III stars can thus be an instrument to constrain the lower mass limit of the PopIII IMF. Using a statistically approach, Hartwig et al. (2015) have, however, recently shown that is not so straightforward to exclude the existence of any PopIII stars survivors and set PopIII  $m_{\min} \geq 0.8 M_\odot$ . Indeed, a large observed stellar sample is required:  $\sim 10^7$  stars for the Milky Way's halo and  $\sim 10^{10}$  for the bulge. However, these numbers are much larger than what can be expected in the largest spectroscopic surveys in the next decade (e.g. 4MOST: de Jong et al. 2019).

At the moment therefore we have only a few observational-driven constraints on the IMF of the PopIII stars and we are not even able to exclude that it was equal to present-day stellar IMF. This brings us to the question: How can we further constrain the lower mass end of the PopIII IMF?

Among all environments hosting ancient stars, ultra-faint dwarf (UFD) galaxies ( $L_{\text{bol}} < 10^5 L_\odot$ ) are the best places to look for the most metal-poor stars. This has been shown both theoretically (e.g. Salvadori & Ferrara 2009; Salvadori et al. 2010; Magg et al. 2017) and observationally (e.g. Ishigaki et al. 2014; Webster, Bland-Hawthorn & Sutherland 2015; Frebel et al. 2016; Pakhomov et al. 2019). By modelling early star formation, Magg et al. (2017) estimate the numbers of surviving PopIII stars in the Milky Way and its dwarf companions. Their results show that UFD satellites are the best systems to look for these stellar relics, since they typically contain a larger fraction of Pop III star survivors, up to 0.74 per cent of the total stellar mass. Furthermore, many theoretical studies have shown that UFD galaxies have the simplest assembly histories with no major merger events. Unlike essentially all larger systems, they underwent little to no further evolution and have survived to the present day as pristine relics from the early Universe. For this reason they are generally considered the ‘fossils’ of the minihaloes that hosted the first stars (e.g. Ricotti & Gnedin 2005; Bovill & Ricotti 2009; Salvadori & Ferrara 2009; Bland-Hawthorn, Sutherland & Webster 2015; Salvadori et al. 2015).

Observationally, UFD galaxies are the most common dwarf galaxies in the Local Group, representing more than 50 per cent of the total number of dwarf satellites (McConnachie 2012). They are the oldest, most dark matter-dominated, most metal-poor, least luminous, and least chemically evolved stellar systems known

(Simon 2019). Since they have the smallest dark matter haloes they are more sensitive to feedback processes, representing the extreme limit of the galaxy formation processes. Most UFD galaxies formed more than 75 per cent of their stars in the first Gyr of evolution and hence have old stellar populations,  $\gtrsim 10$  Gyr (e.g. Brown et al. 2014). Recent *Hubble Space Telescope* (HST) observations of Eridanus II revealed that in this distant UFD galaxy  $> 75\%$  of the stars have ages  $> 13$  Gyr and possibly formed in a very short burst, lasting  $< 200$  Myr (Gallart et al. 2021). Finally, UFDs are the galaxies that contain the highest fraction of extremely metal-poor stars,  $[\text{Fe}/\text{H}] < -3$  (e.g. Kirby et al. 2013; Salvadori et al. 2015).

In this work, we investigate the frequency of PopIII star survivors in UDF galaxies to constrain the minimum mass,  $m_{\min}$ , and the characteristic mass,  $m_{\text{ch}}$ , of the first stars. To this end we develop a semi-analytical model that follows the formation and evolution of an isolated UFD galaxy. The ideal target for our study is Boötes I since this is the UFD galaxy with the best observational data (e.g. Frebel 2010; Norris et al. 2010; Lai et al. 2011; Gilmore et al. 2013; Frebel et al. 2016), and it has also been studied extensively from a theoretical point of view (e.g. Romano et al. 2014; Vincenzo et al. 2014). The colour–magnitude diagram (CMD) of Boötes I reveals that its stellar population is old and metal-poor, with a mean age  $\sim (13.3 \pm 0.3)$  Gyr (Brown et al. 2014). Given the higher luminosity of Boötes I ( $L_\star \sim 10^{4.5} L_\odot$ ) with respect to other UFDs, this is also the only UFD galaxy for which we can derive an almost complete metallicity distribution function (MDF; Norris et al. 2010; Lai et al. 2011; Gilmore et al. 2013).

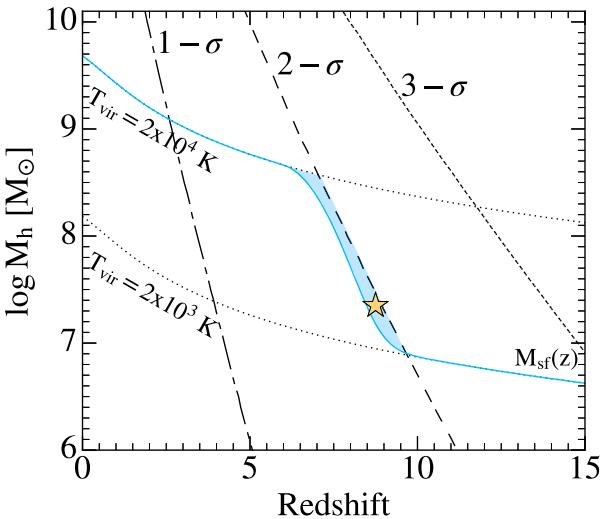
The article is structured as follows: Section 2 describes the main features of the semi-analytic model whose data-calibration process is presented in Section 3. In Section 4, it is explained why it is fundamental to account for the incomplete sampling of the stellar IMF in the modelling of poorly star-forming UFDs, in particular to compute the expected fraction of PopIII stars survivors. The observationally driven constraints for the minimum mass of PopIII stars are presented in Section 5, while in Section 6 we demonstrate how to get stronger constraints by computing the synthetic CMD, and thus compare model predictions with truly observed stars. Finally the summary and the discussion of results are in Section 7.

## 2 MODEL DESCRIPTION

Our semi-analytic model follows the star formation and the chemical enrichment history of a UFD galaxy from the epoch of its virialization ( $z = z_{\text{vir}}$ ) until present day ( $z = 0$ ). We model the evolution of present-day UFD galaxies by assuming that they experienced a ‘quiet’ assembling history, i.e. without major merger events. The evolution in isolation is a good approximation for these tiny systems, as has been shown by cosmological models and simulations for the Local Group formation (e.g. Salvadori & Ferrara 2009; Salvadori et al. 2015; Safarzadeh et al. 2018), as well as recent observations (Gallart et al. 2021).

### Initial conditions:

The nature of UFD galaxies is an ongoing debate. However, their observed properties suggest that they could be the living relics of the first star-forming minihaloes that hosted PopIII stars. This scenario has been investigated, for example, by Salvadori & Ferrara (2009) that explored the origin of UDF galaxies using a cosmological model for the formation of the Milky Way and its dwarf satellites, including UFD galaxies. These authors selected Milky Way's satellite candidates among star-forming haloes that correspond to  $< 2 - \sigma$



**Figure 1.** Evolution of the minimum halo mass to form stars,  $M_{\text{sf}}(z)$  (solid blue line), of the halo mass corresponding to  $1\sigma$ ,  $2\sigma$ ,  $3\sigma$  fluctuations of the density field (long-short, long, and short dashed lines), and of the halo mass with virial temperature  $T_{\text{vir}} = 2 \times 10^3 \text{ K}$  and  $T_{\text{vir}} = 10^4 \text{ K}$  (dotted lines) adapted from Salvadori & Ferrara (2009). The light blue shaded area shows the condition to form satellite dwarf galaxies,  $M_{\text{sf}}(z) < M_h < M_{2\sigma}$ . The yellow star represents our choice for the initial conditions:  $M_h = 10^{7.35} M_{\odot}$  and  $z_{\text{vir}} = 8.7$ .

density fluctuations<sup>1</sup> because these objects are the most likely to become satellites (Diemand, Madau & Moore 2005). By comparing their model results with observations, they concluded that UFD galaxies are associated with minihaloes formed prior to the end of reionization,  $z_{\text{rei}} > 6$ , which have not experienced mergers, i.e. they evolved in isolation (see also Salvadori et al. 2015). Following this approach we select the dark matter haloes that correspond to  $<2 - \sigma$  fluctuations of the density field<sup>2</sup> and that are able to form stars. The latter condition is described by the minimum halo mass required to allow star formation,  $M_{\text{sf}}(z)$ , see Fig. 1, whose evolution accounts for the increasing Lyman Werner and ionizing radiation (see section 2.1 of Salvadori & Ferrara 2009). In Fig. 1, our UFD galaxies candidates are those with dark matter halo mass  $M_{\text{sf}}(z) < M_h < M_{2\sigma}$  (indicated with the shaded light blue area). Following the results of Salvadori et al. (2015), we individuated the progenitor of Boötes I in the halo with dark matter halo  $M_h = 10^{7.35} M_{\odot}$  that virializes at  $z_{\text{vir}} = 8.7$  (yellow star in Fig. 1) and adopt these values as initial condition of our model.

We assume that at the epoch of virialization the dark matter halo contains a total amount of gas equal to  $\Omega_b/\Omega_m M_h$ . The star formation becomes possible when the gas begins to infall in the central part of the halo and to cool down. Considering the redshift of virialization, the infalling gas can be assumed to be metal-free (e.g. Salvadori & Ferrara 2012).

#### Infall rate:

Following Salvadori, Ferrara & Schneider (2008) and Salvadori et al. (2015), we exploit the results of simulations presented in Kereš

<sup>1</sup>The quantity  $\sigma(M, z)$  represents the linear rms density fluctuation smoothed with a top-hat filter of mass  $M$  at redshift  $z$ .

<sup>2</sup>We adopt a cold dark matter ( $\Lambda$ CDM) with  $h = 0.669$ ,  $\Omega_m = 0.3103$ ,  $\Omega_\Lambda = 0.6897$ ,  $\Omega_b h^2 = 0.02234$ ,  $\sigma_8 = 0.8083$ ,  $n = 0.9671$ , according to the latest Planck results (Planck Collaboration VI 2018).

et al. (2005) to compute the infall rate as

$$\frac{dM_{\text{inf}}}{dt} = A \left( \frac{t}{t_{\text{inf}}} \right)^2 \exp \left( -\frac{t}{t_{\text{inf}}} \right), \quad (1)$$

where  $A$  is the normalization constant, which is set to be  $A = 2M_h \frac{\Omega_b}{\Omega_m} \frac{1}{t_{\text{inf}}}$  so that for  $t \rightarrow \infty$  the accreted gas reaches the initial value,  $M_{\text{inf}}(\infty) = \Omega_b/\Omega_m M_h$ . In addition,  $t$  is the time since virialization and  $t_{\text{inf}}$  is the time-scale on which the gas cools, which is a free parameter of the model.

**Star formation rate:** It has been assumed that star formation occurs in a free-fall time  $t_{\text{ff}} = (3\pi/32G\rho(z))^{1/2}$ , where  $G$  is the gravitational constant and  $\rho$  the total mass density inside the halo. Stars are assumed to form in a single burst at each time-step,  $dt$ , and the star formation rate,  $\Psi$ , is given by

$$\Psi = \frac{dM_{\star}}{dt} = \epsilon_{\star} \frac{M_{\text{gas}}}{t_{\text{ff}}(z)}, \quad (2)$$

where  $M_{\star}$  is the total stellar mass formed,  $M_{\text{gas}}$  is the mass of the cold gas, and  $\epsilon_{\star}$  is the star formation efficiency, a second free parameter of our model.

#### PopIII and PopII/I stars:

The metallicity of the cold gas of the interstellar medium (ISM),  $Z_{\text{gas}} = M_{\text{Z}}^{\text{ISM}}/M_{\text{gas}}$ , determines whether the stars forming are PopIII or PopII/I stars. PopIII stars only form if the metallicity of the gas is lower than the critical metallicity,  $Z_{\text{cr}} = 10^{-4.5} Z_{\odot}$  (de Bennassuti et al. 2017), instead PopII/I stars form if  $Z_{\text{gas}} > Z_{\text{cr}}$ . At each time-step we compute the stellar mass formed, and we assume that the star formation occurs only if a minimum value is reached, i.e. if  $M_{\star}(t_i) \geq M_{\star}^{\min} = 100 M_{\odot}$  (Klessen, Krumholz & Heitsch 2011).

#### Stellar IMF:

The stellar mass formed at each time-step is distributed according to an IMF. As a first approximation we assume that the stellar mass formed is enough to fully populate the IMF. In other words, the total mass of stars formed at each time-step is distributed throughout the overall stellar mass range. As stellar IMF for PopII/I stars we choose a Larson-type (Larson 1998):

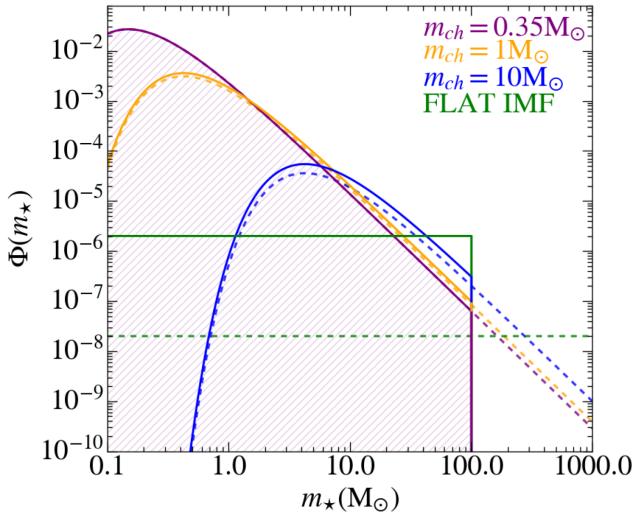
$$\phi(m_{\star}) = \frac{dN}{dm_{\star}} \propto m_{\star}^{-2.35} \exp \left( -\frac{m_{\text{ch}}}{m_{\star}} \right), \quad (3)$$

where the stellar mass  $m_{\star}$  varies in the mass range  $[0.1-100] M_{\odot}$  and  $m_{\text{ch}} = 0.35 M_{\odot}$  is the characteristic mass. Since the PopIII IMF is unknown, we initially assume it to be equal to the IMF of present-day stars (3, see Fig. 2) and we explore the impact of different minimum mass of PopIII stars,  $m_{\min}$ . Then (Section 5) we vary both the maximum mass of PopIII stars,  $m_{\max}$ , and the PopIII IMF shape by assuming three different  $m_{\text{ch}}$  along with a flat IMF. All our explored PopIII IMFs are summarized in Fig. 2.

#### Stellar evolution:

The model accounts for the finite stellar lifetimes, i.e. for the evolution of stars with different masses. We use the relation derived by Raiteri, Villata & Navarro (1996) in which the stellar lifetime depends not only on the stellar mass but also on the stellar metallicity,  $t_{\star}(m_{\star}, Z_{\star})$ , which is settled by the ISM out of which stars formed. All stars whose lifetime is greater than the difference between the age of the Universe at  $z = 0$  ( $\sim 13.8$  Gyr) and the age of the Universe when they formed,  $t_{\star} > (t_{z=0} - t_{z_{\text{form}}})$ , can survive until today (see details in Salvadori et al. 2008). In particular, according to the adopted lifetime-mass-metallicity relations, all stars with  $m_{\star} \leq 0.8 M_{\odot}$  have a lifetime greater than the age of the Universe and can therefore survive until today independent on their formation redshift.

(a) *Returned fraction and metal yields:*



**Figure 2.** Normalized PopIII IMFs explored in our model (see labels). The solid lines represent IMFs in the mass range  $[0.1\text{--}100]\text{M}_\odot$ , while dashed lines have maximum mass,  $m_{\max} = 1000\text{M}_\odot$ . The hatched area highlights the IMF of present-day stars (Larson 1998).

Stars return gas into the ISM, via supernova (SN) explosions or via stellar winds in the asymptotic giant branch (AGB) phase. This gas becomes thus newly available for star formation. We compute the rate at which the gas is returned into the ISM as

$$\frac{dR(t)}{dt} = \int_{m_{\text{turn}}(z)}^{100\text{M}_\odot} (m_* - w_m(m_*)) \Phi(m_*) \Psi(t - t_*) dm_*, \quad (4)$$

where  $w_m(m_*)$  is the remnant mass of a star with initial mass  $m_*$ ,  $\Psi$  the star formation rate, and  $t_* = t_*(m_*, Z_*)$  is the stellar lifetime.

