A&A 625, A77 (2019)
https://doi.org/10.1051/0004-6361/201834911

Astronomy & Astrophysics

The ultra-diffuse galaxy NGC 1052-DF2 with MUSE

II. The population of DF2: stars, clusters, and planetary nebulae*

Jérémy Fensch[1], Remco F. J. van der Burg[1], Tereza Jeřábková[1,2,3], Eric Emsellem[1,4], Anita Zanella[1], Adriano Agnello[1,5], Michael Hilker[1], Oliver Müller[6], Marina Rejkuba[1], Pierre-Alain Duc[6], Patrick Durrell[7], Rebecca Habas[8], Sungsoon Lim[9], Francine R. Marleau[8], Eric W. Peng[10,11], and Rubén Sánchez Janssen[12]

[1] European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany
e-mail: jfensch@eso.org
[2] Helmholtz Institut für Strahlen und Kernphysik, Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany
[3] Astronomical Institute, Charles University in Prague, V Holešovičkách 2, 180 00 Praha 8, Czech Republic
[4] Université Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, 69230 Saint-Genis-Laval, France
[5] DARK, Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, 2100 Copenhagen, Denmark
[6] Observatoire Astronomique de Strasbourg (ObAS), Université de Strasbourg – CNRS, UMR, 7550 Strasbourg, France
[7] Youngstown State University, One University Plaza, Youngstown, OH 44555, USA
[8] Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstraße 25/8, Innsbruck 6020, Austria
[9] NRC Herzberg Astronomy and Astrophysics Research Centre, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada
[10] Department of Astronomy, Peking University, Beijing 100871, PR China
[11] Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, PR China
[12] UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

Received 18 December 2018 / Accepted 18 January 2019

ABSTRACT

NGC 1052-DF2, an ultra-diffuse galaxy (UDG), has been the subject of intense debate. Its alleged absence of dark matter, and the brightness and number excess of its globular clusters (GCs) at an initially assumed distance of 20 Mpc suggest a new formation channel for UDGs. We present the first systematic spectroscopic analysis of the stellar body and the GCs in this galaxy (six previously known and one newly confirmed member) using MUSE at the VLT. Even though NGC 1052-DF2 does not show any spatially extended emission lines, we report the discovery of three planetary nebulae (PNe). We conduct full spectral fitting on the UDG and the stacked spectra of all the GCs. The UDG's stellar population is old, 8.9 ± 1.5 Gyr; metal poor, $[M/H] = -1.07 \pm 0.12$; and with little or no α-enrichment. The stacked spectrum of all GCs indicates a similar age of 8.9 ± 1.8 Gyr, but a lower metallicity of $[M/H] = -1.63 \pm 0.09$ and a similarly low α-enrichment. There is no evidence for a variation in age and metallicity in the GC population with the available spectra. The significantly more metal-rich stellar body with respect to its associated GCs, the age of the population, its metallicity, and its α-enrichment are all in line with other dwarf galaxies. NGC 1052-DF2 thus falls on the same empirical mass–metallicity relation as other dwarfs for the full distance range assumed in the literature. We find that both debated distance estimates (13 and 20 Mpc) are similarly likely, given the three discovered PNe.

Key words. galaxies: star clusters: general – galaxies: stellar content – galaxies: dwarf

1. Introduction

Ultra-diffuse galaxies (UDGs) are a particular type of low surface brightness galaxies, defined as having central surface brightnesses of $\mu_{g,0} > 24\,\mathrm{mag\,arcsec^{-2}}$, and sizes of $R_{\rm eff} > 1.5$ kpc (van Dokkum et al. 2015). Galaxies with such properties have been known for several decades (Sandage & Binggeli 1984; Impey et al. 1988; Dalcanton et al. 1997; Conselice et al. 2003), but their particularly high abundance in galaxy clusters drew attention in the last few years (e.g., van Dokkum et al. 2015; Koda et al. 2015; Mihos et al. 2015; Muñoz et al. 2015; van der Burg et al. 2016). UDGs are now also routinely identified in groups and in the field (Román & Trujillo 2017; van der Burg et al. 2017; Shi et al. 2017; Müller et al. 2018).

To explain their high abundance in overdense regions such as the Coma cluster, van Dokkum et al. (2015) proposed that UDGs may be hosted by massive, Milky Way-like, dark matter (DM) halos that could protect them from environmental effects. One UDG in particular, DF44, was measured to have a stellar velocity dispersion consistent with a $10^{12}\,M_\odot$ halo (van Dokkum et al. 2016). In addition, the empirical linear relation observed between the mass of the globular cluster (GC) system and the halo mass (Blakeslee et al. 1997; Peng et al. 2004; Harris et al. 2017) allows this quantity to be used to assess the DM content of UDGs. The high number of GCs around DF44 (~100) would confirm the hypothesis of it being hosted by a very massive DM halo, along with a few other UDGs with a GC excess, but most UDGs have GC systems typical of dwarf galaxy DM halos (Beasley & Trujillo 2016; Amorisco et al. 2018; Lim et al. 2018). This is in line with a stacked weak-lensing study performed by Sifón et al. (2018), showing that not all UDGs can have halo masses similar to those

* Based on observations collected at the European Southern Observatory under ESO programs 2101.B-5008(A) and 2101.B-5053(A).

estimated for DF44. Formation scenarios need to explain how galaxies with similar masses and morphologies may be hosted in a broad variety of DM halo masses.

Di Cintio et al. (2017) suggested the possibility that internal processes (i.e., gas outflows associated with feedback) can, under some circumstances, kinematically heat the distribution of stars and form very extended systems similar to UDGs. An early cessation of star formation at $z \sim 2$ would render their stellar masses, and associated surface brightness, low (cf. Yozin & Bekki 2015). At different quenching times, this scenario results in UDG-like galaxies with low metallicities ($-1.8 \lesssim$ [Fe/H] $\lesssim -1.0$) and a range of ages (Chan et al. 2018). This is supported by photometric (Pandya et al. 2018) and spectroscopic observations (Kadowaki et al. 2017; Gu et al. 2018; Ferré-Mateu et al. 2018; Ruiz-Lara et al. 2018) of UDGs. The general consensus is that UDGs have stellar populations that are typically old (>9 Gyr) and metal poor ([M/H] ~ -0.5 to -1.5). Moreover, these studies found that the UDG stellar masses and stellar metallicities fall on the empirical relation found for dwarf galaxies (Kirby et al. 2013). They conclude that UDGs are most likely the result of both internal processes, such as bursty star formation histories (SFH) or high-spin halos (Amorisco & Loeb 2016; Rong et al. 2017), and environmental effects such as tidal disruption (Collins et al. 2013; Yozin & Bekki 2015). We note that UDGs may also form in tidal debris (see, e.g., Kroupa 2012; Duc et al. 2014; Bennet et al. 2018).

To reconcile both the discovery of UDGs with exceptional characteristics such as DF44 and the average properties of typical UDGs, several different formation channels need to be invoked. Most stellar population studies have targeted "ordinary" UDGs, with typical dwarf galaxy DM halos. While such galaxies may be well represented in current hydrodynamical simulations (see, e.g., Chan et al. 2018), an open question is how more extreme cases (for instance UDGs with an extremely high or extremely low halo mass for their stellar mass) have formed.

Of particular recent interest is the UDG NGC 1052-DF2[1] (hereafter DF2), which may have a special formation channel. Using the velocities of ten GCs associated with DF2, van Dokkum et al. (2018a) claimed a low total mass that is consistent with the stellar mass alone (however, see Martin et al. 2018; Famaey et al. 2018; Laporte et al. 2019, for a re-analysis). Hypotheses put forward by van Dokkum et al. (2018a) suggest that DF2 may have been formed by gas ejected by tides following a merger or quasar winds from the massive elliptical NGC 1052, whose projected distance is only 14′, or ∼80 kpc at a distance of 20 Mpc.

