



HAL
open science

Stars and gas in the most metal-poor galaxies - I. COS and MUSE observations of SBS 0335-052E

Aida Wofford, Alba Vidal-García, Anna Feltre, Jacopo Chevallard, Stéphane Charlot, Daniel P. Stark, Edmund C. Herenz, Matthew Hayes

► **To cite this version:**

Aida Wofford, Alba Vidal-García, Anna Feltre, Jacopo Chevallard, Stéphane Charlot, et al.. Stars and gas in the most metal-poor galaxies - I. COS and MUSE observations of SBS 0335-052E. Monthly Notices of the Royal Astronomical Society, 2021, 500, pp.2908-2927. 10.1093/mnras/staa3365 . insu-03748227

HAL Id: insu-03748227

<https://insu.hal.science/insu-03748227>

Submitted on 14 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stars and gas in the most metal-poor galaxies – I. COS and MUSE observations of SBS 0335–052E

Aida Wofford ^{1,★}, Alba Vidal-García,^{2,3★} Anna Feltre ^{4,5}, Jacopo Chevallard ², Stéphane Charlot,² Daniel P. Stark,⁶ Edmund C. Herenz⁷ and Matthew Hayes ⁸

¹*Instituto de Astronomía, Universidad Nacional Autónoma de México, Unidad Académica en Ensenada, Km 103 Carr: Tijuana–Ensenada, Ensenada 22860, México*

²*Sorbonne Universités, UPMC-CNRS, UMR7095, Institut d’Astrophysique de Paris, F-75014 Paris, France*

³*Laboratoire de Physique de l’Ecole normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France*

⁴*INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via P. Gobetti 93/3, I-40129 Bologna, Italy*

⁵*Scuola Internazionale Superiore di Studi Avanzati (SISSA), Via Bonomea 265, I-34136 Trieste, Italy*

⁶*Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA*

⁷*European Southern Observatory, Av. Alonso de Córdova 3107, 763 0355 Vitacura, Santiago, Chile*

⁸*Department of Astronomy and Oskar Klein Centre for Cosmoparticle Physics, AlbaNova University Centre, Stockholm University, SE-10691 Stockholm, Sweden*

Accepted 2020 October 23. Received 2020 October 23; in original form 2020 May 22

ABSTRACT

Among the nearest most metal-poor starburst-dwarf galaxies known, SBS 0335–052E is the most luminous in integrated nebular He II $\lambda 4686$ emission. This makes it a unique target to test spectral synthesis models and spectral interpretation tools of the kind that will be used to interpret future rest-frame UV observations of primeval galaxies. Previous attempts to reproduce its He II $\lambda 4686$ luminosity found that X-ray sources, shocks, and single Wolf–Rayet stars are not main contributors to the He II-ionizing budget; and that only metal-free single rotating stars or binary stars with a top-heavy IMF and an unphysically low metallicity can reproduce it. We present new UV (COS) and optical (MUSE) spectra that integrate the light of four super star clusters in SBS 0335–052E. Nebular He II, [C III], C III], C IV, and O III] UV emission lines with equivalent widths between 1.7 and 5 Å and a C IV $\lambda\lambda 1548, 1551$ P-Cygni like profile are detected. Recent extremely metal-poor shock + precursor models and binary models fail to reproduce the observed optical emission-line ratios. We use different sets of UV and optical observables to test models of constant star formation with single non-rotating stars that account for very massive stars as blueshifted O V $\lambda 1371$ absorption is present. Simultaneously fitting the fluxes of all high-ionization UV lines requires an unphysically low metallicity. Fitting the P-Cygni like + nebular components of C IV $\lambda\lambda 1548, 1551$ does not constrain the stellar metallicity and time since the beginning of star formation. We obtain $12+\log(\text{O}/\text{H}) = 7.45 \pm 0.04$ and $\log(\text{C}/\text{O}) = -0.45^{+0.03}_{-0.04}$ for the galaxy. Model testing would benefit from higher spatial resolution UV and optical spectroscopy of the galaxy.

Key words: methods: statistical – techniques: spectroscopic – galaxies: dwarf – ultraviolet: ISM – ultraviolet: stars.

1 INTRODUCTION

1.1 Extremely metal-poor massive ($\geq 8 M_{\odot}$) stars

Stars with main-sequence (MS) masses above 20–25 M_{\odot} are important because they dominate the ionizing flux of young stellar populations. Unfortunately, in the extremely low metallicity regime, which is defined as $Z \leq Z_{\odot}/10$, where Z is the fraction by mass of metals to all elements, the evolution of these stars and their chemical, radiative, and mechanical energy outputs are poorly understood. In this metallicity regime, we lack observations of individual stars, in particular far-ultraviolet (FUV) spectroscopy, which enables the characterisation of stellar wind properties and the measurement of atmospheric abundances and ionizing luminosities (García et al. 2014). One looks for metal-deficient massive stars in H II regions with

a low ionized-gas O/H ratio, where O/H is the fraction by number of oxygen to hydrogen atoms. This is because (i) the atmospheres of MS massive stars have metal abundances which are similar to those of the H II regions in which they are embedded (Asplund et al. 2009; García-Rojas, Simón-Díaz & Esteban 2014; Bouret et al. 2015); (ii) oxygen is the most abundant metal in nearby H II regions and optical oxygen lines are luminous in these regions; and (iii) it is relatively easy to measure the O/H content of H II regions from the latter lines. Extremely metal-poor H II regions give their name to extremely metal-poor galaxies (XMPs), which are arbitrarily defined to have $\text{O}/\text{H} \leq (\text{O}/\text{H})_{\odot}/10$ in the ionized gas (Kunth & Östlin 2000), or equivalently, using $12+\log(\text{O}/\text{H})_{\odot} = 8.69$ as the solar reference value (Asplund et al. 2009), $12+\log(\text{O}/\text{H}) \leq 7.56$.¹ Unfortunately,

¹Hereafter, we use the solar reference value of Asplund et al. (2009), which corresponds to $Z_{\odot} = 0.0134$, unless specified otherwise. We sometimes give abundances in parenthesis which correspond solar reference values of Caffau et al. (2011), i.e. $12+\log(\text{O}/\text{H})_{\odot} = 8.83$ and $Z_{\odot} = 0.0152$.

* E-mail: awofford@astro.unam.mx (AW); alba.vidal@phys.ens.fr (AVG)

the three known nearby XMPs where individual massive stars can be observed (Sextans A, SagDIG, and Leo P) are not strongly star-forming (Camacho et al. 2016; Garcia 2018; Evans et al. 2019; Garcia et al. 2019). Thus, their stars do not significantly cover the parameter space of mass, luminosity, effective temperature, age, rotation, dynamical interaction in dense massive clusters, and in the case of stars in binary systems, orbital parameters and mass ratio combinations, which is relevant for either OB-type star astrophysics or as a reference to model distant starburst galaxies. The lack of empirical templates for extremely metal-poor massive stars severely limits our ability to understand progenitors of gravitational-wave events, long-Gamma Ray Bursts and Supernovae of types II/Ib/Ic; the integrated light of unresolved massive star populations; and the radiative, mechanical, and chemical feedback of massive star populations. In this context, rest-frame FUV observations of super star clusters (SSCs), also known as young massive clusters (YMCs), which are located in XMPs, provide critical constraints.

1.2 Extremely metal-poor SSCs

Super star clusters are dense aggregates of young ($\lesssim 100$ Myr) stars with total stellar mass $\geq 10^4 M_{\odot}$, which are considered to be fundamental building blocks of galaxies (Portegies Zwart, McMillan & Gieles 2010). In particular, the most massive SSCs are believed to be the progenitors of today’s globular clusters (Bastian 2016). Ultraviolet observations of SSCs in XMPs, reveal how extreme-low metallicity affects the integrated properties of massive star populations and their surrounding gas, and provide valuable constraints to spectral synthesis models, i.e. models of the integrated spectra of starlight which sometimes also include its reprocessing by interstellar gas and dust. Such models constitute the main tool for interpreting observations of unresolved stellar populations and galaxies, near and far.

1.3 XMPs in the context of future large telescopes

The James Webb Space Telescope (JWST) and Extremely Large Telescopes (ELTs) will obtain rest-frame UV spectra of thousands of galaxies in the re-ionization era, at redshifts between 6 and 15. Spectral synthesis models will be the main tool to interpret such observations, which will challenge our understanding of galaxies with extreme low metal content (Byler et al. 2020). Given that it only takes a few pair instability supernovae to reach a metallicity of 2 per cent by a redshift of $z \sim 7$ (Wise et al. 2012), it is unlikely that we will ever be able to study distant, truly metal-free galaxies in detail. Thus, in order to understand how low metallicity affects the properties of stars and gas, it is essential to study SSCs in XMPs. Of course, XMPs are expected to also host older stellar populations which produced the oxygen that is currently present in the galaxy. This is the case in the prototypical XMP starburst, I Zw 18, which contains Red Giant Branch stars formed at least 1 Gyr ago (Annibali et al. 2013).

1.4 Low metallicity threshold of the nearby Universe

Using SDSS DR7, Sánchez Almeida et al. (2016) found that XMPs have a low metallicity threshold of $12 + \log(\text{O}/\text{H}) = 7.1$. This threshold corresponds to 1/39th (or 1/54th) of the present-day photospheric oxygen abundance of the Sun. Oxygen abundance measurements in AGC 198691 (Hirschauer et al. 2016) and SBS 0335–052W (Izotov et al. 2009) push the threshold even lower to $12 + \log(\text{O}/\text{H}) = 7.0 \pm 0.03$, i.e. to 1/49th (or 1/68th) of the solar

value. Interestingly, the latter two XMPs have a metallicity which is similar to that of some Damped Lyman Alpha Systems at $z \sim 5$ (Rafelski et al. 2014). The low metallicity content of XMPs remains to be explained. It is not clear if it is due to a low time-averaged star formation rate; the preferential escape of oxygen from their low gravitational potential well after type II SN explosions; accretion of metal-poor gas (Filho et al. 2015); or location in sparsely populated groups of galaxies where inflow of IGM gas is low and tidal interactions between galaxies are rare, leading to fewer episodes of star formation.

1.5 Types of XMPs

Galaxies which have been classified as XMPs comprise blue compact galaxies (BCGs), objects caught in a phase of substantial star formation which appear blue and compact because their star-forming regions are large relative to the galaxy’s projected size; blue compact dwarf (BCD) galaxies, which have similar properties to BCGs, but are fainter than $M_B = -18$ mag (Thuan, Izotov & Lipovetsky 1997); dwarf irregular (dIs) galaxies, which can be considered extreme late-type spirals with large gaseous components (Battaner & Florido 2000); and low surface brightness (LSBs) galaxies, which are diffuse galaxies defined as having surface brightnesses at least one order of magnitude lower than the dark night sky (Impey & Bothun 1997). These galaxies span the stellar mass range ($\log M_*/M_{\odot} = 6-10$) and are still forming their stellar disc. Some authors consider as ‘true’ XMPs objects with O/H values below what is expected from their absolute blue-band magnitudes or stellar masses (Ekta & Chengalur 2010; Pustilnik, Perepelitsyna & Kniazev 2016). Here we adopt the broader definition of XMP.

1.6 How to find XMPs?

Multiple methods have been used to identify XMPs. Morales-Luis et al. (2011) and Sánchez Almeida et al. (2016) searched for XMPs using spectra from the seventh data release of the Sloan Digital Sky Survey (SDSS DR7, Abazajian et al. 2009). They selected as candidate XMPs objects with weak [N II] lines relative to $\text{H}\alpha$ or strong [O III] $\lambda 4363$ emission, and then measured the O/H of the objects. Strong nebular [O III] $\lambda 4363$ emission indicates a high electron temperature of the emitting gas, which is attributed to inefficient cooling of the gas via metal emission lines. Based on this method, (Sánchez Almeida et al. 2016) found that the fraction of XMPs relative to all emission line galaxies in SDSS DR7 is only 0.01 per cent. James et al. (2015) used a morphology-based approach to target faint blue systems with H II regions embedded in a diffuse continuum. Of the 12 candidate XMPs found via this approach, two were spectroscopically confirmed as XMPs. Some of the most metal-poor XMPs have been found by the Arecibo Legacy Fast ALFA (ALFALFA, Giovanelli et al. 2005) survey, which is a multiwavelength study of galaxies with HI masses in the range of $10^6 - 10^7 M_{\odot}$. More recently, Hsyu et al. (2018) identified 94 candidate XMPs via photometric colours and morphologies in SDSS DR12 data, and confirmed 45 as XMPs via 3-m Lick Observatory and 10-m W.M. Keck Observatory optical spectroscopy. Using colour-colour diagrams, Senchyna & Stark (2019) uncovered 53 candidate XMPs with uniformly high specific star formation rates in SDSS imaging at magnitudes $16 < i' < 23$, i.e. extending significantly below the completeness limits of the SDSS spectroscopic survey, and confirmed 32 as XMPs via Multiple Mirror Telescope spectroscopy. Finally, Kojima et al. (2020) have used a machine-learning classifier in order find XMPs candidates in Subaru/HyperSuprime-Cam optical

images that are about 100 times deeper than SDSS. Spectroscopic follow-up of 10 candidates confirmed three galaxies as XMPs.

1.7 Why focusing on SBS 0335–052E?

Studies of the most metal-poor nearby XMPs in the UV face two challenges: the scarcity of suitable sources, and the difficulty of obtaining rest-frame UV spectra, as these require space-based observations. SBS 0335–052E is one of the nearest (~ 54.1 Mpc),² most metal-poor (largest oxygen abundance is $12+\log(\text{O}/\text{H}) = 7.338 \pm 0.012$, Izotov et al. 1999) starburst XMP known. It is also one of most well-studied. Although it is farther and more metal-rich than the prototypical, I Zw 18, which is located at ~ 18.2 Mpc (Aloisi et al. 2007) and whose North-Western component has $12+\log(\text{O}/\text{H}) = 7.17 \pm 0.04$ (Skillman & Kennicutt 1993), the integrated He II luminosity of SBS 0335–052E ($1.87e39 \text{ erg s}^{-1}$, Kehrig et al. 2018) is larger than that of I Zw 18 ($1.12e38 \text{ s}^{-1}$, Kehrig et al. 2015). This makes of SBS 0335–052E a unique target to test spectral synthesis models. We recently obtained high-quality UV spectra of SBS 0335–052E, as part of our pilot program aimed at testing the combined stellar population + photoionization models of Gutkin, Charlot & Bruzual (2016). A major goal of our program is to understand the achievements and limitations of these models and various fitting techniques when applied to spatially unresolved observations of XMPs.

1.8 Outline of this paper

In Section 2, we summarize results from previous studies of SBS 0335–052E, and explain the importance of this galaxy for extragalactic studies. In Section 3, we describe the UV and optical observations which are used in this work. In Section 4, we provide evidence for massive stars being a plausible main source of ionizing photons in the observed region of the galaxy. In Section 5, we describe the models and analysis tool. We also present results from comparing the models to the observations. In Section 6, we discuss our findings. Finally, in Section 7, we summarize and conclude.

2 SBS 0335–052E

2.1 Previous observations

Among the nearest and most metal-poor starburst XMPs known, BCD galaxy SBS 0335–052E is one of the most well-studied along with I Zw 18, which held the record of the most metal-poor galaxy known for a long time (Kunth & Östlin 2000). SBS 0335–052E was discovered in the Second Byurakan Survey (Izotov et al. 1990). Previous observations of SBS 0335–052E include, but are not limited to, the following wavelength ranges and telescopes/instruments: in the radio, Very Large Array (VLA; Pustilnik et al. 2001; Johnson, Hunt & Reines 2009), Giant Metrewave Radio Telescope (Ekta, Pustilnik & Chengalur 2009), and Atacama Large Millimeter Array (ALMA; Hunt et al. 2014; Cormier et al. 2017); in the far-infrared, *Herschel* PACS (Rémy-Ruyer et al. 2015); in the mid-infrared, *Spitzer* MIPS and IRS (Wu et al. 2008); in the near-infrared, *Hubble Space Telescope* (HST) NICMOS (Reines, Johnson & Hunt 2008; Thompson et al. 2009); in the optical, Multiple Mirror Telescope (Izotov et al. 1997), Keck II (Izotov et al. 1999), 3.6 m

ESO (Papaderos et al. 2006), Special Astrophysical Observatory (SAO, Moiseev et al. 2010), Very Large Telescope (VLT) GIRAFFE (Izotov et al. 2006), VLT FORS1+UVES (Izotov et al. 2009), VLT Multi Unit Spectroscopic Explorer (MUSE; Herenz et al. 2017); in the ultraviolet, *International Ultraviolet Explorer* (Dufour et al. 1993), *Far Ultraviolet Spectroscopic Explorer* (Grimes et al. 2009), *HST* FOS (Garnett et al. 1995), *HST* GHRS (Thuan et al. 1997), *HST* Cosmic Origins Spectrograph (COS; James et al. 2014); in the X-rays, *Chandra* (Thuan et al. 2004; Prestwich et al. 2013).

