

Conditions for detecting lensed Population III galaxies in blind surveys with the *James Webb Space Telescope*, the *Roman Space Telescope*, and *Euclid*

Anton Vikaeus¹,¹★ Erik Zackrisson¹,¹ Daniel Schaerer^{2,3}, Eli Visbal⁴, Emma Fransson¹,
Sangeeta Malhotra⁵, James Rhoads⁵ and Martin Sahlén^{6,7}

¹Observational Astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden

²Geneva Observatory, University of Geneva, Chemin Pegasi 51, CH-1290 Versoix, Switzerland

³CNRS, IRAP, 14 Avenue E. Belin, F-31400 Toulouse, France

⁴University of Toledo, Department of Physics and Astronomy and Ritter Astrophysical Research Center, 2801 Bancroft, Toledo, OH 43606, USA

⁵Astrophysics Science Division, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA

⁶Swedish Collegium for Advanced Study, Thunbergsvägen 2, SE-752 38 Uppsala, Sweden

⁷Theoretical Astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden

Accepted 2022 February 17. Received 2022 February 17; in original form 2021 July 8

ABSTRACT

Dark matter haloes that reach the H I-cooling mass without prior star formation or external metal pollution represent potential sites for the formation of small – extremely faint – Population III galaxies at high redshifts. Gravitational lensing may in rare cases boost their fluxes to detectable levels, but to find even a small number of such objects in randomly selected regions of the sky requires very large areas to be surveyed. Because of this, a small, wide-field telescope can in principle offer better detection prospects than a large telescope with a smaller field of view. Here, we derive the minimum comoving number density required to allow gravitational lensing to lift such objects at redshift $z = 5–16$ above the detection thresholds of blind surveys carried out with the *James Webb space telescope* (*JWST*), the *Roman space telescope* (*RST*) and *Euclid*. We find that the prospects for photometric detections of Pop III galaxies are promising, and that they are better for *RST* than for *JWST* and *Euclid*. However, the Pop III galaxies favoured by current simulations have number densities too low to allow spectroscopic detections based on the strength of the He II 1640 emission line in any of the considered surveys unless very high star formation efficiencies ($\epsilon \gtrsim 0.1$) are evoked. We argue that targeting individual cluster lenses instead of the wide-field surveys considered in this paper results in better spectroscopic detection prospects, while for photometric detection, the wide-field surveys perform considerably better.

Key words: gravitational lensing: strong – techniques: photometric – techniques: spectroscopic – stars: Population III – Dark ages, reionization, first stars.

1 INTRODUCTION

The very first stars – the so-called Population III (hereafter Pop III) – formed from pristine gas containing only the elements produced during big bang nucleosynthesis (H, He, and trace amounts of Li), and are generically expected to be very massive, with typical masses in the $\sim 10–1000 M_{\odot}$ range (e.g. Hosokawa et al. 2016; Hirano et al. 2018). At the end of their lifetimes, these objects are expected to either enrich the surrounding interstellar medium by means of supernovae or mass lost on the asymptotic giant branch (e.g. Heger & Woosley 2002; Rydberg et al. 2020), or directly collapse into black holes (Latif, Khochfar & Whalen 2020; Whalen et al. 2020). In the case of very low mass (Magg et al. 2018) or low formation redshift (Liu & Bromm 2020), Pop III stars could in principle also survive to the present era.

Detecting Pop III stars remain one of the milestone observations needed to piece together the puzzle of early structure formation, but individual Pop III stars are expected to be too faint, by a large margin, to be directly observed with present or nascent facilities, unless subject to extreme gravitational lensing (Rydberg et al. 2013; Windhorst et al. 2018). The exception would be the rare case of a supermassive ($M \sim 10^5 M_{\odot}$) Population III star (Haemmerlé et al. 2018). These gargantuan stars, considered already in the 1960s as direct sources of the, by then, newly discovered quasi-stellar objects (although argued as dynamically unstable by e.g. Chandrasekhar 1964) now also offer a venue of direct detection of Pop III stars, their supernovae or their resulting direct collapse black holes with only mild magnifications or even without the help of gravitational lensing (Surace et al. 2018, 2019; Whalen et al. 2020).

In this paper, we pursue the alternative avenue of detecting the integrated light from many Pop III stars formed simultaneously within the same Pop III galaxy¹ (e.g. Johnson, Greif & Bromm

* E-mail: anton.vikaeus@physics.uu.se

¹ Sometimes alternatively referred to as Pop III star clusters.

2008; Johnson et al. 2009; Stiavelli & Trenti 2010; Visbal, Bryan & Haiman 2017; Yajima & Khochfar 2017; Inayoshi, Li & Haiman 2018; Johnson & Aykutalp 2019; Kulkarni, Visbal & Bryan 2019), i.e. the summed light from several hundreds to thousands of Pop III stars formed within a time window of just a few million years inside a single, chemically pristine dark matter halo. The detection of such objects could potentially grant us insights into the stellar initial mass function (IMF) of Pop III stars, the number densities of atomic cooling dark matter haloes that host Pop III galaxies, the Pop III star formation efficiency (ϵ), large scale metal-mixing, etc.

The observational signatures typically associated with Pop III galaxies are related to the fact that a top-heavy IMF produces a very hard ionizing flux because of the high effective temperatures of Pop III stars. This outgoing flux will ionize and excite the surrounding interstellar medium, producing strong hydrogen and helium emission lines such as Ly α , He II λ 1640, H α , and H β without corresponding metal lines – indicative of a completely pristine top-heavy stellar population (Schaerer 2002, 2003; Inoue 2011; Zackrisson et al. 2011b).

Summing up the light from several hundreds to thousands of Pop III stars surely improves our prospects for detecting some of the very first stellar structures in the Universe. Unfortunately this, in itself, will generally not suffice for the Pop III galaxies seen in current simulations, unless we at the same time may take advantage of large magnifications due to gravitational lensing from foreground objects (Zackrisson et al. 2012, 2015).

In this paper, we will investigate the prospects for upcoming telescopes to detect Pop III galaxies with the help of gravitational lensing in blind surveys targeting redshifts in the range $z = 5\text{--}16$, i.e. from a cosmic epoch prior to the predicted peak of the Pop III star formation rate density at $z \sim 10$ (e.g. Liu & Bromm 2020) to just after the end of cosmic reionization.

Assessing the detectability of gravitationally lensed Pop III galaxies requires us to determine the probabilities of achieving a certain gravitational magnification, where the minimum required magnification depends on the intrinsic brightness of the high-redshift Pop III galaxies that we are searching for. With this in hand, we can calculate the number densities² (number of Pop III galaxies per comoving Mpc³) required to allow the detection of some minimum number of highly magnified Pop III galaxies in the survey area probed by future space telescope missions. These number densities can then be compared to those predicted in contemporary simulations of Pop III galaxies.

Throughout this paper, we focus exclusively on ‘pure’ Pop III galaxies, i.e. objects without any metal-enriched stars. Simulations indicate that Pop III stars could potentially also form inside pockets of pristine gas within an otherwise metal-enriched galaxy (e.g. Sarmiento, Scannapieco & Cohen 2018; Sarmiento, Scannapieco & Côté 2019), thereby giving rise to a spectrum containing a mixture of both Pop III and Pop II/I spectral signatures, but the detectability of such ‘hybrid’ Pop III galaxies is beyond the scope of this paper.