A second striking feature of this galaxy is its GC system. DF2 has 12 confirmed GCs (Emsellem et al. 2019), an unusually large population when compared to normal dwarf galaxies (Lim et al. 2018; Amorisco et al. 2018). This is at odds with the DM deficiency, as explained above. These GCs are also very luminous: their absolute magnitudes are similar to those of the most massive Milky Way GCs at an assumed distance of 20 Mpc (van Dokkum et al. 2018a). Trujillo et al. (2019) advocated for a closer distance of 13 Mpc, and showed that DF2 and its GCs would then fall on the same empirical relation as other UDGs. The exact distance of DF2 is still debated (van Dokkum et al. 2018b). The new GC candidates associated with DF2 from Trujillo et al. (2019), if confirmed, would move the peak of the GC luminosity function towards fainter magnitudes and alleviate the issue of "too bright" GCs. This would further increase the discrepancy between the DM halo mass estimated through GC kinematics and that from the GC abundance. In a companion paper (Emsellem et al. 2019, hereafter Paper I), we have indeed confirmed one new candidate GC from Trujillo et al. (2019).

In this series of two papers, we study DF2 with MUSE observations taken at the VLT. Thanks to the field of view of this integral field spectrograph, we are able to simultaneously probe the stellar body of the UDG and seven bright associated GCs for the first time. Paper I focuses on the kinematics of the UDG; this paper presents a stellar population analysis of this galaxy and its associated GCs.

In Sect. 2 we present the data reduction, sky removal, and extraction of spectra. We estimate the age and metallicity of the stellar body and the GCs in Sect. 3. We report the discovery of three planetary nebulae in Sect. 4. We discuss the origin and the distance of the UDG and its association with the surrounding GCs in Sect. 5. The conclusions are presented in Sect. 6.

2. Data

The details of the observations, reduction, and flux extraction procedures are detailed in Paper I. In the following we summarize the main points of the procedure.

2.1. Observation and reduction

MUSE observations of NGC 1052-DF2 were conducted via two ESO-DDT programs (2101.B-5008(A) and 2101.B-5053(A), PI: Emsellem) between July and November 2018 amounting to a total of ∼5.1 h on-target integration time. We obtained 28 individual exposures with slight dithers and rotations to account for systematics due to the slicers. We deliberately offset the MUSE field by ∼8″ with respect to the center of the galaxy (see Fig. 1) to include an area where the surface brightness of the UDG is several magnitudes fainter than in the center, which is used for the sky removal.

The OBs were all reduced using the latest MUSE ESOREX pipeline recipes (Sect. 2.4.2). The reduction follows the standard steps. As the object is very faint and standard sky subtraction was not able to recover a signal of sufficient quality, the full sky subtraction was done with the principle component analysis-based software Zurich Atmosphere Purge (ZAP; Soto et al. 2016). The principal components, or eigenspectra, are derived from the outermost regions of the MUSE object cube, where the sky is most dominant: the sky region is defined by excluding the bright sources and an ellipse centered on the UDG (see Paper I). In the following we use as fiducial datacube the output of the ZAP procedure with an ellipse of circularized radius 30″, 45 eigenvalues, and 50 spectral bins for the continuum filter. We discuss the effect of these parameters on the results in Sect. 3.2. The final data set, rendered in a mock HST broadband color image, using the same filters as in van Dokkum et al. (2018a), is shown in Fig. 1.

2.2. Extraction of spectra

The detection of the sources is described in Paper I.

We first created a spatial mask, presented in Paper I, to remove the background and foreground objects surrounding the UDG. We extracted the spectrum of the UDG by summing each

[1] This name was tagged by van Dokkum et al. (2018a). However, this galaxy was already identified under different names (KKSG04, PGC 3097693, [KKS2000]04, see, e.g., Karachentsev et al. 2000; Trujillo et al. 2019). We choose to keep NGC 1052-DF2 as the galaxy has been popularized under this name in the recent literature.

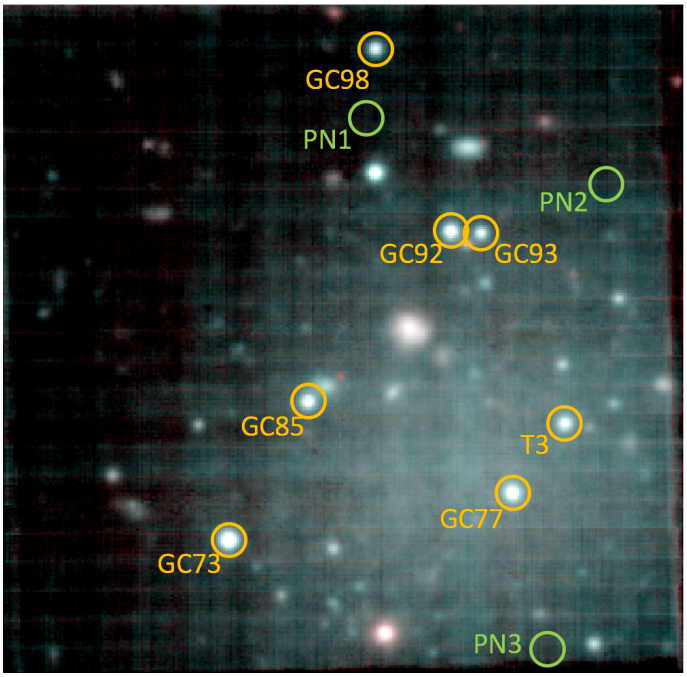

Fig. 1. Mock true color image of the collapsed MUSE cube within the F606W and F814W HST bandpasses. The GCs confirmed in Paper I are shown with orange circles. The locations of the three confirmed PNe are shown with green circles. The PNe do not show up in these mock broadband images. The field of view is $1' \times 1'$.

channel of the masked cube with a spatial weight corresponding to the flux of the UDG in the HST F814W image.

The GC and PN spectra are extracted with a Gaussian weight function to provide a S/N-optimized extraction. The full width at half maximum is set to $\sim 0.8''$ to approximately match the point spread function. The background is measured locally with identical apertures in eight nearby locations that do not overlap with identified sources. In each channel we obtain the source flux by subtracting the median of the sky exposures from the weighted sum of the source spectrum. The dominant source of uncertainty is taken from the scatter in the sky spectrum values. The relative velocities are small (see Paper I), thus we do not correct for the relative velocities of the GCs. Contrarily to van Dokkum et al. (2018c), we do not weight each GC by its S/N. This would provide us with the highest reachable S/N, but the brightest source, GC73 (see Fig. 1), would dominate the stack.

To estimate the physical spread in the different parameters (age, [Fe/H], α-enrichment) in the GC population we also create 100 bootstrapped spectra. These are new spectra constructed by adding together seven GC spectra that are randomly picked from the sample with replacement.

3. Stellar populations of DF2: stellar body and GCs

We show in Fig. 2 the spectrum obtained for the UDG and the stack of all GCs. We note strong Balmer and calcium triplet (CaT) absorption lines, plus shallower absorption lines such as Mg and Fe. We do not detect any emission lines. This is consistent with the nondetection of atomic gas which implies a stringent upper limit on the gas fraction of DF2 (below 2%, see Chowdhury 2019). Around the Hα line, we estimate a signal-to-noise ratio (S/N) of 62 pix^{-1} for DF2 and 72 pix^{-1} for the stack of all GCs (see Paper I).

3.1. Fitting procedure

We use the fitting routine pPXF (Cappellari & Emsellem 2004; Cappellari 2017) combined with the eMILES library (Vazdekis et al. 2016). The details of the fitting procedure are given in Paper I. In the following we summarize the main points of the procedure.

As template spectra we use the eMILES single stellar populations (SSPs) with a Kroupa (2001) initial mass function (IMF) and the Padova 2000 (Girardi et al. 2000) isochrones, which were shown to perform well in the expected regime of old and low-metallicity stellar populations (Conroy et al. 2009). The original range of metallicity values being rather sparse (only 7 values covering [Fe/H] from −2.32 to 0.22 with logarithmic spacing), we linearly interpolate for 16 more metallicity values between [Fe/H] = −2.32 and −0.71, following Kuntschner et al. (2010).

To avoid being biased by the flux calibration differences between our MUSE data and the eMILES library, we make use of multiplicative polynomials during the fit (Cappellari 2017). For the study we chose to allow for a 12-degree Legendre multiplicative polynomial; the impact of changing the degree is discussed in Sect. 3.2.

3.2. Ages and metallicities: fitting method

We estimate the stellar population parameters by fitting single stellar populations (SSPs) to our spectra. The parameter uncertainties are derived from fitting 100 new spectra, constructed by adding the randomly shuffled residuals to the best fit. The best SSP fits for the UDG and GC stack are shown in Fig. 2. It should be noted that the first CaT line is masked during the fit (as in van Dokkum et al. 2018a) because it is located in a region affected by sky residuals. It is nonetheless well recovered by the fits, which independently shows that the sky subtraction did not affect these lines.