The gas that is returned to the ISM carries with it the products of nuclear burning, i.e metals ( $Z$ ). The rate of heavy elements newly synthesized inside stars and re-ejected into ISM is

$$\begin{aligned} \frac{dY_Z(t)}{dt} = & \int_{m_{\text{turn}}(z)}^{100\text{M}_\odot} (m_* - w_m(m_*) - m_Z(m_*, Z_*)) Z(t - t_*) \\ & + m_Z(m_*, Z_*) \Phi(m_*) \Psi(t - t_*) dm_*, \end{aligned} \quad (5)$$

where  $m_Z(m_*, Z_*)$  is the mass of heavy elements produced by a star with initial mass  $m_*$  and metallicity  $Z_*$ , and  $Z(t - t_*)$  is the abundance of metals in the ISM at the time  $(t - t_*)$ . In our model we follow the evolution of the total amount of heavy elements ( $Z$ ) and of iron (Fe). The values used in our model to compute the metal yields and returned fraction of gas are those derived by Woosley & Weaver (1995) for both PopII/I and PopIII stars with masses  $< 100\text{M}_\odot$ , while for PopIII stars with  $140\text{M}_\odot \leq m_{\text{PopIII}} \leq 260\text{M}_\odot$  we adopt the yields by Heger & Woosley (2002).

#### Mechanical feedback:

Stars with masses  $m_* = [8\text{--}40]\text{M}_\odot$  end their life as SN. SNe explosions power a blast wave which, if sufficiently energetic, may overcome the gravitational pull of the host halo leading to expulsion of gas and metals into the intergalactic medium (IGM). We assume the mass of gas ejected into the IGM due to SN explosions to be regulated by the equation

$$\frac{1}{2} M_{\text{ej}} v_{\text{esc}}^2 = E_{\text{SN}}, \quad (6)$$

where  $E_{\text{SN}} = \epsilon_w N_{\text{SN}} \langle E_{\text{SN}} \rangle$  is the total kinetic energy injected by SN-driven winds,  $N_{\text{SN}}$  the number of SNe,  $\langle E_{\text{SN}} \rangle$  the average SN

explosion energy, and  $\epsilon_w$  is a free parameter, which controls the conversion efficiency of the explosion energy into kinetic form. Assuming the IMF described by 3, we get that the number of SNe per unit of stellar mass formed  $\eta_{\text{SN}}$  is

$$\eta_{\text{SN}} = \frac{\int_{8\text{M}_\odot}^{40\text{M}_\odot} m_* \Phi(m_*) dm_*}{\int_{0.1\text{M}_\odot}^{100\text{M}_\odot} \Phi(m_*) dm_*} \approx 10^{-2}. \quad (7)$$

For the average value of the SN energy we assume  $\langle E_{\text{SN}} \rangle \sim 10^{51}$  erg for PopII SNe, which is a typical value for a normal core-collapse SNe. For massive PopIII stars ( $[140\text{--}260]\text{M}_\odot$ ), i.e. Pair Instability SNe (PISN), we use the mass–energy relation of Heger & Woosley (2002), in which the energy increases with the mass of PISN. For less massive PopIII stars ( $[8\text{--}40]\text{M}_\odot$ ) we assume the same energy of PopII SNe.

Following Salvadori et al. (2008) we differentiate equation (6) to get the gas ejection rate:

$$\frac{dM_{\text{ej}}}{dt} = \frac{2 \epsilon_w \langle E_{\text{SN}} \rangle}{v_{\text{esc}}^2} \frac{dN_{\text{SN}}}{dt} \approx \frac{2 \epsilon_w \langle E_{\text{SN}} \rangle}{v_{\text{esc}}^2} 10^{-2} \frac{dM_*}{dt} \quad (8)$$

and compute the escape velocity of the gas as a function of the halo virial radius,  $v_{\text{esc}}^2 = \frac{GM}{r_{\text{vir}}}$  (Barkana & Loeb 2001).

In our model we assume that the metallicity of the ejected gas is the same as that of the ISM. This is an approximation, we are assuming that there is a complete mixing between the gas inside the halo and the heavy elements injected into the ISM by SN explosions.

A limitation of this study is that only core-collapse SNe ( $m_* = [8\text{--}40]\text{M}_\odot$ ) have been considered for PopII/I stars and we have not included the contribution of SN Ia. The choice was guided by the fact that the evolutionary time-scales associated with SN Ia are typically very long,  $>1\text{Gyr}$ , compared to the evolution time-scale of our galaxy ( $\sim 500$  Myr). Although there are studies that have shown that the evolutionary time-scale of SN Ia can be shorter (e.g. Matteucci et al. 2006), cosmological models for UFDs formation that also include prompt formation of SNIa show that their influence on the chemical evolution is extremely limited, and is only relevant at the highest [Fe/H] (Salvadori et al. 2015). In fact, a detailed chemical abundance study of Boötes I did not reveal significant SNIa contribution (Gilmore et al. 2013). Therefore, their contribution will only be considered in future work. Finally, we considered the evolution of an isolated UFD galaxy, however to obtain a more accurate and realistic model the cosmological context should be considered.

#### Galaxy evolution:

To follow the evolution of a UDF galaxy we solve equation (2) along with the subsequent system of differential equations:

$$\frac{dM_{\text{gas}}}{dt} = -\Psi + \frac{dR}{dt} + \frac{dM_{\text{inf}}}{dt} - \frac{dM_{\text{ej}}}{dt} \quad (9)$$

$$\frac{dM_Z}{dt} = -Z^{\text{ISM}} \Psi + \frac{dY_Z}{dt} + Z^{\text{inf}} \frac{dM_{\text{inf}}}{dt} - Z^{\text{ISM}} \frac{dM_{\text{ej}}}{dt}, \quad (10)$$

where  $M_{\text{gas}}$  in equation (9) is the mass of cold gas inside the halo, which increases due to the infall rate (equation 1) and to the amount of gas released by massive and AGB stars (equation 4), and it decreases due to star formation (equation 2) and gas ejection caused by SN explosions (equation 8). The second equation, equation (10), describes the variation of the total mass of metals ( $M_Z$ ) in gas where  $Z^{\text{ISM}}$ ,  $Z^{\text{inf}}$ , and  $Z^{\text{ej}}$  are, respectively, the metallicity in the ISM, in the infalling gas, and in the ejected gas. In our model we assume  $Z^{\text{inf}} = 0$ , i.e. the infalling gas is considered metal (and iron free) and  $Z^{\text{ej}} = Z^{\text{ISM}}$  (see previous paragraph). Thus, we can write the fourth

equation that describes the evolution of the iron mass in the ISM as

$$\frac{dM_{\text{Fe}}}{dt} = -\frac{M_{\text{Fe}}^{\text{ISM}}}{M_{\text{gas}}} \Psi + \frac{dY_{\text{Fe}}}{dt} - \frac{M_{\text{Fe}}^{\text{ISM}}}{M_{\text{gas}}} \frac{dM_{\text{ej}}}{dt}, \quad (11)$$

where  $M_{\text{Fe}}^{\text{ISM}}$  is the total iron mass in the ISM.

## 2.1 Numerical code

To follow the evolution of the UFD galaxy through cosmic time we solve the set of the differential equations (2–11), integrating them step by step. To this end we build a grid of cosmic time values from  $t_{\text{form}} = t(z_{\text{vir}} = 8.7)$  to  $t(z = 0)$  with a constant time-step  $\Delta t = (t_{i+1} - t_i) = 1$  Myr. This time-step is adequate to describe the formation and evolution of the stars and ensures that  $\Delta t < t_f(z)$ , thus allowing us to accurately follow the star formation. Furthermore, it is smaller than the lifetime of the most massive stars contributing to chemical enrichment,  $\Delta t < t_*(260 M_\odot) \sim 4$  Myr.

To make comparison with observations, we derive the cumulative star formation history (*Cumulative SFH*) that is defined as the ratio between the stellar mass formed until  $t_i = t(z_i)$  and the *total* stellar mass formed, i.e. until  $z = 0$ :

$$\text{Cumulative SFH}(i) = \frac{\sum_{t_i=t_{\text{vir}}}^{t(z_i)} M_{\star}(t_i)}{\sum_{t_i=t_{\text{vir}}}^{t(z=0)} M_{\star}(t_i)} \quad (12)$$

At each time-step  $t_i$ , the evolution of the mass of gas,  $M_{\text{gas}}(t_i)$ , metals,  $M_{\text{Z}}(t_i)$ , and iron,  $M_{\text{Fe}}(t_i)$  are computed. These quantities allow us to calculate the metallicity of the stars,  $Z_{\star} = M_{\text{Z}}/M_{\text{gas}}$ , and their iron abundance  $[\text{Fe}/\text{H}]$ :

$$\begin{aligned} [\text{Fe}/\text{H}] &= \log \left( \frac{N_{\text{Fe}}}{N_{\text{H}}} \right) - \log \left( \frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\odot} \\ &\approx \log \left( \frac{M_{\text{Fe}}}{M_{\text{gas}}} \right) - \log \left( \frac{M_{\text{Fe}}}{M_{\text{H}}} \right)_{\odot}, \end{aligned} \quad (13)$$

where  $N_{\text{Fe}}$  ( $N_{\text{H}}$ ) is the number of iron (hydrogen) atoms and we used  $M_{\text{Fe}} = N_{\text{Fe}} \cdot m_{\text{Fe}}$  and  $M_{\text{gas}} \approx M_{\text{H}} = N_{\text{H}} \cdot m_{\text{H}}$ . For the solar values we assume  $\frac{M_{\text{Fe}}}{M_{\text{H}} \odot} = 1.27 \times 10^{-3}$  and  $Z_{\odot} = 0.02$  (Anders & Grevesse 1989).

Finally, at each time-step we keep track of the number, mass and iron abundance of stars that survive until  $z = 0$ , in order to determine the total stellar mass of surviving stars and their MDF. From the MDF we derive the average stellar iron abundance  $\langle [\text{Fe}/\text{H}] \rangle$ , while we get the galaxy luminosity,  $L_{\star}$ , from the total stellar mass of surviving stars by assuming  $M_{\star}/L_{\star} \approx 1$  (e.g. Salvadori et al. 2008).

## 3 MODEL CALIBRATION

The model includes three free parameters: the infall time,  $t_{\text{inf}}$ , the star formation efficiency,  $\epsilon_{\star}$ , and the SN wind efficiency,  $\epsilon_w$ . We calibrate our model by comparing with observations, using a statistical approach. In particular, we fix the free parameters of the model to reproduce the following observed properties of Boötes I:

- (i) The total luminosity of the galaxy,  $\log(L_{\star}/L_{\odot}) = 4.5 \pm 0.1$  (Kirby et al. 2013);
- (ii) The average iron abundance of stars,  $\langle [\text{Fe}/\text{H}] \rangle = -2.58 \pm 0.43$  (Norris et al. 2010; Lai et al. 2011);
- (iii) The time interval, counted starting from the first star formation event, needed to form at least the 50 per cent of the total stellar mass (Cumulative SFH = 0.5,  $t_{50} \approx (600 \pm 400)$  Myr, Brown et al. 2014);

- (iv) The observed MDF<sup>3</sup> (Norris et al. 2010; Lai et al. 2011; Gilmore et al. 2013).

Since PopIII stars have not been observed so far, we calibrated our model by assuming that there are no zero-metallicity stars survivors. Therefore, we set the minimum mass of PopIII stars to  $m_{\min} = 0.9 M_{\odot}$ .

By varying  $t_{\text{inf}}$ ,  $\epsilon_{\star}$ ,  $\epsilon_w$  we find the combination that minimizes the reduced  $\chi^2$ . The  $\chi^2$  distribution presents a minimum for  $(t_{\text{inf}}, \epsilon_{\star}, \epsilon_w) = (14.5 \text{ Myr}, 0.07, 0.03)$  for which we get  $\chi^2 = 0.33$ . In the top panel of Fig. 3 we show the  $\chi^2$  confidence levels in the parameter space of  $\epsilon_{\star}$  and  $\epsilon_w$  at fixed  $t_{\text{inf}} = 14.5$  Myr. We note that degeneracy exists between the parameters, however the minimum of  $\chi^2$  is unique.

From each combination of  $(\epsilon_{\star}, \epsilon_w)$ , we derive  $\langle [\text{Fe}/\text{H}] \rangle$ , and the total luminosity,  $L_{\star}$ . The corresponding Fe– $L_{\star}$  relation is shown in the middle panel of Fig. 3, where we compare our model results with the observed  $\langle [\text{Fe}/\text{H}] \rangle$ – $L_{\star}$  relation for Boötes I and other Local Group UFDs (Kirby, Martin & Finlator 2011; McConnachie 2012; Kirby et al. 2013; Simon 2019; Gallart et al. 2021). As we can see, the  $\chi^2$  99% confidence level contour contains three other UFD galaxies: Hercules, Leo IV, and Eridanus II. This means that we expect these galaxies to have experienced an evolution similar to that of Boötes I and hence to have similar best values of the free parameter  $(\epsilon_{\star}, \epsilon_w)$ . To check this, we recalibrate the model to fit the observed properties of each one of these three additional UFDs (see Table 1 for their physical properties). The  $(\epsilon_{\star}, \epsilon_w)$  values obtained for these UFDs are shown in Fig. 3 (top panel). We see that in all cases the combinations of  $(\epsilon_{\star}, \epsilon_w)$  are enclosed in the 99 per cent  $\chi^2$  confidence level contours of Boötes I.

Finally, the bottom panel of Fig. 3 shows the comparison between the observed MDF of Boötes I and the simulated one, which are in good agreement, while the same comparison for Hercules, Leo IV, and Eridanus II is in Appendix A.

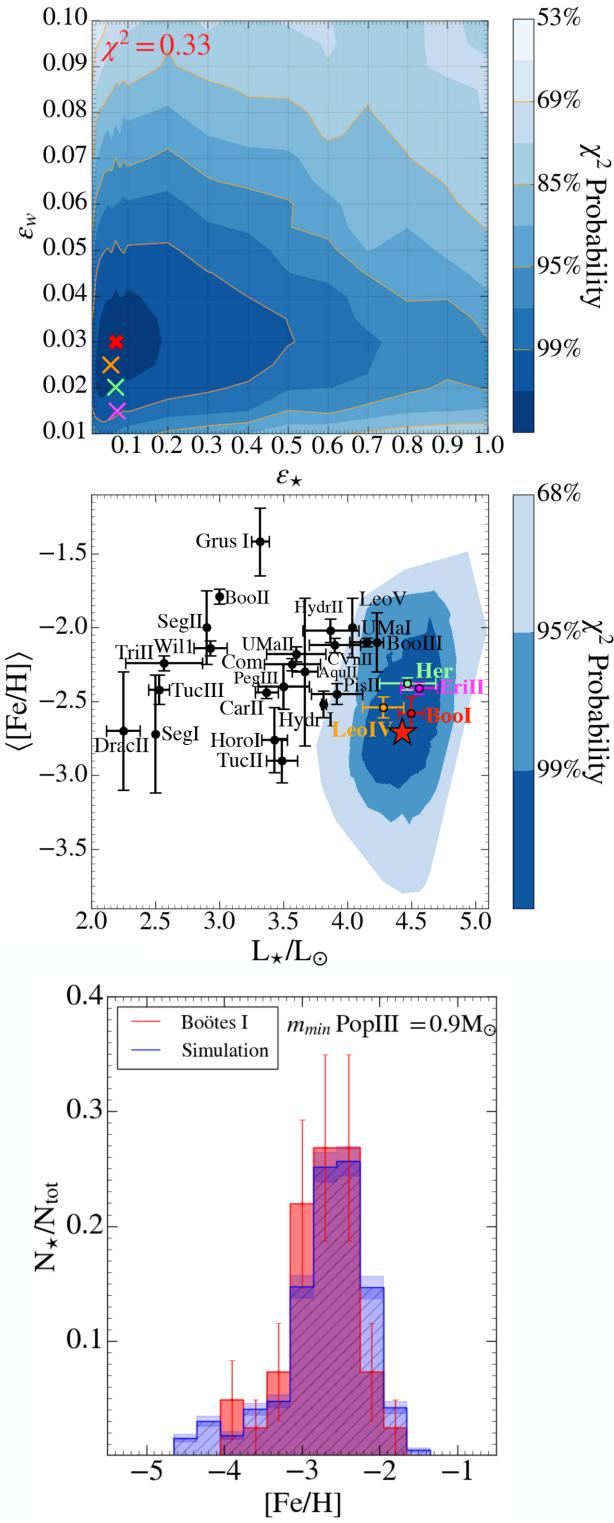
## 4 MODELLING THE IMF RANDOM SAMPLING

The data-calibrated model was obtained by assuming to fully populate the IMF (Section 2). It is customary to treat the stellar mass distribution, i.e. the IMF, like a continuous function where the stellar mass formed is distributed over the whole range of masses. When dealing with a large galaxy, or a high star formation rate in general, it is justified to think of the IMF as a densely sampled probabilistic function. However, when dealing with poorly star-forming systems, the hypothesis of a fully populated IMF breaks down (Kroupa & Weidner 2003; Weidner & Kroupa 2006; Carigi & Hernandez 2008; Kroupa et al. 2011; Leaman 2012; Weidner et al. 2013; de Bennassuti et al. 2017; Applebaum et al. 2018). It is therefore important to understand how the incomplete sampling of the IMF influences the chemical evolution of poorly star-forming UFDs.