Throughout the paper, we use a flat Lambda cold dark matter cosmological model with the following parameters throughout the paper; $H_0 = 67.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.685$, $\Omega_m = 0.315$, and $\Omega_b = 0.0487$ (Planck Collaboration XVI 2014).

The paper is organized as follows; in Section 2, we describe the model framework in which we perform our calculations, and the

assumptions involved. In Section 3, we present our findings and in Section 4 discuss their implications. Our conclusions are summarized in Section 5.

2 METHOD AND MODELS

The vast majority of Pop III galaxies will, due to their low intrinsic luminosities and large distances from us, remain hidden below the detection thresholds of even the largest telescopes currently planned (Schaerer et al. 2020). For every Pop III galaxy, there is, however, a small but non-zero probability for strong gravitational lensing from matter overdensities (galaxies or galaxy clusters) along the line of sight, which could lift the object into the detectable regime. This probability is a function of the magnification and the redshift of the object, although the latter dependence is relatively weak throughout the high-redshift regime considered here (see fit in equation 1). Hence, a given survey volume must contain a minimum number of Pop III galaxies in order to ensure a reasonable probability that at least a few are sufficiently magnified to end up above the detection threshold of that survey. This may be converted into a lower limit on the comoving number density of Pop III galaxies (with a given intrinsic luminosity) at the survey redshift that Pop III galaxies must meet to ensure reasonable detection prospects. This limit may then be compared to the properties (comoving number density and luminosity) of Pop III galaxies predicted by numerical simulations to assess whether gravitational lensing has any chance of rendering them observable in upcoming surveys of the high-redshift Universe. In this section, we present the computational machinery for deriving this limit. The focus here is on large-area surveys (many square degrees) that would inevitably cover both galaxy clusters and void regions, or alternatively small-area surveys (a few square arcminutes) that cover random fields. Small-area surveys that focus exclusively on strong-lensing galaxy cluster fields are outside the scope of this paper (but see Zackrisson et al. 2012, for a study of this type on the detectability of Pop III galaxies).

2.1 Spectral synthesis modelling

The data for the Pop III galaxy emission-line strengths and luminosities used for our predictions are taken from the model by Raiter, Schaerer & Fosbury (2010). In this paper we consider a fiducial top-heavy lognormal stellar IMF of characteristic mass $m_c = 60 M_\odot$ and width $\sigma = 1$, in the mass range $1\text{--}500 M_\odot$ that is forming stars at a constant rate (as in the Tumlinson 2006 lognormal case E IMF). Pop III galaxies are likely to retain their tell-tale spectral signatures only during a short time after the initial starburst, before supernova feedback quenches further star formation and the surrounding nebula is polluted by metals, we therefore consider a maximum lifetime $t_{\text{PopIII}} = 10 \text{ Myr}$ as our fiducial Pop III galaxy ‘lifetime’ and is therefore used as the star-forming age (τ) at which we extract the spectrum.

2.2 Spectral signatures

In this paper, we primarily consider two different criteria for the detection of Pop III galaxies in blind surveys – detection of their rest-frame UV (1500-Å) continuum flux using photometry, or detection of their He II λ 1640 line emission using spectroscopy.

The 1500-Å continuum detection criterion is the less demanding of the two, and simply means that the Pop III galaxy has been rendered sufficiently bright by gravitational lensing to end up above the detection threshold of a photometric survey (typically a multiband

²Throughout this paper, we will refer to the number of Pop III galaxies per comoving cubic megaparsec simply as ‘number density’ unless stated otherwise.

survey from which dropouts can be used to assess redshift) and would hence be expected to appear in its source catalogue. However, meeting this criterion does not necessarily imply that the Pop III galaxy would be identifiable among the other types of far more numerous sources detectable in the same survey. This 1500-Å flux limit none the less serves as an indicator of whether the galaxies could in principle be targeted for followed-up spectroscopy (either with *JWST* or a ground-based large telescope). Identifying a Pop III galaxy based on photometry alone may not be impossible, although this would require special conditions to be met. This includes colour signatures due to strong Balmer line emission yet missing metal lines (requires rest-frame optical data, hence *JWST* photometry; Inoue 2011; Zackrisson et al. 2011b), colour signatures due to very strong Ly α emission (requires high Ly α emission through the intergalactic medium, i.e. that the Pop III galaxy is sitting in an ionized bubble or is at $z \lesssim 6$; Pello & Schaerer 2003; Zackrisson et al. 2011a), colour signatures due to the Lyman bump (requires a redshift $z < 6$ to avoid significant absorption by the intergalactic medium; Inoue 2010) and colour signatures due to a very blue UV continuum slope (requires extreme leakage of ionizing photons to reveal the otherwise subdominant stellar continuum; Raiter et al. 2010; Zackrisson et al. 2011b). We stress that for some of these signatures, it seems unrealistic to assume that the required physical conditions would be met by more than a small fraction of the Pop III galaxies in a given survey volume, which could make the compounded probability of having a galaxy with just the right colours signature for identification and a fortunate sightline for extreme magnification, very low.

The He II $\lambda 1640$ emission line originates from gas regions surrounding sources of highly energetic radiation, due to the high ionization potential of helium. Owing to the expected top-heavy IMF and high temperatures of Pop III stars, one suspects that a strong He II $\lambda 1640$ emission line (i.e. a high equivalent width; bright line with respect to the surrounding continuum) should be indicative of a Pop III galaxy in cases when no other metal lines can be observed (e.g. Schaerer 2002, 2003; Raiter et al. 2010). For our fiducial Pop III galaxy IMF and star-forming age ($\tau \sim 10$ Myr, see Section 2.1), the predicted He II $\lambda 1640$ emission-line luminosity is given by $\sim 6 \times 10^{41}$ erg s $^{-1}$, normalized to SFR = 1 M $_{\odot}$ /yr. At $z \sim 10$ such a Pop III galaxy has an emission-line flux of $\sim 4.5 \times 10^{-22}$ erg s $^{-1}$ cm $^{-2}$ – for a total stellar mass $M_{*} \sim 10^4$ M $_{\odot}$. The equivalent width is found to be about 24.5 Å, which is considerably stronger than that seen in other metal-poor galaxies – the next to lowest metallicity used in Raiter et al. (2010) (i.e. $Z = 5 \times 10^{-6} Z_{\odot}$) dramatically reduces the equivalent width to a mere 2.8 Å. Observations of metal-poor galaxies have typically returned equivalent widths in the range 1–3 Å (Nanayakkara et al. 2019; Feltre et al. 2020; Saxena et al. 2020). In Saxena et al. (2020), at $z = 2.5$ –5, they find equivalent widths reaching at most 8 Å when removing galaxies harbouring a suspected active galactic nuclei (AGN) from the observed sample (one of the seven suspected AGN included in the data set had an equivalent width of 21.4 Å while the other six showed equivalent widths < 7.5 Å). On the other hand, in Nanayakkara et al. (2019), at $z = 2$ –4, one finds a few objects with equivalent widths reaching the levels of the Pop III stellar populations in our fiducial models. That is, one may still find metal-enriched galaxies that are strong He II $\lambda 1640$ emitters with equivalent widths reaching that expected of a pristine Pop III galaxy with a top-heavy IMF. However, taking into account other sources of helium ionizing radiation, such as AGN activity, and removing them from the sample, observations typically return equivalent widths that are significantly lower than for a Pop III galaxy. Hence, one expects this emission line to be a very powerful diagnostic of potential Pop III galaxy candidates.