2.2 Environment

SBS 0335–052E has a companion BCD, known as SBS 0335-052W, which is separated in the east–west direction by 84 arcsec, i.e. by 22 kpc at the distance adopted by Pustilnik et al. (2001) (~ 54 Mpc). The galaxy resides not far from a rather large group, with only ~ 50 kpc in projection from an L_* spiral. The difference in radial velocities is within the group velocity dispersion. This group is at the border of a void (Peebles et al. 2001).

2.3 Ionized- and neutral-gas abundances and abundance ratios

Oxygen abundance measurements in the ionized gas of SBS 0335–052E include those of Izotov et al. (1999), Papaderos et al. (2006), Izotov et al. (2006) and Izotov et al. (2009). The latest measurement found variations over spatial scales of ~ 1 – 2 kpc in the range 7.11 to 7.32 dex. Table 2 of Izotov et al. (2006) gives the relative abundances of N, Ne, S, Cl, Ar, and Fe relative to O, in the ionized gas. The difference between ionized- and neutral-gas metal abundances in SBS 0335–052E is larger than for I Zw 18 (Lebouteiller et al. 2013; see also James et al. 2014). Finally, using the O III] $\lambda 1666$ and C III] $\lambda 1909$ line fluxes measured with *HST*'s Faint Object Spectrograph,³ Garnett et al. (1995) derive $\log(\text{C}/\text{O}) = -0.94 \pm 0.17$.

2.4 Dust and H I gas

Thuan et al. (1997) report that dust is clearly present and spatially mixed with the SSCs in SBS 0335–052E. According to Rémy-Ruyer et al. (2015), SBS 0335–052E has the lowest dust-to-stars mass ratio of the Dwarf Galaxy Survey of Madden et al. (2013) ($\log([M_{\text{dust}}/M_{\text{stars}}] = -4.53)$). It has also been found that SBS 0335–052E has a high H I to dust mass ratio for a BCD (Dale et al. 2001). *Physical properties.* Table 1 of Cormier et al. (2017) summarizes the main physical properties of the galaxy. SBS 0335–052E is a starburst galaxy with a total stellar mass of $5.6 \times 10^6 M_{\odot}$ and a star formation rate of $0.7 M_{\odot} \text{ yr}^{-1}$. It also has the highest specific star formation rate ($\log(sSFR/\text{yr}) = -8.13$) of the Dwarf Galaxy Survey of Madden et al. (2013) according to Rémy-Ruyer et al. (2015). At the distance of SBS 0335–052E, the detection of a red giant branch, and thus an underlying old population, is not possible.

2.5 Morphology

Thuan et al. (1997) identified six compact SSCs and a supershell of radius ~ 380 pc, which delineates a large supernova cavity. The left-hand panel of Fig. 1 shows a false-colour RGB image of SBS 0335–052E where the supershell can be seen to the NE of the overlaid circle representing the COS aperture. Using an unsharp

²Luminosity distance from NED, based on the $H\alpha$ velocity from Moiseev, Pustilnik & Kniazev (2010)

³This instrument is no longer available on *HST*.

Table 1. Observation log of *HST* observations used in this work. Columns: (1) ID of data set. (2) J2000 right ascension in format: hours minutes seconds. (3) J2000 declination in format: degrees minutes seconds. (4) Start time of observations in format: yy-mm-dd hh:mm:ss.00. (5) Exposure time. (6) Configuration. Wavelength ranges in Å of mirrors and gratings: MIRRA & MIRRB, 1650–3200; G130M segments, B[1137–1274], A[1292–1432]; G160M segments, B[1421–1592], A[1612–1784]; G185M segments, A[1720–1753], B[1819–1852], C[1917–1950]. (7) Central wavelength of filter or grating. (8) Program ID and name of PI.

Data set	RA J2000	Dec J2000	Start time y:m:d h:m:s	Exp time (s)	Instrument/Aperture/ Filter or Grating	λ_{cen} (Å)	PID/PI
J9FVA3030	03 37 44.000	−05 02 40.00	2006-08-22 01:14:31	680	ACS/WFC1/FR656N	6651	10575/Oestlin
J9FVA3010	03 37 44.000	−05 02 40.00	2006-08-22 00:44:12	430	ACS/WFC1/F550M	5581	10575/Oestlin
J9FVA3020	03 37 44.003	−05 02 39.86	2006-08-22 00:57:14	680	ACS/WFC1/F435W	4330	10575/Oestlin
LB7H91KEQ	03 37 43.980	−05 02 38.90	2010-03-02 05:11:27	40	COS/PSA/MIRRA	NUV	11579/Aloisi
LB7H91010	03 37 43.980	−05 02 38.90	2010-03-02 05:17:17	9534	COS/PSA/G130M	1291	11579/Aloisi
LCNE03SHQ	03 37 43.980	−05 02 38.90	2015-03-01 06:22:28	3	COS/PSA/MIRRB	NUV	13788/Wofford
LCNE03010	03 37 43.980	−05 02 38.90	2015-03-01 06:28:31	4956	COS/PSA/G160M	1611	13788/Wofford
LB7H91010	03 37 43.980	−05 02 38.90	2015-03-01 09:26:12	5675	COS/PSA/G185M	1835	13788/Wofford
LDN709020	03 37 43.980	−05 02 38.90	2018-08-29 18:52:07	5605	COS/PSA/G160M	1623	15193/Aloisi
LDN759020	03 37 43.980	−05 02 38.90	2019-02-14 06:49:25	5345	COS/PSA/G160M	1623	15193/Aloisi

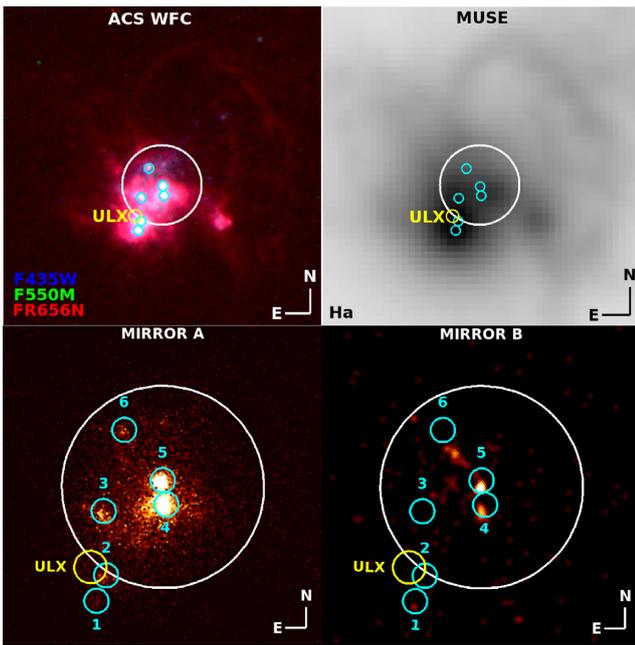


Figure 1. Top-left: *HST* ACS WFC RGB image of SBS 0335–052E composed from images of program 10575 (PI: Öestlin). The colour scale is logarithmic. In order to enhance the northern H α shell, the contrast and bias of the images were adjusted and the images were smoothed with a Gaussian kernel of three pixels in radius. Top-right: MUSE H α image obtained from the observed-frame wavelength range 6647 to 6655 Å. The colour scale is logarithmic. Bottom-left: *HST* COS NUV Mirror A TA image (PID 11579, PI: Aloisi). The colour scale is linear. Bottom-right: COS NUV Mirror B TA image (PID 13788, PI: Wofford). The colour scale is linear and the image is smoothed with a Gaussian kernel of three pixels in radius in order to enhance the signal from SSCs 4 and 5. In all panels: (i) the largest circle (white) represents the COS PSA footprint and has a diameter of 2.5 arcsec, i.e. 656 pc at the adopted distance to the galaxy; (ii) the medium-sized circle (yellow) shows the location of the ULX which is reported in Prestwich et al. (2013) and has a diameter equal to *Chandra*'s positional uncertainty, i.e. 0.4 arcsec; and (iii) the smallest circles (cyan) indicate the positions of SSCs 1 to 6 which are identified in Thuan et al. (1997).

masking technique, Papaderos et al. (1998) discovered several lower-luminosity (and possibly more evolved) SSCs outside of the main conglomerate of the dominant six SSCs. Hereafter, we adopt the SSC IDs of Thuan et al. (1997) and references therein.

2.6 Importance for extra-galactic studies

Izotov et al. (1999) observed a region of SBS 0335–052E which is centred on SSC 5 with the Keck II telescope. Using their emission line measurements, we find $[\text{O III}] \lambda 5007 / [\text{O II}] \lambda 3727 = 15$. Based on this ratio, SBS 0335–052E would be classified as a Green Pea (GP) galaxy (Cardamone et al. 2009).⁴ Green Pea galaxies are compact galaxies that are unresolved in SDSS images and were discovered in the citizen science project Galaxy Zoo.⁵ The green colour of GPs is because $[\text{O III}] \lambda \lambda 4959, 5007$ dominates the flux of the SDSS *r*-band, which is mapped to the green channel in the SDSS false-colour *gri*-band images. As shown in later papers, e.g. Izotov, Guseva & Thuan (2011), similar types of nearby galaxies exist also at redshifts where they do not appear green. Green pea like galaxies are the current best nearby analogues of high- z Lyman-alpha ($\text{Ly } \alpha$) emitters (LAEs), e.g. Yang et al. (2016), Yang et al. (2017), Jaskot & Oey (2014), Henry et al. (2015), and Izotov et al. (2016). LAEs are an important population of star-forming galaxies at $z > 2$, increasing in fraction to constitute 60 per cent of Lyman break galaxies (LBGs) at redshifts $z > 6$ (Stark 2012), and they are used to probe reionization, e.g. Malhotra & Rhoads (2004), Ouchi et al. (2010), Hu et al. (2010), Kashikawa et al. (2011). LAEs resemble some of the faint LBGs that dominate the luminosity density of the universe in the epoch of reionization (Herenz et al. 2019) in size (Malhotra et al. 2012), dust, and stellar mass (Finkelstein et al. 2010). They have high line equivalent widths, e.g. Malhotra & Rhoads (2002), Zheng et al. (2014), indicating intense star formation despite their comparatively low stellar masses.⁶ In this context, understanding the origin of the high $[\text{O III}]/[\text{O II}]$ ratio of SBS 0335–052E could shed light on some of the spectral properties of GP galaxies. It is thus important to understand the dominant ionizing source in SBS 0335–052E.

2.7 Ly α properties

Interestingly, Thuan et al. (1997) and Kunth et al. (2003) report an absence of strong H I Ly α emission from SBS 0335–052E, which is attributed to a combination of dust attenuation, redistribution of the Ly α photons by resonant scatterings over the whole area of the

⁴Note that in the latter paper, the authors just call them Pea galaxies.

⁵<http://zoo1.galaxyzoo.org/>

⁶Note that there exist bright LBGs that show strong Ly α as well as faint LBGs with no Ly α (Dunlop 2013).

surrounding H I cloud, and the geometry of the cloud. Note that the column density of H I is very high, i.e. $\log[N(\text{H I}) \text{ cm}^{-2}] = 20.15$ (James et al. 2014). The UV spectrum of *HST* program 11579 (PI: Aloisi) confirms the absence of significant Ly α emission from SBS 0335–052E, as reported in James et al. (2014). The latter spectrum is reproduced in the top panel of Fig. 2. In addition, extended diffuse Ly α is undetected around the galaxy (Östlin et al. 2009). Wofford, Leitherer & Chandar (2011) present Ly α line observations of 20 star-forming galaxies with a range of morphologies, oxygen abundances, and SFRs. All the galaxies in their sample with Ly α in emission have outflows of H I gas. Rivera-Thorsen et al. (2015) present far-UV spectroscopy of 14 strongly star-forming galaxies at low redshift ($0.028 < z < 0.18$). They find that all galaxies in their sample with Ly α escape show bulk outflow velocities of $\geq 50 \text{ km s}^{-1}$, although a number of their galaxies with similar velocities show little or no Ly α escape. James et al. (2014) find that in SBS 0335–052E the H I gas is outflowing at $\sim 30 \text{ km s}^{-1}$ relative to the UV background in SBS 0335–052E. In summary, SBS 0335–052E is a strongly star-forming galaxy with the oxygen line ratio of a GP galaxy, with outflowing H I gas, but with no Ly α emission.

3 OBSERVATIONS

In this section, we present *HST* COS and *VLT* MUSE spectra of SBS 0335–052E which we use to (i) discuss plausible ionizing sources for the interstellar gas in the galaxy, (ii) test spectral synthesis models, and (iii) derive physical properties of the stars and gas.

3.1 COS spectroscopy

As part of program 13788 (PI: Wofford), we obtained integrated observations of SBS 0335–052E’s SSCs 3, 4, 5 and 6, centred on SSC 5, using *HST*’s COS medium resolution gratings, G160M and G185M ($\lambda/FWHM \sim 16\,000\text{--}20\,000$), which cover the wavelength range from 1421 to 1950 Å (observed-frame), with some gaps due to unfilled space between the detector segments. The spectra were obtained with the circular Primary Science Aperture (PSA), which is 2.5 arcsec in diameter, and whose footprint on the galaxy is shown in the left-hand panel of Fig. 1. The latter figure also shows the locations of the six SSCs identified by Thuan et al. (1997), and the ULX which is reported in Prestwich et al. (2013). Our data extend the wavelength coverage of *HST* program 11579 (PI: Aloisi), which uses COS G130M, also centred on SSC 5, to cover the wavelength range $\sim 1130\text{--}1420$ Å (observed-frame). James et al. (2014) use UV absorption lines in the G130M spectrum to derive metal abundances in the neutral gas. In this work, we use the G130M spectrum to look for signatures of Very Massive Stars (VMSs), which are stars with masses above $100 M_{\odot}$ (Vink 2012). We also use the G130M spectrum to obtain a better fit to the stellar continuum and check that the flux levels of all data sets match, including the optical spectrum.

We used the following observing sequence. (1) Perform target-acquisition (TA) procedures in order to ensure that the target is well centred in the PSA. This avoids throughput losses, produces reliable wavelength zero points, and outputs an image of the target over an NUV band of wavelength range, 1700–3200 Å (observed-frame). In order to preserve the NUV detector, the TA image was obtained with MIRRORB, which offers an attenuation factor of ~ 25 relative to MIRRORA. For reference, the plate scale of the NUV detector is 0.075 arcsec. (2) Obtain spectra in TIME-TAG mode using the FP-POS = ALL setting and FLASH = YES so that the spectra can be corrected for flat-field anomalies and OSM drifts. (3) Obtain additional spectra during subsequent orbits in order to achieve

the desired signal-to-noise ratio (SNR). Note that for the G130M observations, MIRRORA was used for the TA. The TA images of programs 11579 and 13788 are shown in the middle and right panels of Fig. 1 with the PSA footprint overlaid and the location of the SSCs identified by the IDs of Thuan et al. (1997). In spite of the lower SNR of our MIRRORB image, clusters 4 and 5 are clearly detected. This is expected as they are the brightest in the UV, which can also be seen in fig. 1 of Kunth et al. (2003), which shows *HST* ACS F140LP and F122M UV images. According to table 2 of Thuan et al. (1997), SSCs 1 and 2, which are located just outside of the COS aperture, are the brightest in the optical.

We retrieved the individual G160M and G185M spectral data sets from the Mikulski Archive for Space Telescopes and initially processed them on-the-fly with version 3.2.1 (2017-04-28) of the CalCOS pipeline. We also retrieved the G130M data set, which was processed through Calcos version 3.0 (2014-10-30). Table 1 provides a summary of the observations which are presented in Fig. 1 and this work. CalCOS processes the data for detector noise, thermal drifts, geometric distortions, orbital Doppler shifts, count-rate nonlinearity, and pixel-to-pixel variations in sensitivity. A standard wavelength scale is applied using the on board wavelength calibration. The G160M data fall on two detector segments, A and B. The two FUV segments are processed independently. The G185M data fall on three detector segments, A, B, and C. For each grating, multiple exposures with the same grating and central wavelength are contained within a visit, these are combined into a single summed spectrum. The final products are one-dimensional, flux calibrated, heliocentric-velocity corrected, background subtracted, and combined spectra which are stored in x1dsum files. We worked directly with the x1dsum files. By fitting a model for the stellar continuum of a Simple Stellar Population (SSP) to the G130M + G160M + G185M data, we noted a mismatch by a factor of 1.27 between the continuum levels of the G160M and G185M observations at 1715 Å, in the wavelength region of overlap between these gratings. The Space Telescope Science institute informed us that a likely cause for this mismatch is that the COS calibration files are overestimating the loss in sensitivity in the NUV gratings, which is why the flux in the NUV G185M data is slightly lower than in the G160M data. The COS team is working on this, and hopefully will release new reference files soon. For this reason, we multiplied the G185M flux array by 1.27.