The location of the UDG and the GCs in the age-metallicity plane is shown in Fig. 3 with the estimation of age and metallicity for the stack of GCs from van Dokkum et al. (2018c), along with their 1σ error bars. For the UDG we find a best fitting age of 8.9 ± 1.5 Gyr and metallicity [M/H] $= -1.07 \pm 0.12$. For the full stack of GCs, we find a best fitting age of 8.9 ± 1.4 Gyr and metallicity of [M/H] $= -1.63 \pm 0.09$. The parameters of the best fits are not sensitive to a change in the degree of the multiplicative polynomial between 11 and 15, nor to a change of parameters in the ZAP procedure (masked radius of $30''$ or $36''$, number of eigenvalues of 30, 45, or 50, and continuum filter window size of 30 and 50 Å).

It should be noted that our method does not consider the detailed continuum shape to derive the parameters because of the use of multiplicative polynomials. In order to check the consistency of our estimates with the broadband colors, we compute the AB magnitude color of the eMILES templates in F606W−F814W. The colors of the two best fit templates, 0.40 mag for the UDGs and 0.35 mag for the GC stacks, agree with the colors computed by van Dokkum et al. (2018c): respectively 0.37 ± 0.05 mag and 0.35 ± 0.02 mag for the UDG and the GC stack.

The age and metallicity estimated for the GC stack are consistent within 1σ for the ages and 2σ for the metallicity with the values obtained by van Dokkum et al. (2018c): age of $9.3^{+1.3}_{-1.2}$ Gyr and [Fe/H] $= -1.35 \pm 0.12$. We obtain a lower metallicity for our GC stack, but it should be noted that we do not use the same stellar libraries for the fits, and the IMF they assume

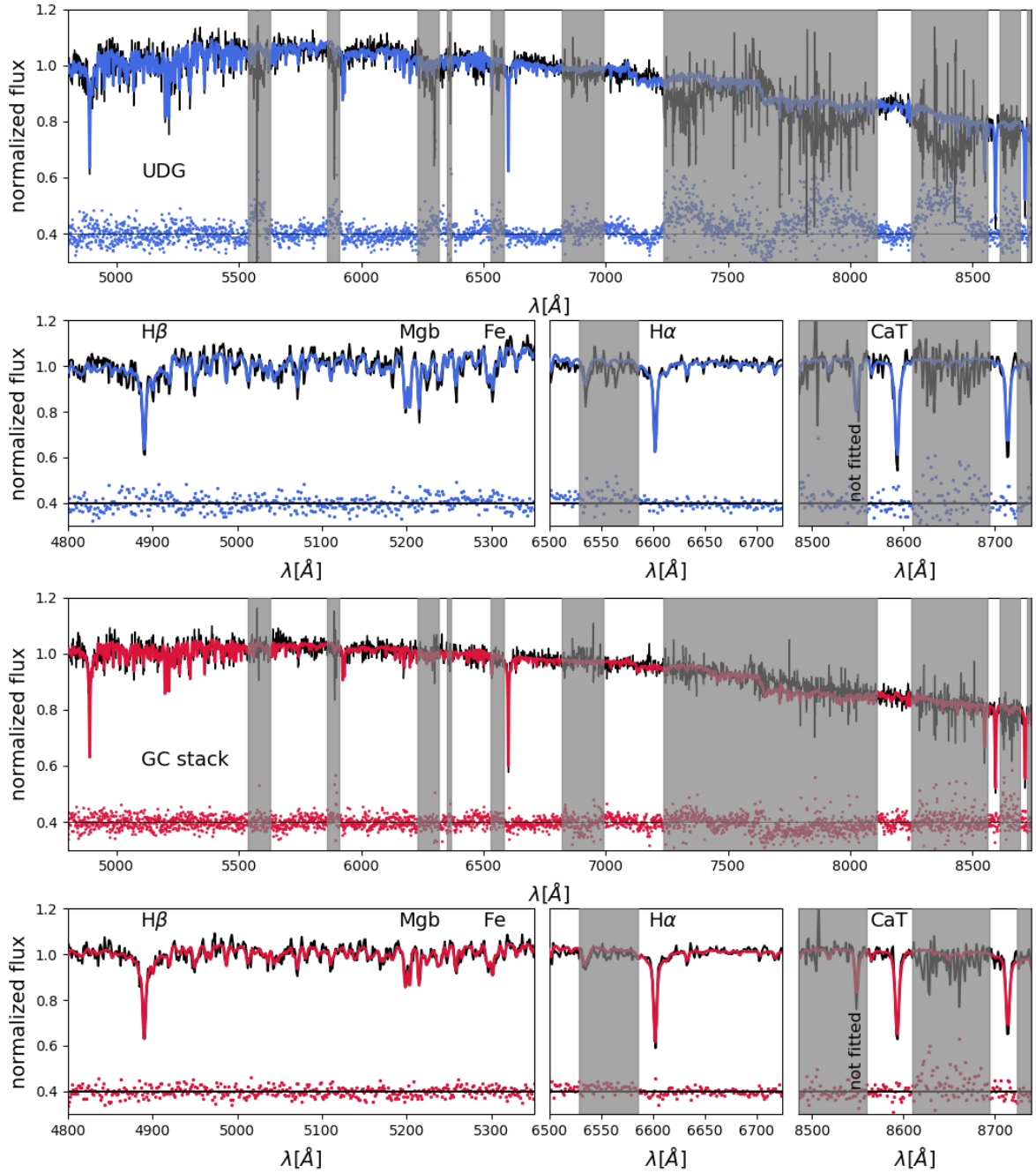

Fig. 2. Comparison between the spectrum and the best fit from pPXF for the UDG and the GC stack spectra. The three plots in the bottom part of each panel show zooms on the important absorption lines. The gray regions are not taken into account for the fit. The scatter points show the residuals.

is not described. Furthermore, the spectral region studied in van Dokkum et al. (2018c) extends further into the blue compared to our MUSE data, where different spectral diagnostics contribute to the fit. Even though these differences may drive systematic shifts between the parameters measured in different studies, we note that in this work we make a direct comparison between the GCs and the stellar body from a single data set, which have similar S/N and assumptions. Even though the exact age and metallicity may be affected by different systematics, the relative differences between GCs and stellar body are significant and robust.

To quantify the spread in age and metallicity inside the GC population, we use our method on the bootstrapped spectra. The median of the 100 realizations has the same age and metallicity as the GC stack. The error bar shows the propagation of the measurement error and the dispersion of the results in the bootstrap sample. We obtain an age of 8.9 ± 2.1 and a metallicity of −1.63 ± 0.11. We see that it is of the same order of magnitude as the error on the estimation of the parameters of the bootstrap, meaning that there is no significant evidence for a spread in properties between the individual clusters. Finally, we used pPXF on three radial sectors of the UDG: inside 0.5 R_e, between 0.5 and 1 R_e, and between 1 and 1.5 R_e, where R_e is the effective radius of DF2 (see Paper I). The best fits all have the same age, but the central sector's metallicity estimate is higher: [Fe/H] = −1.07 ± 0.12 compared to [Fe/H] = −1.19 ± 0.12 and −1.19 ± 0.14 for the two

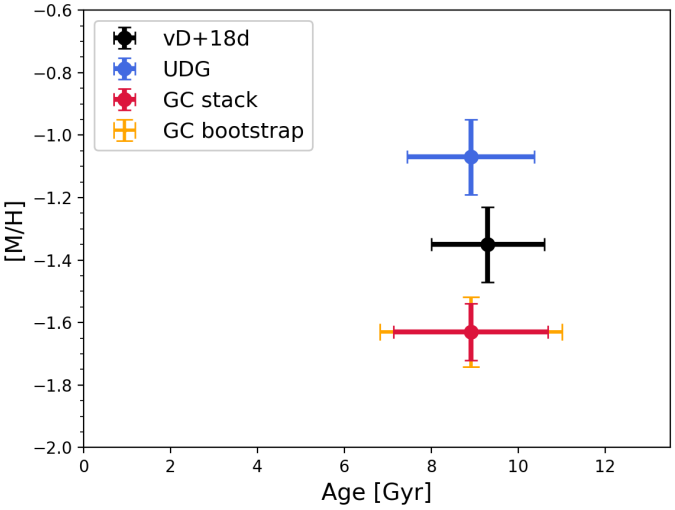

Fig. 3. Location of the best fit and 1σ error bars in the age–metallicity plane for the UDG and the GC stack. The result of the study by van Dokkum et al. (2018c) is shown in black. The location of the median age and metallicity of the GC bootstrap sample is shown in orange. The orange error bar is both a measure of the error of the fit and an estimate of the physical parameter spread intrinsic to the GC sample.

outer sectors. We thus note that the higher metallicity found in the center of DF2 is not significant, and that the metallicity gradient in DF2 is consistent with being flat.