One should note, however, that there are still considerable uncertainties attached to the expected 1500-Å continuum fluxes, He II $\lambda 1640$ emission-line fluxes and He II $\lambda 1640$ equivalent widths for Pop III galaxies, even in the case where the Pop III IMF is assumed to be known. As discussed by Raiter et al. (2010), the continuum fluxes at 1500 and 1640 Å (the latter of which affects the He II $\lambda 1640$ equivalent width) are not dominated by direct star light, but by nebular two-photon continuum, which is sensitive to the assumed gas density (with high density leading to a reduction in flux), whereas the He II $\lambda 1640$ line flux is sensitive to the assumed ionization parameter of the gas (with low-ionization parameter leading to a reduction in He II line flux due to absorption of He $^{+}$ -ionizing photons by H atoms, unlike what is expected from simplified recombination models). The Raiter et al. (2010) models are, moreover, based on non-rotating Pop III stars, and as shown in simulations by Stacy, Bromm & Loeb (2011), Stacy et al. (2013), one typically expects Pop III stars to be rapid rotators, unless significantly affected by magnetic breaking (Hirano & Bromm 2018). Yoon, Dierks & Langer (2012) argue that rotation may significantly boost the He $^{+}$ -ionizing fluxes of Pop III stars by factors of several, which could lead to a corresponding increase in He II $\lambda 1640$ line emission. At low total stellar masses ($\lesssim 10^4$ M $_{\odot}$), the He II $\lambda 1640$ line flux may furthermore be significantly affected by IMF sampling effects (Mas-Ribas, Dijkstra & Forero-Romero 2016), which increases the variance of this quantity throughout the Pop III galaxy population. As argued by Vikaesus, Zackrisson & Binggeli (2020), even slightly metal-enriched galaxies with star formation rates (SFR) $\lesssim 1$ M $_{\odot}$ yr $^{-1}$ may show elevated He II $\lambda 1640$ Å equivalent widths due to IMF sampling issues, to the point that a low-metallicity galaxy without Pop III stars may come across as a Pop III hybrid galaxy (although not likely a pure Pop III galaxy of the type we discuss here). In summary, both the 1500-Å continuum flux and the He II $\lambda 1640$ emission line are likely uncertain by a factor of a few, even for a fixed functional form of the Pop III stellar IMF. While our fiducial models are based on Raiter et al. (2010) predictions, in Section 3, we also discuss how of factor of ≈ 3 departures from these predictions would affect our conclusions.

2.3 Magnification distribution

The magnification data were taken from Zackrisson et al. (2015), where magnification (μ) probability functions ($P(> \mu)$) were calculated using a ray-tracing algorithm on a sequence of lens planes populated with both dark matter from N -body simulations, and galaxies based on a semi-analytical model. Robertson et al. (2020) presented a comparison of high-magnification ($\mu > 10$) probabilities based on similar models from different groups, and at the redshift limit of their comparison ($z \approx 5$), our results agree well (differing only on the order of a few per cent) with those of Hilbert et al. (2008) and Tinker et al. (2008). At lower redshifts ($z \gtrsim 1$), our model agrees well with Robertson et al. (2020) and Hilbert et al. (2008) – at most reaching relative variations ~ 20 per cent, while Tinker et al. (2008) consistently have probabilities ~ 50 per cent higher than our model at these low redshifts. In this paper, we consider Pop III galaxy redshifts in the range $z = 5$ –16, hence the low-redshift discrepancies are not directly relevant for the stringency of our results.

We find that, for magnifications $\mu > 10$ and redshifts in the range $z = 3$ –16, the functional form of our magnification data can be approximated by a power law in μ with a weak polynomial dependence on redshift (in the range considered):

$$P(> \mu) \approx (-4.2 \times 10^{-5} z^2 + 1.3 \times 10^{-3} z + 1.5 \times 10^{-4}) \mu^{-2}. \quad (1)$$

Table 1. The relevant survey parameters for the three telescopes considered (5σ point source detection limits). The *JWST* numbers for NIRCcam and the deep NIRSpec (0.013 deg² survey with ~ 28 h exposures) are quoted from planned surveys (Rieke et al. 2019), while the NIRSpec values for a single field of view ~ 100 hr NIRSpec exposure (very deep) are estimated using the exposure time calculator. We include two different setups with *RST* (Akeson et al. 2019) and *Euclid* (Laureijs et al. 2011; Marchetti et al. 2017) in order to highlight the utility of using either wide surveys or deep surveys.

Telescope	Survey area (deg ²)	Photo. depth (AB mag)	Spec. depth (erg s ⁻¹ cm ⁻²)
Deep NIRCcam/ NIRSpec	0.013	30.6	2.9×10^{-19}
Med. NIRCcam	0.053	29.7	–
Very deep NIRSpec	0.0034	–	1.3×10^{-19}
Deep <i>RST</i> (photo.)	40	29	–
Deep <i>RST</i> (spec.)	12	–	1×10^{-17}
Wide <i>RST</i>	2000	27	1×10^{-16}
Deep <i>Euclid</i>	40	26	5×10^{-17}
Wide <i>Euclid</i>	15 000	24	3×10^{-16}

As argued in Zackrisson et al. (2015), Peacock (1982) we cannot, however, apply the magnification probability function to arbitrary magnifications (μ) for extended objects such as star clusters or galaxies. As we consider larger magnifications, the size of the region in the lens plane responsible for the magnification maps to an increasingly small area in the source plane. For $\mu > 1000$ and a source plane redshift of $z = 10$, we begin to encounter differential magnification on scales $\lesssim 10$ pc, for which reliable predictions would require detailed models for the spatial distribution of the stars in the galaxy and the surface brightness profile of the ionized gas. This in particular limits the usage of the *RST* and *Euclid* surveys to a minimum galaxy mass ($M_{*,\min}$), while for *JWST* this will only be an issue when doing spectroscopy or photometry on very low-mass galaxies ($\lesssim 1000 M_{\odot}$) at high redshift. The minimum galaxy mass applicable for a given telescope can be estimated by using equation (7) or equation (6) while setting $\mu_{\min} \geq 1000$ and solving for the mass. For the wide (deep) *RST* and *Euclid* surveys, using $z = 10$ and star-forming age $\tau \approx 10$ Myr, this puts us at minimum total stellar masses; 3×10^4 ($4.8 \times 10^3 M_{\odot}$) and 4.8×10^5 ($7.5 \times 10^4 M_{\odot}$), respectively, for photometry, and at 2.2×10^6 ($2.2 \times 10^5 M_{\odot}$) and 6.7×10^6 ($1.1 \times 10^6 M_{\odot}$), respectively, for spectroscopy. Similarly, for *JWST* one finds; 1.1×10^3 and $2.9 \times 10^3 M_{\odot}$ for photometry and spectroscopy, respectively (based on the deep NIRCcam and very deep NIRSpec observing modes presented in Table 1). We indicate this minimum allowed galaxy mass in Fig. 1 for all the different telescopes and survey modes by a star-shaped marker that is colour coded to match the respective telescope. In particular, the survey modes with low depth are limited from below in the galaxy masses that they can detect, even with strong gravitational lensing. Throughout the paper, a redshift $z = 10$ and a star-forming age $\tau \approx 10$ Myr are chosen as fiducial values due to the star formation rate density of Pop III stars peaking around $z = 10$ (e.g. Liu & Bromm 2020) and the spectra of the galaxies simulated in this paper reaching stable values after a couple of million years which will be retained throughout the defined lifetime of 10 Myr before forming distinctive spectral features from metal enrichment. The predictions in this paper which relate to total stellar mass therefore refers to Pop III galaxy that has formed their stars during a 10 Myr window – which can then be compared to simulations with predicted Pop III total stellar masses forming over a similar time window. If we detect a Pop III galaxy we should note it may very well have any age within the time