For a point-source well-centred on the PSA, the instrumental *FWHM* values of the G160M and G185M gratings are 0.09 Å ($\sim 18 \text{ km s}^{-1}$ at 1500 Å) and 0.11 Å ($\sim 22 \text{ km s}^{-1}$ at 1500 Å), respectively, as derived by multiplying the dispersions of the gratings (0.012 Å pix^{-1} and 0.037 Å pix^{-1} , respectively) by the sizes of the spectral resolution elements of their associated detectors (6 pixels for FUV XDL and 3 pixels for NUV MAMA). Since CalCOS outputs oversampled data, we resample the data to the nominal dispersions of the gratings. Wavelength zero-point shifts can occur due to imperfect centring and the flux profiles of the galaxies in the PSA. For the data taken with G160M, it is possible to check for wavelength offsets by determining the centroids of Galactic UV absorption lines towards our targets; with the caveat that negative offsets of $\leq -50 \text{ km s}^{-1}$ are likely due to intermediate or high-velocity clouds (IVC or HVC, respectively; e.g. Wofford et al. 2011). To check for offsets, we use the Si II $\lambda 1527$ and Al II $\lambda 1671$ Milky Way (MW) absorption lines. We find velocity blueshifts of -20 and -25 km s^{-1} , respectively. We do not correct the observations for these negative offsets which might be explained by Milky Way (MW) infalling clouds. We use the *FWHM* of the above two Galactic UV absorptions as the effective spectral resolution of the observations. We find *FWHM* values of 0.54 Å ($\sim 106 \text{ km s}^{-1}$) and ~ 0.49 Å ($\sim 89 \text{ km s}^{-1}$) for Si II $\lambda 1527$

and $\text{Al II } \lambda 1671$, respectively, i.e. an average of $FWHM$ value of 0.51 \AA ($\sim 97 \text{ km s}^{-1}$ at 1600 \AA). We use the IDL routine `gauss_smooth.pro` to smooth the data in order to match the effective spectral resolution given by the MW lines.

Fig. 2 shows the SBS 0335–052E COS G130M + G160M + G185M spectra from programs 11579 and 13788. When looking at the G130M spectra, one can see the three contributions to the $\text{Ly } \alpha$ $\lambda 1216$ profile: the MW $\text{Ly } \alpha$ absorption, the Earth’s geocoronal $\text{Ly } \alpha$ emission, and the $\text{Ly } \alpha$ absorption which is intrinsic to SBS 0335–052E and is reported in James et al. (2014). The G130M data also show blueshifted $\text{O V } \lambda 1371$ absorption, which has been attributed to the presence of VMSs in three metal-poor galaxies (Wofford et al. 2014; Crowther et al. 2016; Smith et al. 2016). The spectral morphology of $\text{O V } \lambda 1371$ is different than in the above galaxies owing to the much lower metallicity of SBS 0335–052E. Since $\text{O V } \lambda 1371$ is observed in early MS (Bouret et al. 2013) and early O If stars (Walborn, Nichols-Bohlin & Panek 1985; Walborn, Parker & Nichols 1995) as well as VMS, the association of blueshifted $\text{O V } \lambda 1371$ with VMS in SBS 0335–052E is not definitive. By looking at the G160M and G185M data, one can see that the high-ionization UV emission lines $\text{C IV } \lambda \lambda 1549, 1551$, $\text{He II } \lambda 1640$, $\text{O III } \lambda \lambda 1661, 1666$, and $[\text{C III}] \lambda \lambda 1907, 1909$ are clearly detected. Previous lower spectral resolution observations of the C IV and O III doublets in SBS 0335–052E are shown in fig. 1 of Dufour et al. (1993) and fig. 1 of Garnett et al. (1995), respectively. The first observation (PI: Terlevich, PID NE077), was obtained with the SWP camera on board the *International Ultraviolet Explorer* using a $10 \text{ arcsec} \times 20 \text{ arcsec}$ oval aperture. It covers the wavelength range from ~ 1150 to $\sim 2000 \text{ \AA}$ (observed frame). The second observation was obtained with *HST*’s Faint Object Spectrograph using a 1 arcsec circular aperture (PI: Skillman, PID 3840). It covers the wavelength range from ~ 1600 to $\sim 2000 \text{ \AA}$ (observed frame). With our G160M observation (0.51 \AA at 1600 \AA), we resolve for the first time the C IV and O III doublets. Finally, we detect $\text{S V } \lambda 1502$ absorption from the photospheres of massive O and B stars (Walborn et al. 1985, 1995; de Mello, Leitherer & Heckman 2000) at the 2σ level.

3.2 MUSE spectroscopy

SBS 0335–052E was observed with VLT/MUSE (Bacon et al. 2010) in service mode on November 16 and 17, 2015. Sky conditions were clear to photometric and the DIMM seeing $FWHM$ varied between 0.9 and 1.2 arcsec . The total open shutter time on target was 5680 s , split into $8 \times 710 \text{ s}$ individual exposures. As recommended in the MUSE users’ manual, the spectrograph’s field of view was rotated by 90° and small dither offsets were applied for each exposure. For flux calibration the white dwarf GD 71 (Bohlin, Colina & Finley 1995) was observed on November 17. On the same date, a twilight flat-field was taken. Each open-shutter exposure is associated to a set of calibration exposures (bias frames, lamp flats, illumination flat, and arc-lamp frames) taken closest in time.

Data reduction of the MUSE observations was performed with the MUSE data reduction system (DRS; Weilbacher et al. 2014), version 1.6.2. We first used the calibration tasks `muse_bias`, `muse_flat`, `muse_wavecal`, `muse_lsf`, and `muse_twilight` to create master bias frames, trace tables and master flats, wavelength calibration tables, line spread function images, and a twilight flat datacube, respectively. With those calibration data products, we then reduced all open-shutter exposures into so-called pixtables with the DRS task `muse_scibasic`. Since the instrument illumination has been shown to be temperature dependent, we en-

sured that we always applied the illumination flats taken closest in time to the actual observation. The standard star pixtable was then fed into the task `muse_standard` to generate a response curve for absolute flux calibration. Thereafter, we fed the response curve and individual science exposure pixtables into the task `muse_scipost`, which calibrates the flux, subtracts the sky spectrum, and astrometrically calibrates the pixtables. Lastly, using the DRS tasks `muse_exp_alignmuse_exp_combine`, the fully calibrated pixtables were combined and resampled on to a common $317 \times 317 \times 3800$ grid (a so-called datacube), where the first two axis represent a spatial position on the sky (regularly sampled at $0.2 \text{ arcsec} \times 0.2 \text{ arcsec}$), and the third axis represents the spectral axis from 4600 \AA to 9160 \AA linearly sampled with 1.2 \AA steps.

The MUSE datacube of SBS 0335–052E was already exploited by Herenz et al. (2017) to discover large extended $\text{H}\alpha$ and $[\text{O III}]$ filaments in the halo of this galaxy. Owing to the unexpected large extend of nebular emission around this galaxy, the MUSE DRS sky-subtraction algorithm initially oversubtracted $\text{H}\alpha$ and $[\text{O III}]$ emission. As described in Herenz et al. (2017), to correct for this, the resulting output sky-spectra from the DRS were first modified by interpolation over the affected wavelengths and subsequently fed into a second iteration of `muse_scipost`, as input sky-spectra. The uncorrected MUSE datacube was exploited by Kehrig et al. (2018) to show that the total He II flux of the galaxy, including regions outside of our aperture, can only be produced by either single, rotating metal-free stars or a binary population with $Z \sim 10^{-5}$ and a ‘top-heavy’ IMF.

For the present study, we use the corrected datacube to extract a spectrum at the position of *HST*/COS. Therefore, we summed all spectral pixels of the MUSE datacube within the COS PSA aperture. We also utilize a continuum subtracted datacube to pinpoint spatially the faint Wolf–Rayet (WR) emission features (sec. 4.4). Therefore we subtract an in spectral direction median-filtered (filter width = 180 \AA) continuum-only cube (see Section 4.1 in Herenz & Wisotzki 2017).

In Fig. 3, we show the MUSE spectrum which corresponds to the location and rough area of the COS aperture. The spectrum is H II -region like. Optical high-ionization emission lines of $[\text{Ar IV}] \lambda 4740$, $\text{He II } \lambda 4686$ and $[\text{Ar III}] \lambda 7135$ are clearly detected. Also detected are the weak $[\text{N II}] \lambda \lambda 6548, 6584$ doublet around $\text{H}\alpha$ and $[\text{N II}] \lambda 5755$. The discussion of the WR features can be found in sec. 4.4.

In order to check if the flux levels of the co-spatial COS and MUSE spectra match, we fitted a low spectral resolution stellar + nebular continuum model (SSP; $Z = 0.002$, 3 Myr , non-rotating stars), which was computed with Starburst99 (Leitherer et al. 2010) to these data. The observations were previously corrected for foreground reddening, redshift and intrinsic stellar reddening using the slope of the UV stellar continuum and the SMC extinction curve. The result is shown in Fig. 4. The model reproduces very well the continuum in the rest-frame wavelength range from 1150 to 1909 \AA and the red portion of the optical spectrum, but shows a mismatch with the observations in the blue part of the optical spectrum. We checked that the mismatch is not due to: i) differences in coordinates between the COS and MUSE spectral extractions (they are the same); ii) the MUSE observing conditions (which were clear); iii) vignetting of COS which reduces the NUV flux (if we add UV flux, the discrepancy worsens); iv) the radius of the MUSE spectral extraction aperture; v) the reddening correction (the same correction is applied to the UV and optical); or vi) binning. Since the COS aperture is fairly large compared to the $FWHM$ of the MUSE PSF, we think that the MUSE wavelength-dependent PSF is not a big factor for the clusters near the centre of the aperture. Unfortunately, one of the four SSCs in the

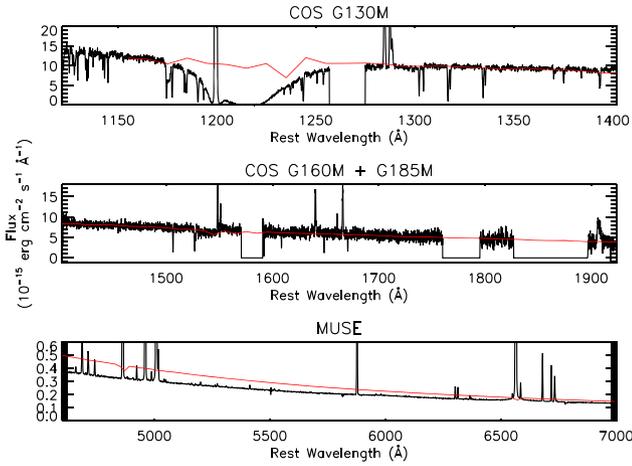


Figure 4. Co-spatial COS and MUSE observations from 1120 to 7000 Å corrected for foreground reddening, redshift, and intrinsic reddening (black curve). We overlay a Starburst99 (Leitherer et al. 2010) low spectral resolution stellar + nebular continuum model in red (see text for more details). There is a significant mismatch between the stellar + nebular continua of the COS and MUSE observations.

COS aperture is close to the edge of the PSA making the correction for the mismatch difficult. Due to this mismatch between the optical and UV continua, we use the UV and optical separately to test our models.

3.3 Spectral line measurements

The spectra are corrected for foreground extinction using a colour excess of $E(B - V) = 0.047$, which was obtained from the infrared-based Galactic dust map of Schlegel, Finkbeiner & Davis (1998), by using the reddening law (A_λ/A_V) of Mathis (1990). They are also corrected for the redshift of the intrinsic Al II $\lambda 1671$ line ($z = 0.013475$, $v \sim 4038 \text{ km s}^{-1}$). Note that we do not use the intrinsic redshift of the Si II $\lambda 1526$ absorption because it is contaminated with a MW C IV absorption. Our redshift measurement is in reasonable agreement with the average from several species in the neutral gas which was measured by James et al. (2014) using the G130M data ($z = 0.013491$, $v \sim 4044 \text{ km s}^{-1}$). The measurements of the intrinsic UV lines are done after the spectra have been re-binned and smoothed to match the effective spectral resolution, and corrected for foreground reddening and redshift. The measurements of the intrinsic optical lines are done after the spectra have been corrected for foreground reddening and redshift. The UV and optical line measurements are uncorrected for intrinsic attenuation due to dust.

Table 2 gives the rest-frame wavelengths of the main lines which are analysed in this work along with the ionization potentials of the ions which are required to produce the lines. We obtained the centroid and *FWHM* of the intrinsic nebular emission and ISM absorption lines with a custom-made IDL routine which was developed by COS science team member K. France. The routine fits a first-order polynomial to the continuum and a Gaussian to each line. The radial velocities corresponding to the centroids of the lines are relative to the intrinsic Al II $\lambda 1671$ absorption line. The *FWHM* and radial velocities of the intrinsic lines are reported in Table 2.

In order to enhance the SNR of weak components to C IV $\lambda\lambda 1548, 1551$ and He II $\lambda 1640$, we obtained the average of the G160M data from programs 13788 (PI: Wofford) and 15193 (PI: Aloisi; see Table 1 for details), weighted by the exposure times. This

Table 2. Line list used in this work and their observed kinematical properties. Columns: (1) Line ID. (2) Rest-frame wavelength in vacuum (UV lines) or air (optical lines). (3) Ion need to produce the line. (4) Ionization energy of the preceding ionization state except for the recombination He II $\lambda 1640$ and $\lambda 4686$ lines, for which we give the ionization energy of the next ionization level, which is what matters. (5) Radial velocity of intrinsic lines (relative to Al II $\lambda 1671$ for the UV and relative to H α $\lambda 6563$ in the optical). (6) Full width half maximum of intrinsic line.

Line ID (1)	λ_{rest} (Å) (2)	Ion needed (3)	eV needed (4)	V_r (km s^{-1}) (5)	<i>FWHM</i> km s^{-1} (6)
Si II $\lambda 1527$	1526.71	Si ⁺	8.15	−2	92
C IV $\lambda 1548$	1548.19	C ³⁺	47.89	24	149
C IV $\lambda 1551$	1550.77	C ³⁺	47.89	56	128
He II $\lambda 1640$	1640.42	He ²⁺	54.42	−15	158
O III] $\lambda 1661$	1660.81	O ²⁺	35.12	7	145
O III] $\lambda 1666$	1666.15	O ²⁺	35.12	−13	141
Al II $\lambda 1671$	1670.79	Al ⁺	5.99	1	83
[C III] $\lambda 1907$	1906.68	C ²⁺	24.38	−	−
C III] $\lambda 1909$	1908.73	C ²⁺	24.38	−	−
He II $\lambda 4686$	4685.71	He ²⁺	54.42	0	209
H β $\lambda 4861$	4861.33	H ⁺	13.60	9	174
[O III] $\lambda 4959$	4958.91	O ²⁺	35.12	8	175
[O III] $\lambda 5007$	5006.84	O ²⁺	35.12	6	170
He I $\lambda 5876$	5875.59	He ⁺	24.6	13	144
[O I] $\lambda 6300$	6300.30	O ⁰	0	10	120
[S III] $\lambda 6310$	6312.10	S ²⁺	23.33	4	137
[N II] $\lambda 6548$	6548.04	N ⁺	14.53	18	177
H α $\lambda 6563$	6562.80	H ⁺	13.60	10	137
[N II] $\lambda 6584$	6583.46	N ⁺	14.53	14	123
[S II] $\lambda 6717$	6716.44	S ⁺	10.36	11	122
[S II] $\lambda 6731$	6731.18	S ⁺	10.36	1	128
[Ar III] $\lambda 7135$	7135.80	Ar ²⁺	27.63	9	126
[S III] $\lambda 9068$	9068.60	S ²⁺	23.33	22	115

was primarily done for two purposes, to better see the P-Cygni like component to C IV $\lambda\lambda 1548, 1551$, and to determine if weak broad He II $\lambda 1640$ emission is present. Note that broad stellar He II emission is not necessarily expected at the metallicity of SBS 0335−052E (Gräfener & Vink 2015). The result is shown in the left and middle panels of Fig. 5. The C IV $\lambda\lambda 1548, 1551$ doublet is composed of a weak P-Cygni like profile and nebular-like emission which is redshifted relative to Al II $\lambda 1671$. The redshifted C IV lines are reminiscent of the redshifted Ly α lines which are sometimes observed in Ly α emitting galaxies and are explained by radiative transfer effects Verhamme et al. (2012). He II $\lambda 1640$ and O III] $\lambda 1666$ are in pure nebular emission (*FWHM* $\sim 158 \text{ km s}^{-1}$) and at rest relative to Al II $\lambda 1671$. Note that when comparing the profiles of He II $\lambda 1640$ in Fig. 2 and the middle panel of Fig. 5, one can observe a slight increase in the amplitude of the He II which is consistent with the observational errors. No increase in the nearby continuum of the averaged spectrum is observed.