3.3. Ages and metallicities: spectral indices

We use a complementary method to estimate ages and metallicities based on the measurement of spectral line indices. In the following we work in the standardized Lick/IDS system (Worthey et al. 1994), and we list several key diagnostics in terms of age and metallicity in Table 1. Two diagnostics shown in Fig. 4, with overplotted grids of theoretical Lick indices of SSPs and based on the MILES spectral library, were obtained from Thomas et al. (2010).

To study the α-enrichment of the GCs and the UDG, we plot in the left panel of Fig. 4 Mg b, as a probe of the α elements, and $\langle$Fe$\rangle$ (the average of Fe λ5270 and Fe λ5335, following Evstigneeva et al. 2007). The α-enrichment of the GC stack and the UDG are not well constrained due to the small separation of iso-[α/Fe] lines in the metal-poor regime. Still, the diagnostics infer slight α-enrichment: from 0 to 0.15 for the UDG and from 0 to 0.3 for the stack of GCs. The latter value is consistent with the value derived by van Dokkum et al. (2018c) ([α/Fe] = 0.16 ± 0.17).

The right panel of Fig. 4 shows the age-sensitive index Hβ versus [MgFe]$' = [\mathrm{Mg}\, b \times (0.72\mathrm{Fe}\,\lambda5270 + 0.28\mathrm{Fe}\,\lambda5335)]^{1/2}$ (which probes the total metallicity, following Evstigneeva et al. 2007). The Lick index suggests a slightly higher metallicity than our full spectral fitting method indicates. However, it confirms the trend given by the first method that DF2 and the stack of all GCs have similar ages, but that the UDG has a higher metallicity.

Even though measurements of Lick indices from individual GC spectra are imprecise given the noise in our spectra, we indicate the cluster-to-cluster variation by showing a distribution of bootstrapped stacks. The scatter of these realizations is represented by the second error quoted in Table 1. We note that the bootstrap uncertainties are similar to the formal statistical uncertainties (first errors quoted). Since the bootstrap error, by construction, also includes the statistical uncertainty on the measurement, their similar magnitude confirms that there is no significant spread in the properties of the individual clusters.

The two different methods indicate that the DF2 stellar population is as old as the GC population, around 9 Gyr, and is significantly more metal-rich, by around 0.5 dex.

4. The DF2 planetary nebulae

The spectra of the three detected PNe are shown in Fig. 5. Their kinematic association with the stellar body of the UDG is confirmed in Paper I. We see strong emission from the [O III] doublet and Hα lines. However, Hβ and [N II] are not detected for any of the PNe, which prevents us from computing their intrinsic extinction or metallicity. The measurement of their apparent 5007 Å magnitude is given in Table 2. It is defined[2] as

$$m_{5007} = -2.5 \log F_{5007} - 13.74, \quad (1)$$

where F_{5007} is the integrated flux in the second [O III] line in erg s^{-1} cm^{-2}. We check our flux calibration by comparing the flux of our GCs with those presented in Trujillo et al. (2019) with their HST observations. We found that the flux from the MUSE cubes are brighter by 0.064 ± 0.079 mag. We neglect this calibration difference and use the flux calibration from MUSE.

We assume a foreground extinction[3] of 0.076 ± 0.006 mag, corresponding to mean of the computed extinction for the line of sight of NGC 1052 (Schlegel et al. 1998; Schlafly & Finkbeiner 2011). The uncertainty includes a propagation of the uncertainty on the foreground extinction, on the flux calibration, and on the flux measurement. The flux measurement uncertainty is obtained by re-noising the spectrum with a Gaussian noise with a dispersion measured in the continuum redward of the [O III] line. Hβ is not detected in any of the PNe, and the S/N prevents us from inferring a meaningful lower limit on the extinction. Thus, we did not correct for internal extinction.

5. Discussion

van Dokkum et al. (2018a) argued that DF2 is DM-deficient and is a very different system from other galaxies, in particular from UDGs that were routinely shown to be hosted by dwarf to Milky Way-sized DM halos (see Sect. 1). In the following we discuss whether DF2 also stands out in terms of its stellar populations.

5.1. How does DF2 compare with other UDGs?

5.1.1. In terms of stellar populations

In Sect. 3.2, we estimated the age, the metallicity, and the α-enrichment of DF2 and its GCs.

We found that the stellar population of DF2 is old, around 9 Gyr. We note that our age estimate should be taken as a lower-limit as blue horizontal branch stars could bias our age estimate to lower ages (Schiavon 2007; Conroy et al. 2018). This age estimate, however, is similar to those obtained for other quiescent UDGs (Gu et al. 2018; Ruiz-Lara et al. 2018; Ferré-Mateu et al. 2018).

The few UDGs with α-enrichment measurements, which are located in the Coma cluster, have [α/Fe] estimated to be between

[2] The standard magnitude zeropoint for PNe is set at 13.74 to approximate the absolute V-band magnitude one would observe if all the [O III] line emission were distributed over the V-band (Allen 1973).

[3] Model accessible at https://irsa.ipac.caltech.edu/applications/DUST/

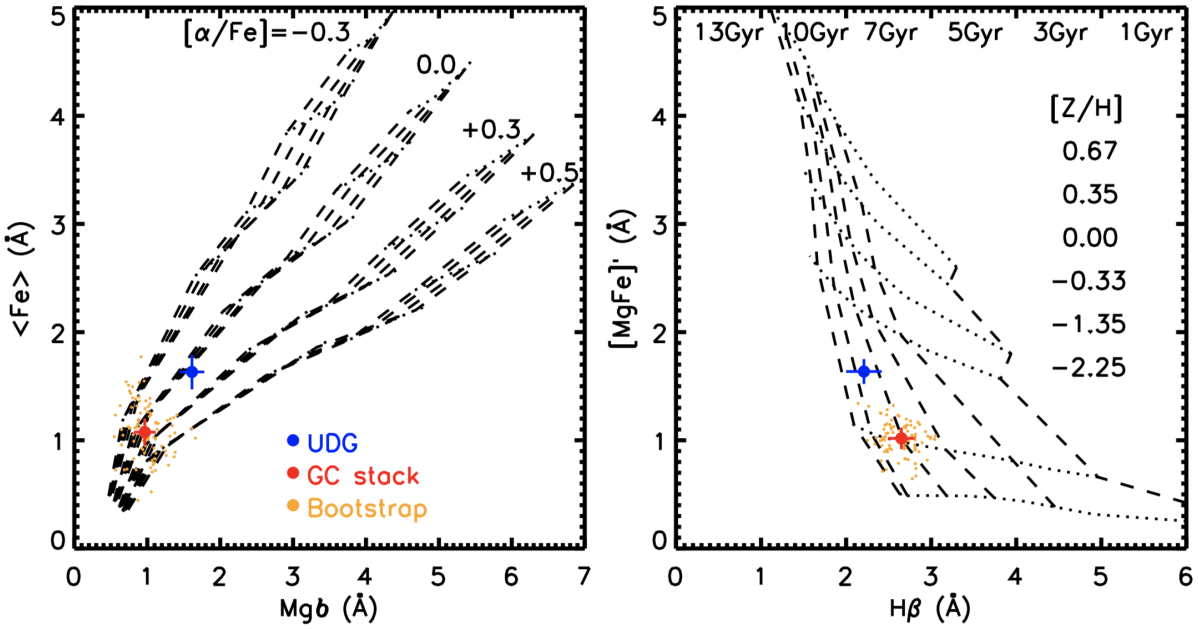

Fig. 4. Lick/IDS indices measured from our spectra of the UDG (blue) and the stacked GC spectrum (red). The cloud of 100 points (orange) shows the bootstrap runs of the GCs, as detailed in the text, and indicates the scatter intrinsic to the stack. The underlying model grid is based on Thomas et al. (2010). A grid with ages of 1, 3, 5, 7, 10, and 13 Gyr (dashed lines), and [Z/H] of −2.25, −1.35, −0.33, 0.0, +0.35, and +0.67 dex (dotted lines) is overplotted in both panels. For clarity, lines for 1 and 3 Gyr are omitted in the *left panel*. *Left panel*: comparison of Mg *b* and ⟨Fe⟩ as a measure of α-enrichment. *Right panel*: comparison of [MgFe]′, a measure of total metallicity, and Hβ, which is primarily an age indicator.