window between its formation and the upper limit of 10 Myr. Since the luminosity of the galaxy is determined by the SFR (for a given IMF) we may find Pop III galaxies with total stellar masses that are lower than the mass at which we set the detection limit, and they would still be observable. In other words, the SFR determines the detectability while the allowed time window sets the maximum total stellar mass that we expect to observe.

2.4 Population III galaxy simulations

As the first stars form through gravitational collapse out of pristine gas, devoid of any metals that enhance cooling, we must consider environments in which hydrogen can cool effectively via, either atomic line transitions or molecular hydrogen cooling. Without such cooling mechanisms, the collapse of pristine gas will halt and resist star formation for very extended times (Bromm 2013). We therefore require the formation of dark matter haloes of sufficient mass in which the virial temperature of the pristine gas reaches $T_{\text{vir}} \approx 10^4$ K whereupon H I atomic cooling may commence and thus ensure a promising formation site for a Pop III galaxy. As star formation ramps up in the Universe, a background of Lyman–Werner (LW) radiation is built up. This background radiation is able to suppress star formation in lower mass dark matter haloes – at redshift $z = 10$, the LW background is sufficiently high that essentially all star formation occurs in haloes with a mass corresponding to the atomic cooling limit, or in more massive haloes (Visbal, Bryan & Haiman 2020). This provides us with a typical dark matter halo mass, which, in turn – depending on the average cosmic density of baryonic matter with respect to dark matter and the star formation efficiency ϵ of pristine gas – gives us an estimate of the typical stellar mass of a Pop III galaxy across different redshifts.

To make proper use of our estimates for the required minimum number density of Pop III galaxies, we make comparisons with a number of independent theoretical simulations that attempt to predict the number density of Pop III galaxies at various redshifts. We use the predictions made by Stiavelli & Trenti (2010), Inayoshi et al. (2018), and Visbal et al. (2020) in order to determine whether the required number densities are indeed reproduced in contemporary simulations. Stiavelli & Trenti (2010) calculates the formation rate of pristine atomic cooling haloes per cMpc³ in the temperature range $10^4 \leq T_{\text{vir}} \leq 2 \times 10^4$ K subject to a homogeneous LW background as the mechanism delaying star formation, and adopt the assumption of inhomogeneous metal mixing to estimate the efficiency with which haloes are enriched and thus lose their pristine Pop III observational signatures. The range in virial temperature implies a range in the total stellar mass of Pop III galaxies, given an assumed star formation efficiency, which can be seen in Figs 1 and 2. Visbal et al. (2020) similarly provide us with number densities and total Pop III stellar masses but for different assumed star formation efficiencies, taking into account an inhomogeneous H₂ dissociating LW background, H I ionizing background from inhomogeneous reionization and inhomogeneous metal enrichment feedback processes. This provides a more dynamical and realistic formation scenario for Pop III galaxies, yielding number densities that are lower than Stiavelli & Trenti (2010) by about an order of magnitude. The discrepancy between the models can be explained by the more refined physical treatment in Visbal et al. (2020) which has the effect of suppressing the number density. Variations in the parameters underlying the mechanisms for delayed star formation was shown in Visbal et al. (2020) to only affect the predicted abundance of Pop III stars by a factor of a few, therefore underscoring the robustness of the models. Furthermore, the star formation efficiency is a free parameter in these models