### 4.1 Stochastic IMF sampling procedure

The Monte Carlo procedure that is used to determine the masses of stars formed in a single burst,  $M_{\star}$ , can be described as follows: we choose a normalized IMF (e.g. Larson, equation 3), and then we build up the probability function, i.e. the cumulative IMF that ranges between 0 and 1. For each randomly extracted number between 0 and 1, we determine the corresponding mass through the probability

<sup>3</sup>we use 8 points in the observed MDF



**Figure 3.** Top panel:  $\chi^2$  confidence level contours in the parameter space ( $\epsilon_*$ ,  $\epsilon_w$ ). The red, orange, light green, and magenta crosses indicate the minimum of the reduced  $\chi^2$  for Boötes I, Leo IV, Hercules, and Eridanus II, respectively. Middle panel: iron–luminosity relation, colours trace the  $\chi^2$  confidence level contours. Bottom panel: comparison between the simulated (blue), and the observed (red) MDF of Boötes I.

**Table 1.** Observed properties of UFDs in the 99% confidence level contour of Boötes I. The observed MDFs are shown in Appendix A.

Galaxy	$\log(L_*/L_\odot)$	$\langle [Fe/H] \rangle$	$t_{50}(\text{Myr})$
Boötes I	$4.5 \pm 0.10$	$-2.58 \pm 0.43$	$(600 \pm 400)$
Hercules	$4.46 \pm 0.14$	$-2.41 \pm 0.04$	$(600 \pm 500)$
Leo IV	$4.28 \pm 0.16$	$-2.54 \pm 0.07$	$(400 \pm 300)$
Eridanus II	$4.70 \pm 0.22$	$-2.38 \pm 0.04$	$(100 \pm \dots)$

function. When the extracted random number falls in the mass interval between  $m_*^i$  and  $m_*^{i+1}$  we assume to form a star with mass  $m_*^i = (m_*^i + m_*^{i+1})/2$ . Random numbers are generated until the total mass of stars equals the stellar mass formed,  $M_* = \sum_i m_*^i N_*^i$ , where  $N_*^i$  is the number of stars with mass  $m_*^i$ . Thus at each time-step, we obtain the *effective IMF* that can be compared with the theoretical one. Our results of the random sampling are in good agreement with those of de Bennassuti et al. (2017).

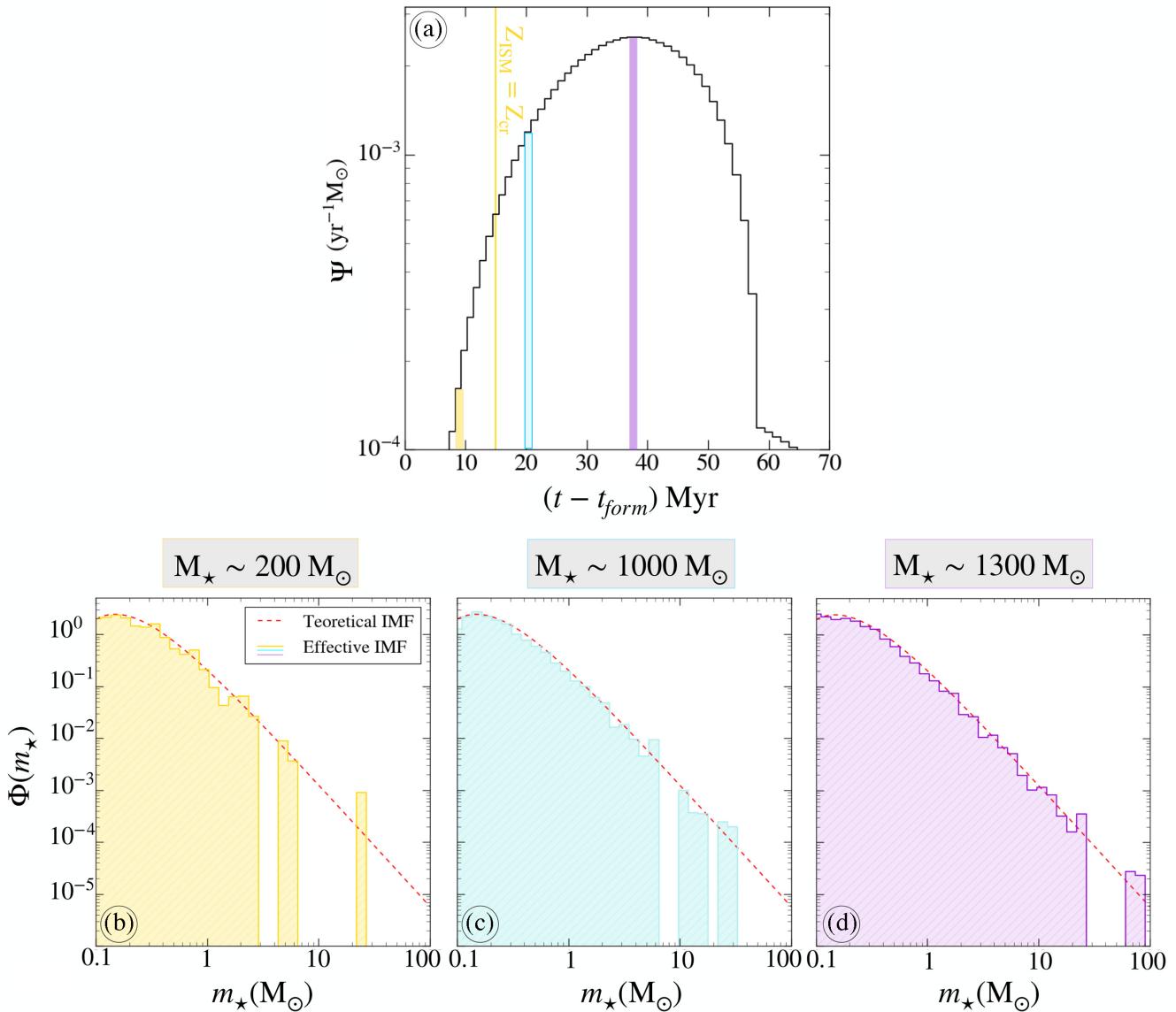
In Fig. 4, we illustrate how this random IMF sampling affects poorly star-forming UFDs when a normal Larson IMF is assumed for both PopIII and PopII/I stars. The top panel shows the predicted SFH of Boötes I, which illustrates that UFDs have extremely low star formation rates across their whole evolution ( $\Psi < 3 \times 10^{-3} M_\odot \text{yr}^{-1}$ ). The bottom panels show the comparison between the theoretical IMF and the effective one, resulting from the Monte Carlo procedure, for three different evolutionary phases. The lower is  $\Psi$ , the worse is the sampling of the theoretical IMF, in particular at the high-mass end. Furthermore, the overall star formation rate is so low in these small systems, that when the IMF is shifted towards higher masses,  $m_{ch} \geq 1 M_\odot$ , we can only form a few stars around the peak. In those cases, the IMF is poorly populated both at high and low masses (Appendix B). As we will see in Section 5 this effect is extremely important for the surviving PopIII fraction.

It should be noted that these results might be affected by the choice of  $\Delta t$ , since the IMF random sampling strongly depends on the total stellar mass formed in a single time-step. However, we are limited in choosing the time-step resolution: if we want to adequately follow the stellar evolution we need to have  $\Delta t < 4 \text{ Myr}$  (see Section 2.1). On the other hand, the typical time-scales for star formation are  $\Delta t \geq 0.1 \text{ Myr}$ . We investigated how different choices of  $\Delta t$ , in the allowed range [0.1–4]Myr, affect the chemical evolution of the galaxy. In order to match the global properties of Boötes I for larger (smaller) time-steps we are obliged to choose lower (higher) star formation efficiencies, so the resulting  $\Psi$  and fraction of PopIII stars are exactly the same. In other words, the system is ‘self-regulated’ by feedback processes, and the results of the IMF random sampling are robust.

In conclusion, our results show that even at the peak of the star formation in Boötes I (Fig. 4) there is not a complete convergence between the effective IMF and the theoretical one. In general, we find that the overall mass range can be fully sampled only when  $\Psi \gtrsim 10^{-1} M_\odot \text{yr}^{-1}$ , which is never the case in UFDs. Thus, our findings demonstrate that to study UFDs it is fundamental to model the incomplete IMF sampling, which not only affects PopIII star formation but also the formation of ‘normal’ PopII/I stars.

#### 4.2 Impact of the IMF random sampling

Accounting for the random sampling of the IMF produces two main differences with respect to a fully populated stellar mass distribution. Firstly, the statistical sampling of the IMF produces a more realistic, discrete number of stars for each mass bin.



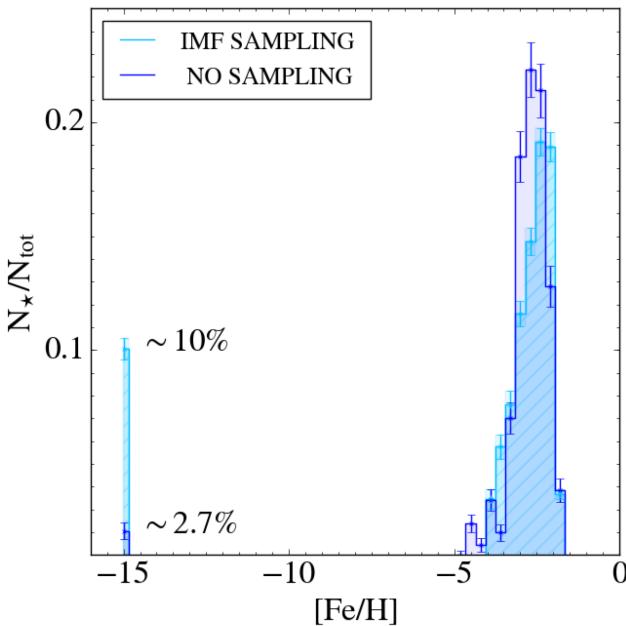
**Figure 4.** Panel a: star formation history of Boötes I according to our fiducial model. The vertical yellow line shows the transition between PopIII and PopII stars. The coloured shaded histogram bins underline the star formation rate in three different evolutionary phases: the early star formation,  $(t - t_{\text{form}}) \approx 10$  Myr, when PopIII star form at a very inefficient rate ( $\Psi \approx 2 \times 10^{-4} M_{\odot} \text{ yr}^{-1}$ , yellow); and two other stages,  $(t - t_{\text{form}}) \approx 20-40$  Myr, when PopII stars form and the star formation rate is at the average ( $\Psi \approx 10^{-3} M_{\odot} \text{ yr}^{-1}$ , blue) and the maximum ( $\Psi \approx 2.5 \times 10^{-3} M_{\odot} \text{ yr}^{-1}$ , purple) value. Panels (b)–(d): for the three evolutionary phases highlighted in panel (a) we compare the normalized theoretical IMF (dashed line) and the effective IMF (histograms) resulting from our Monte Carlo procedure (Section 4.1). Note that the IMFs have been normalized to one.

Secondly, when the star formation rate is small ( $\lesssim 10^{-3} M_{\odot} \text{ yr}^{-1}$ ), the effective IMF is only populated with a few stars beyond  $8 M_{\odot}$  (see Fig. 4). Since stars with  $8 M_{\odot} \leq m_{\star} \leq 40 M_{\odot}$  are those that explode as SNe and are responsible for outflows and enrich the ISM with metals, the chemical enrichment history of the UFDs with and without IMF random sampling could be very different.

Due to the stochastic nature of the IMF sampling, every time that a star formation event occurs the effective stellar mass distributions can be differently populated, especially at the higher masses. For this reason we exploit a statistical approach and derive the main properties of Boötes I (see Section 3) by averaging among the

results of 50 runs of the code and quantifying the scatter among them.

In Fig. 5 we compare the MDF obtained by assuming a fully populated IMF to that achieved with stochastic IMF sampling. In both cases we assume the same free parameters (Section 3). The most important difference between the two MDFs is the number of surviving zero-metallicity stars, which is more than three times higher in the random IMF sampling case, i.e.  $\approx 10$  per cent of the total number of stars instead of 2.7 per cent. As we can see in Fig. 5, the resulting MDFs are also different in the position and amplitude of the peak. In fact, the MDF obtained with the random IMF sampling has a less pronounced peak that is shifted towards higher



**Figure 5.** Comparison between the MDFs obtained with (light blue) and without (blue, same as Fig. 3) IMF random sampling for which we, respectively, get:  $\log(L_*/L_\odot) = 4.1$ ,  $\langle [\text{Fe}/\text{H}] \rangle = -2.6$ , and  $\log(L_*/L_\odot) = 4.4$ ,  $\langle [\text{Fe}/\text{H}] \rangle = -2.7$ . The MDF obtained by stochastically sampling the IMF has been derived by averaging over 50 runs. The error bars represent the  $\pm\sigma$  dispersion among different runs. For the case without the random sampling we show the Poissonian errors.

iron abundance,  $[\text{Fe}/\text{H}] \approx -2$ . As a consequence the observed properties ( $L_*$ , and average  $[\text{Fe}/\text{H}]$ ) are not well reproduced in the random IMF sampling model due to the lower star formation rate (for more details see Appendix C).

We can therefore conclude that when we account for the IMF random sampling in the modelling of poorly star-forming UFD galaxies, their evolution and global properties change. Hence, the model needs to be recalibrated, using the procedure described in Section 3. The new free parameters that minimize the reduced  $\chi^2$  are  $(t_{\text{inf}}, \epsilon_*, \epsilon_w) = (14.5 \text{ Myr}, 0.9, 0.03)$ . Note that  $\epsilon_*$  is much larger in this case.<sup>4</sup> Because of the poor IMF sampling at high masses, the effect of mechanical feedback driven by SNe is not continuous but intermittent. The impact of SNe explosions is therefore stronger since it is produced by a finite number of SNe instead of fractional numbers (Appendix C) indeed, following the chemical evolution of the galaxy it emerged that the explosion of one, at most two SNe, is enough for the ISM metallicity to exceed  $Z^{\text{ISM}} = 10^{-3.5} Z_\odot$ . As a consequence, the star formation rate is more easily damped. Thus to reproduce the total luminosity of Boötes I, a higher star formation efficiency,  $\epsilon_*$ , is required. With the recalibrated model, using these new parameters, we get that the fraction of zero-metallicity stars is  $\sim 12$  per cent with respect to the total survivor stars.