The right-hand panel of Fig. 5 shows a close-up view of [C III], C III] $\lambda\lambda 1907, 1909$, for which only G185M data from program 13788 is available. The individual components of [C III], C III] $\lambda\lambda 1907, 1909$ are barely resolved. For densities lower than or equal to $10^{3.5} \text{ cm}^{-3}$, the maximum flux ratio, $F(1907)/F(1909)$, is expected to be $\sim 5/3$ (Ferland 1981). We find that the observed value of this ratio fluctuates around $\sim 5/3$ depending on the sampling used and how we fit the continuum.

Since the G160M data from program 15193 became publicly available when we had completed most of the analysis of the present work, and since the G160M data from our program (13788) has

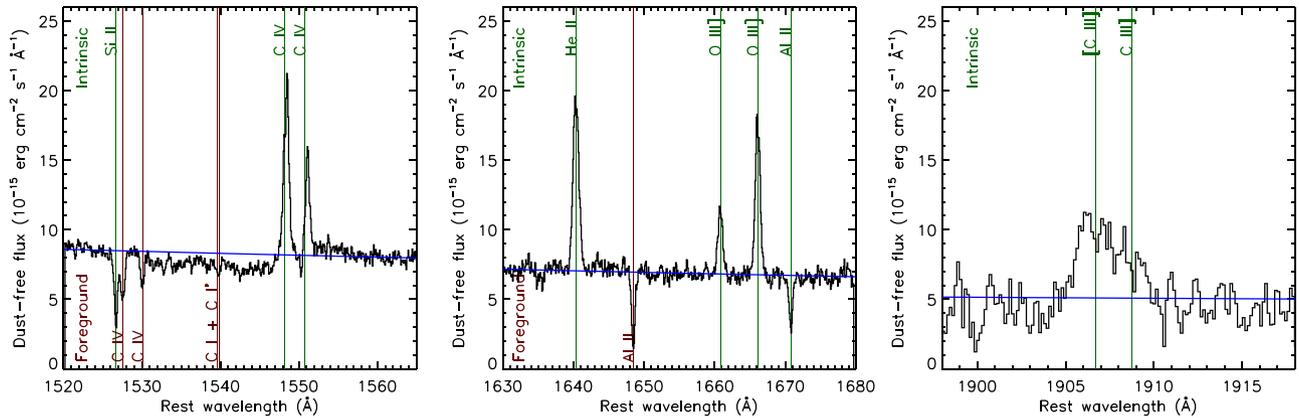


Figure 5. Close-up views of high-ionization UV emission lines. In the left and middle panels, we show the weighted average of the G160M data from programs 13788 (PI: Wofford) and 15193 (PI: Aloisi), while in the right-hand panel we show the G185M data from program 13788. Intrinsic lines are identified at the top, while foreground lines are identified at the bottom. All spectra are smoothed to the effective spectral resolution ($\text{FWHM} = 0.57 \text{ \AA}$) and corrected for: foreground reddening using $E(B - V) = 0.047 \text{ mag}$ and the Milky Way extinction law; redshift using $z_{1671} = 0.01349$; and intrinsic reddening using $E(B - V) = 0.017 \text{ mag}$ and the SMC extinction law. We overlay a power-law fit to the continuum, which is given by $F = 1.72 \times 10^9 \times \lambda^{-2.6}$ (blue curve). Left-hand panel: Note that the intrinsic Si II $\lambda 1527$ line is offset by only 4 km s^{-1} while the intrinsic C IV $\lambda\lambda 1548, 1551$ nebular emission lines are redshifted by 44 and 68 km s^{-1} , respectively. Also note the P-Cygni like profile of C IV. Middle panel: Note that the He II $\lambda 1640$ and O III] $\lambda\lambda 1661, 1666$ nebular emission lines are at rest relative to the intrinsic Al II $\lambda 1671$. Also note the tentative detection of excess emission on both sides of He II. Right-hand panel: Similar but for the region around [C III], C III] $\lambda\lambda 1907, 1909$. The components of the latter doublet are barely resolved.

excellent SNR, the models tested in this work and the physical properties which we derive use the data from our program alone. The SBS 0335–052E data from program 15193 are analysed in Hernandez et al. (2020) in the context of neutral-gas chemical abundances.

Moving to the optical, the [O III] and H α lines show narrow and broad components, but the flux of the broad component is negligible relative to the flux of the narrow component in both cases. The models which we test have a spectral sampling of 1 \AA . For comparing models with observations: (i) we re-sampled the COS spectra to match the sampling of the models; and we (ii) re-computed the emission line fluxes, using the coarser sampling and Trapezoidal numerical integration for profiles which were non-Gaussian. The corresponding fluxes are shown in Table 3. For the comparison of models with observations we define the C IVe flux which includes the stellar and nebular emission, and the C IVa flux, which includes the stellar absorption only. Note that the spectral resolution of the MUSE data is comparable to that of the models. Thus we do not recompute the optical fluxes previous to the comparison of these with the model fluxes.

4 DOMINANT SOURCE OF IONIZING PHOTONS

Izotov et al. (1997) observed SBS 0335–052E with the MMT using a $1 \text{ arcsec} \times 181 \text{ arcsec}$ slit. They found that the intensity of the He II $\lambda 4686$ emission, integrated over the $1 \text{ arcsec} \times 6 \text{ arcsec}$ central region, is 3 per cent that of H β and several orders of magnitude larger than the theoretical values predicted by models of photoionized H II regions. They also found that in SBS 0335–052E the peak of the He II $\lambda 4686$ emission is 200 pc to the NW relative to the peak of the emission of other emission lines (H β , [O III] $\lambda\lambda 5007, 4363$, and He I $\lambda 5876$). This led them to suggest that the He⁺ ionization is not caused by main-sequence O stars.

Izotov et al. (2006) observed SBS 0335–052E in the spectral range $\lambda 3620\text{--}9400 \text{ \AA}$ at $\text{FWHM} = 0.5 - 1 \text{ \AA}$, with the imaging

spectrograph GIRAFFE on the UT2 of the VLT. Their observations cover a region of $11.4 \text{ arcsec} \times 7.3 \text{ arcsec}$. These authors produced images of the galaxy in the continuum and in emission lines of different stages of excitation, with a spatial scale of 0.52 arcsec per pixel; and arrived to the same conclusion as Izotov et al. (1997). Indeed, they found that while the maximum of emission in the majority of lines, including the strong lines H β $\lambda 4861$, H α $\lambda 6563$, [O III] $\lambda\lambda 4363, 5007$, [O III] $\lambda\lambda 3726, 3729$, coincides with the youngest south-eastern SSCs 1 and 2, the emission of He II $\lambda 4686$ line is offset to the more evolved north-west SSCs 4 and 5. This finding and the fact that the velocity dispersion of the He II $\lambda 4686$ line is systematically higher, by $\sim 50\text{--}100$ per cent, relative to the other lines led the latter authors to suggest that the hard ionising radiation which is responsible for the He II $\lambda 4686$ emission is not related to the most massive youngest stars, but rather to fast radiative shocks. In this section, we investigate if HMXB, accreting intermediate-mass black holes (IMBHs), fast-radiative shocks, WR stars, or simple stellar populations which account for massive star evolution in close binaries could be responsible for the He II $\lambda 4686$ and He II $\lambda 1640$ emission towards the older SSCs 4 and 5.

4.1 High-mass X-ray binaries

Garnett et al. (1991) suggested that X-rays produced by massive X-ray binaries could be a source of ionization in H II regions. According to Izotov et al. (1997), ~ 4000 HMXBs would be required to explain the high luminosity of He II $\lambda 4686$ ($5.74 \times 10^{38} \text{ erg s}^{-1}$) in SBS 0335–052E. They obtained this number of HMXBs by directly scaling from the observed luminosity of the nebula surrounding LMC X-1, at He II $\lambda 4686$, which is $1.5 \times 10^{36} \text{ erg s}^{-1}$ (Pakull & Angebault 1986). The equivalent number of O7 stars that these authors inferred from the H β luminosity ($2.06 \times 10^{40} \text{ erg s}^{-1}$) is ~ 5000 . However, in a deep *Chandra* observation of the galaxy, Thuan et al. (2004) detect a faint X-ray point source (29 counts, $L_X = 3.5 \times 10^{39} \text{ erg s}^{-1}$) lying $\sim 0.3 \text{ arcsec}$ north of SSC 2. This source is consistent with only two HMXB according to table 1 of Douma et al. (2015).

Table 3. Observed fluxes and equivalent widths of high-ionization UV emission lines. Columns: (1) Line ID. (2) Wavelength limits of regions defining the left continuum, line, and right continuum. (3) Flux corrected for foreground reddening as described in Section 3.1. (4) Flux error. (5) Rest-frame equivalent width. (6) Equivalent-width error. (7) Method used for obtaining the flux (T = Trapezoidal numerical integration, G = Gaussian fitting). Method T is adopted for non-Gaussian line profiles. The errors were obtained using error propagation.

Line ID	Wavelength ranges Å	F (erg s ⁻¹ cm ⁻²)	σ_F (erg s ⁻¹ cm ⁻²)	EW (Å)	σ_{EW} (Å)	Method
(1)	(2)	(3)	(4)	(5)	(6)	(7)
C Iva 1550	1516.4,1522.4,1540.4,1547.4,1561.4,1567.4	-6.14E-15	4.58E-16	-0.9	0.1	T
C I ve 1550	1516.4,1522.4,1547.4,1552.0,1561.4,1567.4	1.15E-14	4.63E-16	1.7	0.2	T
He II λ 1640	1627.0,1632.0,1632.0,1650.0,1650.0,1655.0	1.02E-14	5.21E-16	1.7	0.3	G
O III] λ 1661	1653.0,1659.0,1659.5,1662.1,1673.0,1677.0	3.90E-15	4.59E-16	0.7	0.1	G
O III] λ 1666	1653.0,1659.0,1664.8,1667.5,1673.0,1677.0	9.02E-15	4.93E-16	1.5	0.1	G
[C III], C III] $\lambda\lambda$ 1907, 1909	1897.0, 1903.0,1903.0,1912.0,1911.0,1920.0	1.83E-14	1.35E-15	5.0	0.8	T
He II λ 4686	4675.0, 4680.0, 4680.0, 4690.0, 4690.0, 4695.0	1.62E-15	4.12E-17	4.5	0.0	G
[Ar IV] λ 4740	4720.0, 4725.0, 4737.5, 4743.5, 4745.0, 4750.0	3.89E-16	1.41E-17	1.0	0.0	G
H β λ 4861	4845.0, 4850.0, 4855.0, 4868.0, 4870.0, 4875.0	4.20E-14	8.72E-16	130.3	0.5	G
[O III] λ 4959	4946.0, 4951.0, 4952.0, 4966.0, 4967.0, 4972.0	4.64E-14	9.60E-16	148.1	0.6	G
[O III] λ 5007	4990.0, 4995.0, 5001.0, 5012.0, 5020.0, 5025.0	1.39E-13	2.86E-15	461.7	1.8	G
He I λ 5876	λ 5860.0, 5865.0, 5871.5, 5880.5, 5883.0, 5888.0	4.23E-15	9.23E-17	21.2	0.1	G
[O I] λ 6300	6280.0, 6290.0, 6298.0, 6303.0, 6320.0, 6330.0	2.90E-16	9.15E-18	1.6	0.0	G
[S III] λ 6310	6290.0, 6295.0, 6309.5, 6315.0, 6320.0, 6325.0	2.77E-16	8.78E-18	1.6	0.0	G
[N II] λ 6548	6520.0, 6525.0, 6545.0, 6550.0, 6600.0, 6605.0	1.17E-16	6.09E-18	0.8	0.0	G
H α λ 6563	6520.0, 6525.0, 6552.0, 6570.0, 6600.0, 6605.0	1.23E-13	2.53E-15	777.2	3.0	G
[N II] λ 6584	6520.0, 6525.0, 6581.0, 6587.0, 6600.0, 6605.0	3.40E-16	9.99E-18	2.4	0.0	G
[S II] λ 6717	6708.0, 6713.0, 6713.5, 6720.5, 6721.0, 6725.0	8.27E-16	2.00E-17	5.3	0.0	G
[S II] λ 6731	6721.0, 6725.0, 6727.5, 6734.5, 6735.0, 6740.0	6.65E-16	1.66E-17	4.2	0.0	G
[Ar III] λ 7135	7120.0, 7125.0, 7132.0, 7141.0, 7150.0, 7155.0	8.22E-16	2.00E-17	6.5	0.0	G
[S III] λ 9068	9057.0, 9062.0, 9063.0, 9074.0, 9075.0, 9080.0	1.45E-15	3.37E-17	21.6	0.2	G

Although the position of the X-ray maximum in the galaxy is no more accurate than 0.42 arcsec, it is consistent with the X-ray source being physically associated with SSC 2. The time-scale for HMXB formation of 3–10 Myr after the onset of the starburst is consistent with the ages of SSCs 1 and 2 (3 Myr, Reines et al. 2008; Adamo et al. 2010). By re-analysing the above Chandra observations, Prestwich et al. (2013) confirm that the X-ray point source which is mentioned above corresponds to an ultraluminous X-ray source (ULX), i.e. a source having $L_X > 10^{39}$ erg s⁻¹. We use the coordinates of table 5 in Prestwich et al. (2013) to find the location of the ULX relative to our COS aperture. The location is indicated by the yellow circles at the southern edge of the COS footprint in fig. 1. If the ULX is responsible for the extended He II emission in SBS 0335–052E, one needs to explain why the estimated location of the ULX is a few hundred parsecs away from the peak of the He II emission and completely discard other sources of ionization.

Izotov et al. (1997) note that the elongated shapes of the faint X-ray contours in the Chandra image, to the north of SBS 0335–052E, suggest that there may also be X-ray emission associated with SSC 3 and the pair 4 + 5. There is also a very faint X-ray source to the northeast that appears to be associated with the supernova cavity seen in the optical image. Excluding the source near SSC 2, the total X-ray emission of the fainter features is ≤ 23 counts, making it virtually impossible to distinguish between point-like and truly extended.

More recently, Kehrig et al. (2018) re-analyzed the above Chandra data. By covolving model-dependent observed X-ray luminosities with the energy-dependent cross-section of He⁺, they estimate the effective X-ray ionizing power. When Kehrig et al. extrapolate from the X-Ray to 4 Rydberg, they infer $\log Q(\text{He II}) \sim 36$, while the observed optical He II line flux requires $\log Q(\text{He II}) \sim 51$. This is a discrepancy of 15 orders of magnitude. The discrepancy is not resolved, if one considers only the diffuse extended He II.

4.2 Accreting IMBHs

Intermediate-mass black holes ($10^2 - 10^6 M_\odot$) are expected to reside in low-mass, star-forming dwarf galaxies that have not significantly grown through mergers/accretion. The detection of unresolved X-ray emission, in some cases spatially coincident with jet/core radio emission, constitutes the most compelling signature of an accreting BH in the absence of dynamical mass measurements (typically limited to the Local Group in the case of dwarf galaxies) and has provided evidence for their presence in a few tens of low-mass galaxies (Mezcua et al. 2016). There is no evidence from the combination of radio (Johnson et al. 2009) and X-ray (Thuan et al. 2004; Prestwich et al. 2013) data that there is an IMBH in SBS 0335–052E. But where does SBS 0335–052E fall on optical diagnostic diagrams for separating star-forming galaxies from AGN?

The position of SBS 0335–052E in the optical diagnostic diagrams of Baldwin, Phillips & Terlevich (1981) and Veilleux & Osterbrock (1987) is given by the red circle in Fig. 6. We overlay models of giant H II regions ionized by simple stellar populations (SSPs) computed by Gutkin et al. (2016, stars), and narrow-line region AGN models computed by Feltre, Charlot & Gutkin (2016, dashed-curves). The upper mass limit of the IMF is $M_{\text{up}} = 100 M_\odot$ for the upper row and $300 M_\odot$ for the bottom row. There are no significant differences between the H II-region models of the top and bottom panels. Most importantly, at extreme-low metallicity (dark-blue and cyan models), the AGN models (dashed lines) are located on the star-forming zone set by Kewley et al. (2001). Furthermore, the AGN and star-forming models overlap in the first diagnostic diagram (first column) at $Z = 0.001$, and the last diagnostic diagram (last column) at $Z = 0.0005$. Thus, at the metallicities of XMPs, these optical diagrams are not useful for separating AGN from star-forming regions. Note that the SSP + H II region models of $Z = 0.001$ approach the SBS 0335–052E observation in the three optical diagnostic diagrams.