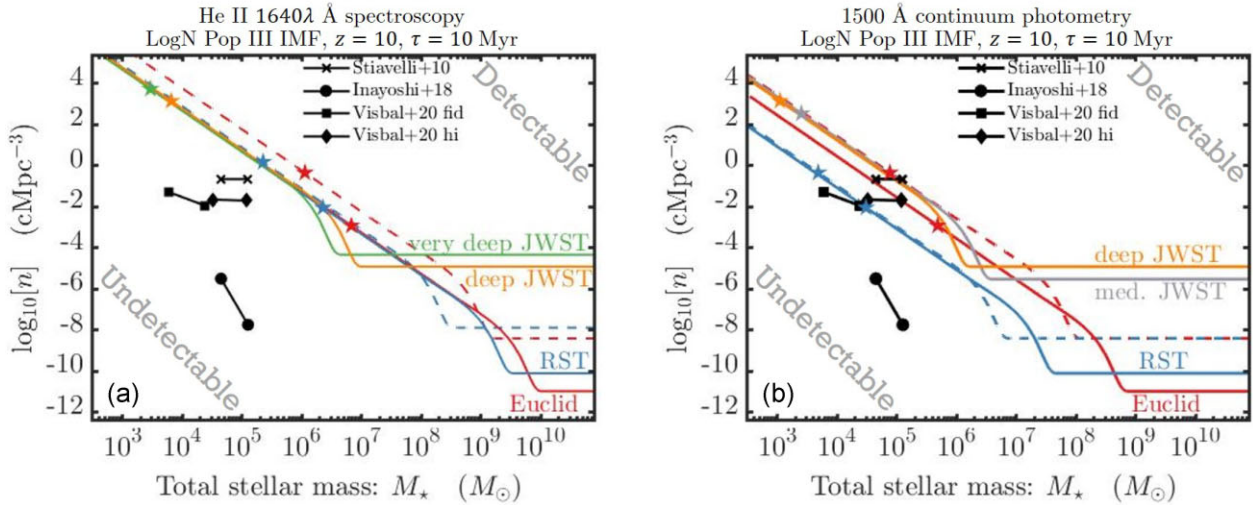


Figure 1. Panel (a): The minimum number density (n_{\min}) of Pop III galaxies required for spectroscopic detection of the He II $\lambda 1640$ Å emission line as a function of the total stellar mass of the Pop III galaxy, at redshift $z = 10$. Panel (b): The minimum number density (n_{\min}) required for photometric detection of the continuum at 1500 Å as a function of the total stellar mass of the Pop III galaxy, at redshift $z = 10$. The stellar population synthesis model used (Raiter et al. 2010) in the above figures is based on a top-heavy lognormal IMF ($m_c = 60 M_\odot$, $\sigma = 1$, with mass range $1 - 500 M_\odot$) with a constant star formation rate where the spectrum is extracted when the galaxy has reached a star-forming age of $\tau = 10$ Myr. The lines correspond to different telescopes with survey parameters as specified in Table 1, the dashed blue and red lines are for deep RST and deep Euclid surveys, respectively. The black marked points are the results of simulations predicting the number density of atomic cooling haloes (with an allowed temperature range $1 \times 10^4 - 2 \times 10^4$ K) which may host Pop III galaxies with masses calculated using $\epsilon = 0.01$ for Stiavelli & Trenti (2010), Inayoshi et al. (2018) and $\epsilon = \{0.001, 0.005\}$ for the Visbal et al. (2020) fiducial (fid) and high (hi) efficiency model, respectively. The lower masses in the Visbal et al. (2020) fiducial and high-efficiency model corresponds to the atomic cooling limit while the slightly higher masses indicates haloes that have star formation suppressed to higher masses due to photoheating as a result of cosmic reionization. We have indicated the minimum mass ($M_{*,\min}$) allowed in order not to breach the $\mu_{\max} = 1000$ limit for each telescope by the star symbols, colour coded to their respective telescopes. The figures can be crudely separated by the coloured lines into two regions, the detectable and the undetectable region. Any Pop III galaxy comoving number density located in the upper-right region, above the required minimum comoving number density set by the telescopes, would have a good chance of ending up above the detection limits of future surveys. On the contrary, Pop III galaxy comoving number densities located underneath the required minimum, would have a small likelihood of being detected. At $z = 10$, the figure reveals that all the telescopes have almost identical prospects for spectroscopic detection of the He II $\lambda 1640$ emission line in galaxies with mass $M_* \lesssim 10^6 M_\odot$ – the only exception being the deep Euclid survey (red dashed line). JWST has an advantage in very deep exposures when it comes to the most massive Pop III galaxies of total stellar mass $M_* \gtrsim 10^6 M_\odot$. When it comes to photometric detection of the UV continuum at 1500 Å; the best telescope is a wide or deep RST survey for all Pop III galaxy masses in the relevant mass range. Trailing the wide and deep RST surveys is the wide Euclid survey, we also find that there is a notable area around $M_* \approx 10^6 M_\odot$ where JWST can be competitive in deep and medium-deep surveys. Based on the minimum required number density of the different surveys as compared to the simulations, we conclude that for the given star formation efficiency, spectroscopic detection is unlikely. When it comes to photometry, the Stiavelli & Trenti (2010) model very likely provide detections with RST and likely also the wide Euclid survey. The Visbal et al. (2020) high-efficiency model is most likely to provide detections with RST and in rare cases also Euclid.

which, in the Visbal et al. (2020) fiducial case, is assumed to not reach significantly above $\epsilon = 0.001$. We also include a high-efficiency model from Visbal et al. (2020) which allow for the formation of slightly more massive Pop III galaxies ($\epsilon = 0.005$). The true value for the Pop III star formation efficiency is debated and has not been constrained well yet due to the absence of large high- z ($z \gtrsim 10$) data sets. In addition to our fiducial star formation efficiency, we have therefore allowed for several alternatives in Fig. 3 to visualize the need to achieve high star formation efficiencies. We have also included Inayoshi et al. (2018), which simulates the number density of pristine atomic cooling haloes forming through frequent mergers in areas of high baryonic streaming motion as the process delaying star formation. These kinds of formation processes produce fewer suitable atomic cooling haloes in which pristine Pop III galaxies can form but might allow for higher star formation efficiencies similar to what is found in Regan et al. (2020) which simulates rapidly forming haloes. As a bottom line, we regard the Stiavelli & Trenti (2010) and Visbal et al. (2020) simulations as viable estimates of number density of pristine Pop III galaxies while Inayoshi et al. (2018) provide lower bounds through more rare formation channels.

Using these predictions, we estimate the comoving number density of these atomic cooling haloes for the corresponding total stellar mass forming in such haloes via equations (2) and (3) for the Stiavelli & Trenti (2010) and Inayoshi et al. (2018) simulations. Visbal et al. (2020) predict similar stellar masses that differ slightly due to modulation by the local LW flux. The virial mass of the atomic cooling halo for the given cosmology is taken to be, as given by Bryan & Norman (1998) and Barkana, Haiman & Ostriker (2001):

$$M_{\text{vir}} \approx 32.7 \times T_{\text{vir}}^{3/2} \left(\frac{1+z}{10} \right)^{-3/2} \quad (2)$$

for a neutral primordial gas, resulting in a stellar mass given by

$$M_* = M_{\text{vir}} \frac{\Omega_b}{\Omega_m} \epsilon, \quad (3)$$

where ϵ designates the star formation efficiency (the fraction of the baryonic mass of the halo converted into stars during a single star formation episode). Expressing n_{\min} as a function of the total stellar mass of the Pop III galaxy enables us to make comparisons with contemporary simulations predicting the number density of atomic cooling haloes which has a mass according to equation (3). The

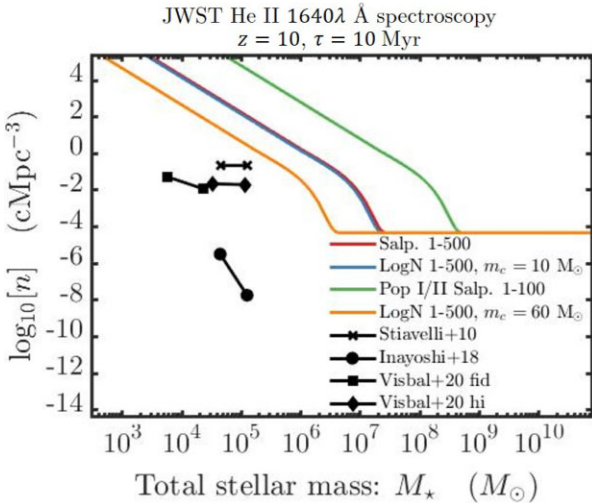


Figure 2. The impact of varying the Pop III IMF on the minimum required number density for the very deep NIRSpec survey mode. The baseline (the flat, lower part of the lines) stay the same irrespective of the IMF since it only affects the luminosity of the source, resulting in a horizontal shift. Note that the lognormal, $m_c = 10 M_\odot$, $\sigma = 1$, in the mass range $1\text{--}500 M_\odot$ (blue line) and Salpeter $1\text{--}500 M_\odot$ (red line) IMFs are almost completely overlapping. We have also included a standard, solar metallicity Pop I/II Salpeter IMF (green line) as a comparison. From this result, we conclude that the lognormal IMF, $m_c = 60 M_\odot$, $\sigma = 1$, in the mass range $1\text{--}500 M_\odot$ that is used in Fig. 1 is indeed the most optimistic – lowering the Pop III stellar mass required for detection by an order of magnitude when compared to the less top-heavy IMFs. The simulations included are the same as in Fig. 1.