<sup>4</sup>The fraction of gas that at each time-step is converted in stars is given by the ratio between star formation efficiency and the free-fall time, i.e.  $\epsilon_* \cdot dt/t_{\text{ff}}$  ( $z$ ) (see Section 2)

## 5 CONSTRAINING THE POPIII MINIMUM MASS

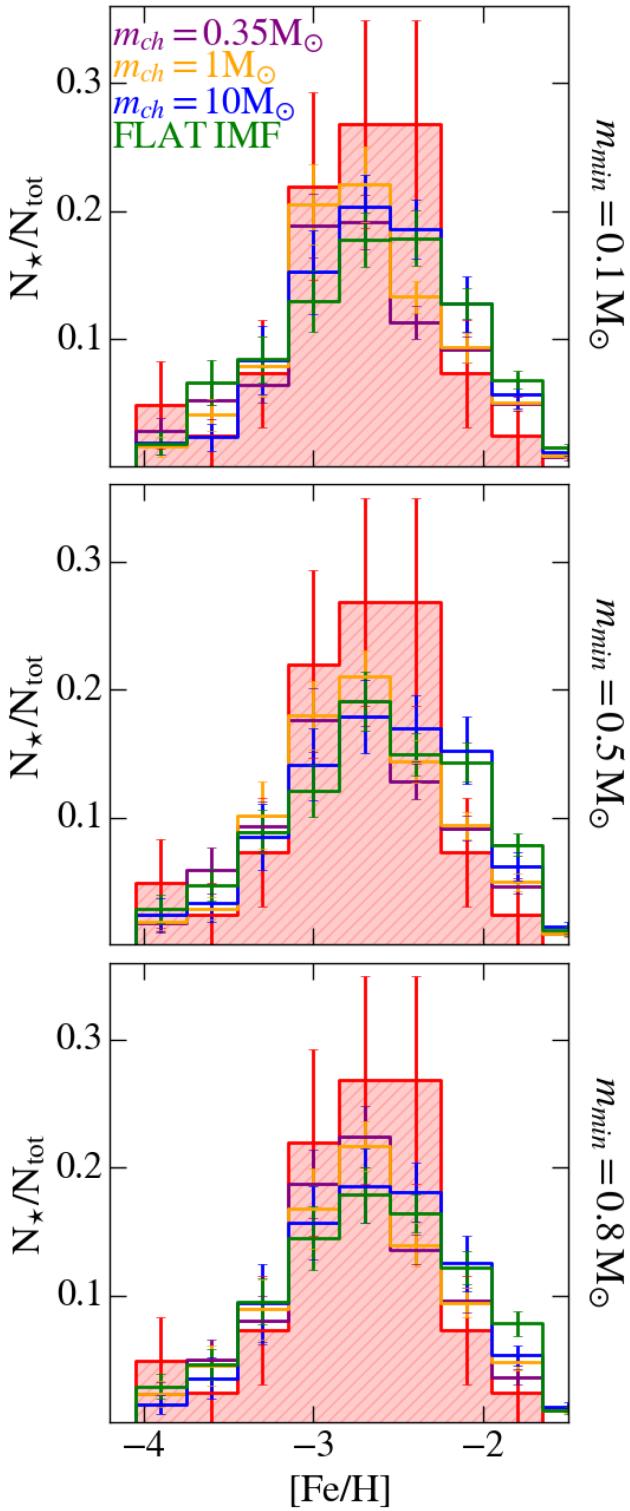
So far we have assumed that PopIII stars form according to the present-day stellar IMF (see equation 3). With this assumption we estimated that the expected fraction of long-lived PopIII stars in Boötes I is  $\sim 12$  per cent of the total. We explore how this result changes by varying the PopIII IMF shape, minimum, and maximum mass. As illustrated in Fig. 2 we here consider three Larson IMF with different characteristic masses,  $m_{\text{ch}}$ , along with a flat IMF case (see also Section 2).

### 5.1 Impact of the PopIII IMF

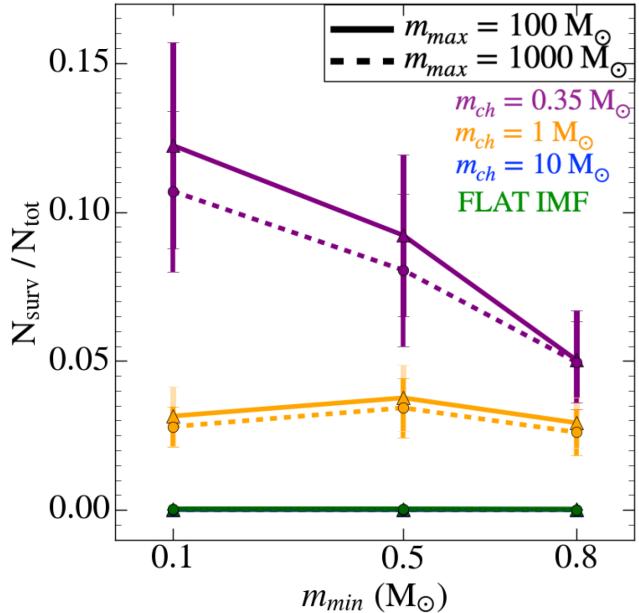
Our results show that the observed global properties of Boötes I are almost entirely determined by PopII/I stars, and that the PopIII IMF has negligible effect on the average  $[\text{Fe}/\text{H}]$  and  $L_*$  of the galaxy. PopIII stars form during the first evolutionary phase of Boötes I,  $(t - t_{\text{form}}) \sim 15 \text{ Myr}$ , when the star formation rate is low,  $\Psi \lesssim 10^{-3.5} \text{ yr}^{-1} M_\odot$  (see Fig. 4). Their contribution to the total stellar mass formed in Boötes I, and therefore to the total  $L_*$ , is only  $\approx 10\%$ . Consequently, very few PopIII SNe are formed, both due to their low star formation rate, and the incomplete sampling of the IMF. The overall ISM metal enrichment is therefore dominated by PopII stars. Similarly, changing the PopIII IMF only slightly affects the MDF of Boötes I. This is illustrated in Fig. 6, where we show the MDFs obtained by varying both the shape and lower mass limit of the PopIII IMF. All the simulated MDFs are in good agreement with that of Boötes I and consistent among each other (i.e. within the dispersion of different runs). This is because the shape of the MDF in the range  $-4 < [\text{Fe}/\text{H}] < 1$  is driven by the PopII/I stars whose IMF has not been changed.

Conversely, the expected number of PopIII stars changes considerably by varying the PopIII IMF. This is illustrated in Fig. 7, where we show the number of surviving PopIII stars,  $N_{\text{surv}}$ , with respect to the total number of stars in Boötes I,  $N_{\text{tot}}$ , for different  $m_{\text{min}}$ ,  $m_{\text{max}}$ , and shape of the PopIII IMF. At fixed  $m_{\text{min}}$  we can see that  $N_{\text{surv}}/N_{\text{tot}}$  decreases as  $m_{\text{ch}}$  increases. This is expected since when  $m_{\text{ch}}$  increases it becomes less likely to form stars with  $m_* \leq 0.8 M_\odot$ , especially when the star formation rate is low ( $\Psi \lesssim 10^{-3} M_\odot \text{ yr}^{-1}$ ). For this reason, both the Flat PopIII IMF and a Larson PopIII IMF with  $m_{\text{ch}} = 10 M_\odot$  give very low  $N_{\text{surv}}/N_{\text{tot}} < 0.0001$ .

By varying  $m_{\text{min}}$  at fixed  $m_{\text{ch}}$ , we can identify two different trends: for  $m_{\text{ch}} = 0.35 M_\odot$ ,  $N_{\text{surv}}/N_{\text{tot}}$  decreases with increasing  $m_{\text{min}}$ , while for all the others IMFs the increase of  $m_{\text{min}}$  has essentially no effect on the fraction of long-lived PopIII stars. These trends can be explained as a combination of two effects: the shape of the theoretical IMF, i.e. the higher is the characteristic mass, the more unlikely it is to form low-mass stars; and the incomplete IMF sampling. For  $m_{\text{ch}} = 0.35 M_\odot$  the low-mass end of the PopIII IMF, i.e. where the maximum resides, is well populated even when the star formation rate is very low (see Fig. 4). So when we increase  $m_{\text{min}}$  we are automatically decreasing the number of low-mass PopIII stars that can form and survive until the present day. Increasing  $m_{\text{min}}$  therefore leads to a decrease in the percentage of survivor PopIII stars. Instead for  $m_{\text{ch}} > 1 M_\odot$  and a flat IMF, when  $\Psi \lesssim 10^{-3} M_\odot \text{ yr}^{-1}$  it is unlikely to form stars with  $m_* \leq 0.8 M_\odot$ . Hence, by increasing  $m_{\text{min}}$  the probability to form low-mass stars does not change significantly and therefore the percentage of long-lived PopIII stars is roughly constant.



**Figure 6.** The simulated MDFs, obtained by varying the PopIII IMF, in comparison with that of Boötes I (red histogram). In the column we show MDFs obtained by varying the minimum mass,  $m_{\min}$ , of the PopIII stars, while the different colours identify different IMF shapes, i.e. the different characteristic mass,  $m_{ch}$ .



**Figure 7.** The number of PopIII survivors,  $N_{\text{surv}}$  ( $Z_{\star} \leq Z_{\text{cr}}$ ), with respect to the total,  $N_{\text{tot}}$ , as a function of the minimum mass of PopIII stars,  $m_{\min}$ , assuming a maximum mass of  $m_{\max} = 100, 1000 M_{\odot}$  (solid, dashed lines). The colours represent the different choices of the PopIII IMFs (see Fig. 2).

In Fig. 7, we also explore how the fraction of PopIII stars changes if we extend the maximum mass of PopIII stars,  $m_{\max}$ , up to  $1000 M_{\odot}$ . As we can see,  $N_{\text{surv}} / N_{\text{tot}}$  is roughly independent of  $m_{\max}$ . In fact, as we have previously discussed, PopIII stars formed in the first evolutionary phases of the galaxy, when  $\Psi \lesssim 10^{-3} M_{\odot} \text{yr}^{-1}$ . It follows that the PopIII IMF is badly populated, especially for  $m_{\star} \geq 10 M_{\odot}$ , whatever IMF is chosen (Fig. 4 panel a). Hence, the probability to form low-mass stars does not change. In conclusion, extending  $m_{\max}$  up to  $1000 M_{\odot}$  has very little effect on the fraction of PopIII stars that we expect to find in Boötes I.

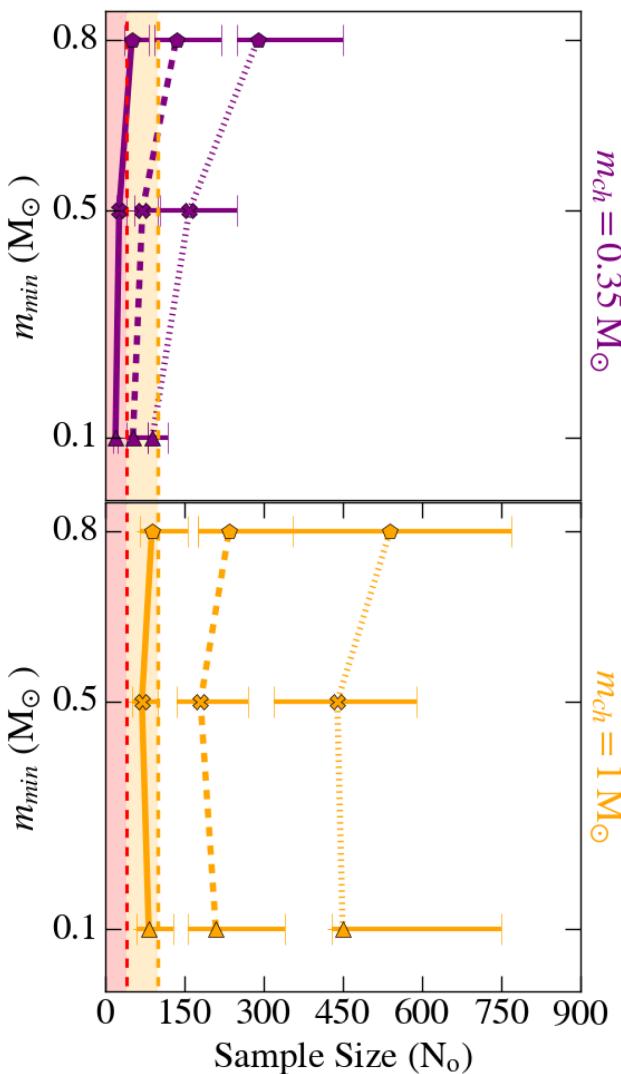
Finally, we made the same analysis for the other UFDs that have global properties consistent with our model (Fig. 3). The fractions of PopIII star survivors that we expect to find in Hercules, Leo IV, and Eridanus II are all consistent with those of Boötes I, within the error bars.

## 5.2 Constraining the low-mass end of PopIII stars

We can now exploit the non-detections of metal-free stars in Boötes I to put constraints on the minimum mass of PopIII stars. The key question is: how many stars do we need to observe to have a probability<sup>5</sup> of 68%, 95%, and 99% to catch PopIII stars for a given IMF?

Following the statistical approach of Hartwig et al. (2015), we estimate the minimum stellar sample ( $N_o$ ) needed to be observed in Boötes I, as a function of  $m_{\min}$  for different PopIII IMFs to

<sup>5</sup> Assuming that all stars have the same probability of being observed, the probability of not having detected any PopIII survivor stars in the observed sample is:  $p_0 = \frac{(N_{\text{tot}} - N_{\text{surv}})! (N_{\text{tot}} - N_o)!}{N_{\text{tot}}! (N_{\text{tot}} - N_{\text{surv}} - N_o)!}$ , where  $N_{\text{tot}}$  is the total number of stars in the galaxy at  $z = 0$ ,  $N_{\text{surv}}$  the expected number of survivor PopIII stars, and  $N_o$  is the number of observed stars.  $1 - p_0$  is therefore the probability to observe PopIII stars.



**Figure 8.** The sample size,  $N_o$ , needed to constrain the PopIII IMF as a function of  $m_{min}$  of the PopIII IMF. The different colours represent the PopIII characteristic mass, purple  $m_{ch} = 0.35 M_{\odot}$ , and orange  $m_{ch} = 1 M_{\odot}$ . The lines mark the different confident levels: 68 per cent (solid), 95 per cent (dashed), and 99 per cent (dotted).

constrain  $m_{min}$  (Fig. 8). Since no PopIII stars have been discovered so far, all models that predict a  $N_o$  smaller than the total number of stars observed in Boötes I can be used to constrain  $m_{min}$ .

For  $m_{ch} = 0.35 M_{\odot}$ ,  $N_o$  increases with  $m_{min}$  (Fig. 8). As discussed in the previous Section 5.1 (Fig. 7), the fraction of PopIII survivors,  $N_{surv}$ , decreases with increasing  $m_{min}$  and so a larger stellar sample is needed to exclude their existence. Instead,  $N_o$  is roughly constant for  $m_{ch} = 1 M_{\odot}$  (Fig. 8), and this reflects the roughly constant trend of  $N_{surv}$  with  $m_{min}$  (Fig. 7). In Fig. 8, we only show our findings for  $m_{ch} \leq 1 M_{\odot}$  because in the case  $m_{ch} = 10 M_{\odot}$  and the flat IMF the sample size required to put constraints is  $\sim 10^4$ , independently of  $m_{min}$  and  $m_{max}$ .

We can now try to get some observationally driven limits on  $m_{min}$  for different PopIII IMFs. So far there are only 41 stars in Boötes I with measured iron abundances. By using these stars (red area in Fig. 8) we can already exclude that PopIII stars formed according to a present-day stellar Larson-type IMF at 68% confidence level. In

other words, either  $m_{ch} > 0.35 M_{\odot}$  or  $m_{min} > 0.8 M_{\odot}$ . We recall that our model is also able to reproduce the properties of the UFD galaxies Hercules, Leo IV, and Eridanus II (Fig. 3). If we also consider stars that have been observed in these UFD galaxies, the actual sample size goes up to  $N_o = 96$  (orange area in Fig. 8). Using these additional data we get even tighter constraints, and the present-day stellar IMF is excluded at 95% of confidence level. Furthermore we can assert that  $m_{min} > 0.8 M_{\odot}$  or  $m_{ch} > 1 M_{\odot}$  at 68% of confidence level.