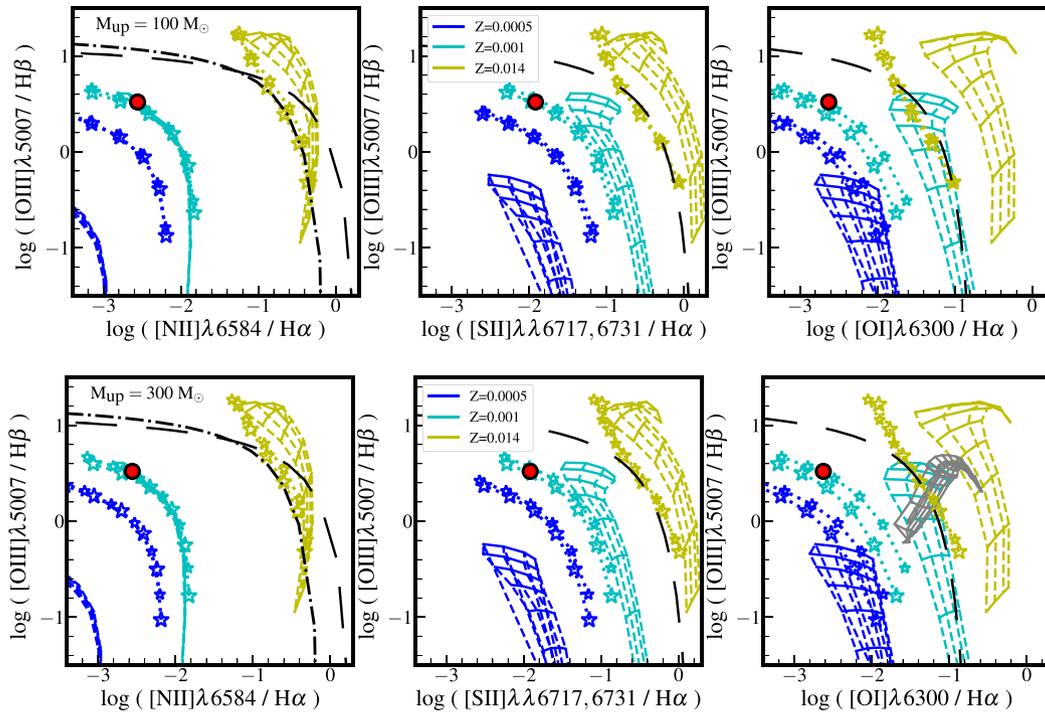


Figure 6. Location of SBS 0335–052E (red filled circle) in the optical diagrams of Baldwin et al. (1981) and Veilleux & Osterbrock (1987) for separating narrow-line AGN from H II region-like galaxies. The observational errors are smaller than the symbol size. In the left-panel diagrams, the main excitation mechanism is star formation if the galaxy lies below the dotted-dashed black curve (Kauffmann et al. 2003), an AGN if the galaxy lies above the long-dashed black curve (Kewley et al. 2001), and a mixture of star formation and Seyfert nucleus, or star formation and LINER, if the galaxy lies between these curves (Kewley et al. 2006). We overlay: (i) narrow-line AGN models of Feltre et al. (2016) (dashed lines intersected by solid lines); and (ii) SSP H II region models of Gutkin et al. (2016) (stars connected with dotted lines). Models (i) and (ii) correspond to ionization parameters from -1.5 to -4 (H II regions) and -1.5 to 5 (AGN, increasing from bottom to top), three values of the interstellar metallicity (as colour-coded by the legends of the middle panels), a dust-to-metal mass ratio of $\xi_d = 0.3$, and hydrogen gas densities of $n_H = 10^2$ (H II regions) and 10^3 cm^{-3} (AGN). We show H II region-like galaxies with: ages of 10^6 and $10^{6.5}$ yr (small and big stars, respectively); IMF upper mass limits of 100 and $300 M_\odot$ (top and bottom rows, respectively); and a C/O ratio of $0.72 (C/O)_\odot$.

Feltre et al. (2016), computed predictions of high-ionization UV emission lines from narrow-line region models of AGN and proposed diagnostic diagrams for separating AGN from star-forming galaxies which are based on UV emission-line ratios. These diagrams separate better star-forming galaxies from AGN at extreme low metallicities, as can be seen in the diagrams which are shown in Fig. 7. In the latter figure, the left- and right-hand panels show SSPs with $M_{\text{up}} = 100 M_\odot$ and $300 M_\odot$, respectively. Although the association of blueshifted $O \text{ V } \lambda 1371$ with VMSs is not definitive, models with VMSs are more likely to reproduce the fluxes of UV high-ionization lines than models with an upper mass limit of $100 M_\odot$ (Gutkin et al. 2016), which is why we test them in this work. Only the $Z = 0.0005$ and $M_{\text{up}} = 300 M_\odot$ SSP + H II region model approaches the UV line ratios of SBS 0335–052E. However, the closest model has nebular + broad He II $\lambda 1640$ emission. As already mentioned, broad He II emission is not necessarily expected at the low metallicity of SBS 0335–052E, and in this galaxy it is not observed (see Fig. 5).

We now discuss some of the IR properties of SBS 0335–052E in the context of this subsection. In Fig. 8 we show that SBS 0335–052E falls within the wedges of Lacy et al. (2004) and Donley et al. (2012) that select AGN in the *Spitzer*/IRAC colour–colour diagram. The position of SBS 0335–052E in Fig. 8 is attributed to an IR excess due to warm dust. Fig. 1 of Houck et al. (2004) shows the *Spitzer* IRS spectrum of SBS 0335–052E. The latter authors say that if the mid-IR spectrum of SBS 0335–052E is representative of star-forming galaxies at higher redshifts, it may be difficult to distinguish between galaxies whose emission is dominated by star

formation and AGN, as the latter case is associated with a relatively featureless flat spectra in the mid-IR.

4.3 Fast radiative shocks from SNRs

Garnett et al. (1991) suggested that radiative shocks produce strong He II emission under certain conditions. The strength of such emission is sensitive mostly to the velocity of the shock, reaching a maximum for $V_{\text{shock}} \sim 120 \text{ km s}^{-1}$, dropping rapidly at higher velocities.

The *FUSE* spectrum of SBS 0335–052E which is presented in Grimes et al. (2009) shows no O VI in emission at $1032\text{--}1038 \text{ \AA}$, which at higher metallicities is generally interpreted as an indication of the presence of shocked gas from SNRs (Rasmussen & Martin 1992; Allen et al. 2008). For reference, the size of the *FUSE* aperture is $25 \text{ arcsec} \times 25 \text{ arcsec}$. It is not clear if the absence of O VI in emission is due to a combination of the low metallicity of the galaxy and the strong nearby Ly β absorption.

Kehrig et al. (2018), analysed the MUSE spectra of SBS 0335–052E but did not correct for the fact that the MUSE DRS sky-subtraction algorithm initially oversubtracted H α and [O III] emission (see our Section 3.2). Thus, we expect their flux measurements of the above lines to suffer from this oversubtraction. They find no high ratios of [S II]/H α and/or [O I]/H α in SBS 0335–052E. At higher metallicities, high values of the latter ratios indicate the presence of shock-excited gas.

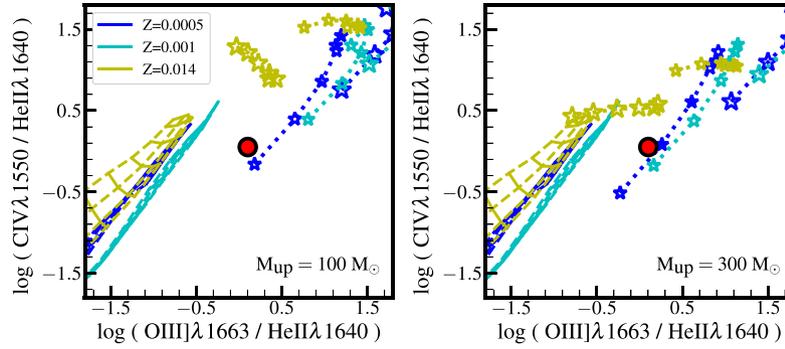


Figure 7. Location of SBS 0335–052E (red filled circle) in the UV diagnostic diagram, C IV ($\lambda 1549 + \lambda 1551$)/He II $\lambda 1640$ versus O III] ($\lambda 1661 + \lambda 1666$)/He II $\lambda 1640$. The observational error bars are smaller than the symbol size. We use the same models and colour scheme as in Fig. 6. The upper mass limit of the stellar IMF is $100 M_{\odot}$ in the left-hand panel and $300 M_{\odot}$ in the right-hand panel.

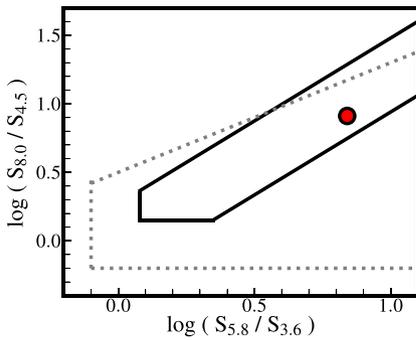


Figure 8. Wedges of Lacy et al., 2004 and Donley et al., 2012 (dashed grey and continuous black lines, respectively) that select AGN in the *Spitzer*/IRAC colour-colour diagram. The non-dashed line represents a more strict cut. SBS 0335–052E (red symbol) would be selected as an AGN based on this diagram. We use the data from the IRAC 3.8 arcsec aperture which is provided by NASA/IPAC Extragalactic Database (NED).

Fig. 6 of Alarie & Morisset (2019), shows optical emission line ratios for shock + precursor models of extremely low metallicity in the right-hand panel. Note that what is actually plotted in that figure is $\log([\text{O III}]\lambda 5007/\text{H}\beta)$ and $\log([\text{N II}]\lambda 6584/\text{H}\alpha)$ (private communication). At SMC metallicity, the models of Alarie & Morisset (2019) are in excellent agreement with the similar models of Allen et al. (2008). However, SMC metallicity is the lowest metallicity which is available for the models of Allen et al. (2008) and this is higher than the known metallicity of SBS 0335–052E. Thus, we compare against the models of Alarie & Morisset (2019). We find that the optical ratios of SBS 0335–052E, $\log([\text{O III}]\lambda 5007/\text{H}\beta) = 0.52$ and $\log([\text{N II}]\lambda 6584/\text{H}\alpha) = -2.56$, are incompatible with the values of the extremely low metallicity shock + precursor models of Alarie & Morisset (2019).

4.4 Wolf–Rayet stars

WR stars are the evolved descendants of stars with initial masses of $>25 - 30 M_{\odot}$. They are divided into two classes, those with strong lines of He and N (WN class) and those with strong He, C, and O (WC and WO class). Depending on the relative intensities of their N or C lines, these classes are further divided into subtypes. WN spectral subtypes follow a scheme involving line ratios of N III-V and He I-II, ranging from WN2 to WN5 for early WN (WNE) stars, and WN7 to WN9 for late WN (WNL) stars. WC spectral subtypes depend

on the line ratios of C III and C IV lines along with the appearance of O III-V, spanning WC4 to WC9 subtypes, for which WC4 – 6 stars are early (WCE) and WC7 – 9 are late (WCL). Rare, O-rich WO stars form an extension of the WCE sequence, exhibiting strong O VI $\lambda\lambda 3811 - 34$ emission. A review of the spectral and physical properties of these stars is provided in Crowther (2007).

Schaerer & Vacca (1998) modelled the nebular and WR He II $\lambda 4686$ emission in young starbursts with metallicities between 0.2 and 1 times Z_{\odot} , and predicted a strong nebular He II $\lambda 4686$ emission due to a significant fraction of WC stars in the early WR phases of the burst. Based on the fact that observations of extreme-low metallicity young starbursts lack broad WR-like emission features, Izotov et al. (1997) excluded WR stars as the explanation for the intense He II $\lambda 4686$ emission in SBS 0335–052E. However, more recent works have reported tentative detections of WR stars in SSC 3, which is within the COS aperture.

Papaderos et al. (2006) observed the galaxy with a 1.2 arcsec wide slit at the 3.6 ESO telescope. They report the tentative detection of WC4 (early carbon-type WR) stars in SSC 3, which is within our COS aperture. The evidence is a broad emission feature which is coincident with the $[\text{Fe III}]\lambda 4658$ nebular emission line but has a width much larger than those of other nebular emission lines in the spectrum of SSC #3. (Papaderos et al. 2006) interpret this broad feature as due to C IV $\lambda 4658$, which is a signature of WC4 stars. From this feature, corrected for $[\text{Fe III}]\lambda 4658$ emission, they estimate a total number of ~ 22 WC4 stars. Given the derived number of O7V stars (~ 800), the ratio of WC4 to O7V stars they find is consistent with the stellar population synthesis models at solar metallicity. Papaderos et al. (2006) point out that Crowther (2007) showed that the line luminosities of extremely low-metallicity WC4 stars are 3–6 times lower than those predicted by the above models, which would increase the number of WC4 stars to 70–130. However, the latter authors say that in order to confirm the detection of WC4 stars and to better determine their number, better observations of the C IV $\lambda 5808$ emission line are required. Finally, Papaderos et al. (2006) found: weak N III $\lambda 4640$ and He II $\lambda 4686$ emission in SSC 3, thus the determination of the number of WNL stars was not possible; and no WR features in regions 1, 2, 4, and 5.

We used the MUSE spectra to further investigate the presence of WR stars. For this purpose, we extracted a smaller datacube around SSC 3. The results are shown in Fig. 9. The data show that WC4 stars could be present in the vicinity of SSC 3. However, we find no evidence of C IV $\lambda 5808$ emission in the vicinity of this SSC. As previously mentioned, Kehrig et al. (2018) independently analysed the MUSE data (uncorrected for the initial MUSE DRS

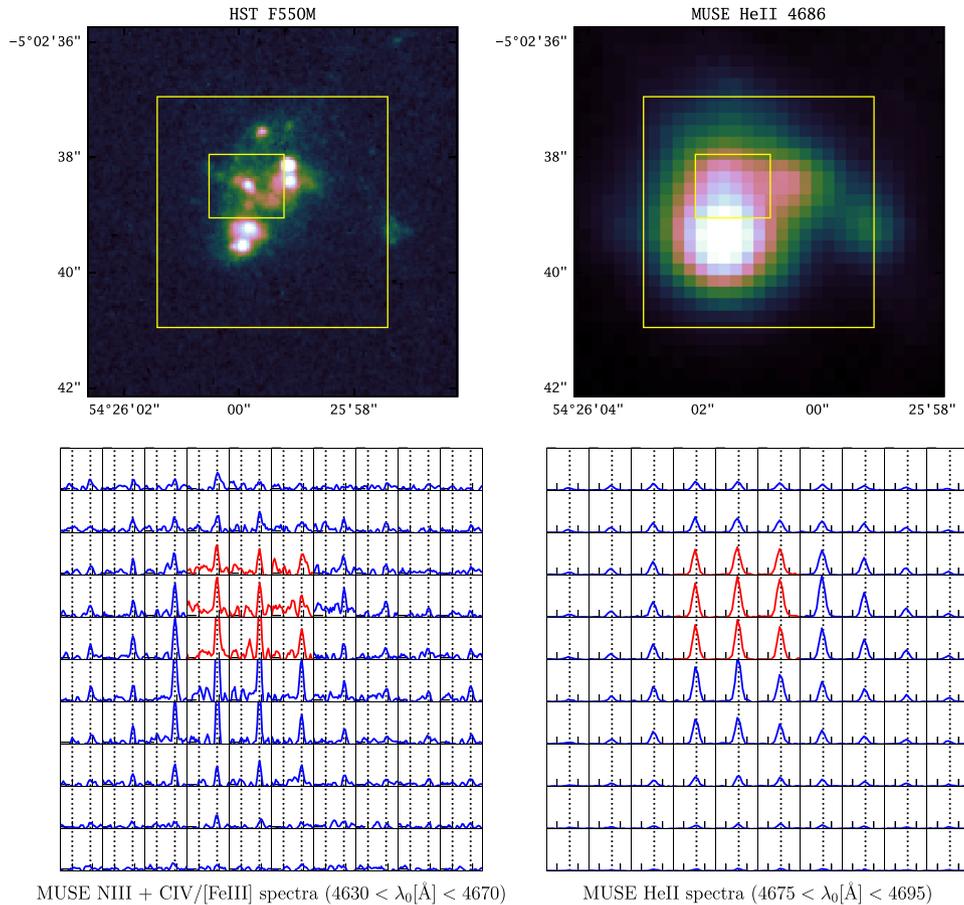


Figure 9. Tentative evidence for broad WR emission lines N III $\lambda 4640$ and C IV/[Fe III] $\lambda 4658$ in MUSE observations. Top left: *HST* F550M image of SBS 0335–052E. The spectral window of this medium band filter does not coincide with any emission lines from the galaxy, therefore it represents an ideal representation of the stellar light. The big yellow box represents the $4 \text{ arcsec} \times 4 \text{ arcsec}$ region for which the spectra from the MUSE cube (shown in the bottom panels) have been extracted, while the small yellow box indicates a $1.2 \text{ arcsec} \times 1.2 \text{ arcsec}$ region around SSC 3. Top right: Continuum subtracted He II $\lambda 4686$ image from MUSE data shown for the same field of view as in the top right-hand panel. The image is displayed using an arcsinh-stretch, with the brightest pixels corresponding to flux densities of $10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$. Bottom left: MUSE continuum subtracted spectra, each extracted over 2×2 spaxels (i.e. quadratic apertures of $0.4 \text{ arcsec} \times 0.4 \text{ arcsec}$), showing the region around the broad WR signatures N III $\lambda 4640$ and C IV $\lambda 4658$ (indicated by vertical dotted lines). The C IV $\lambda 4658$ emission feature is blended with nebular [Fe III] $\lambda 4658$ emission. Shown in red are the spectra tracing emission from cluster 3, i.e. within the region outlined by the smaller yellow box in the top panels. The scale on the y-axis is linear from 0 to $8 \times 10^{-18} \text{ erg s}^{-1} \text{ \AA}^{-1}$ and the spectra have been smoothed by a 3.6 \AA (3 spectral pixel) wide boxcar filter in order to enhance the visibility of the faint features. Bottom right: Same as the bottom-left panel, but for He II $\lambda 4686$ emission with the y-axis extending from 0 to $10^{-16} \text{ erg s}^{-1} \text{ \AA}^{-1}$ and with unsmoothed spectra.

sky-oversubtraction). They found a WR knot in SBS 0335–052E. However, their analysis of the data discards single WR stars as the dominant source of He II ionization.