minimum halo mass which can host Pop III star formation is a function of redshift and spatial location and is set by the local LW flux as well as feedback from reionization of the local intergalactic medium. In Fig. 3, we consider a range of star formation efficiencies $\epsilon = \{0.001, 0.005, 0.01, 0.1\}$ in order to calculate the expected total stellar mass of atomic cooling haloes across redshifts, which can then be used to estimate the minimum required number density of Pop III galaxies with that mass. Furthermore, we use a fiducial value $\epsilon = 0.01$ to calculate the masses of the atomic cooling haloes simulated in Stiaivelli & Trenti (2010) and Inayoshi et al. (2018) that goes into Figs 1 and 2. The correct value for the star formation efficiency is indeed a matter of debate and is also important for the results of this paper – high values produce bright objects that will be more easily detected since less magnification is required, which in turn relaxes the required minimum number density that we calculate in this paper. The opposite is of course true for low star formation efficiencies. Stiaivelli & Trenti (2010) argue that $\epsilon \gtrsim 0.15$ is very unlikely, Visbal et al. (2020) explore a range of $\epsilon = (0.0001\text{--}0.005)$ but uses 0.001 as their fiducial value. Some earlier work have suggested somewhat higher ranges $\epsilon = 0.01\text{--}0.1$ (see e.g. Wise & Cen 2009) while Jaacks et al. (2018) suggest $\epsilon = 0.05$. In Skinner & Wise (2020), they suggest H_2 self-shielding Pop III star-forming haloes typically having $\epsilon = 0.0001\text{--}0.001$. Also, Zackrisson et al. (2012) showed, using the haloes simulated in Stiaivelli & Trenti (2010), Trenti, Stiaivelli & Shull (2009), that high Pop III galaxy number densities with star formation efficiencies $\epsilon \gtrsim 0.1$ are inconsistent with the observed UV luminosity function (see Bouwens et al. 2011; e.g. Oesch et al. 2012) at redshifts $z \approx 7\text{--}10$.

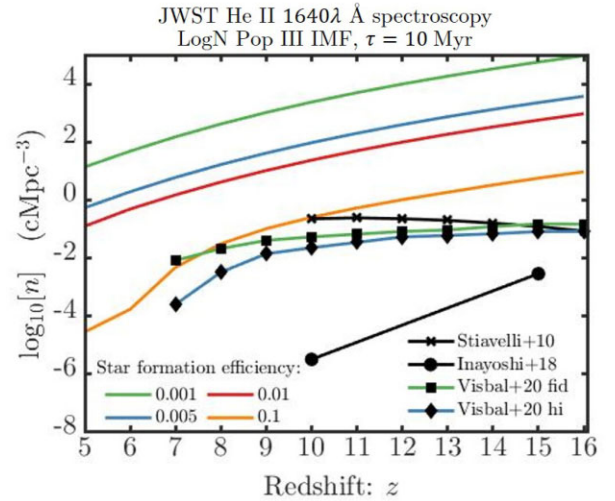


Figure 3. The minimum number density (n_{\min}) required for spectroscopic detection of the He II $\lambda 1640 \text{ \AA}$ emission line in suitable haloes (atomic cooling) using the very deep *JWST* survey, as a function of redshift. The lines correspond to different star formation efficiencies (ϵ) in the range $0.001\text{--}0.1$. The IMF used is the fiducial lognormal top-heavy, $m_c = 60 M_\odot$, $\sigma = 1$ in the mass range $1\text{--}500 M_\odot$. The virial mass of these atomic cooling haloes ($T_{\text{vir}} = 10^4 \text{ K}$) are redshift dependent which, given a star formation efficiency, implies a range of total stellar masses for each line in the figure. The virial mass range is $M_{\text{vir}} \approx 1.5\text{--}7 \times 10^7 M_\odot$, increasing from high redshift to low redshift. The total stellar mass for the plotted lines are then given by $M_* = M_{\text{vir}} \frac{\Omega_b}{\Omega_m} \epsilon \approx M_{\text{vir}} 0.15 \epsilon$ (cf. equation 2 and 3). The included simulations (black markers) provide the number densities of atomic cooling haloes. Note that the Visbal et al. (2020) simulations here are predictions for $\epsilon = \{0.001, 0.005\}$ – therefore the colour on the lines for these simulations have been matched with the colour of the minimum required number density with the corresponding star formation efficiency. Furthermore, the total Pop III stellar masses predicted in Visbal et al. (2020) vary slightly from the mass at the atomic cooling limit at the higher redshifts as star formation is then allowed in haloes with lower masses. Notice the different behaviour for the $\epsilon = 0.1$ line which produces a galaxy which is massive enough (i.e. bright enough) for it to not require gravitational lensing at the lowest redshift, but begin to require magnification around $z = 6.5$. All the other star formation efficiencies require gravitational lensing throughout the redshift range. The $\epsilon = 0.1$ line is the only realization which produces Pop III galaxies that are even close to some of the simulations. This plot can also be applied to *RST* and *Euclid* as long as the total stellar mass of the Pop III galaxy is not too low – which restricts the redshifts at which it is valid. At $\epsilon = 0.01$, the deep *RST* survey have enough sensitivity to spectroscopically detect atomic cooling haloes without breaching $M_{*,\min}$ for $z \lesssim 7$, while deep *Euclid* surveys cannot even reach $z \sim 5$ – pushing the star formation efficiency to $\epsilon = 0.1$ puts these limits at $z \lesssim 12$ and $z \lesssim 8$ for deep *RST* and *Euclid* surveys, respectively. The very deep NIRSpec survey is only affected by this limit in the case of $\epsilon = 0.001$, where we are restricted to $z \lesssim 11.5$.

2.5 Telescopes and survey parameters

The observational facilities best suited to conduct blind searches for rare and faint objects such as lensed Pop III galaxies at high redshift are space telescopes, due to their combination of large survey area, angular resolution, and high sensitivity. In this paper, we focus on the expected capabilities of the *James Webb space telescope*, the *Roman space telescope*, and *Euclid*. We emphasize that this does not mean that these are the only viable options for studying Pop III signatures in the early Universe. For example, Grisdale et al. (2021) recently demonstrated the potential for detecting He II $\lambda 1640$ emission line from Pop III galaxies using the ground-based

ELT/HARMONI spectrograph. However, due to the small field of view of this instrument, the success of such an approach hinges on knowing where to point the *ELT*, as we have no a priori information on the location of Pop III galaxies on the sky.³ The selected targets would hence first have to be culled from detections in a larger-area survey of the type we consider here.

In Table 1, we summarize the assumed survey parameters, based on surveys described in Rieke et al. (2019), Marchetti, Serjeant & Vaccari (2017), Akeson et al. (2019), and Laureijs et al. (2011). As can be seen, *Euclid* provides the largest survey areas of 15 000 or 40 deg² in the wide or deep surveys, respectively – this is followed by *RST* which has 2000 or 40 deg² in its wide and deep surveys, respectively, although the deep spectroscopic survey covers only 12 deg². *JWST*, on the other hand, only reaches survey areas of around 0.05 deg² in the surveys included. The small survey area of *JWST* is greatly compensated by its depth which is considerably better than *Euclid* and *RST*. However, since rare, high-magnification lines of sight are also required to detect Pop III galaxies with properties similar to those seen in current simulations, it is a priori not clear which telescope will perform the best.