Table 1. Lick/IDS indices for the UDG spectra and stacked GC spectra in Å.

	Hβ	Mg *b*	Fe 5270	Fe 5335	⟨Fe⟩	[MgFe]′
UDG	2.21 ± 0.21	1.61 ± 0.17	1.69 ± 0.20	1.57 ± 0.25	1.63 ± 0.16	1.64 ± 0.11
GCs	2.65 ± 0.16 ± 0.19	0.97 ± 0.15 ± 0.22	1.07 ± 0.18 ± 0.18	1.08 ± 0.18 ± 0.30	1.08 ± 0.13 ± 0.23	1.02 ± 0.10 ± 0.14

Notes. The first errors represent statistical uncertainties measured from the stacked spectra. Also listed is the scatter obtained from bootstrapping the GCs that end up in the stack.

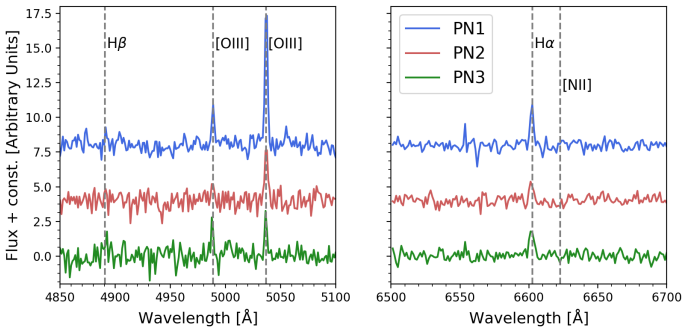

Fig. 5. Zoom-in on the spectra of the three PNe. *Left panel*: region around the [O III] doublet and *right panel*: region around the Hα line. The fluxes are arbitrarily shifted for clearer visualization. The location of the brightest expected lines for PNe, observed for $z = 0.06$, are shown with vertical dashed lines.

Table 2. 5007 Å apparent and absolute magnitude for the three PNe.

	m_{5007}	M_{5007} at 13 Mpc	M_{5007} at 20 Mpc
PN1	28.4 ± 0.05	−2.24 ± 0.05	−3.18 ± 0.05
PN2	29.32 ± 0.14	−1.32 ± 0.14	−2.26 ± 0.14
PN3	29.91 ± 0.16	−0.73 ± 0.16	−1.67 ± 0.16

Notes. The absolute magnitude is obtained for two assumed distances to DF2: 13 and 20 Mpc.

0 and 0.6 dex (Ruiz-Lara et al. 2018; Ferré-Mateu et al. 2018). Our estimation of [α/Fe] for DF2 between 0 and 0.3 is located within this range of values.

To study the metallicity of DF2, we show in Fig. 6 the location of the UDG in the mass–metallicity plane along with data from previous studies of quiescent UDGs. We indicate two different stellar masses for DF2, one inferred for a distance of 20 Mpc ($2-3\times10^8\ M_\odot$; van Dokkum et al. 2018a) and the other for a distance of 13 Mpc ($6\pm3\times10^7\ M_\odot$; Trujillo et al. 2019). We see that DF2 has a similar metallicity to the other UDGs previously studied, and falls on the empirical relation for dwarf galaxies from Kirby et al. (2013) for both mass estimates. We note that our data provide us with a much tighter metallicity estimate than most of the values available in the literature.

This stellar mass–metallicity relation is interpreted as an effect of self-enrichment. The more massive a galaxy, the less metal is lost to galactic winds launched by star formation feedback (Kirby et al. 2013). The mass–metallicity relation may also result from the galaxy-wide stellar IMF becoming systematically lighter at the top with decreasing baryonic mass or star

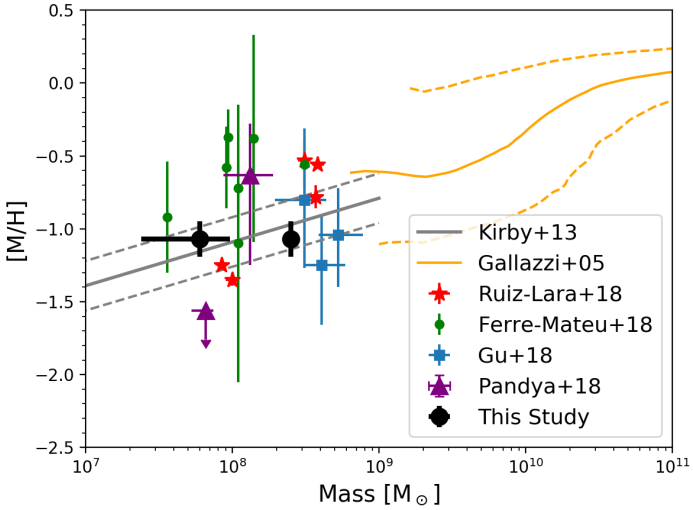

Fig. 6. Stellar mass and metallicity of DF2 and UDGs from the literature (Gu et al. 2018; Ferré-Mateu et al. 2018; Ruiz-Lara et al. 2018; Pandya et al. 2018). We show the location of DF2 for two mass estimates, corresponding to the two distance estimates of 13 and 20 Mpc (see text). The empirical mass–metallicity relations for low-mass and high-mass systems (from Kirby et al. 2013 and Gallazzi et al. 2005, respectively) are shown in gray and orange.

formation rate as shown to be the case using the IGIMF theory (Köppen et al. 2007; Recchi et al. 2015).

If we assume that DF2 is DM-deficient (van Dokkum et al. 2018a), we would then expect DF2 to be an outlier of the relation with a lower metallicity than galaxies with the same stellar mass, which typically have a halo mass of $10^{10}\,M_\odot$ (see, e.g., Read et al. 2017). However we see that for the assumed distance of 20 Mpc, corresponding to the DM-deficiency hypothesis, DF2 lies within the scatter of the relation. Even though the scatter of the relation is quite large (∼1 dex), DF2 has a higher metallicity than DF44 ([Fe/H] = −1.3 ± 0.4, see Gu et al. 2018), which has a similar stellar mass and an extremely massive DM halo (∼$10^{12}\,M_\odot$, see van Dokkum et al. 2016).

A first possibility could be that DF2 had a larger stellar mass than today and gradually lost part of it due to stripping. This stripping would not modify the metallicity of DF2, but would move its location in this plot horizontally towards lower stellar mass and thus closer to the relation. This process, which could explain the location of some dwarfs above the stellar mass–metallicity relation (see the case of Antlia2; Torrealba et al. 2019), could also move a metal-deficient UDG closer to the relation. Furthermore, Trujillo et al. (2019) note a significant brightening of DF2 in the northern region in ultra-deep g-band Gemini data, which might be a trace of a past stripping event. We note that the stripping of the stars only begins when most of the DM mass is already lost (around 90%; see, e.g., Peñarrubia et al. 2008). Stellar mass stripping could then fit in the hypothesis of a DM-deficient galaxy. However, such a stripping scenario should also affect the GC system which should be stripped, or at least heated kinematically (Smith et al. 2013), which does not seem consistent with both the number (see Sect. 5.1.2) or the low velocity dispersion of the GCs associated with DF2.

A second possibility is that the gas of DF2 was already enriched in metals. This could be the case if DF2 was formed through tidally stripped material. We discuss this possibility in detail in Sect. 5.3.

Overall, DF2 shows a stellar population typical of quiescent UDGs. Its location in the mass–metallicity plane, which is very similar to that of dSphs, is not what we would expect for a DM-free galaxy. This could be a hint to the origin of this galaxy.