4.5 SSPs with massive stars in close-binary systems

There is observational evidence that massive stars are in binary systems with close to 70 per cent interacting over the course of their evolution (Massey et al. 2009; Chini et al. 2012; Sana et al. 2012, 2013). BPASS v2.1 SSP models which account for massive star evolution in close binaries are presented in Eldridge et al. (2017, stellar component) and Xiao, Stanway & Eldridge (2018, nebular component). Note that at the metallicity of SBS 0335–052E ($Z \sim 6 \times 10^{-4}$), there are no observational constraints for individual close-binary systems.

Kehrig et al. (2018) used models which account for single rotating stars and binaries to interpret the integrated He II $\lambda 4686$ luminosity of SBS 0335–052E and found that it can only be reproduced with

single rotating stars if the stars are metal-free and with binaries if the IMF is top-heavy and the stellar metallicity is $Z \sim 10^{-5}$. Fig. 12 of Stanway & Eldridge (2019) confirms that BPASS v2.1 binary models with a metallicity of $Z = 10^{-5}$ can reproduce the integrated He II $\lambda 4686$ luminosity of SBS 0335–052E, which is reported in Kehrig et al. (2018). However, $Z = 10^{-5}$ is unrealistically low for SBS 0335–052E.

Here we investigate if the BPASS v2.1 binary models with nebular emission can reproduce the optical and UV emission line ratios which we measured for SBS 0335–052E. Fig. 10 shows optical and UV diagnostic diagrams similar to those of Figs 6 (bottom row) and 7 (right-hand panel). In addition, in the bottom left panel of Fig. 10 we show a plot of He II $\lambda 4686/H\beta$ versus age. In Fig. 10, the position of SBS 0335–052E (red filled symbol) is shown relative to BPASS v2.1 models.

We find that that the binary models which are closest to the observations are of different metallicities in each diagram. In particular, the top-right panel shows that only models with a metallicity higher than

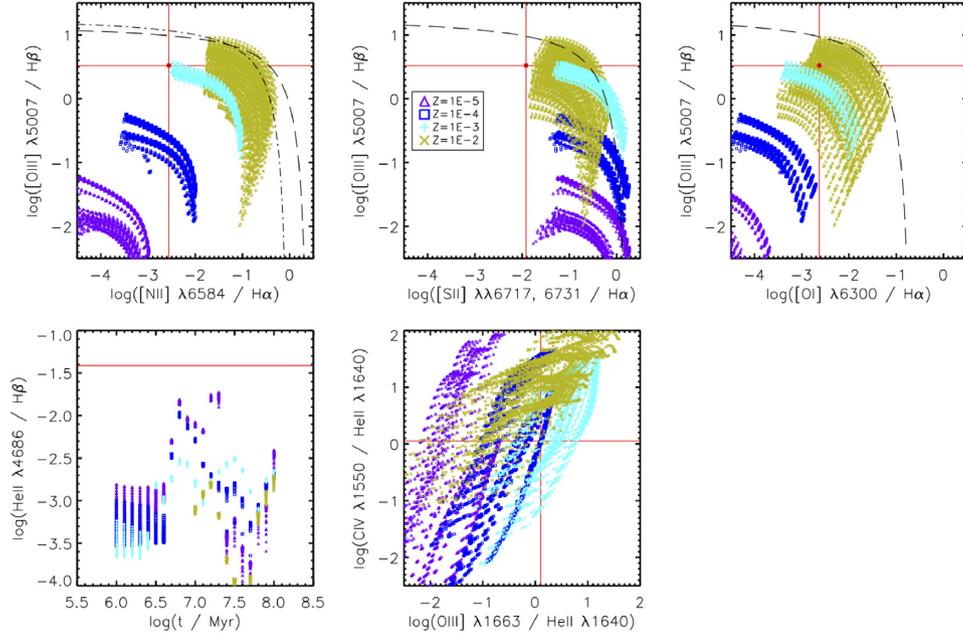


Figure 10. Top row: Same diagrams as in Fig. 6 but we show the location of SBS 0335–052E (red filled circle at intersection of vertical and horizontal red lines) relative to the SSP BPASS v2.1 models which are described in Section 4.5. The radius of the red circle is larger than or equal to the observational error bar. Bottom-left: Location of SBS 0335–052E (horizontal red line) relative to BPASS v2.1 models in the He II $\lambda 4686/H\beta$ versus age diagram. The y-axis error bar is of the size of the thickness of the red vertical line. Bottom-middle: Location of SBS 0335–052E (red filled circle at intersection of vertical and horizontal red lines) relative to BPASS v2.1 models in the UV diagnostic diagram. We show models of $M_{up} = 300 M_{\odot}$ and metallicities given in the legend of the top-middle panel. We include models with $\log(U)$ values ranging from -3.5 to -1.5 in steps of -0.5 and $\log(n[H])$ values ranging from 0 to 3 in steps of 1.

that of SBS 0335–052E can reproduce the observed emission line ratios. Furthermore, none of the binary models are able to reproduce the observed He II $\lambda 4686/H\beta$ ratio, as can be seen in the bottom-left panel of Fig. 10. On the other hand, some of the binary models with $Z = 0.001$ approach the UV emission line ratios of SBS 0335–052E, as shown in the bottom-right panel of the same Figure. Thus, we conclude that the SSP + H II region BPASS v2.1 binary models fail to reproduce the optical emission line ratios of SBS 0335–052E.

5 COMPARISON OF MODELS WITH OBSERVATIONS

In this section, we present the models and spectral interpretation tool which we use in this work. We also present results from five tests where we compare the models to different sets of observables.

5.1 Models and Tool

We compare models to the co-spatial COS and MUSE observations separately. For this purpose, we use the spectral interpretation tool, BEAGLE (Chevallard & Charlot 2016), which incorporates in a flexible and consistent way the production of radiation from stars and its transfer through the interstellar medium. The version of BEAGLE which we use relies on the models of Gutkin et al. (2016), who follow the prescription of Charlot & Longhetti (2001) to describe the emission from stars and the interstellar gas. In particular, the models are computed combining the latest version of the Bruzual & Charlot (2003) stellar population synthesis model with the standard photoionization code CLOUDY (Ferland et al. 2013). The latest implementation of Bruzual & Charlot (2003) use the most up-to-date

single non-rotating massive-star evolution models from the Padova group (Tang et al. 2014).

Using single non-rotating models, as we do in the present work, sets a reference for comparison with more comprehensive models. Also note that before interacting and after merging, massive stars evolve as single stars. The ionized-gas O/H and C/O which are obtained via standard methods disagree with the metal-free or unrealistically low metallicity scenarios.

In the models of Gutkin et al. (2016), the main adjustable parameters of the photoionized gas are the interstellar metallicity, Z_{ISM} ; the typical ionization parameter of newly ionized H II regions, U_S , which characterizes the ratio of ionizing-photon to gas densities at the edge of the Strömgren sphere; and the dust-to-metal mass ratio, ξ_d , which characterizes the depletion of metals on to dust grains. We let the C/O number abundance ratio vary, but only consider models with hydrogen density $n_H = 100 \text{ cm}^{-3}$. In their figures 5 (all panels) and 10 (panels b and d), Gutkin et al. (2016) show that at low metallicity, changing n_H from 10 to 1000 cm^{-3} has a negligible effect on the ratios of the optical and UV lines which are used in our paper. For the attenuation by dust, we adopt the 2-component model of Charlot & Fall (2000) to account for the effects linked to dust/star geometry (including ISM clumpiness). Table 4 gives the ranges of the main adjustable parameters. In the table, $\hat{\tau}_V$ is the V-band dust attenuation optical depth and μ the fraction of $\hat{\tau}_V$ arising from the diffuse ISM rather than from giant molecular clouds. We adopt the same metallicity for stars and star-forming gas ($Z_{stars} = Z_{ISM}$) and assume that all stars in a galaxy have the same metallicity. By fitting the *HST* photometry of the four SSCs which are located within the COS aperture, Reines et al. (2008) and Adamo et al. (2010) find an age range of 6–12 Myr. Given this age range, in this work, we test models of constant star formation. The adopted ranges for the galaxy mass, $\log(\text{mass}/M_{\odot})$, and time since the beginning of star

Table 4. Grid sampling of main adjustable parameters of photoionization models which are described in Section 5.1.

Parameter	Sampled values
$\log(Z_{\text{ISM}}/Z_{\odot})^a$	−2.2 to 0.24
$\log(U_S)$	−4.0, −3.5, −3.0, −2.5, −2.0, −1.5, −1.0
ξ_d	0.1, 0.3, 0.5
$\tilde{\tau}_V$	0 to 5
μ	0 to 1
$n_{\text{H}} \text{ (cm}^{-3}\text{)}$	100
$(\text{C/O})/(\text{C/O})_{\odot}^b$	0.14, 0.52, 1.0
$\log(t/\text{yr})^c$	6.0 to 10.0
$\log(\text{mass}/M_{\odot})^d$	3.0 to 11.0

^a $Z_{\odot} = 0.01524$. See Gutkin et al. (2016) for metallicity steps.

^b $(\text{C/O})_{\odot} = 0.44$.

^cTime since the beginning of star formation (max_stellar_age in BEAGLE).

^d $M_{\odot} = 1.989 \times 10^{33} \text{ g}$.

formation, $\log(t/\text{yr})$, are provided in Table 4. We adopt a Chabrier (2003) IMF with an upper mass limit of $M_{\text{up}} = 300 M_{\odot}$, i.e. we account for the presence of VMSSs, given that we found a blueshifted O V 1371 absorption.

With this parameterization, we use BEAGLE to fit the available observational constraints. Note that we do not fit any observables which depend on the velocity dispersion of the emission lines. As output, we obtain the posterior probability distributions of the above adjustable model parameters. By fitting the UV and optical data independently, we obtain independent measures of the time since the beginning of star formation and the gas metallicity.

5.2 Test 1: fit to P-Cygni + nebular C IV 1550 profile

When the integrated spectrum of a star-forming galaxy shows a P-Cygni like C IV $\lambda\lambda$ 1548, 1551 profile, the profile is generally interpreted as originating in the winds of O-type MS stars (section 3.4 Wofford et al. 2011). Fig. 24 of Leitherer et al. (2010) shows the C IV profile of a simple stellar population. The C IV emission component is strongest a few Myr after the burst of star-formation, when O-type MS stars are still present. Fig. 29 of Leitherer et al. (2010) shows the C IV profile in the case of constant star formation, 50 Myr after the beginning of star formation, i.e. when there is an equilibrium between birth and death of stars. In this case, C IV still shows a P-Cygni like profile. Fig. 29 of Leitherer et al. (2010) also shows that the C IV absorption component is sensitive to the metallicity of the stellar population. Figs 16–19 of Vidal-García et al. (2017) show similar behaviours of C IV for the models which are used in the present work.

As test 1, we attempt to fit the P-Cygni like plus nebular C IV profile of SBS 0335–052E and nothing else. Fig. 11 shows that the models are able to reproduce the combined stellar + nebular profile.

In Fig. 12, we show the corresponding two-dimensional joint probability distribution functions (PDFs) and one-dimensional marginal posterior PDFs of physical quantities. Hereafter, we refer to this type of plot as a triangle plot. The properties which are included in Fig. 12 are the time since the beginning of star formation, the ionized-gas oxygen abundance, the dust-to-gas metal mass ratio, and the ionization parameter. We find that just fitting the stellar + nebular C IV profile does not constrain any of the latter four quantities. In particular, the distributions are broad and/or double-peaked. When considering the median values of fitted parameters and comparing them to values of the physical properties which are reported in the literature, we find that the median-model time since the beginning of

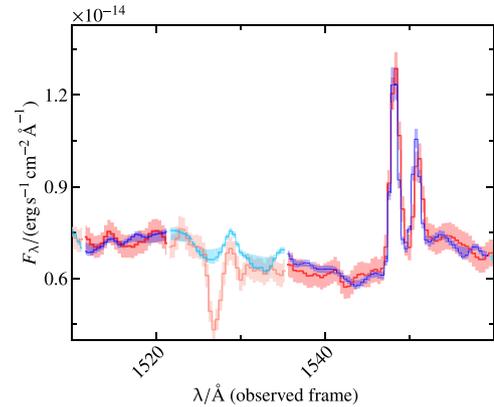


Figure 11. Fit to observed C IV $\lambda\lambda$ 1548, 1551 P-Cygni like + nebular profile. The red curve and shadow correspond to the observations and observational uncertainties, respectively. The dark-blue curve and shadow represents the posterior median of models and 68 per cent posterior credible interval, respectively. The cyan portion of the spectrum was excluded from the fit as it contains interstellar absorptions which are not accounted for in the models.

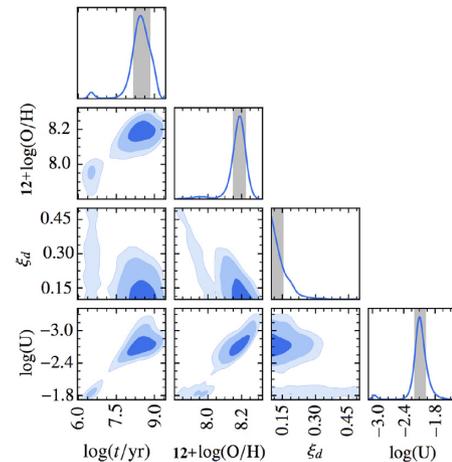


Figure 12. Triangle plot from fit to C IV profile. Table 4 and the text which describes it explain what the properties which are plotted are. The shaded blue clouds show values which are one, two, and three σ from the median. The shaded grey rectangles are values which are one σ from the median.

star formation, which is $\log(\text{age}/\text{yr}) \sim 7.8$ (~ 63 Myr), is higher than the range of photometrically derived ages for the individual SSCs within the COS aperture (Reines et al. 2008; Adamo et al. 2010); and that the median-model oxygen abundance, which is $12+\log(\text{O}/\text{H}) \sim 8.2$, is significantly higher than values obtained via the standard method by Papaderos et al. (2006), Izotov et al. (2006) and Izotov et al. (2009).

5.3 Tests 2 and 3: fits to UV fluxes and equivalent widths

As test 2, we try to simultaneously fit: (i) the fluxes of all nebular high-ionization UV emission lines and (ii) the equivalent widths (EWs) of the absorption and nebular emission components of C IV $\lambda\lambda$ 1548, 1551. The wavelength limits which we use to obtain the flux and EWs are given in Table 3. As test 3, we repeat test 2 but excluding the He II λ 1640 flux from the fit, as this is the highest ionization UV emission line which we detect. The results from tests 2 and 3 are summarized in Table 5, which gives: line ID (column 1), observed

Table 5. Observed versus median-model UV fluxes and equivalent widths. Columns: (1) Line ID. (2) Either the logarithm of the observed flux ($\text{erg s}^{-1} \text{cm}^{-2}$) corrected for foreground reddening only (top five rows), or the EWs (\AA) of the C IV components which are described in Section 5.3 (last two rows). The values in this column are the same as the ones in Table 3. (3) Either the median-model flux including intrinsic reddening, or the median-model EW, both for test 2. (4) Same as (3) but for test 3. (5) Residual between (2) and (3). (6) Residual between (2) and (4).