Considering *JWST*, we can compare the assumed survey parameters to the by now established cycle 1 general observer, guaranteed time observations and early release science programs. A non-exhaustive selection of programs that are relevant for the kind of surveys suggested in this paper are given by; JADES (see e.g. Eisenstein, Ferruit & Rieke 2017), CEERS (Finkelstein et al. 2017), WDEEP (reaching emission-line fluxes of $\sim 10^{-18}$ erg s⁻¹ cm⁻² and imaging to $m_{AB} \sim 30.6$ –30.9; Finkelstein et al. 2021), FRESCO (covering ~ 0.017 deg² while reaching emission-line fluxes $\sim 3.3 \times 10^{-18}$ erg s⁻¹ cm⁻²; Oesch et al. 2021), UNCOVER (targeting Abell 2744 with imaging to $m_{AB} \sim 29.5$ –30 where emission lines can be measured for all sources with $m_{AB} < 30$; Labbe et al. 2021), the WIDE MOS survey (reaching $m_{AB} \sim 24$ while taking spectra for objects with H α emission-line fluxes $> 10^{-17}$ erg s⁻¹ cm⁻² over an area of ~ 0.1 deg², see e.g. Ferruit 2017, for one of the targeted fields). Note that the survey parameters taken from Rieke et al. (2019) and used in Table 1 summarizes imaging and spectroscopy from the JADES and CEERS programs, as such, we use these as representative parameters to assess *JWST*'s capabilities. However, as we see, there are multiple programs operating at different depths and over different survey areas which should in principle be summed to assess *JWST*'s overall likelihood of detecting pristine Pop III galaxies in any of the first cycle programs. In reality, this implies slightly better prospects for detection with *JWST* than the results cited in this paper, where we focus on the summarized JADES and CEERS survey presented in Rieke et al. (2019). Furthermore, several approved programs that include targets of individual cluster lenses, e.g. Windhorst et al. (2017) and Treu et al. (2017) will introduce competitive alternatives to the wider fields considered in this paper.

2.6 Minimum comoving density for the detection of lensed Pop III galaxies

We derive the comoving volumes spanned by the survey areas of the different telescopes by integrating the comoving differential volume element over the considered redshift and survey area. The depth in redshift of the comoving volumes is set by the typical redshift

range ($\Delta z = 1$) in which galaxies are picked using the Lyman dropout technique using broadband filters. The comoving volume, defined here as V_c , is a function of the cosmological model used, the redshift (z), redshift bin-size (Δz) and the survey area (units of square degrees, defined as A). In a flat Friedmann–Lemaître–Robinson–Walker spacetime, the comoving volume is given by

$$V_c = \frac{4\pi}{3} \frac{A}{41253} [D_c(z + \Delta z/2)^3 - D_c(z - \Delta z/2)^3], \quad (4)$$

where $D_c(z \pm \Delta z/2)$ is the comoving distance at redshift $z \pm \Delta z/2$. Given the comoving survey volume $V_c(A, z, \Delta z)$ probed at each redshift, we can calculate the minimum comoving number density of chemically pristine Pop III galaxies (n_{\min} required in order to detect at least one Pop III galaxy candidate in one single survey, by applying the constraint

$$n_{\min} = \frac{1}{P(> \mu_{\min}, z) \times V_c(A, z, \Delta z)}. \quad (5)$$

A time window ($t_{\text{PopIII}} = 10$ Myr) is later applied to the simulations where necessary, in order to make sure that we calculate the number of chemically pristine Pop III galaxies inside the comoving volume, that is, galaxies that would be identifiable as Pop III before losing its unique diagnostic observational signatures.

Notice the dependency on μ_{\min} , which, in turn, is a function of the emission-line flux $F_{\text{He,II } 1640}$ for spectroscopy, or, the continuum flux $F_{\text{UV } 1500}$ in the case of photometry. The minimum magnification μ_{\min} determines the probability for gravitational lensing to boost the intrinsic flux of a galaxy above the detection threshold of the considered telescope/survey. This is calculated through the magnification probability function $P(> \mu, z)$ – which determines the probability of getting a magnification greater than μ .

The logic goes as follows; for a given comoving number density (n), we can find the total number of objects inside a comoving volume V_c by multiplying the 2. Whether we can detect e.g. the He II $\lambda 1640$ emission line from objects inside this volume depends on the telescope's detection limit, the emission-line flux ($F_{\text{He,II } 1640}$) and the magnification (μ) from gravitational lensing. The required magnification (μ_{\min}) is clearly a decisive parameter when it comes to estimating n_{\min} . We should therefore ask ourselves what minimum magnifications, we would require to detect a Pop III galaxy, either spectroscopically, or photometrically. In the case of spectroscopy, the answer is given by

$$\mu_{\min} = \frac{4\pi D_L(z)^2 F_{\text{limit}}}{\mathcal{L}_{\text{He,II } 1640}} \frac{\tau}{M_*}, \quad (6)$$

where F_{limit} is the spectroscopic detection limit of the telescope considered (see Table 1), $D_L(z)$ is the luminosity distance to the galaxy, $\mathcal{L}_{\text{He,II } 1640}$ is the He II $\lambda 1640$ emission-line luminosity per solar mass per year, M_* the total stellar mass of the galaxy and τ is the age of the galaxy after onset of star formation in the galaxy. Note that what we have above is simply a ratio between the telescope flux detection limit and the emission-line flux, scaled up or down with the average (constant) star formation rate; $\text{SFR}^{-1} = \frac{\tau}{M_*}$.

In the case of photometry, we use the rest frame luminosity density at 1500-Å to estimate the detectability of Pop III galaxies. Using AB magnitudes and a factor $(1+z)$ for the K -correction, we have the following:

$$\mu_{\min} = \frac{c}{(1+z)\lambda_0^2} 4\pi D_L(z)^2 10^{-(m_{\text{limit}}+48.6-25)/2.5} \frac{\tau}{M_*} \frac{1}{\mathcal{L}_{\text{UV } 1500}}, \quad (7)$$

where instead we now have the dependence on the photometric detection limit (m_{limit} – in AB magnitudes), the rest frame wavelength λ_0 (in our case $\lambda_0 = 1500$ Å) and $\mathcal{L}_{\text{UV } 1500}$ which now designates the

³But see Johnson & Aykutaalp (2019) for a scenario in which Pop III galaxies form in the vicinity of high-redshift quasars, which would allow clues to their preferred location.

luminosity density at 1500 Å (units of $\text{erg s}^{-1} \text{Å}^{-1}$), normalized to $1 M_{\odot} \text{yr}^{-1}$, and c designates the speed of light.

The fiducial star-forming age at which we extract the spectrum of the Pop III galaxies is given by $\tau = 10 \text{ Myr}$. In reality, the star-forming age of the observed galaxy would form a distribution of values that for the sake of observational availability of the object would have to be bounded by $\tau < t_{\text{popIII}} = 10 \text{ Myr}$. Since the luminosity of the galaxy is dependant on the star-forming age (the UV/line luminosity generally ramps up with time quickly as star formation has commenced) we could, in principle, observe the galaxy at a very early stage as it has not yet reached its equilibrium luminosity – making it somewhat harder to detect. For example, between 1 and 2 Myr, the luminosity in the He II $\lambda 1640$ emission increases by ~ 22 per cent while the continuum luminosity at 1500 Å increases by a factor of ~ 2.3 in the same time-span. For the Pop III, top-heavy IMFs considered here, we find that after $\tau = 2 \text{ Myr}$ the galaxy has reached a stable luminosity after which it does not change significantly and therefore defines an age from which the luminosity that the galaxy has will be retained for a very long period, given a constant star formation rate – making $\tau = 10 \text{ Myr}$ a suitable fiducial value. We will discuss the effects of looking at lower star-forming ages further in Section 3.3

2.7 Detection probability

The number density limit in equation (5) is valid for a certain probability P_{det} to end up with at least one Pop III galaxy above the detection threshold. The probability of lifting one pristine Pop III galaxy, out of N possible, for a given redshift, above the detection limit is given by; $P_{\text{det}} = 1 - [1 - P(> \mu_{\text{min}})]^N$, where N is the number of pristine galaxies observable during any time window t_{PopIII} within the Δt (the cosmic time spanned by the surveyed volume) which require the magnification μ_{min} for detection.