5.1.2. In terms of GC systems

In Fig. 3 we see that the metallicity of the GCs surrounding DF2 is significantly lower than that of the UDG, by around 0.5 dex. Lotz et al. (2004) found that field stars in 45 local dE are typically 0.1−0.2 mag redder than their GCs, which they interpreted as a legacy of different star formation events and/or different metallicities. This color mismatch seems to be lower for UDGs (less than 0.05 mag for DF17 and DF2 Beasley et al. 2016; van Dokkum et al. 2018c). In the case of DF2, we can show that it is driven by the stellar body being more metal-rich than the GCs. This is typical for dwarf galaxies of similar masses, including the Fornax dSph, which has an excess of GCs (see, e.g., Cole et al. 2012; Larsen et al. 2014).

In the left panel of Fig. 4 we see that the α-enrichment of the GCs is between [α/Fe] = 0 and 0.3. These are also typical values for GCs in dwarfs, whose GCs are known to be less α-enriched than those in more massive galaxies (Sharina et al. 2010). Thus, the stellar populations of GCs around DF2 do not seem to deviate from previous known systems.

DF2 seems to have a rather high specific frequency[4] of GCs compared to other UDGs (above 11, see van Dokkum et al. 2018c). Studies have shown that the S_N of UDGs varies dramatically from galaxy to galaxy and is on average higher than in dwarf galaxies (Amorisco et al. 2018; Lim et al. 2018). Moreover, we note that the fraction of light that is in GCs for DF2 is similar to that of other UDGs (such as DF17; see van Dokkum et al. 2015; Peng & Lim 2016). The only feature in the GC system of DF2 that differs from other GC systems, and which remains unexplained, is that the peak magnitude of the GC luminosity function is unusually high if we assumes a 20 Mpc distance (van Dokkum et al. 2018c). We note that Trujillo et al. (2019) found that the GC luminosity function of DF2 is standard, if located at a distance of 13 Mpc.

5.1.3. In terms of PNe

It is the first time that PNe have been discovered around a UDG, thanks to the use of an integral field unit (IFU) spectrograph with good spatial resolution. It is possible to compare our number of detections with an estimate of the expected number of PNe for such a system.

The total number of PNe per bolometric luminosity of the host galaxy is parametrized as $\alpha = N_{PN}/L_{bol}$. We define $\alpha_{2.5}$ as the number of PNe in the brightest 2.5 mag of the planetary nebula luminosity function (PNLF) per bolometric luminosity. While stellar evolutionary models still have difficulties in reproducing the constancy of the PNLF bright cut-off in galaxies of different morphology (see, e.g., Marigo et al. 2004), the study of the luminosity-specific PN numbers (the α parameter) in external galaxies (Buzzoni et al. 2006) provides a way of estimating the expected number of PNe in a galaxy. A typical α for metal-poor populations is $\sim 3\times10^{-7}$ PN per $L_{bol}/L_\odot$. The three detected PNe are probably in the brightest 2.5 mag of the PNLF, and, using the standard PNLF, $\alpha_{2.5} \approx \alpha/10$. So if $L_{bol} \approx 6\times10^6$ to $10^8\,L_\odot$ for DF2, then our three PNe imply $\alpha \sim 3\times10^{-8}$ to 5×10^{-7}, in reasonable agreement with expectations for a metal-poor stellar population (Buzzoni et al. 2006). We note that our field of view

[4] Number of GC (N_{GC}) per 15 absolute magnitude in the V band (M_V): $S_N = N_{GC}\,10^{0.4(M_V+15)}$ (Harris & van den Bergh 1981).

does not cover all the outskirts of DF2 where other PNe may be found. Thus, DF2 does not seem to have a different PNe formation rate than other systems.

5.2. What is the distance to DF2?

The distance of DF2 is the subject of an as-yet-unsettled debate. Indeed, as noted in van Dokkum et al. (2018a), a shorter distance would give a smaller stellar mass and increase the DM mass needed to recover the velocity dispersion measured.

van Dokkum et al. (2018a) computed a distance of 19.0 ± 1.7 Mpc from the surface brightness fluctuations (SBF) of the stellar body of DF2 and adopted a nominal distance of 20 Mpc. This distance was confirmed by an independent team, using the same technique (Blakeslee & Cantiello 2018). Trujillo et al. (2019) claim that the calibration used by van Dokkum et al. (2018a) is only valid for colors redder than that of DF2, and that the extrapolation to bluer colors is not trivial. They use five different redshift-independent methods to compute the distance of DF2 which all give consistent result of ∼13 Mpc. For such a distance, the measured velocity dispersion cannot be achieved without a significant DM content. van Dokkum et al. (2018b) demonstrated that the tip of the red giant branch (TRGB) stars may be blended in the HST images. By using a megamaser-TRGB-SBF distance ladder they find a new estimate of the distance of 18.7 ± 1.7 Mpc, which is consistent with their first distance estimate.

Another reliable distance estimator at these distances is the bright abrupt cut-off of the PNLF, whose absolute magnitude is almost independent of galaxy type, at around $M^\star = -4.51$ mag (see Ciardullo 2012, for a recent review). However, a trend towards a fainter cut-off magnitude in low-metallicity galaxies is expected from theoretical models (Dopita et al. 1992; Schönberner et al. 2010), which is confirmed by observations (see, e.g., Ciardullo 2012). Unfortunately, low-metallicity objects are usually not very massive and do not have enough PNe to sample the PNLF well. Hence, the metallicity dependence of $M^\star$ is hard to probe at the low-metallicity end. In particular, the Dopita et al. (1992) theoretical relation was not confirmed at metallicities lower than that of the SMC. The cut-off magnitude of the low-metallicity SMC and NGC 55 are estimated to be around $M^\star = -4.10$ (see review by Ciardullo 2012). The S/N does not allow us to detect the [NII] line in the spectra of the PNe for a direct metallicity estimate of the PNe. If we extrapolate the Dopita et al. (1992) relation to the stellar metallicity of 1/10 solar for DF2, derived in Sect. 3.2, we would expect a cut-off magnitude of $M^\star = -3.67$ for the PNLF of DF2.

Our IFU observations allowed us to find three PNe. To quantify how much information these three PNe provide us on the distance estimate, we perform a maximum likelihood estimation (MLE), using the PNLF from Ciardullo et al. (1989). For a cut-off magnitude $M^\star$, the number of PN with absolute magnitude M is proportional to

$$N(M) \propto \mathrm{e}^{0.307\,M}(1 - \mathrm{e}^{3(M^\star - M)}). \quad (2)$$

The likelihood function L can be written as

$$L = \prod_{i=1}^{3} \frac{N(m_i - \mu)}{\int_{M^\star}^{m_1 - \mu} N(m)\,\mathrm{d}m}, \quad (3)$$

with μ the distance modulus, m_i the apparent magnitude of each PN, and m_1 the completeness limit. We set as completeness limit a [O III] emission line peaking at three times the local rms measured for the PNe. This gives $m_1 = 30.64$ mag. We minimize $-\ln(L)$ by varying μ and $M^\star$. We define respectively the 1σ and 3σ error range by the range of parameters for which respectively, $\Delta \ln(L) < 0.5$ and 4.5. In Fig. 7 we show the result of the MLE. For $M^\star = -4.51$, the distance that maximizes the likelihood is 19.0 Mpc, and the 1σ and 3σ upper limits are respectively 33.3 and 38.1 Mpc. These values decrease for fainter $M^\star$. In particular, for $M^\star = -3.67$, which is the value expected from Dopita et al. (1992) for 1/10 solar metallicity, the distance that maximizes the likelihood is 12.9 Mpc, and the 1σ and 3σ upper limits are respectively 22.6 and 25.9 Mpc.