Line ID (1)	Observed (2)	Test 2 (3)	Test 3 (4)	Δ_1 (5)	Δ_2 (6)
C IVe 1550	-13.94 ± 0.02	-13.95	-13.95	-0.01	-0.01
He II λ 1640	-13.99 ± 0.02	-14.00	-	-0.01	-
O III] λ 1661	-14.41 ± 0.05	-14.45	-14.46	-0.04	-0.05
O III] λ 1666	-14.04 ± 0.02	-14.02	-14.03	-0.02	-0.01
C III] λ 1909	-13.74 ± 0.03	-13.72	-13.72	+0.02	+0.02
C IVe 1550	1.70 ± 0.20	1.70	1.83	0.00	0.13
C IVa 1550	-0.90 ± 0.10	-0.29	-0.28	0.61	0.62

Table 6. Median physical properties derived from the UV data. For each property, we give the one-sigma errors around the median. Columns: (1) Property. (2) Results from test 1. (3) Results from test 2.

Line ID (1)	Test 2 (2)	Test 3 (3)
$\log(M_\star / M_\odot)$	$+6.24^{+0.13}_{-0.09}$	$+6.31^{+0.40}_{-0.15}$
t/Myr	$+3.90^{+0.40}_{-0.40}$	$+5.60^{+12.4}_{-1.6}$
$\log(Z_{\text{ISM}})$	$-3.99^{+0.01}_{-0.01}$	$-3.95^{+0.07}_{-0.03}$
$\log(U_S)$	$-2.01^{+0.06}_{-0.05}$	$-1.90^{+0.16}_{-0.09}$
ξ_d	$+0.47^{+0.03}_{-0.04}$	$+0.44^{+0.04}_{-0.09}$
$12+\log(\text{O}/\text{H})$	$+6.45^{+0.03}_{-0.02}$	$+6.52^{+0.08}_{-0.05}$
$A_{1500} \text{ (mag)}$	$+0.32^{+0.40}_{-0.22}$	$+0.28^{+0.59}_{-0.20}$
$(\text{C}/\text{O})/(\text{C}/\text{O})_\odot$	$+0.79^{+0.05}_{-0.05}$	$+0.81^{+0.07}_{-0.06}$

flux or EW (column 2), median-model flux or EW corresponding to test 2 (column 3), median-model flux or EW corresponding to test 3 (column 4), difference between columns 2 and 3 (column 5), and difference between columns 2 and 4 (column 6). From test 2 we conclude that the residuals of all observables with the exception of the C IV absorption EW, are within the observational uncertainty. From test 3 we conclude that excluding the He II flux from the fit does not significantly impact the results from test 2. In summary, we are unable to fit the C IV absorption EW within the observational uncertainties. A main difficulty with trying to measure the C IV EW is that it is weak and very broad. Thus, it is not clear where to put the wavelength limits for the numerical integration of this feature.

Table 6 gives the median values of the physical parameters corresponding to tests 2 and 3. We do not expect our masses to be accurate since we are not fitting the stellar continuum. For reference, for the clusters which are within the COS aperture, Adamo et al. (2010) obtained a total stellar mass of $4.5 \times 10^6 M_\odot$. For the time since the beginning of star formation we obtain 3.9 ± 0.4 Myr via test 2 and $5.60^{+12.4}_{-1.6}$ via test 3. For reference, Adamo et al. (2010) finds that the youngest cluster within our aperture is 6 ± 1 Myr. For the ionized-gas oxygen abundance, we find that it is unrealistically low, which is in agreement with the work presented in Kehrig et al. (2018), where the authors fitted the flux of the integrated He II 4686 emission-line of SBS 0335–052E. Finally, for $\log(\text{C}/\text{O})$, we find similar values via test 2 and 3, i.e. ~ -0.45 dex. This is higher than

Table 7. Observed versus median-model optical emission-line fluxes. Columns: (1) Line ID. (2) Logarithm of the observed flux ($\text{erg s}^{-1} \text{cm}^{-2}$) corrected for foreground reddening only. (3) Median-model flux corresponding to test 4. (4) Median-model flux corresponding to test 5. (5) Difference between columns 2 and 3. (6) Difference between columns 2 and 4.

Line ID (1)	Observed (2)	Test 4 (3)	Test 5 (4)	Δ_1 (5)	Δ_2 (6)
[Ar IV] λ 4740	-15.41 ± 0.02	-15.63	-15.54	0.22	0.13
H β λ 4861	-13.38 ± 0.01	-13.38	-13.38	0.00	0.00
[O III] λ 4959	-13.33 ± 0.01	-13.39	-13.35	0.06	0.02
[O III] λ 5007	-12.86 ± 0.01	-12.91	-12.88	0.05	0.02
He I λ 5876	-14.37 ± 0.01	-14.36	-14.36	-0.01	-0.01
[O I] λ 6300	-15.54 ± 0.01	-15.58	-	0.04	-
[S III] λ 6310	-15.56 ± 0.01	-15.40	-	-0.16	-
[N II] λ 6548	-15.93 ± 0.02	-16.22	-	0.29	-
H α λ 6563	-12.91 ± 0.01	-12.92	-12.91	0.01	0.00
[N II] λ 6584	-15.47 ± 0.01	-15.76	-	0.29	-
[S II] λ 6717	-15.08 ± 0.01	-15.12	-15.16	0.04	0.08
[S II] λ 6731	-15.18 ± 0.01	-15.22	-15.26	0.04	0.08
[Ar III] λ 7135	-15.09 ± 0.01	-15.14	-15.15	0.05	0.06
[S III] λ 9068	-14.84 ± 0.01	-14.70	-14.70	-0.14	-0.14

the value obtained by Garnett et al. (1995) via the standard method, which is $\log(\text{C}/\text{O}) = -0.94 \pm 0.17$.

5.4 Tests 4 and 5: fits to optical emission-line fluxes

In the optical, we exclude He II λ 4686 from all fits, as none of the models which we tried were able to reproduce the flux of this line. As test 4, we try to simultaneously reproduce all of the optical lines which are shown in the first column of Table 7. For the latter fit, columns (2) and (3) of Table 7 give the observed and median-model flux values. The residual between (2) and (3) is given in column (5). The latter column shows that neither the high-ionization emission lines of [Ar IV] λ 4740, [S III] λ 6310 and [S III] λ 9068 nor the [N II] λ 6548, 6584 lines are well reproduced. In fact, only the Balmer and He I lines are reproduced within the 2 per cent observational errors. As test 5 we try fitting only the fluxes of lines with fluxes stronger than or equal to that of the [S II] λ 6717 line. The median-model fluxes corresponding to the latter fit are shown in column (4) of Table 7. The residual between (2) and (4) is given in column (6). The result is very similar although the fits to [O III] λ 4959 and [O III] λ 5007 slightly improve.

The median physical properties derived from simultaneous fits to optical emission-line fluxes are summarized in Table 8. We do not expect the masses to be accurate since we are not fitting the stellar continuum. The ages are unrealistically young, which is required to produce hard ionizing photons and approach the observed fluxes of the high-ionization optical lines. The attenuation at 1500 \AA is smaller than that derived from the UV. We do not expect the C/O value to be accurate since there are no carbon lines in the optical.

6 DISCUSSION

6.1 EWs of high-ionization UV emission lines: nearby versus distant systems

Senchyna et al. (2017) present *HST* COS UV observations of 10 nearby star-forming galaxies with $12+\log(\text{O}/\text{H})$ between 7.81 and 8.48. They find that objects with $12+\log(\text{O}/\text{H})$ above 8.1 have: (i) P-Cygni like C IV λ 1548, 1551 profiles with broad absorptions forming in the wind of O-type MS stars and (ii) broad He II λ 1640

Table 8. Median physical properties derived from optical data. For each property, we give the one-sigma errors around the median. Columns: (1) Property. (2) Results from test 4. (3) Results from test 5.

Line ID (1)	Test 4 (2)	Test 5 (3)
$\log(M_*/M_\odot)$	$+5.93^{+0.02}_{-0.01}$	$+5.92^{+0.01}_{-0.01}$
t / Myr	$+0.13^{+0.11}_{-0.03}$	$+0.11^{+0.01}_{-0.00}$
$\log(Z_{\text{ISM}})$	$-3.06^{+0.01}_{-0.01}$	$-3.25^{+0.01}_{-0.01}$
$\log(U_S)$	$-2.23^{+0.01}_{-0.01}$	$-2.22^{+0.01}_{-0.01}$
ξ_d	$+0.27^{+0.02}_{-0.03}$	$+0.11^{+0.01}_{-0.00}$
$12+\log(\text{O}/\text{H})$	$+7.48^{+0.01}_{-0.01}$	$+7.42^{+0.01}_{-0.01}$
$A_{1500} \text{ (mag)}$	$+0.05^{+0.08}_{-0.04}$	$+0.01^{+0.02}_{-0.00}$
$(\text{C}/\text{O})/(\text{C}/\text{O})_\odot$	$+0.98^{+0.02}_{-0.02}$	$+0.13^{+0.06}_{-0.02}$

emission forming in the winds of WR stars. On the other hand, they find that objects with $12+\log(\text{O}/\text{H}) \leq 8.0$ have weak stellar C IV $\lambda\lambda$ 1548, 1551 and He II λ 1640 profiles and are dominated by nebular emission in [C III], C III] $\lambda\lambda$ 1907, 1909, C IV $\lambda\lambda$ 1548, 1551, and He II λ 1640. They find seven objects with [C III], C III] $\lambda\lambda$ 1907, 1909 emission, some of them reaching extremely high equivalent width values ($\sim 10 - 15 \text{ \AA}$). Finally, they find that C IV $\lambda\lambda$ 1548, 1551 and He II λ 1640 have equivalent widths below 2 \AA and conclude that the detection of C IV $\lambda\lambda$ 1548, 1551 at $>20 \text{ \AA}$ may require extremely metal poor stars and gas. In SBS 0335–052E, which is more metal-poor than the previous sample, C IV $\lambda\lambda$ 1548, 1551 shows a weak P-Cygni profile and nebular components. The EWs of the nebular [C III], C III] $\lambda\lambda$ 1907, 1909 and C IV doublet emissions are 5 and 1.7 \AA , respectively. The largest C IV $\lambda\lambda$ 1548, 1551 EW which has been reported for a nearby starburst galaxy is $\sim 10 \text{ \AA}$. This is for galaxy J104457, which has $12+\log(\text{O}/\text{H}) = 7.45$ (Berg et al. 2019b) and is at a very similar redshift (0.013) compared to SBS 0335–052E.

Senchyna & Stark (2019) present *HST* COS UV observations of six star-forming XMPs with $12+\log(\text{O}/\text{H})$ between 7.4 and 7.7, i.e. lower than in the work of Senchyna et al. (2017). They find stellar C IV $\lambda\lambda$ 1548, 1551 in five of the galaxies and high [C III]+C III] equivalent widths ($\sim 11 \text{ \AA}$) in two objects. The highest EWs of C IV $\lambda\lambda$ 1548, 1551 and He II which they find are 4.4 and 1.7 \AA , respectively. For SBS 0335–052E we find EWs of 1.7 \AA for C IV $\lambda\lambda$ 1548, 1551 and He II λ 1640. Although nebular C IV $\lambda\lambda$ 1548, 1551 emission is present in some very high sSFR nearby systems, nearby systems do not reach the equivalent widths of C IV which are observed in individual lensed systems at $z > 6$. Undisputed detections of nebular emission in He II in predominantly star-forming systems at $z > 6$ remain elusive (Senchyna & Stark 2019). Objects located at $z = 2-4$ with high values of EWs(C III)], i.e. $\geq 10 \text{ \AA}$, are analysed in Nakajima et al. (2018), who find that EWs(C III)] $> 20 \text{ \AA}$ can only be modelled with a combination of photoionization by stars and an AGN.

6.2 C/O

Berg et al. (2016) and Berg et al. (2019a) studied the C/O ratio derived via UV collisionally excited emission lines of a sample of 12 low metallicity galaxies with significant detections of carbon and oxygen lines. At low metallicity, ($12+\log(\text{O}/\text{H}) < 8.0$), they find that no clear trend is evident in C/O versus O/H. We find $\log(\text{C}/\text{O}) \sim -0.45$ via tests 2 and 3, and $12+\log(\text{O}/\text{H}) = 7.48 / 7.42$ via tests 4/5, respectively. This is very similar to what Berg et al. (2016) report for galaxy J120122, which is $\log(\text{C}/\text{O}) \sim -0.45$ and $12+\log(\text{O}/\text{H}) = 7.45$.

6.3 JWST observation

As a simple exercise, we placed SBS 0335–052E at redshift $z = 10$, and used the COS G130M + G160M + G185M spectrum to simulate an observation of this galaxy with the near-infrared spectrograph (NIRSpec) onboard *JWST*. For this purpose, the spectrum’s flux was diluted using the luminosity distance for cosmological parameters $H_0 = 69.6$, $\Omega_M = 0.286$, and $\Omega_{\text{vac}} = 0.714$. In addition, we accounted for IGM absorption of UV photons using Madau (1995) attenuation. We degraded the spectrum to match NIRSpec’s prism resolution, i.e. $\Delta\lambda = 15.5 \text{ \AA}$ or $R \sim 100$ at $1550 \times (1+z) \text{ \AA}$. We used the exposure time calculator of the NIRSpec Microshutter Array (MSA) prototype. The latter assumes a point source centred in an open MSA shutter of $0.2 \text{ arcsec} \times 0.45 \text{ arcsec}$, with background from two adjacent open shutters. Detecting the C IV $\lambda\lambda$ 1548, 1551 doublet with an SNR of 6 per resolution element would require an exposure time of 10^5 s , similar to what is adopted for ‘deep’ observations of the *JWST*/NIRSpec Guaranteed Time Observations (GTO) program (Chevallard et al. 2019).

7 SUMMARY AND CONCLUSIONS

We present new *HST*/COS UV spectroscopy and VLT/MUSE optical spectroscopy of one of the nearest most metal-poor starburst galaxies known, SBS 0335–052E. The main results from the analysis of these and archival data follow.

- (i) The C IV $\lambda\lambda$ 1548, 1551 doublet is composed of a P-Cygni like profile plus nebular components which are redshifted relative to the Al II λ 1670 line (Section 3.3).
- (ii) In spite of its high sSFR ($\log[s\text{SFR}/\text{yr}] = -8.13$), and extremely low metallicity, we do not detect equivalent widths in excess of 10 \AA for the high-ionization UV emission lines of He II, O III], [C III], C III], or C IV. Thus, the combination of the above two properties does not guarantee large equivalent widths of these UV lines (Section 3.3).
- (iii) We report the detection of blueshifted O V λ 1371 which could indicate the presence of VMSs in the galaxy (Section 3.1), as well as photospheric S V λ 1502 absorption at the 2σ level. In addition, we report the tentative detection of broad N III λ 4640 and C IV/[Fe III] λ 4658 emission lines from WR stars and no detection of C IV λ 5808 (Section 4.4) in the vicinity of SSC3. Higher spatial resolution UV observations of SBS 0335–052E would help to better constrain the populations of WR and potential VMSs in the observed region.
- (iv) It is unlikely that the ULX which is reported in Prestwich et al. (2013, our Section 4.1), an IMBH (Section 4.2), or an extremely low metallicity shock + precursor (Section 4.3) are the explanation for the high-ionization UV and optical emission lines of SBS 0335–052E. In addition, the SSP BPASS v2.1 binary models with $M_{\text{up}} = 300 M_\odot$ and nebular emission cannot reproduce the optical emission line ratios of SBS 0335–052E, while the similar single non-rotating models can.
- (v) We use different sets of UV and optical observables to test constant star formation models with single non-rotating stars (we note that BPASS v2.1 binary models with constant star formation and nebular emission are not publicly available). We use different sets of UV and optical observables to test constant star formation models with single non-rotating stars. The models include VMSs and account for the integrated light of stars, gas and dust. We could not combine the UV and optical observations for this purpose due to a flux mismatch between the UV and optical. Single non-rotating models which include VMSs are able to simultaneously reproduce

all UV emission lines but only if the metallicity is unrealistically low (Section 5.3). The same models are unable to reproduce the fluxes of the high-ionization optical emission lines. This is likely because in the models, the abundances of some of the corresponding elements might be wrong (Section 5.4).

(vi) We find that $12+\log(\text{O}/\text{H}) = 7.45 \pm 0.04$ and $\log(\text{C}/\text{O}) = -0.45^{+0.03}_{-0.04}$ (Section 6.2), which is similar to the values of these abundance ratios for galaxy J120122 (Berg et al. 2016).

(vii) Galaxies with the properties of SBS 0335–052E would be detected with *JWST*/NIRSpec deep observations.

ACKNOWLEDGEMENTS

We acknowledge the referee for useful comments that improved the quality of this paper. AW acknowledges the support of UNAM via grant agreement PAPIIT no. IA105018. AW, AVG, AF, JC, and SC acknowledge the support of the ERC via an Advanced Grant under grant agreement no. 321323-NEOGAL. AVG acknowledges the support of the ERC via an Advanced Grant under grant agreement no. 742719-MIST. AF acknowledges the support from grant PRIN MIUR2017–20173ML3WW_001. MH acknowledges the support of the Swedish Research Council, Vetenskapsrådet, and the Swedish National Space Agency (SNSA), and is a fellow of the Knut and Alice Wallenberg Foundation. Based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, under NASA Contract NAS5-26555. These observations are associated with Programs 10575, 11579, 13788, and 15193. Support for Program 13788 was provided by NASA through a grant from the Space Telescope Science Institute. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Based on observations collected at the European Southern Observatory under ESO programme 96.B-0690(A). We are extremely thankful to Christophe Morisset for his help with the shock + precursor models of Alarie & Morisset (2019); to J. J. Eldridge and Elizabeth Stanway for their help with the BPASS models.