## 6 STRONGER POPIII IMF CONSTRAINTS

By using the non-detection of zero-metallicity stars in Boötes I, we have constrained the shape and minimum mass of the PopIII IMF (Section 5). However, even if PopIII stars with  $m_{*} \leq 0.8 M_{\odot}$  were able to form and survive until today it does not necessarily imply that they are readily identifiable. Currently, we can obtain the spectra necessary for iron abundance determination, only for the most luminous red giant branch stars in UFD galaxies. Therefore, to quantify and identify PopIII stars that are on the red giant branch, CMDs are required. This will allow us to determine how many PopIII stars are bright enough to be observed with current telescopes, but also to make testable predictions for the new-generation telescopes and instruments (e.g. ELT with the MOSAIC spectrograph; Hammer et al. 2014; Evans et al. 2015).

### 6.1 Simulating the Boötes I CMD

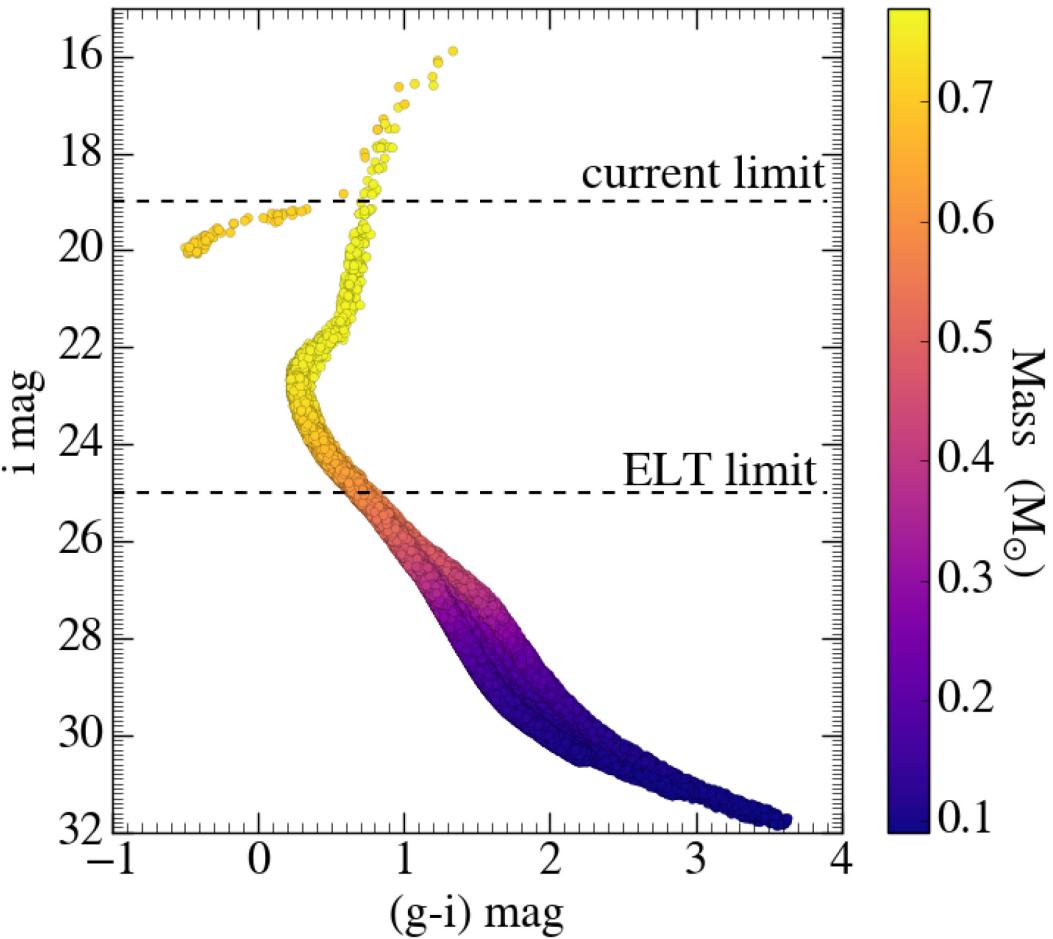
To simulate the CMD of Boötes I we used the PARSEC isochrones (Bressan et al. 2012) and CMD generator<sup>6</sup> adopting the SDSS *ugriz* photometric system. For the evolutionary tracks we use PARSEC version 1.2S plus COLIBRI S35 (Pastorelli et al. 2019) that add the TP-AGB evolution. For dust we assumed the scaling relations from Marigo et al. (2008), the extinction curve of Cardelli, Clayton & Mathis (1989) plus O'Donnell (1994), and we applied extinction coefficients computed star-by-star. Finally, we selected the Kroupa IMF that, especially in the low-mass end, is the most similar to the Larson IMF assumed in our model (see Section 2). Note that PARSEC is limited in metallicity, therefore we assign to metal-free stars the minimum available metallicity of  $Z_{*} = 10^{-7} Z_{\odot}$ . The synthetic CMD is shown in Fig. 9, where *i* magnitude has been corrected, taking into account the distance of Boötes I (McConnachie 2012). For each star in the CMD we assign a random error in magnitude and colour<sup>7</sup> in order to simulate real observations.

### 6.2 Can we really catch zero-metallicity stars?

The synthetic CMD in Fig. 9 shows a clear general trend of stellar mass and luminosity, i.e. the smaller is  $m_{*}$ , the fainter is the star. The spectroscopic data, currently available for Boötes I, reach a magnitude  $i \sim 19$ , for which 41 stars have measured [Fe/H]. At the moment we are thus only able to measure iron and other chemical elements in the more massive, and therefore more luminous stars,  $m_{*} \geq 0.7 M_{\odot}$ . However, to catch potential zero-metallicity stars with  $m_{*} = 0.1 M_{\odot}$ , observations should reach extremely deep magnitude,  $i = 32$ . Indeed, as shown in Fig. 9,

<sup>6</sup>available at <http://stev.oapd.inaf.it/cmd>

<sup>7</sup>We adopt the errors estimated in: <http://classic.sdss.org/dr4/algorithms/sdssUBVRITransform.html>



**Figure 9.** The Boötes I simulated CMD, derived using PARSEC CMD generator with SDSS photometric bands. The colour bar represents the mass of the stars in the CMD, while the two dashed horizontal lines indicate, respectively, the limit magnitude of the current observations of Boötes I,  $i = 19$ , and the limit given for MOSAIC on ESO ELT,  $i = 25$ .

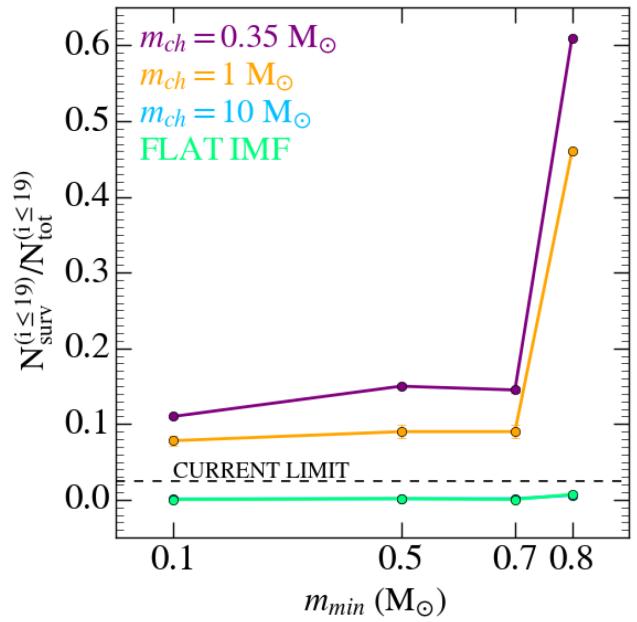
the planned MOSAIC instrument on the ESO ELT, will be able to reach  $V \approx 25$  at  $R \approx 15\,000$ – $20\,000$  (Evans et al. 2015). This means that most of these very low-mass PopIII stars, if they ever existed, will be invisible even for future generation telescopes.

### 6.3 How to get tighter PopIII IMF constraints

In Section 5.2, we estimated the fraction of PopIII stars that survive until  $z = 0$  with respect to the total number of surviving stars in Boötes I, i.e. all the stars present in the CMD (Fig. 9). What happens if we compute the number of PopIII star survivors within the sample of stars that are realistically observable?

In Fig. 10, we show the fraction of PopIII survivors in Boötes I, as a function of  $m_{\min}$ , for the current magnitude limit,  $i \leq 19$ , i.e.  $N_{\text{surv}}^{(i \leq 19)} / N_{\text{tot}}^{(i \leq 19)}$ . For the two cases,  $m_{\text{ch}} = 0.35$  and  $1 M_{\odot}$ , the new fractions of zero-metallicity stars increase with  $m_{\min}$ , i.e. the trend is inverted with respect to what was found in Fig. 7. Furthermore at  $m_{\min} = 0.8 M_{\odot}$ , the fractions for  $m_{\text{ch}} = 0.35 M_{\odot}$  ( $m_{\text{ch}} = 1 M_{\odot}$ ) are one order of magnitude higher than in Fig. 7, reaching 60% (45%) of the total, instead of 5% (3%).

The reason for these significant differences has to be found in the stellar mass–magnitude relationship, shown in Fig. 9. In the case of  $m_{\min} = 0.1 M_{\odot}$  and  $m_{\text{ch}} = 0.35 M_{\odot}$ , long-lived PopIII stars are distributed over all magnitudes according to their masses, in



**Figure 10.** The fraction of zero-metallicity stars with respect the truly observable sample of stars as a function of the PopIII minimum mass. Colours identify the different PopIII IMFs, where  $m_{\max} = 100 M_{\odot}$  in all cases.

the same way as PopII/I stars. The fraction of PopIII survivors at  $i < 19$  is thus comparable to the total fraction (e.g. 12% for  $m_{ch} = 0.35M_{\odot}$  in both Figs 7 and 10). However, as  $m_{min}$  increases, a higher number of PopIII star survivors is more massive and thus more luminous. This results in a higher fraction of long-lived PopIII stars in the observable region, i.e.  $i \leq 19$  (corresponding to  $m_{*} \gtrsim 0.7M_{\odot}$ ). Going to deeper magnitudes, in this case, will thus only result in a higher contamination by PopII stars and thus a lower  $N_{surv}/N_{tot}$ . On the other hand, if  $m_{ch} = 10M_{\odot}$  or if the IMF is flat, the fractions of PopIII star survivors do not vary significantly with respect to Fig. 7 since the probability to form long-lived  $m_{*} < 0.8M_{\odot}$  PopIII stars is always very low, independently on  $m_{min}$  (see Fig. 2 and Appendix B).

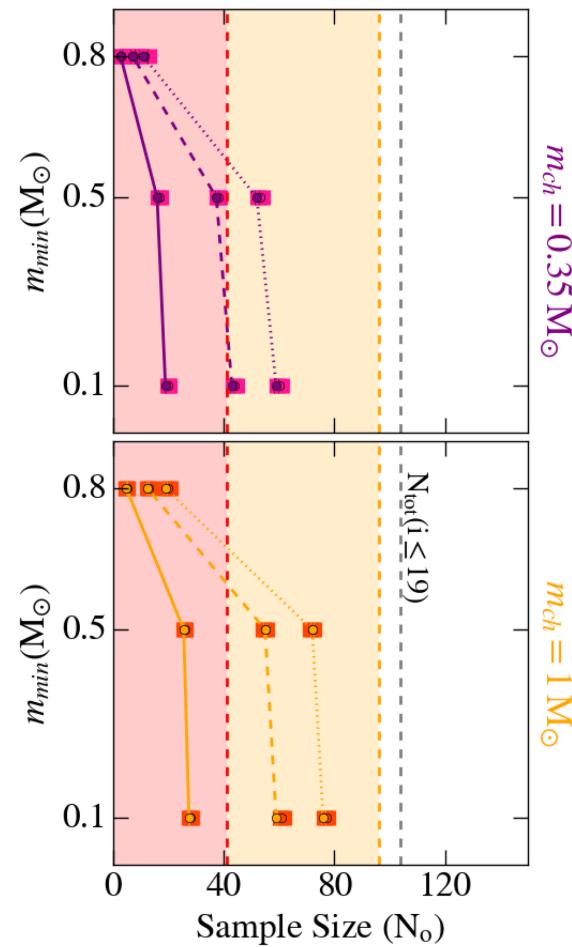
Among the 41 stars spectroscopically observed in Boötes I, no long-lived PopIII stars ( $Z_{*} < Z_{cr} = 10^{-4.5}Z_{\odot}$ ) have been identified. We can thus assert that, according to current observations, the probability to have long-lived PopIII stars in Boötes I should be  $P_{obs} < 1/N_{tot} = 1/41 \sim 2\%$ . Since  $P_{obs}$  represents the upper limit of the probability to detect PopIII stars, all models that predict a probability larger than  $P_{obs}$  can be discarded. This current (upper) limit is shown in Fig. 10, where we can appreciate the power of this simple approach: models with  $m_{*} \leq 0.8M_{\odot}$  or  $m_{ch} \leq 1M_{\odot}$  can be already excluded, which implies that we can put even tighter constraints on the PopIII IMF.

In Fig. 11, we show the sample size,  $N_o$ , needed to constrain the low-mass end of PopIII IMF at different confidence levels, using the approach from Section 5.2, but now only focusing on stars with  $i \leq 19$ . First, we notice that the required stellar sample size is smaller ( $N_o \leq 100$ ) with respect to the previous case ( $N_o \sim 800$ , Fig. 8). In fact,  $N_o$  strongly depends on both  $N_{surv}^{(i \leq 19)}$  and  $N_{tot}^{(i \leq 19)}$ , the higher is the fraction of zero-metallicity stars the lower is  $N_o$ . By using the current sample of stars observed in Boötes I, we are able to exclude  $m_{min} = 0.8M_{\odot}$  at 99% of confidence level for  $m_{ch} = 0.35M_{\odot}$  and  $m_{ch} = 1M_{\odot}$ . Furthermore, by including the data for Hercules, Leo IV, and Eridanus II, we can conclude that PopIII stars should have  $m_{min} > 0.8 M_{\odot}$  or  $m_{ch} > 1M_{\odot}$  at 99% of confidence level. Note that we are not showing models with  $m_{ch} = 10M_{\odot}$  and a flat IMF since  $N_{surv}^{(i \leq 19)}/N_{tot}^{(i \leq 19)} \approx 0.2$  per cent (see Fig. 10) hence even by observing all stars with  $i \leq 19$  in Boötes I ( $N_{tot}^{(i \leq 19)} \approx 100$ ) we are unable to constrain  $m_{min}$ .

#### 6.4 How deep should we go?

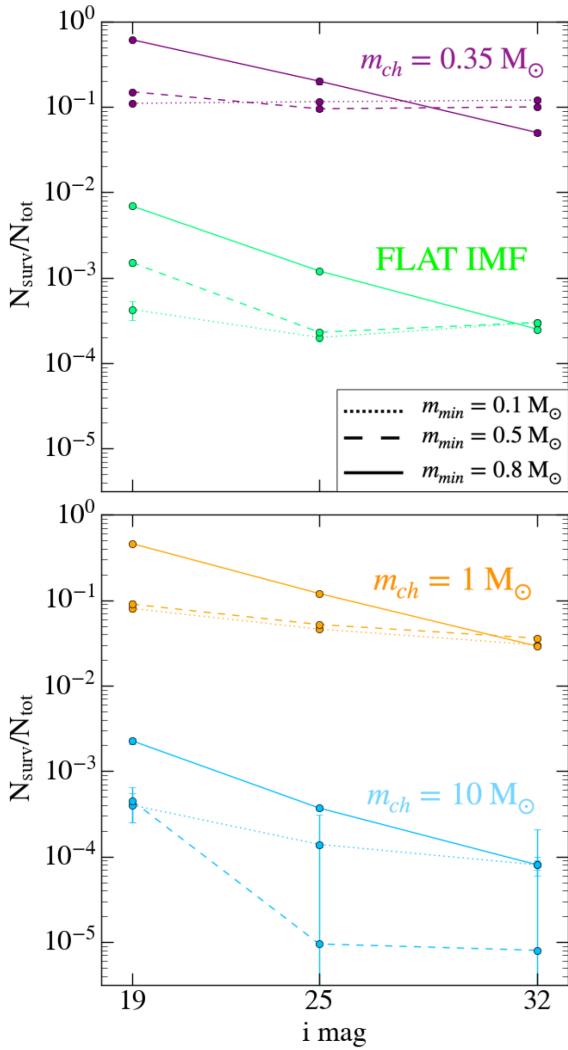
Fig. 12 shows how the fraction of PopIII stars changes as a function of the observed magnitude. In most cases we see, somewhat counter-intuitively, that the fraction of PopIII survivor stars *decreases as we reach fainter magnitudes*. This is because most of the PopIII IMFs tested here are more top-heavy compared to the adopted PopII/I IMF. This means that as we go to fainter magnitudes (and thus reach lower mass stars), the stellar sample becomes more and more dominated by the normal low-mass PopII/I stars. An exception is when the PopII/I and PopIII IMF is assumed to be the same ( $m_{ch} = 0.35M_{\odot}$  and  $m_{min} = 0.1M_{\odot}$ ). In that case the trend with limiting magnitude is flat (Fig. 12) since both populations are equally distributed in mass (and thus magnitude).