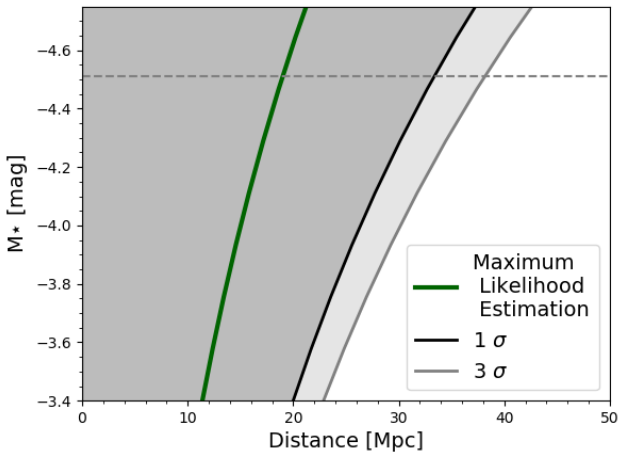

Fig. 7. Maximum likelihood estimation and error range on the distance and cut-off magnitude. The distance that maximizes the likelihood is shown in green. The 1σ error sector is in dark gray with a thick dark line border. The 3σ error sector is in light gray with a thick gray line border. The horizontal dashed gray line shows $M^\star = -4.51$.

Thus, we note that none of the two former distance estimates is significantly more likely, given the three discovered PNe. Given the off-centered field of view that we chose, we may have missed a brighter PN in the southwest part of DF2. In order to give strong constraints on the distance to DF2 the potential brightest PNe would need to be 1.5 mag brighter than PN1, a magnitude for which a 20 Mpc distance would be ruled out by 3σ.

5.3. What is the origin of DF2?

From a kinematic study of ten GCs surrounding DF2, van Dokkum et al. (2018a) inferred a low (projected) velocity dispersion, which they interpreted as DF2 "lacking" DM. This claim has been heavily scrutinized (Trujillo et al. 2019; Famaey et al. 2018; Laporte et al. 2019; Kroupa et al. 2018) and is revisited in Paper I. If confirmed by other independent tracers, this lack of DM calls for an additional formation channel to explain the existence of both DM-deficient and DM-dominated UDGs (such as DF44, van Dokkum et al. 2016).

As a first hypothesis, van Dokkum et al. (2018a) propose that the claimed lack of dark matter in DF2 may be explained if DF2 is a tidal dwarf galaxy (TDG), i.e., a galaxy formed from material that was expelled from a massive galaxy host during a galactic interaction (see review by Duc & Mirabel 1999). The proximity of the massive galaxy NGC 1052 and the peculiar radial velocity of DF2 (+293 km s^{-1} if at 20 Mpc) would support this hypothesis. Moreover, galaxies with typical morphological parameters of UDGs were observed to be still connected by a stellar stream to a massive host (Bennet et al. 2018). Unfortunately, no measurement of the stellar populations of those systems has been performed so far.

Because of this particular mode of formation, TDGs are indeed expected to be dark-matter free (Bournaud & Duc 2006; Wetzstein et al. 2007; Lelli et al. 2015) and almost devoid of stars from the host (Boquien et al. 2010). Furthermore, old TDGs enter in the category of UDGs with their low central surface brightness and large effective radii (Duc et al. 2014). Interestingly, we note that the cluster formation efficiency of TDGs, which is the fraction of SFR that happens in bound clusters, is seen to be very high (50%) compared to other systems (Fensch et al. 2018). Last but not least, they inherit their metal-enrichment from the more massive host. All observed tidal dwarf galaxies, which have stellar ages of typically less than 1 Gyr, deviate from the luminosity–metallicity diagram and have a significantly higher metallicity than other dwarfs for a similar luminosity, with a typical metallicity of around half solar, independent of their mass (Duc et al. 2000; Weilbacher et al. 2003). They are thus outliers of the stellar mass–metallicity relation.

Given the age of the stellar population of DF2, if it is a TDG the interaction must have happened at around $z = 2$ where the metal-enrichment of the gas in the outskirts of the host galaxy could still be quite low (see, e.g., Jones et al. 2013). Unfortunately, there is not much data on old TDGs as their low surface brightness makes them difficult to study (but see Duc et al. 2014), unless some or all of the Milky Way and Andromeda satellite galaxies are very old TDGs (Metz & Kroupa 2007; Pawlowski et al. 2011; Yang et al. 2014). In Sect. 5.1.1 we noted that DF2 is on the stellar mass–metallicity relation, contrarily to the young TDGs. If the level of pre-enrichment is between 0.001 and 0.01 $Z_\odot$, it is possible that the TDGs would not reach the mass–metallicity relation of young TDGs after many Gyr (Recchi et al. 2015). Under the DM-deficiency hypothesis, a small pre-enrichment of DF2 could then explain the location of DF2 in the stellar mass–metallicity diagram (see discussion in Sect. 5.1.1). Moreover, we note that most GCs in "normal" dwarf galaxies with spectroscopic metallicity measurements are very metal-poor (e.g., [Fe/H] ~ -2 dex, for GCs in the Fornax dSph, see de Boer & Fraser 2016). A coeval formation of the UDG and its clusters in pre-enriched gas ejected from a massive galaxy could explain how the GCs of DF2 have been enriched to [Fe/H] ~ -1.6 dex. Thus, the metallicity of DF2 and its GCs could be consistent with the TDG origin hypothesis.

6. Conclusions

We present the first simultaneous analysis of the stellar population of a UDG and its surrounding globular clusters.

We fit SSPs to the starlight component of the stellar body and the stack of all GCs using the empirical stellar library eMILES with the fitting routine pPXF.

We find that the UDG stellar populations are consistent with an old age, 8.9 ± 1.5 Gyr; low metallicity, [M/H] $= -1.19 \pm 0.11$; and little to no α-enrichment, i.e., formed on a timescale larger than 1 Gyr. The GC spectra are consistent with the same age, 8.9 ± 1.4 Gyr, but have a lower metallicity than DF2 ([Fe/H] $= -1.55 \pm 0.09$). This result is consistent with the Lick indices diagnostics and the broadband colors of DF2 and its clusters.

The stellar mass and metallicity of the UDG fall on the empirical relation found for old dwarf galaxies. In particular, DF2 has a comparable metallicity to DF44, which has the same stellar mass, but was shown to have a Milky Way-like DM halo. This relation is a consequence of the self-enrichment of galaxies and thus depends on the total mass of the galaxy. Under the DM-deficiency hypothesis we would then expect DF2 to have lower metallicity than galaxies with similar stellar mass. We note that stellar mass loss due to stripping could move a metal-deficient galaxy back to the relation, but this would affect its GC system, which does not seem to be the case for DF2. Another hypothesis is that DF2 has a tidal origin and was formed by gas pre-enriched in metals.

We also report the discovery of the first three PNe in a UDG. That number is consistent with the number of PNe in other galaxies with similar luminosity and metallicities. We find that distance estimates of 13–20 Mpc are similarly likely, given the three discovered PNe.

Acknowledgements. We thank the anonymous referee for the constructive comments which helped improve this paper. We thank Lodovico Coccato for his extremely useful help on the use of ZAP. We thank Simon Conseil for his very reactive answers to our questions on ZAP. We thank Harald Kuntschner for the useful discussions on the measurement and interpretation of Lick indices, and for letting us use his code. We thank Magda Arnaboldi for the very useful discussions on PNe. We thank Souradeep Bhattacharya and Glenn van de Ven for the helpful discussions. O.M. is grateful to the Swiss National Science Foundation for financial support. E.W.P. acknowledges support from the National Natural Science Foundation of China through Grant No. 11573002