DATA AVAILABILITY

The *HST* data underlying this article are available in the Mikulski Archive for Space Telescopes at <https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>, and can be accessed with the data set identifiers which are provided in column 1 of Table 1. The reduced MUSE datacube used here and in Herenz et al. (2017) is available online via the CDS at <http://vizier.unistra.fr/viz-bin/cat/J/A+A/606/L11>.

REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543
 Adamo A., Zackrisson E., Östlin G., Hayes M., 2010, *ApJ*, 725, 1620
 Alarie A., Morisset C., 2019, *Rev. Mex. Astron. Astrofis.*, 55, 377
 Allen M. G., Groves B. A., Dopita M. A., Sutherland R. S., Kewley L. J., 2008, *ApJS*, 178, 20
 Aloisi A. et al., 2007, *ApJ*, 667, L151
 Annibali F. et al., 2013, *AJ*, 146, 144
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
 Bacon R. et al., 2010, in McLean I. S., Ramsa S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773508
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
 Bastian N., 2016, in Moraux E., Lebreton Y., Charbonnel C., eds, EAS Publications Series, Vol. 80, EAS Publications Series, p. 5
 Battaner E., Florido E., 2000, *Fundam. Cosm. Phys.*, 21, 1
 Berg D. A., Chisholm J., Erb D. K., Pogge R., Henry A., Olivier G. M., 2019b, *ApJ*, 878, L3
 Berg D. A., Erb D. K., Henry R. B. C., Skillman E. D., McQuinn K. B. W., 2019a, *ApJ*, 874, 93
 Berg D. A., Skillman E. D., Henry R. B. C., Erb D. K., Carigi L., 2016, *ApJ*, 827, 126
 Bohlin R. C., Colina L., Finley D. S., 1995, *AJ*, 110, 1316
 Bouret J.-C., Lanz T., Hillier D. J., Martins F., Marcolino W. L. F., Depagne E., 2015, *MNRAS*, 449, 1545
 Bouret J. C., Lanz T., Martins F., Marcolino W. L. F., Hillier D. J., Depagne E., Hubeny I., 2013, *A&A*, 555, A1
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Byler N., Kewley L. J., Rigby J. R., Acharyya A., Berg D. A., Bayliss M., Sharon K., 2020, *ApJ*, 893, 1
 Caffau E., Ludwig H.-G., Steffen M., Freytag B., Bonifacio P., 2011, *Sol. Phys.*, 268, 255
 Camacho I., García M., Simón-Díaz S., 2016, *A&A*, 585, A82
 Cardamone C. et al., 2009, *MNRAS*, 399, 1191
 Chabrier G., 2003, *PASP*, 115, 763
 Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
 Charlot S., Longhetti M., 2001, *MNRAS*, 323, 887
 Chevallard J., Charlot S., 2016, *MNRAS*, 462, 1415
 Chevallard J. et al., 2019, *MNRAS*, 483, 2621
 Chini R., Hoffmeister V. H., Nasserí A., Stahl O., Zinnecker H., 2012, *MNRAS*, 424, 1925
 Cormier D. et al., 2017, *MNRAS*, 468, L87
 Crowther P. A., 2007, *ARA&A*, 45, 177
 Crowther P. A. et al., 2016, *MNRAS*, 458, 624
 Dale D. A., Helou G., Neugebauer G., Soifer B. T., Frayer D. T., Condon J. J., 2001, *AJ*, 122, 1736
 de Mello D. F., Leitherer C., Heckman T. M., 2000, *ApJ*, 530, 251
 Donley J. L. et al., 2012, *ApJ*, 748, 142
 Douna V. M., Pellizza L. J., Mirabel I. F., Pedrosa S. E., 2015, *A&A*, 579, A44
 Dufour R. J., Skillman E. D., Garnett D. R., Shields G. A., Peimbert M., Torres-Peimbert S., Terlevich E., Terlevich R., 1993, *Rev. Mex. Astron. Astrofis.*, 27, 115
 Dunlop J. S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, *Astrophysics and Space Science Library*, Vol. 396, The First Galaxies, p. 223
 Ekta B., Chengalur J. N., 2010, *MNRAS*, 406, 1238
 Ekta B., Pustilnik S. A., Chengalur J. N., 2009, *MNRAS*, 397, 963
 Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *Publ. Astron. Soc. Aust.*, 34, e058
 Evans C. J. et al., 2019, *AAP*, 622, 129
 Feltre A., Charlot S., Gutkin J., 2016, *MNRAS*, 456, 3354
 Ferland G. J., 1981, *ApJ*, 249, 17
 Ferland G. J. et al., 2013, *Rev. Mex. Astron. Astrofis.*, 49, 137
 Filho M. E., Sánchez Almeida J., Muñoz-Tuñón C., Nuza S. E., Kitaura F., Heß S., 2015, *ApJ*, 802, 82
 Finkelstein S. L., Papovich C., Giavalisco M., Reddy N. A., Ferguson H. C., Koekemoer A. M., Dickinson M., 2010, *ApJ*, 719, 1250
 García M., 2018, *MNRAS*, 474, L66
 García M., Herrero A., Najarro F., Camacho I., Lorenzo M., 2019, *MNRAS*, 484, 422
 García M., Herrero A., Najarro F., Lennon D. J., Alejandro Urbaneja M., 2014, *ApJ*, 788, 64
 García-Rojas J., Simón-Díaz S., Esteban C., 2014, *A&A*, 571, A93
 Garnett D. R., Kennicutt R. C., Jr, Chu Y.-H., Skillman E. D., 1991, *ApJ*, 373, 458
 Garnett D. R., Skillman E. D., Dufour R. J., Peimbert M., Torres-Peimbert S., Terlevich R., Terlevich E., Shields G. A., 1995, *ApJ*, 443, 64
 Giovanelli R. et al., 2005, *AJ*, 130, 2598
 Grimes J. P. et al., 2009, *ApJS*, 181, 272
 Gräfenr G., Vink J. S., 2015, *A&A*, 578, L2

- Gutkin J., Charlot S., Bruzual G., 2016, *MNRAS*, 462, 1757
- Henry A., Scarlata C., Martin C. L., Erb D., 2015, *ApJ*, 809, 19
- Herenz E. C., Hayes M., Papaderos P., Cannon J. M., Bik A., Melinder J., Östlin G., 2017, *A&A*, 606, L11
- Herenz E. C., Wisotzki L., 2017, *A&A*, 602, A111
- Herenz E. C. et al., 2019, *A&A*, 621, A107
- Hernandez S., Aloisi A., James B. L., Ferland G. J., Fox A. J., Tosi M., Tumlinson J., 2020, *ApJ*, 892, 19
- Hirschauer A. S. et al., 2016, *ApJ*, 822, 108
- Houck J. R. et al., 2004, *ApJS*, 154, 211
- Hsyu T., Cooke R. J., Prochaska J. X., Bolte M., 2018, *ApJ*, 863, 134
- Hu E. M., Cowie L. L., Barger A. J., Capak P., Kakazu Y., Trouille L., 2010, *ApJ*, 725, 394
- Hunt L. K. et al., 2014, *A&A*, 561, A49
- Impey C., Bothun G., 1997, *ARA&A*, 35, 267
- Izotov I. I., Guseva N. G., Lipovetskii V. A., Kniazev A. I., Stepanian J. A., 1990, *Nature*, 343, 238
- Izotov Y. I., Chaffee F. H., Foltz C. B., Green R. F., Guseva N. G., Thuan T. X., 1999, *ApJ*, 527, 757
- Izotov Y. I., Guseva N. G., Fricke K. J., Papaderos P., 2009, *A&A*, 503, 61
- Izotov Y. I., Guseva N. G., Thuan T. X., 2011, *ApJ*, 728, 161
- Izotov Y. I., Lipovetsky V. A., Chaffee F. H., Foltz C. B., Guseva N. G., Kniazev A. Y., 1997, *ApJ*, 476, 698
- Izotov Y. I., Schaerer D., Blecha A., Royer F., Guseva N. G., North P., 2006, *A&A*, 459, 71
- Izotov Y. I., Schaerer D., Thuan T. X., Worseck G., Guseva N. G., Orlitová I., Verhamme A., 2016, *MNRAS*, 461, 3683
- James B. L., Aloisi A., Heckman T., Sohn S. T., Wolfe M. A., 2014, *ApJ*, 795, 109
- James B. L., Kaposov S., Stark D. P., Belokurov V., Pettini M., Olszewski E. W., 2015, *MNRAS*, 448, 2687
- Jaskot A. E., Oey M. S., 2014, *ApJ*, 791, L19
- Johnson K. E., Hunt L. K., Reines A. E., 2009, *AJ*, 137, 3788
- Kashikawa N. et al., 2011, *ApJ*, 734, 119
- Kauffmann G. et al., 2003, *MNRAS*, 346, 1055
- Kehrig C., Vilchez J. M., Pérez-Montero E., Iglesias-Páramo J., Brinchmann J., Kunth D., Durret F., Bayo F. M., 2015, *ApJ*, 801, L28
- Kehrig C., Vilchez J. M., Guerrero M. A., Iglesias-Páramo J., Hunt L. K., Duarte-Puertás S., Ramos-Larios G., 2018, *MNRAS*, 480, 1081
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, 556, 121
- Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, *MNRAS*, 372, 961
- Kojima T. et al., 2020, *ApJ*, 898, 142
- Kunth D., Leitherer C., Mas-Hesse J. M., Östlin G., Petrosian A., 2003, *ApJ*, 597, 263
- Kunth D., Östlin G., 2000, *A&AR*, 10, 1
- Lacy M. et al., 2004, *ApJS*, 154, 166
- Lebouteiller V., Heap S., Hubeny I., Kunth D., 2013, *A&A*, 553, A16
- Leitherer C., Ortiz Otálvaro P. A., Bresolin F., Kudritzki R.-P., Lo Faro B., Pauldrach A. W. A., Pettini M., Rix S. A., 2010, *ApJS*, 189, 309
- Leitherer C., Tremonti C. A., Heckman T. M., Calzetti D., 2011, *AJ*, 141, 37
- Madau P., 1995, *ApJ*, 441, 18
- Madden S. C. et al., 2013, *PASP*, 125, 600
- Malhotra S., Rhoads J. E., 2002, *ApJ*, 565, L71
- Malhotra S., Rhoads J. E., 2004, *ApJ*, 617, L5
- Malhotra S., Rhoads J. E., Finkelstein S. L., Hathi N., Nilsson K., McLinden E., Pirzkal N., 2012, *ApJ*, 750, L36
- Massey P., Zangari A. M., Morrell N. I., Puls J., DeGioia-Eastwood K., Bresolin F., Kudritzki R.-P., 2009, *ApJ*, 692, 618
- Mathis J. S., 1990, *ARA&A*, 28, 37
- Mezcua M., Civano F., Fabbiano G., Miyaji T., Marchesi S., 2016, *ApJ*, 817, 20
- Moiseev A. V., Pustilnik S. A., Kniazev A. Y., 2010, *MNRAS*, 405, 2453
- Morales-Luis A. B., Sánchez Almeida J., Aguerri J. A. L., Muñoz-Tuñón C., 2011, *ApJ*, 743, 77
- Nakajima K. et al., 2018, *A&A*, 612, A94
- Östlin G., Hayes M., Kunth D., Mas-Hesse J. M., Leitherer C., Petrosian A., Atek H., 2009, *AJ*, 138, 923
- Ouchi M. et al., 2010, *ApJ*, 723, 869
- Pakull M. W., Angebault L. P., 1986, *Nature*, 322, 511
- Papaderos P., Izotov Y. I., Fricke K. J., Thuan T. X., Guseva N. G., 1998, *A&A*, 338, 43
- Papaderos P., Izotov Y. I., Guseva N. G., Thuan T. X., Fricke K. J., 2006, *A&A*, 454, 119
- Peebles P. J. E., Phelps S. D., Shaya E. J., Tully R. B., 2001, *ApJ*, 554, 104
- Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, *ARA&A*, 48, 431
- Prestwich A. H., Tsantaki M., Zezas A., Jackson F., Roberts T. P., Foltz R., Linden T., Kalogera V., 2013, *ApJ*, 769, 92
- Pustilnik S. A., Brinks E., Thuan T. X., Lipovetsky V. A., Izotov Y. I., 2001, *AJ*, 121, 1413
- Pustilnik S. A., Perepelitsyna Y. A., Kniazev A. Y., 2016, *MNRAS*, 463, 670
- Rafelski M., Neeleman M., Fumagalli M., Wolfe A. M., Prochaska J. X., 2014, *ApJ*, 782, L29
- Rasmussen A., Martin C., 1992, *ApJ*, 396, L103
- Reines A. E., Johnson K. E., Hunt L. K., 2008, *AJ*, 136, 1415
- Rivera-Thorsen T. E. et al., 2015, *ApJ*, 805, 14
- Rémy-Ruyer A. et al., 2015, *A&A*, 582, A121
- Sana H. et al., 2012, *Science*, 337, 444
- Sana H. et al., 2013, *A&A*, 550, A107
- Schaerer D., Vacca W. D., 1998, *ApJ*, 497, 618
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Senchyna P., Stark D. P., 2019, *MNRAS*, 484, 1270
- Senchyna P. et al., 2017, *MNRAS*, 472, 2608
- Skillman E. D., Kennicutt R. C., Jr, 1993, *ApJ*, 411, 655
- Smith L. J., Crowther P. A., Calzetti D., Sidoli F., 2016, *ApJ*, 823, 38
- Stanway E. R., Eldridge J. J., 2019, *A&A*, 621, A105
- Stark D., 2012, *Nature*, 489, 370
- Sánchez Almeida J., Pérez-Montero E., Morales-Luis A. B., Muñoz-Tuñón C., García-Benito R., Nuza S. E., Kitaura F. S., 2016, *ApJ*, 819, 110
- Tang J., Bressan A., Rosenfield P., Slemer A., Marigo P., Girardi L., Bianchi L., 2014, *MNRAS*, 445, 4287
- Thompson R. I., Sauvage M., Kennicutt R. C., Engelbracht C., Vanzi L., Schneider G., 2009, *ApJ*, 691, 1068
- Thuan T. X., Bauer F. E., Papaderos P., Izotov Y. I., 2004, *ApJ*, 606, 213
- Thuan T. X., Izotov Y. I., Lipovetsky V. A., 1997, *ApJ*, 477, 661
- Veilleux S., Osterbrock D. E., 1987, *ApJS*, 63, 295
- Verhamme A., Dubois Y., Blaizot J., Garel T., Bacon R., Devriendt J., Guiderdoni B., Slyz A., 2012, *A&A*, 546, A111
- Vidal-García A., Charlot S., Bruzual G., Hubeny I., 2017, *MNRAS*, 470, 3532
- Vink J. S., 2012, in Roming P., Kawai N., Pian E., eds, Proc. IAU Symp. Vol. 7, Death of Massive Stars: Supernovae and Gamma-Ray Bursts, Kluwer, Dordrecht, p. 29
- Walborn N. R., Nichols-Bohlin J., Panek R. J., 1985, NASA Reference Publication, 1155
- Walborn N. R., Parker J. W., Nichols J. S., 1995, VizieR Online Data Catalog, III/188
- Weilbacher P. M., Streicher O., Urrutia T., Pécontal-Rousset A., Jarno A., Bacon R., 2014, in Manset N., Forshay P., eds, ASP Conf. Ser. Vol. 485, Astronomical Data Analysis Software and Systems XXIII. Astron. Soc. Pac., San Francisco, p. 451
- Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, *ApJ*, 745, 50
- Wofford A., Leitherer C., Chandar R., 2011, *ApJ*, 727, 100
- Wofford A., Leitherer C., Chandar R., Bouret J.-C., 2014, *ApJ*, 781, 122
- Wu Y., Charmandaris V., Houck J. R., Bernard-Salas J., Lebouteiller V., Brandl B. R., Farrah D., 2008, *ApJ*, 676, 970
- Xiao L., Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 477, 904
- Yang H., Malhotra S., Gronke M., Rhoads J. E., Dijkstra M., Jaskot A., Zheng Z., Wang J., 2016, *ApJ*, 820, 130
- Yang H. et al., 2017, *ApJ*, 844, 171
- Zheng Z.-Y., Wang J.-X., Malhotra S., Rhoads J. E., Finkelstein S. L., Finkelstein K., 2014, *MNRAS*, 439, 1101