In conclusion, a higher limiting magnitude does not imply a higher fraction of PopIII survivors. So why should we go deeper in magnitude? The key point is that for  $m_{ch} > 1M_{\odot}$  the predicted fraction of PopIII star survivors is so low ( $< 0.5\%$ ) that if we want to constrain  $m_{min}$  to at least the 68% of confidence level we need larger stellar samples, which can only be obtained with deeper observations.



**Figure 11.** The sample sizes,  $N_o$ , needed to constrain the minimum mass,  $m_{min}$ , of PopIII stars, for  $m_{max} = 0.35M_{\odot}$  (purple) and  $m_{max} = 1M_{\odot}$  (orange), at different confidence levels, for stars with  $i \leq 19$ . The shaded areas display the currently observed stars in Boötes I (red) and Boötes I+Hercules+Leo IV+Eridanus II (orange). The vertical dashed line in grey delineates the total number of stars with  $i \leq 19$  in Boötes I. The filled circles represent the results assuming  $m_{max} = 100M_{\odot}$  for PopIII stars, while filled squares correspond to  $m_{max} = 1000M_{\odot}$ .

We use the simulated CMD (Fig. 9) to count the number of stars at a given magnitude in Boötes I. The results are illustrated in Fig. 13, where we show the limiting magnitude (sample size) required to constrain  $m_{min}$  at the 68% confidence level for different  $m_{ch}$ . Future instruments and telescopes, such as MOSAIC on the ELT, will allow us to measure the metallicity of  $10^4$  stars in Boötes I (see Fig. 13) and thus either discover a bona fide zero-metallicity star or convert the persisting non-detection into even stronger PopIII IMF constraints:  $m_{min} > 0.8M_{\odot}$  or  $m_{ch} > 5M_{\odot}$  at 68% confidence level. To get constraints for  $m_{ch} = 10M_{\odot}$  (or a flat IMF) we need  $N_o \approx 10^{4.3}$  stars, which implies to observe  $\sim 90\%$  of the stars in Boötes I and thus to reach extremely deep magnitudes,  $i = 26$ . Finally, after simulating the theoretical CMDs for Hercules, Leo IV, and Eridanus II, thereby accounting for the different distances, we find that with MOSAIC on the ELT we can collect a sample size of  $10^{4.2}$  stars at  $i \leq 25$ , and therefore constrain the PopIII minimum mass,  $m_{min} > 0.8M_{\odot}$ , for  $m_{ch} = 7M_{\odot}$  at 68% of confidence level. In conclusion, by targeting UFDs with next-generation instruments and telescopes we will be able to tightly constrain the minimum mass of PopIII stars,  $m_{min} > 0.8M_{\odot}$ , independent of the shape of their IMF.



**Figure 12.** Fractions of PopIII survivor stars in Boötes I as a function of the  $i$  magnitude for PopIII IMFs with different shapes and  $m_{\min}$ , assuming  $m_{\max} = 100 M_{\odot}$ .

### 6.5 Deeper or wider?

Combining observations and theoretical predictions for Boötes I allows us to derive the galaxy's *real* stellar spatial distribution. To achieve this goal, we take advantage of the SDSS archive to obtain the observed CMD of Boötes I. By comparing the simulated Boötes I CMD to a real one, we can select member stars from the observed one, and assign RA and Dec. to the stars of the simulated CMD.

For each star in the observed CMD we identify the best '*twin*' in the simulated one, i.e. we look for the simulated star with colour and magnitude closest to the observed one by minimizing the distance in the  $g-i$  and  $i$  magnitude space.<sup>8</sup> We then assign the positions of the observed stars (RA, DEC) to the simulated ones (Fig. 14, left-hand panel). Since the errors assigned to the stars in the simulated CMD are random (Section 6.1), in every run the simulated stars move around their positions in the CMD. Therefore, we identify the best twins for each run and then obtain the final results by averaging over 50 runs. The Boötes I SDSS data are limited in magnitude,  $i = 24$ , above which the observed CMD is not well populated. Therefore,

<sup>8</sup> $d = \sqrt{((g-i)_{\text{obs}} - (g-i)_{\text{sim}})^2 + (i_{\text{obs}} - i_{\text{sim}})^2}$

the spatial distribution of the simulated stars, shown in Fig. 14, is limited to  $i \leq 24$ .

Starting from the centre of Boötes I, we count the number of stars with different limiting magnitudes as a function of the radius and field of view (FOV) of the targeted area, see Fig. 14 (right). Currently we have spectroscopic data of 41 Boötes I stars with a limiting magnitude  $i \sim 19$ . By covering an area  $\sim 400$  arcmin $^2$ , it is possible to observe up to  $\approx 100$  stars at  $i \leq 19$ . However, we can obtain the same sample size by going deeper in magnitude. For example, to reach 100 stars at  $i = 21$  (23), the FOV should be  $\approx 50$  arcmin $^2$  (10 arcmin $^2$ ). Finally, by reaching limiting magnitude  $i = 24$ , and covering the entire FOV of Boötes I, we can obtain a sample of  $> 10^3$  observed stars, which implies that we can constrain  $m_{\min}$  for  $m_{ch} > 2 M_{\odot}$  at 68% of confidence level (see Fig. 13).

## 7 SUMMARY AND DISCUSSION

The aim of this work is to give observationally driven constraints on the IMF of the first (PopIII) stars and to understand if zero-metallicity low-mass stars ( $m < 0.8 M_{\odot}$ ) are able to form. We developed, from scratch, a semi-analytical model that describes the evolution of an isolated UFD galaxy, from the epoch of its formation until present day. To set the initial conditions of our model we used the results of cosmological simulations (Salvadori & Ferrara 2009; Salvadori et al. 2015) and we follow their findings by assuming that UFD evolve in isolation, i.e. that they did not experience major merger events. The model is data-calibrated, using a statistical approach based on  $\chi^2$  analysis. The free parameters of our model have been fixed in order to reproduce the global observed properties of Boötes I, i.e. the total luminosity,  $L_*$ , the average stellar iron abundance, [Fe/H], the MDF and the star formation history.

Our key results and the main implications arising from them are:

(i) Our analysis of the random sampling of the stellar IMF for  $m_* = [0.1 - 100/1000] M_{\odot}$  showed that in poorly star-forming systems like UFD galaxies,  $\langle \Psi \rangle \approx 10^{-3} M_{\odot} \text{yr}^{-1}$ , the assumption of a fully populated IMF breaks down (Section 4.1).

(ii) Including the IMF random sampling has a strong impact on the fraction of long-lived PopIII stars in Boötes I: it increases by a factor of  $\sim 3$  with respect to the model in which we assumed to fully populate the IMF (Section 4.2).

(iii) We explored how the number of surviving PopIII stars depends on the underlying PopIII IMF, where we tested three Larson-type IMFs ( $m_{ch} = [0.35, 1, 10] M_{\odot}$ ) and a flat IMF, with different low/high-mass limits (Section 5.1). Our results showed that:

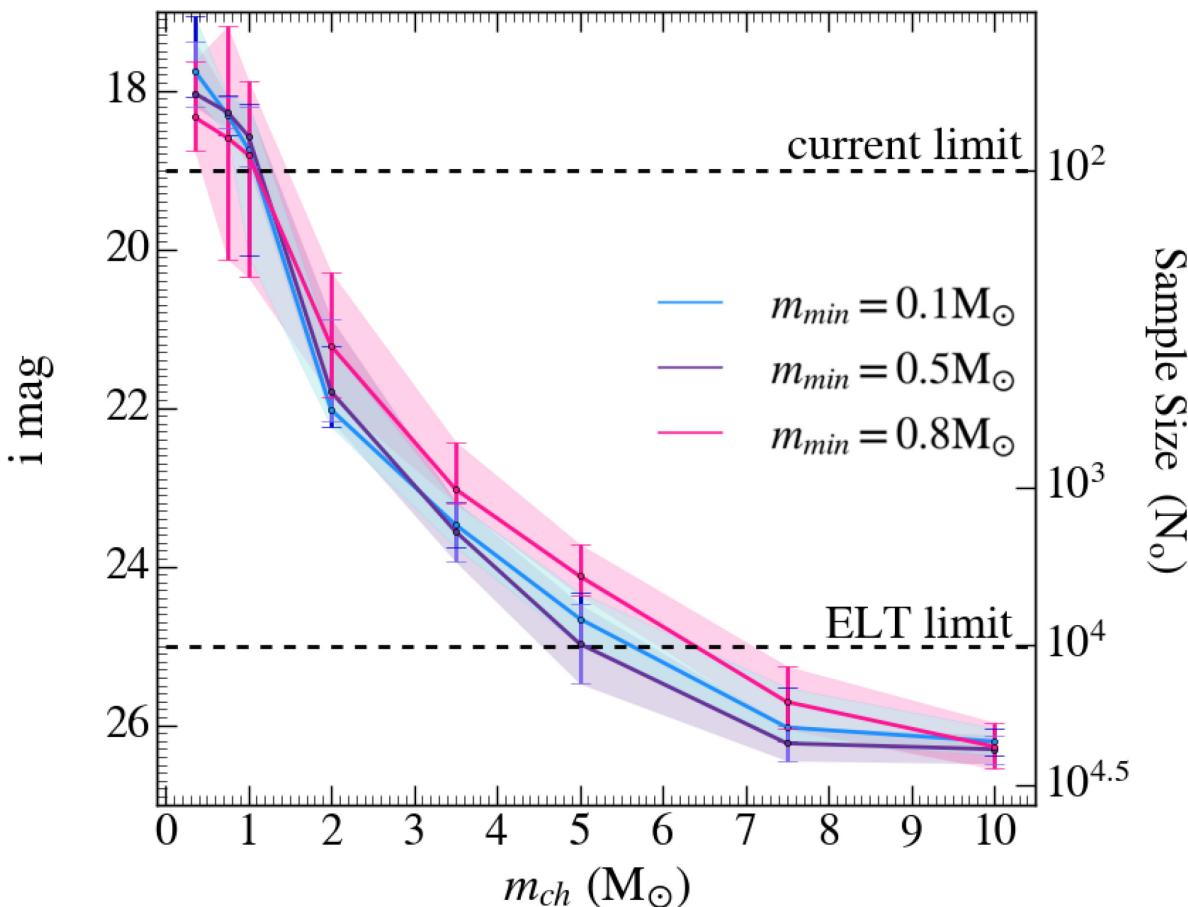
(a) The higher is the characteristic mass,  $m_{ch}$ , the lower is the number of expected PopIII star survivors.

(b) For  $m_{ch} = 0.35 M_{\odot}$  the fraction of PopIII stars at  $z = 0$  strongly depends on the minimum mass,  $m_{\min}$ .

(c) For  $m_{ch} > 1 M_{\odot}$  and for a flat IMF the number of expected PopIII star survivors is very low (< 0.5%) and roughly independent of  $m_{\min}$ .

(d) The fraction of zero-metallicity stars survivors is not significantly affected by the choice of the maximum mass ( $m_{\max} = 100, 1000 M_{\odot}$ ) of PopIII stars.

(iv) By taking advantage of the non-detection of zero-metallicity stars among the 41 stars observed in Boötes I, we can already exclude that PopIII stars form according to a present-day stellar Larson-type IMF, with  $m_{ch} = 0.35 M_{\odot}$  and  $m_{\min} \leq 0.8 M_{\odot}$ , at 68% confidence level. When data for three other UFDs are taken into



**Figure 13.** The magnitude or equivalently the total number of observed stars in Boötes I required to constrain the minimum mass of PopIII stars,  $m_{\min}$ , at 68% of confidence level, as a function of  $m_{\text{ch}}$ . Colours delineate models with different  $m_{\min}$ , while the shaded areas represent the  $\pm\sigma$  errors.

account (Hercules, Leo IV, and Eridanus II, also reproduced by our model), we conclude that PopIII stars should have  $m_{\min} > 0.8M_{\odot}$  or  $m_{\text{ch}} > 1M_{\odot}$  at 68% confidence level (Section 5.2).

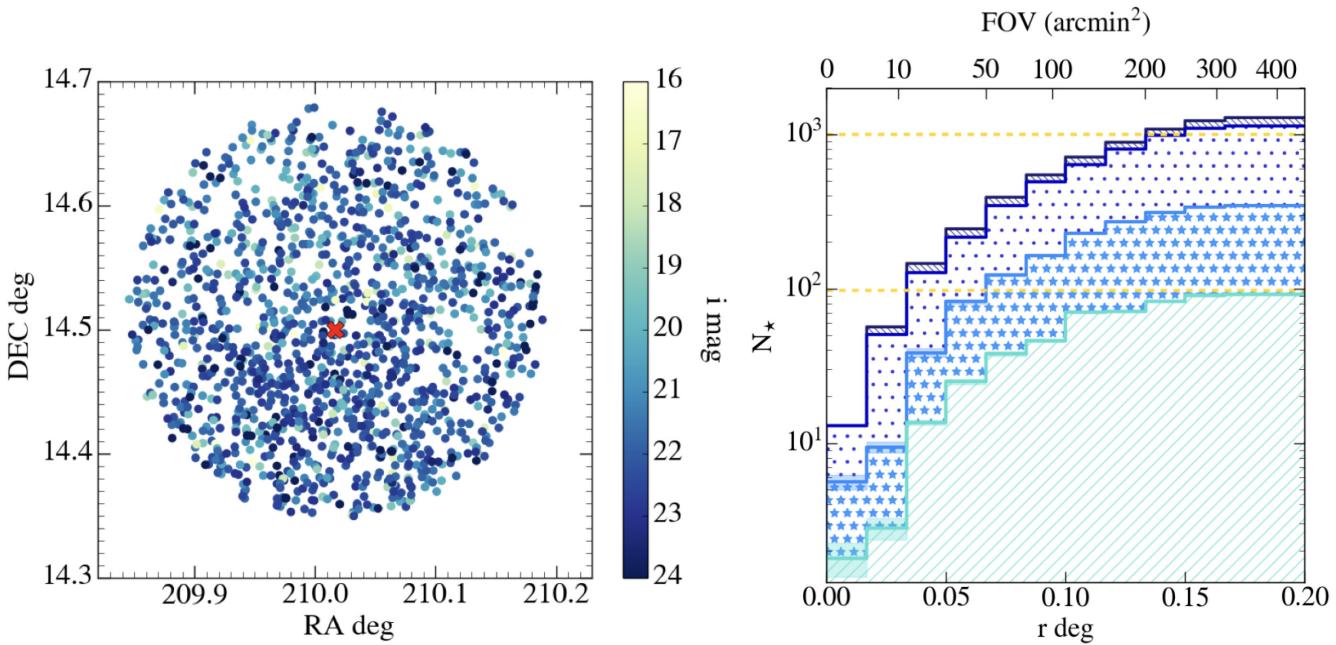
(v) Exploiting a synthetic CMD of the Boötes I (Section 6.1), we were able to take into account the mass range of the observed sample. Using Boötes I data only, this allowed us to conclude that PopIII should not have  $m_{\min} = 0.8M_{\odot}$  at 99% of confidence level for  $m_{\text{ch}} = 0.35M_{\odot}$  and  $m_{\text{ch}} = 1M_{\odot}$ . If the data of Hercules, Leo IV, and Eridanus II are also taken into account we can conclude that PopIII stars should have  $m_{\min} > 0.8M_{\odot}$  or  $m_{\text{ch}} > 1M_{\odot}$  at 99 per cent of confidence level (Sections 6.2 and 6.3).