References

Allen, C. W. 1973, Astrophysical Quantities (London: University of London, Athlone Press)
Amorisco, N. C., & Loeb, A. 2016, MNRAS, 459, L51
Amorisco, N. C., Monachesi, A., Agnello, A., & White, S. D. M. 2018, MNRAS, 475, 4235
Beasley, M. A., & Trujillo, I. 2016, ApJ, 830, 23
Beasley, M. A., Romanowsky, A. J., Pota, V., et al. 2016, ApJ, 819, L20
Bennet, P., Sand, D. J., Zaritsky, D., et al. 2018, ApJ, 866, L11
Blakeslee, J. P., & Cantiello, M. 2018, Res. Notes Am. Astron. Soc., 2, 146
Blakeslee, J. P., Tonry, J. L., & Metzger, M. R. 1997, AJ, 114, 482
Boquien, M., Duc, P.-A., Galliano, F., et al. 2010, AJ, 140, 2124
Bournaud, F., & Duc, P.-A. 2006, A&A, 456, 481
Buzzoni, A., Arnaboldi, M., & Corradi, R. L. M. 2006, MNRAS, 368, 877
Cappellari, M. 2017, MNRAS, 466, 798
Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138
Chan, T. K., Kereš, D., Wetzel, A., et al. 2018, MNRAS, 478, 906
Chowdhury, A. 2019, MNRAS, 482, L99
Ciardullo, R. 2012, Ap&SS, 341, 151
Ciardullo, R., Jacoby, G. H., Ford, H. C., & Neill, J. D. 1989, ApJ, 339, 53
Cole, D. R., Dehnen, W., Read, J. I., & Wilkinson, M. I. 2012, MNRAS, 426, 601
Collins, M. L. M., Chapman, S. C., Rich, R. M., et al. 2013, ApJ, 768, 172
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Conroy, C., Villaume, A., van Dokkum, P. G., & Lind, K. 2018, ApJ, 854, 139
Conselice, C. J., Gallagher, III., J. S., & Wyse, R. F. G. 2003, AJ, 125, 66
Dalcanton, J. J., Spergel, D. N., Gunn, J. E., Schmidt, M., & Schneider, D. P. 1997, AJ, 114, 635
de Boer, T. J. L., & Fraser, M. 2016, A&A, 590, A35
Di Cintio, A., Brook, C. B., Dutton, A. A., et al. 2017, MNRAS, 466, L1
Dopita, M. A., Jacoby, G. H., & Vassiliadis, E. 1992, ApJ, 389, 27
Duc, P. A., & Mirabel, I. F. 1999, in Galaxy Interactions at Low and High Redshift, eds. J. E. Barnes, & D. B. Sanders, IAU Symp., 186, 61
Duc, P.-A., Brinks, E., Springel, V., et al. 2000, AJ, 120, 1238
Duc, P.-A., Paudel, S., McDermid, R. M., et al. 2014, MNRAS, 440, 1458
Emsellem, E., van der Burg, R. F. J., Fensch, J., et al. 2019, A&A, 625, A76
Evstigneeva, E. A., Gregg, M. D., Drinkwater, M. J., & Hilker, M. 2007, AJ, 133, 1722
Famaey, B., McGaugh, S., & Milgrom, M. 2018, MNRAS, 480, 473
Fensch, J., Duc, P. A., Boquien, M., et al. 2018, A&A, submitted [arXiv:1903.10789]
Ferré-Mateu, A., Alabi, A., Forbes, D. A., et al. 2018, MNRAS, 479, 4891
Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41
Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, A&AS, 141, 371
Gu, M., Conroy, C., Law, D., et al. 2018, ApJ, 859, 37
Harris, W. E., & van den Bergh, S. 1981, AJ, 86, 1627
Harris, W. E., Blakeslee, J. P., & Harris, G. L. H. 2017, ApJ, 836, 67
Impey, C., Bothun, G., & Malin, D. 1988, ApJ, 330, 634
Jones, T., Ellis, R. S., Richard, J., & Jullo, E. 2013, ApJ, 765, 48

Kadowaki, J., Zaritsky, D., & Donnerstein, R. L. 2017, ApJ, 838, L21
Karachentsev, I. D., Karachentseva, V. E., Suchkov, A. A., & Grebel, E. K. 2000, A&AS, 145, 415
Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102
Koda, J., Yagi, M., Yamanoi, H., & Komiyama, Y. 2015, ApJ, 807, L2
Köppen, J., Weidner, C., & Kroupa, P. 2007, MNRAS, 375, 673
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P. 2012, PASA, 29, 395
Kroupa, P., Haghi, H., Javanmardi, B., et al. 2018, Nature, 561, E4
Kuntschner, H., Emsellem, E., Bacon, R., et al. 2010, MNRAS, 408, 97
Laporte, C. F. P., Agnello, A., & Navarro, J. F. 2019, MNRAS, 484, 245
Larsen, S. S., Brodie, J. P., Forbes, D. A., & Strader, J. 2014, A&A, 565, A98
Lelli, F., Duc, P.-A., Brinks, E., et al. 2015, A&A, 584, A113
Lim, S., Peng, E. W., Côté, P., et al. 2018, ApJ, 862, 82
Lotz, J. M., Miller, B. W., & Ferguson, H. C. 2004, ApJ, 613, 262
Marigo, P., Girardi, L., Weiss, A., Groenewegen, M. A. T., & Chiosi, C. 2004, A&A, 423, 995
Martin, N. F., Collins, M. L. M., Longeard, N., & Tollerud, E. 2018, ApJ, 859, L5
Metz, M., & Kroupa, P. 2007, MNRAS, 376, 387
Mihos, J. C., Durrell, P. R., Ferrarese, L., et al. 2015, ApJ, 809, L21
Muñoz, R. P., Eigenthaler, P., Puzia, T. H., et al. 2015, ApJ, 813, L15
Müller, O., Jerjen, H., & Binggeli, B. 2018, A&A, 615, A105
Pandya, V., Romanowsky, A. J., Laine, S., et al. 2018, ApJ, 858, 29
Pawlowski, M. S., Kroupa, P., & de Boer, K. S. 2011, A&A, 532, A118
Peñarrubia, J., McConnachie, A. W., & Navarro, J. F. 2008, ApJ, 672, 904
Peng, E. W., & Lim, S. 2016, ApJ, 822, L31
Peng, E. W., Ford, H. C., & Freeman, K. C. 2004, ApJS, 150, 367
Read, J. I., Iorio, G., Agertz, O., & Fraternali, F. 2017, MNRAS, 467, 2019
Recchi, S., Kroupa, P., & Ploeckinger, S. 2015, MNRAS, 450, 2367
Román, J., & Trujillo, I. 2017, MNRAS, 468, 4039
Rong, Y., Guo, Q., Gao, L., et al. 2017, MNRAS, 470, 4231
Ruiz-Lara, T., Beasley, M. A., Falcón-Barroso, J., et al. 2018, MNRAS, 478, 2034
Sandage, A., & Binggeli, B. 1984, AJ, 89, 919
Schiavon, R. P. 2007, ApJS, 171, 146
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schönberner, D., Jacob, R., Sandin, C., & Steffen, M. 2010, A&A, 523, A86
Sharina, M. E., Chandar, R., Puzia, T. H., Goudfrooij, P., & Davoust, E. 2010, MNRAS, 405, 839
Shi, D. D., Zheng, X. Z., Zhao, H. B., et al. 2017, ApJ, 846, 26
Sifón, C., van der Burg, R. F. J., Hoekstra, H., Muzzin, A., & Herbonnet, R. 2018, MNRAS, 473, 3747
Smith, R., Sánchez-Janssen, R., Fellhauer, M., et al. 2013, MNRAS, 429, 1066
Soto, K. T., Lilly, S. J., Bacon, R., Richard, J., & Conseil, S. 2016, MNRAS, 458, 3210
Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, MNRAS, 404, 1775
Torrealba, G., Belokurov, V., & Koposov, S. E. 2019, MNRAS, 484, 2181
Trujillo, I., Beasley, M. A., Borlaff, A., et al. 2019, MNRAS, 486, 1192
van der Burg, R. F. J., Muzzin, A., & Hoekstra, H. 2016, A&A, 590, A20
van der Burg, R. F. J., Hoekstra, H., Muzzin, A., et al. 2017, A&A, 607, A79
van Dokkum, P. G., Abraham, R., Merritt, A., et al. 2015, ApJ, 798, L45
van Dokkum, P., Abraham, R., Brodie, J., et al. 2016, ApJ, 828, L6
van Dokkum, P., Danieli, S., Cohen, Y., et al. 2018a, Nature, 555, 629
van Dokkum, P., Danieli, S., Cohen, Y., Romanowsky, A. J., & Conroy, C. 2018b, ApJ, 864, L18
van Dokkum, P., Cohen, Y., Danieli, S., et al. 2018c, ApJ, 856, L30
Vazdekis, A., Koleva, M., Ricciardelli, E., Röck, B., & Falcón-Barroso, J. 2016, MNRAS, 463, 3409
Weilbacher, P. M., Duc, P. A., & Fritze-v. Alvensleben, U. 2003, A&A, 397, 545
Wetzstein, M., Naab, T., & Burkert, A. 2007, MNRAS, 375, 805
Worthey, G., Faber, S. M., Gonzalez, J. J., & Burstein, D. 1994, ApJS, 94, 687
Yang, Y., Hammer, F., Fouquet, S., et al. 2014, MNRAS, 442, 2419
Yozin, C., & Bekki, K. 2015, MNRAS, 452, 937