(vi) By combining the observed CMD for Boötes I (based on SDSS data) with the simulated one, we obtained the spatial distribution of stars within the galaxy. By reaching  $i \leq 24$ , and covering the entire Boötes I, we can obtain an observed sample of  $>10^3$  stars. As a consequence, if no metal-free star were discovered, we could conclude that PopIII stars should have  $m_{\min} > 0.8M_{\odot}$  or  $m_{\text{ch}} > 2M_{\odot}$  at 68% of confidence level (Section 6.5).

(vii) Targeting the four UFDs Boötes I, Hercules, Leo IV, and Eridanus II with future generations instruments (such as ELT/MOSAIC) it will be possible to collect a sample of  $>10000$  stars. If the non-detection of metal-free stars will be persistent, we will limit  $m_{\min} > 0.8M_{\odot}$  independent of the PopIII IMF at 68% confidence level (Section 6.4).

In addition, we explored how our results are affected by the choice of critical metallicity value,  $Z_{\text{cr}}$  that established the transition between PopIII-to-PopII/I stars. The exact value of  $Z_{\text{cr}}$  is largely uncertain, it is estimated that it can vary as  $Z_{\text{cr}} = (10^{-4.5 \pm 1})Z_{\odot}$  (Bromm et al. 2001; Schneider et al. 2003; Omukai et al. 2005). However, we have not found any dependence of the fractions of zero-metallicity stars on critical metallicity. In fact, following the chemical evolution of the galaxy it emerged that the explosion of one, at most two SNe, is enough for the metallicity to exceed  $10^{-3.5} Z_{\odot}$ . Therefore, we move from a metal-free environment with  $Z_{\text{ISM}} = 0$ , to one with  $Z_{\text{ISM}} > 10^{-3.5} Z_{\odot}$ . Hence, our results are not affected by the choice of the  $Z_{\text{cr}}$  value.

There have been several theoretical works (e.g. Romano et al. 2014; Vincenzo et al. 2014) that model the evolution of Boötes I since it is one of the observationally best-studied UFD galaxy. However, none of these models take into account the stochastic sampling of the IMF. Furthermore, none of these works have investigated the expected fraction of PopIII star survivors in Boötes I. In agreement with our findings, Vincenzo et al. (2014) concluded that the star formation efficiency of Boötes I was very small. This is also consistent with studies of Boötes I in cosmological context (e.g. Bovill & Ricotti 2009; Salvadori & Ferrara 2009; Romano et al. 2014; Bland-Hawthorn et al. 2015; Salvadori et al. 2015).



**Figure 14.** Left-hand panel: the Boötes I spatial distribution of stars from the simulated CMD. The colour bar highlights the different magnitudes, while the red cross identifies the centre of mass,  $x_{\text{cm}} = 210.017 \text{ deg}$ , and  $y_{\text{cm}} = 14.5 \text{ deg}$ . Right-hand panel: the number of stars enclosed in a circumference of radius  $r$  (bottom) and the corresponding FOV (top), for different limiting magnitudes. These trends were obtained by averaging over 50 realizations of the Boötes I simulated CMD and the corresponding spatial distributions, and the shaded areas represent the  $\pm\sigma$  Poisson errors.

One of the key points in this work was to estimate the fraction of PopIII star survivors in UFD galaxies. The same was done by Magg et al. (2017), who derived the percentage of surviving PopIII that are expected to be found in our Milky Way and its satellites. Unfortunately, we cannot make a direct comparison, since they used a logarithmically flat IMF. However our results for  $m_{\text{ch}} = 10 M_{\odot}$  and the case of flat IMF predict a PopIII fraction consistent with their findings.

Our methods allow us to estimate the minimum stellar sample size required to put observationally driven constraints on the PopIII IMF. This approach was used by Hartwig et al. (2015), for halo/bulge stars of our Milky Way. They found that the expected number of PopIII star survivors decreases as  $m_{\text{min}}$  increases assuming a logarithmically Flat PopIII IMF. Our findings have the same trend for IMF with  $m_{\text{ch}} = 0.35 M_{\odot}$ . However, for the other IMFs ( $m_{\text{ch}} = 10 M_{\odot}$  and flat IMF) the fraction of PopIII survivors is extremely low ( $< 0.01\%$ ) regardless of the minimum mass. The reason is twofold: (i) the IMF shape, which makes it very unlikely to form low-mass stars (Fig. 2); (ii) the random sampling of the IMF which strongly affects the population of the lower mass end (see Fig. B1), i.e. the number of stars with  $m_{\star} \leq 0.8 M_{\odot}$ . In fact, for  $m_{\text{ch}} \geq 1 M_{\odot}$  and for flat IMF it is unlikely to form long-lived PopIII stars independently of the minimum mass. Furthermore, Hartwig et al. (2015) found that to conclude that PopIII stars should have  $m_{\star} \geq 0.8 M_{\odot}$  at a 68% confidence level, we need to observe  $N_0 \approx 10^6$  stars in the Milky Way stellar halo. However, with our method, even tighter constraints can be put on the PopIII IMF by only using  $N_0 \approx 10^2$  stars.

For the first time, we have performed an in-depth analyses of four UFDs, Boötes I, Hercules, Leo IV, and Eridanus II with the goal to constrain the IMF of PopIII stars. By developing a chemical evolution model of isolated UFDs, and combining observational constraints with a statistical approach we have shown that old and pristine UFDs

are the ideal systems to understand the physical properties of the first stars. Our study has revealed that it is crucial to take into account the incomplete sampling of the IMF in these poorly star-forming systems, to realistically represent their stellar populations. This is especially relevant for the earliest star formation, where PopIII stars formed very inefficiently for a limited time. Our novel approach of connecting our models with real data through synthesizing the CMD of Boötes I has revealed that taking the mass distribution of the observed stars into account allows us to put even stronger constraints on the  $m_{\text{min}}$  and the shape of the PopIII IMF. In this work, with only 96 observed stars in Boötes I, Hercules, Leo IV, and Eridanus II, we can derive at a 99% confidence level that  $m_{\text{min}} > 0.8 M_{\odot}$  or  $m_{\text{ch}} > 1 M_{\odot}$ . Thus, we conclude that PopIII stars were typically more massive than those that form today, and the mass distribution of the first stars was fundamentally different from stars at later times. Finally, we showed that future telescopes will be able to get much larger samples ( $> 10^4$  stars), which will provide even better understanding of the nature of the very first stars in the universe.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for the useful and constructive comments. We thank Eline Tolstoy and the Cosmology group at the Scuola Normale Superiore of Pisa for insightful comments on the early version of this work. SS thanks Piercarlo Bonifacio for inspiring discussions, which triggered the study of the synthetic CMD of Boötes I along with the theoretical predictions for future generation telescopes. We acknowledge support from the ERC Starting Grant NEFERTITI H2020/808240. SS acknowledges support from the PRIN-MIUR2017, The quest for the first stars, prot. n. 2017T4ARJ5.

## DATA AVAILABILITY

The authors confirm that the data supporting the findings of this study are available within the articles: Norris et al. (2010), Lai et al. (2011), Gilmore et al. (2013) for the Boötes I MDF; Kirby et al. (2011, 2013) for  $\langle \text{[Fe/H]} \rangle - L_*$  relation for UFDs. The data that support the findings of this study are available on request from the corresponding author MR.

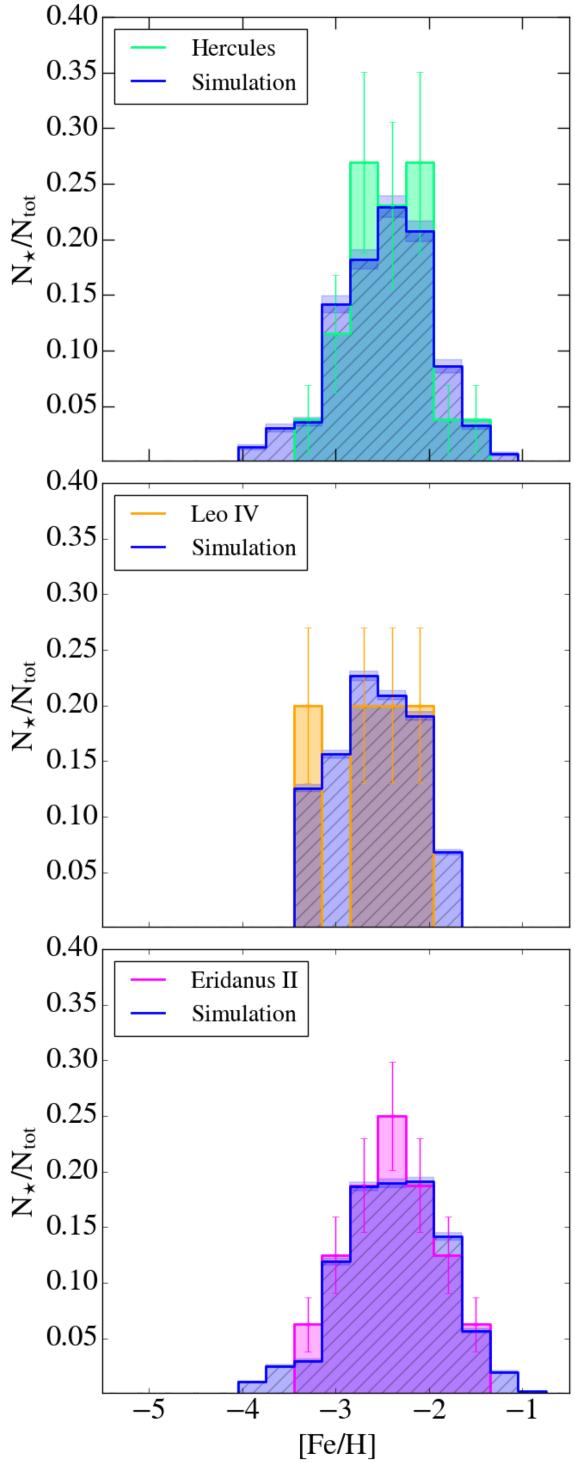
## REFERENCES

- Abel T., Bryan G. L., Norman M. L., 2002, *Science*, 295, 93
- Aguado D. S., González Hernández J. I., Allende Prieto C., Rebolo R., 2019, *ApJ*, 874, L21
- Anders E., Grevesse N., 1989, *Geochim. Cosmochim. Acta*, 53, 197
- Aoki W., Tominaga N., Beers T., Honda S., Lee Y., 2014, *Science*, 345, 912
- Applebaum E., Brooks A. M., Quinn T. R., Christensen C. R., 2020, *MNRAS*, 492 preprint ([arXiv:1811.00022](https://arxiv.org/abs/1811.00022))
- Barkana R., Loeb A., 2001, *Phys. Rep.*, 349, 125
- Beers T. C., Christlieb N., 2005, *ARA&A*, 43, 531
- Bland-Hawthorn J., Sutherland R., Webster D., 2015, *ApJ*, 807, 154
- Bonifacio P. et al., 2018, *A&A*, 612, A65
- Bonifacio P. et al., 2019, *MNRAS*, 487, 3797
- Bovill M. S., Ricotti M., 2009, *ApJ*, 693, 1859
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127
- Bromm V., 2013, *Rep. Prog. Phys.*, 76, 112901
- Bromm V., Ferrara A., Coppi P., Larson R., 2001, *MNRAS*, 328, 969
- Bromm V., Coppi P. S., Larson R. B., 2002, *ApJ*, 564, 23
- Brown T. M. et al., 2014, *ApJ*, 796, 91
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245
- Carigi L., Hernandez X., 2008, *MNRAS*, 390, 582
- Chiti A., Mardini M. K., Frebel A., Daniel T., 2021, *ApJL*, submitted
- Clark P. C., Glover S. C., Smith R. J., Greif T. H., Klessen R. S., Bromm V., 2011, *Science*, 331, 1040
- de Bennassuti M., Salvadori S., Schneider R., Valiante R., Omukai K., 2017, *MNRAS*, 465, 926
- de Jong R. S. et al., 2019, *The Messenger*, 175, preprint ([arXiv:1903.02464](https://arxiv.org/abs/1903.02464))
- Diemand J., Madau P., Moore B., 2005, *MNRAS*, 364, 367
- Dopcke G., Glover S. C. O., Clark P. C., Klessen R. S., 2013, *ApJ*, 766, 103
- Evans C. et al., 2015, Ground-based and Airborne Instrumentation for Astronomy V, Vol. 9908, preprint ([arXiv:1501.04726](https://arxiv.org/abs/1501.04726))
- Frebel A., 2010, *Astron. Nachr.*, 331, 474
- Frebel A., Norris J. E., 2015, *ARA&A*, 53, 631
- Frebel A., Norris J. E., Gilmore G., Wyse R. F., 2016, *ApJ*, 826, 110
- Freeman K., Bland-Hawthorn J., 2002, *ARA&A*, 40, 487
- Gallart C. et al., 2021, *ApJ*, 09, preprint ([arXiv:2101.04464](https://arxiv.org/abs/2101.04464))
- Gilmore G., Norris J. E., Monaco L., Yong D., Wyse R. F., Geisler D., 2013, *ApJ*, 763, 61
- Gratton R., Sneden C., Carretta E., 2004, *ARA&A*, 42, 385
- Greif T., Springel V., White S., Glover S., Clark P., Smith R., Klessen R., Bromm V., 2011, *ApJ*, 737, 75
- Hammer F. et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. SPIE, Bellingham, p. 914727
- Hartwig T., Bromm V., Klessen R. S., Glover S. C., 2015, *MNRAS*, 447, 3892
- Heger A., Woosley S. E., 2002, *ApJ*, 567, 532
- Hirano S., Bromm V., 2017, *MNRAS*, 470, 898
- Hirano S., Hosokawa T., Yoshida N., Umeda H., Omukai K., Chiaki G., Yorke H. W., 2014, *ApJ*, 781, 60
- Hosokawa T., Omukai K., Yoshida N., Yorke H. W., 2011, *Science*, 334, 1250
- Ishigaki M. N., Aoki W., Arimoto N., Okamoto S., 2014, *A&A*, 562, A146
- Ishigaki M. N., Tominaga N., Kobayashi C., Nomoto K., 2018, *ApJ*, 857, 46
- Ji A. P., Frebel A., Ezreddine R., Casey A. R., 2016, *ApJ*, 832, L3
- Kereš D., Katz N., Weinberg D. H., Davé, 2005, *MNRAS*, 363, 2
- Kirby E. N., Martin C. L., Finlator K., 2011, *ApJ*, 742, L25
- Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, *ApJ*, 779, 102
- Klessen R. S., Krumholz M. R., Heitsch F., 2011, *Adv. Sci. Lett.*, 4, 258
- Kroupa P., Weidner C., 2003, *ApJ*, 598, 1076
- Kroupa P., Weidner C., Pflamm-Altenburg J., Thies I., Dabringhausen J., Marks M., Maschberger T., 2013, *Planets, Stars and Stellar Systems. Vol. 5: Galactic Structure and Stellar Populations*, p. 115
- Lai D. K., Lee Y. S., Bolte M., Lucatello S., Beers T. C., Johnson J. A., Sivarani T., Rockosi C. M., 2011, *ApJ*, 738, 51
- Larson P. L., 1998, *Gaia*, 15, 389
- Leaman R., 2012, *AJ*, 144, 183
- Li H., Tan K., Zhao G., 2018, *ApJS*, 238, 16
- Machida M. N., Matsumoto T., Inutsuka S.-I., 2008, *ApJ*, 685, 690
- Magg M., Hartwig T., Agarwal B., Frebel A., Glover S. C., Griffen B. F., Klessen R. S., 2017, *MNRAS*, 473, 5308
- Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, *A&A*, 482, 883
- Matteucci F., Panagia N., Pipino A., Mannucci F., Recchi S., Della Valle M., 2006, *MNRAS*, 372, 265
- McConnachie A. W., 2012, *AJ*, 144, 4
- McWilliam A., 1998, *AJ*, 115, 1640
- Norris J. E., Yong D., Gilmore G., Wyse R. F., 2010, *ApJ*, 711, 350
- O'Donnell J. E., 1994, *ApJ*, 422, 158
- O'Shea B. W., 2007, in American Astronomical Society Meeting Abstracts. p. 79.01
- Oey M. S., 2003, *MNRAS*, 339, 849
- Omukai K., Palla F., 2001, *ApJ*, 561, L55
- Omukai K., Tsuribe T., Schneider R., Ferrara A., 2005, *ApJ*, 626, 627
- Pakhomov Y. V., Mashonkina L. I., Sitnova T. M., Jablonka P., 2019, *Astron. Lett.*, 45, 259
- Pastorelli G. et al., 2019, *MNRAS*, 485, 5666
- Planck Collaboration VI, 2018, *A&A*, 641, A6
- Raiteri C., Villata M., Navarro J., 1996, *A&A*, 315, 105
- Reichert M., Hansen C. J., Hanke M., Skúladóttir Á., Arcones A., Grebel E. K., 2020, *A&A*, 641, A127
- Ricotti M., Gnedin N. Y., 2005, *ApJ*, 629, 259
- Romano D., Bellazzini M., Starkenburg E., Leaman R., 2014, *MNRAS*, 446, 4220
- Safarzadeh M., Ji A. P., Dooley G. A., Frebel A., Scannapieco E., Gómez F. A., O'Shea B. W., 2018, *MNRAS*, 476, 5006
- Salvadori S., Ferrara A., 2009, *MNRAS*, 395, L6
- Salvadori S., Ferrara A., 2012, *MNRAS*, 421, L29
- Salvadori S., Schneider R., Ferrara A., 2007, *MNRAS*, 381, 647
- Salvadori S., Ferrara A., Schneider R., 2008, *MNRAS*, 386, 348
- Salvadori S., Ferrara A., Schneider R., Scannapieco E., Kawata D., 2010, *MNRAS*, 401, L5
- Salvadori S., Skuladóttir A., Tolstoy E., 2015, *MNRAS*, 454, 1320
- Salvadori S., Bonifacio P., Caffau E., Korotin S., Andreevsky S., Spite M., Skúladóttir Á., 2019, *MNRAS*, 487, 4261
- Schaerer D., Pelló R., 2002, in Scientific Drivers for ESO Future VLT/VLTI Instrumentation. Springer, Springer, p. 48
- Schneider R., Ferrara A., Salvaterra R., Omukai K., Bromm V., 2003, *Nature*, 422, 869
- Sestito F. et al., 2019, *MNRAS*, 484, 2166
- Simon J. D., 2019, *ARA&A*, 57, 375
- Skúladóttir Á., Tolstoy E., Salvadori S., Hill V., Pettini M., Shetrone M. D., Starkenburg E., 2015, *A&A*, 574, A129
- Smith R. J., Clark P. C., Glover S., Klessen R. S., 2010, in Whalen D. J., Bromm V., Yoshida N., eds, AIP Conf. Proc. Vol. 1294, First Stars and Galaxies: Challenges for the Next Decade. Am. Inst. Phys., New York, p. 285
- Stacy A., Bromm V., 2014, *ApJ*, 785, 73
- Stacy A., Greif T. H., Klessen R. S., Bromm V., Loeb A., 2013, *MNRAS*, 431, 1470
- Stacy A., Bromm V., Lee A. T., 2016, *MNRAS*, 462, 1307
- Starkenburg E. et al., 2018, *MNRAS*, 481, 3838
- Susa H., 2019, *ApJ*, 877, 99
- Tan J. C., McKee C. F., 2004, *ApJ*, 603, 383

- Tarumi Y., Hartwig T., Magg M., 2020, *ApJ*, 897, 58  
 Tolstoy E., Hill V., Tosi M., 2009, *ARA&A*, 47, 371  
 Tumlinson J., 2007, *ApJ*, 665, 1361  
 Turk M. J., 2009, PhD thesis, Stanford University  
 Venn K. A., Irwin M., Shetrone M. D., Tout C. A., Hill V., Tolstoy E., 2004, *AJ*, 128, 1177  
 Vincenzo F., Matteucci F., Vattakunnel S., Lanfranchi G. A., 2014, *MNRAS*, 441, 2815  
 Webster D., Bland-Hawthorn J., Sutherland R., 2015, *ApJ*, 799, L21  
 Weidner C., Kroupa P., 2006, *MNRAS*, 365, 1333  
 Weidner C., Kroupa P., Pfleiderer-Altenburg J., Vazdekis A., 2013, *MNRAS*, 436, 3309  
 Wise J. H., Abel T., Turk M. J., Norman M. L., Smith B. D., 2012, *MNRAS*, 427, 311  
 Wollenberg K. M. J., Glover S. C. O., Clark P. C., Klessen R. S., 2020, *MNRAS*, 494, 1871  
 Woosley S. E., Weaver T. A., 1995, *ApJS*, 101, 181  
 Yoshida N., Abel T., Hernquist L., Sugiyama N., 2003, *ApJ*, 592, 645  
 Yoshida N., Omukai K., Hernquist L., Abel T., 2006, *ApJ*, 652, 6

## APPENDIX A: HERCULES, LEO IV, AND ERIDANUS II

Fig. A1 shows the comparison between the observed and the simulated MDFs for those UFDs that reside in the 99 per cent confidence level  $\chi^2$  contour of Boötes I, i.e. Hercules, Leo IV, and Eridanus II (see Fig. 3, Section 3). We can see that in all cases we find a good agreement between model results and observations.



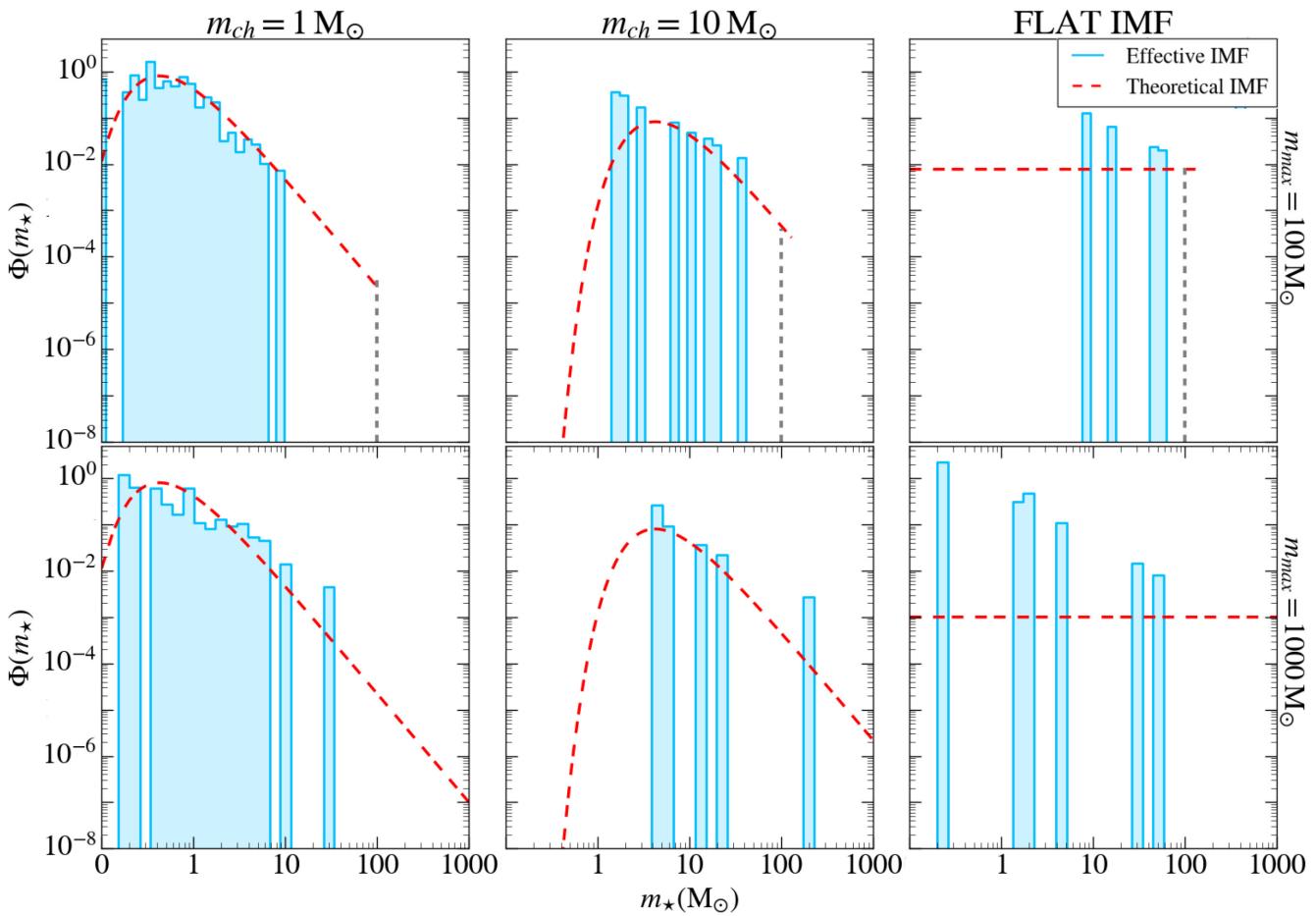
**Figure A1.** The comparison between the observed and the simulated (blue) MDFs for UDFs in the 99 per cent of confidence level  $\chi^2$  contour of Boötes I: Hercules (light green), Leo IV (orange), and Eridanus II (magenta).

## APPENDIX B: POPIII IMF RANDOM SAMPLING

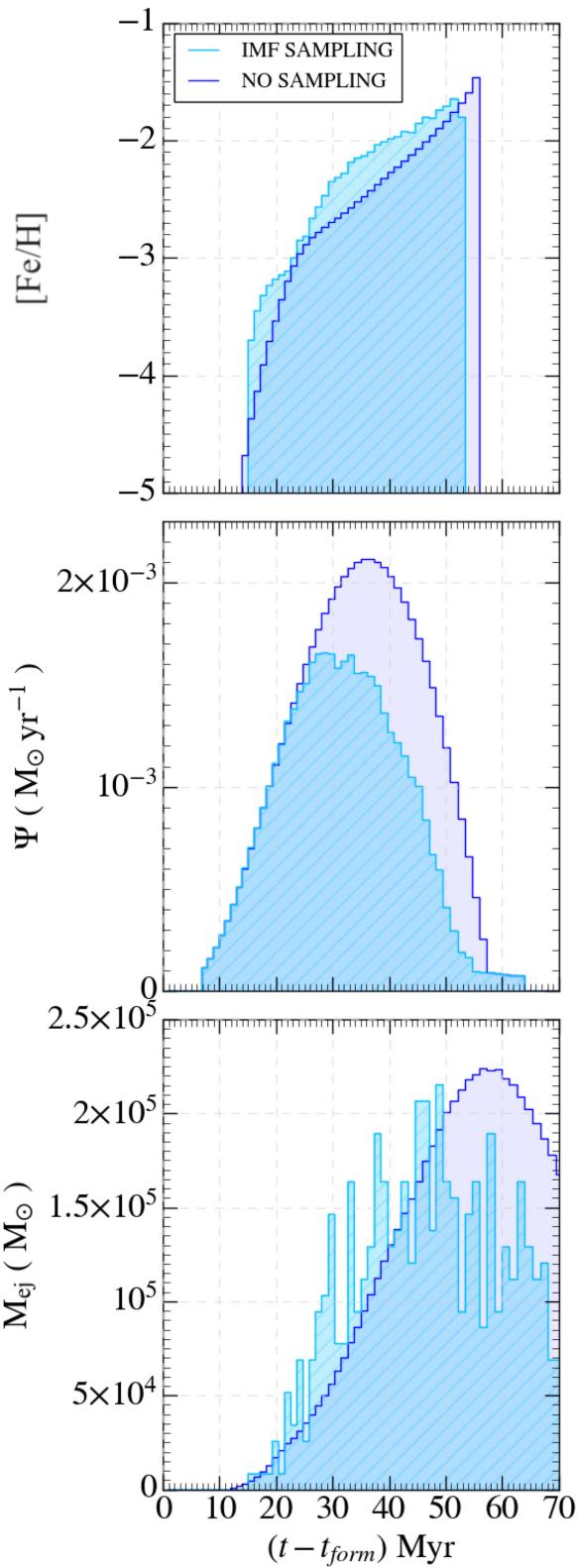
Fig. B1 shows the comparison between the theoretical PopIII IMF (see Fig. 2) and the effective one obtained with the random sampling procedure, when we assume to form a total stellar mass of  $\approx 200M_{\odot}$ . This is the typical mass formed in a single burst during the early evolutionary phases of Boötes I, i.e. when  $\Psi \approx 2 \times 10^{-4} M_{\odot} \text{yr}^{-1}$  (Fig. 4). We see that the theoretical PopIII IMF is poorly sampled independent of its shape, the  $m_{\text{ch}}$ , and the  $m_{\text{max}}$ . Furthermore, while for  $m_{\text{ch}} = 1M_{\odot}$  we still have a good sampling of the PopIII IMF at the lowest mass end,  $m_{\star} < 1M_{\odot}$ , when the peak of the IMF is shifted towards higher masses ( $m_{\text{ch}} = 10M_{\odot}$ ) or there is not a preferential mass scale, i.e. the IMF is flat, the lowest mass end becomes less and less populated with respect to the theoretical IMF. This implies that the number of surviving PopIII stars is much lower than expected. These results are even more extreme for  $m_{\text{max}} = 1000M_{\odot}$  (bottom panels).

## APPENDIX C: IMF SAMPLING: BOÖTES I EVOLUTION

In Fig. C1, we show the evolution of the ISM iron abundance, [Fe/H], the star formation rate,  $\Psi$ , and the ejected gas mass,  $M_{\text{ej}}$ , of Boötes I when we consider models with and without the random sampling of the IMF. In the bottom panel of Fig. C1, we see that [Fe/H] in the random sampling model starts with a time-delay of  $\sim 2$  Myr. This is because the more massive SNe,  $m_{\star} = 40M_{\odot}$ , that explode first are not produced in this early burst of star formation, which has  $\Psi \approx 10^{-4} M_{\odot} \text{yr}^{-1}$  (middle panel). Still, as soon as low-mass SNe stars to explode, [Fe/H] in the IMF sampling model rapidly grows, exceeding the value of the model without sampling (top) in spite of the equal rate of star formation (middle panel). This is because the random sampling of the IMF provide us with a finite number of SNe instead of a fraction. Although delayed, the effect of feedback from SNe is stronger on both the ISM enrichment (top) and the mass of gas ejected (bottom panel). When  $(t - t_{\text{form}}) \approx 25$  Myr, i.e. at the peak of



**Figure B1.** The comparison between the theoretical IMF (red dashed line) and the effective one (blue histograms) for different shapes and maximum mass,  $m_{\text{max}}$ , of the PopIII IMF. We show a Larson IMF with  $m_{\text{ch}} = 1M_{\odot}$  (left-hand panels),  $m_{\text{ch}} = 10M_{\odot}$  (middle panels), and a flat IMF (right-hand panels). The top (bottom) panels show results for  $m_{\text{max}} = 100M_{\odot}$  ( $m_{\text{max}} = 1000M_{\odot}$ ).



**Figure C1.** The evolution of iron-abundance (top panel), star formation rate (medium panel), and the ejected gas mass (bottom panel) for the model with (light blue) and without (dark blue) IMF random sampling.

the star formation for the IMF sampling model, we find that  $[\text{Fe}/\text{H}]$

$\approx -2.6$ , which corresponds to the peak of the MDF (Fig. 5). In the model without IMF sampling the star formation rate reaches higher values but this maximum appears at later times,  $(t - t_{\text{form}}) \approx 35$  Myr. Yet, because of the lower metal enrichment  $[\text{Fe}/\text{H}] \approx -2.7$ . This explains why the MDF is higher but shifted towards lower  $[\text{Fe}/\text{H}]$  values in the model without IMF sampling (Fig. 5).