Demanding a specific probability to lift one galaxy above the detection limit therefore sets a relation between N and P for the given required magnification. For example, to get a 95 per cent probability ($P_{\text{det}} = 0.95$, i.e. 2σ level) of detecting the He II $\lambda 1640$ emission line from a single galaxy with total stellar mass $M_{\star} \approx 4.5 \times 10^4 M_{\odot}$ using *JWST* with the depth $1.3 \times 10^{-19} \text{ erg s}^{-1} \text{cm}^{-2}$, at $z = 10$, our computational machinery indicates that we need a magnification of $\mu_{\text{min}} \approx 64$, which has a probability $P(> \mu_{\text{min}}) \approx 2 \times 10^{-6}$ of occurring. This results in $N \approx \frac{\ln(1-P_{\text{det}})}{\ln(1-P(> \mu_{\text{min}}))}$ – implying that we need approximately 1.4 million pristine galaxies (i.e. with a star-forming age $\tau < 10 \text{ Myr}$) inside the whole survey volume considered by *JWST*, in order to make a single detection of a lensed galaxy.

The relation in equation (5) is not as stringent, instead having a scaling $N = \frac{1}{P}$, leading to a requirement of $N \approx 460\,000$ pristine galaxies in the survey volume considered – about a factor of 3 less than that required for the 2σ level. We therefore consider equation (5) as a lower bound on the minimum comoving number densities presented in Section 3.

3 RESULTS

In Fig. 1, we show that all included telescopes require more or less identical minimum required number densities for spectroscopic detection of the He II $\lambda 1640$ Å emission line. However, due to limitations in the maximum magnifications ($\mu_{\text{max}} \approx 1000$) that are applicable, *JWST* is the only telescope that has the required sensitivity to detect atomic cooling haloes with lower star formation efficiencies (e.g. $\epsilon = 0.001$) up to high redshifts ($z \lesssim 11.5$). We find that *JWST* does better than the other telescopes for the most massive Pop III

galaxies ($M_{\star} \gtrsim 10^6 M_{\odot}$). We also find that the *Euclid* wide survey perform better than its deep survey when it comes to detecting the He II $\lambda 1640$ Å emission line. However, the magnification limit (μ_{max}) at $z = 10$ suggests that the wide *Euclid* survey can only be applied to total stellar masses $\gtrsim 6.7 \times 10^6 M_{\odot}$, which is notably larger than predicted by the simulations included. *RST*, on the other hand, can reach total stellar masses of $M_{\star} \approx 2 \times 10^5 M_{\odot}$ in its deep survey without breaching the magnification limit at $z = 10$. In the case of photometric detection of the UV continuum at 1500 Å; at $z = 10$, the best telescope is *RST* for all the total Pop III galaxy stellar masses. However, we find a notable area around $M_{\star} \gtrsim 7 \times 10^5 M_{\odot}$ where *JWST* can be competitive in deep and medium-deep exposures. The photometric wide *RST* survey can reach Pop III galaxies forming in atomic cooling haloes with star formation efficiency $\epsilon = 0.01$ for $z \lesssim 11.5$, without breaching the magnification limit – for *Euclid* on the other hand, the deep survey is necessary, which reaches $z \lesssim 8$ without breaching the magnification limit.

The individual lines in Fig. 1 display segments with different slope in the different mass ranges, depending on the survey. The flat, horizontal part of the lines indicates the region where the telescope does not require gravitational lensing to detect the galaxy. The location that the horizontal line has on the y-axis is determined by the survey area of the telescope. The location where the lines bend upwards is set by the depth of the telescope and therefore also marks the mass below which gravitational lensing is required – and as we move to the left along the lines, the magnification (μ) becomes larger. The slope of the line is steeper to begin with but levels off at a certain point – this occurs when the required magnification reaches $\mu \approx 10$ where the magnification probability function enters the $P(\mu) \propto \mu^{-2}$ regime, as is described by the polynomial fit in equation (1). We separate the figures into a ‘detectable’ region and a ‘undetectable’ region, where the separation is determined by the lines belonging to the different surveys. Any simulation that predicts Pop III galaxy number densities above these lines implies that there is a good chance for detection – placing it in the detectable region. Likewise, if the predicted number density is below the lines, prospects for detection are poor – therefore placing it in the undetectable region.

Apart from the telescope survey area and depth, the relevant parameters for the minimum Pop III galaxy number density for detection are; the age of Pop III burst (τ), the Pop III IMF and the redshift of the observed galaxy. Varying the survey area (A) will shift the lines in Figs 1 and 2 up (if A is decreased) or down (if A is increased), while the depth of the telescope shift the lines to the left (for higher depth, i.e. higher sensitivity) or right (for lower depth). The survey volume scales linearly with survey area (equation 4), implying that n_{min} scales inversely with A , while the detection limit of the telescope enters quadratically into the magnification probability function (see equation 1) due to its dependence on μ_{min} (see equation 6 and 7).

3.1 The detectability of simulated galaxies

Looking again at Fig. 1(a) indicates that, at $z = 10$, the prospects for spectroscopic detection of the He II $\lambda 1640$ Å emission line in the surveys included in this paper are not ideal. For example, even the simulations in Stiavelli & Trenti (2010) suggests number densities that are 1–2 orders of magnitude lower than what is required from our calculations (depending on which of the two total stellar masses that is considered). We keep in mind though that this prediction hinges on the value of the star formation efficiency, which in Fig. 1 is set to 0.01 for the Stiavelli & Trenti (2010) and Inayoshi

et al. (2018) simulations. The calculations also show that all the included surveys require very similar minimum number densities for detection, however, *JWST* is the only telescope that does not breach the maximum allowed magnification at this redshift. When it comes to photometry, the prospects for detection are quite good when comparing to the predictions of the simulations. *RST* and *Euclid* surveys are predicted to detect Pop III galaxies using photometry, given the number densities suggested by Stiavelli & Trenti (2010). We also see that both the wide and deep *RST* surveys should detect the more massive galaxies predicted by Visbal et al. (2020) in their high-efficiency ($\epsilon = 0.005$) simulations while the galaxies simulated in their fiducial model ($\epsilon = 0.001$) has number densities that are at the very edge minimum requirements. For the Stiavelli & Trenti (2010) simulations, we also find that *JWST* may be able to detect about one Pop III galaxy if its total stellar mass is $\gtrsim 10^5 M_\odot$. Looking at Fig. 3, we can see the redshift evolution of the minimum required comoving number density for spectroscopic detection of the He II $\lambda 1640$ Å emission line. Here, we can once again see that our fiducial $\epsilon = 0.01$ model requires number densities that are higher than the predictions from simulations. However, this figure suggests that if the star formation efficiency can be pushed up to $\epsilon = 0.1$, Stiavelli & Trenti (2010) predict number densities that are sufficient at $z \sim 10$ – but such high star formation efficiencies together with their predictions for the number density of atomic cooling haloes were shown to be inconsistent with the UV luminosity functions at $z \approx 7$ –10 by Zackrisson et al. (2012). Note that Visbal et al. (2020) has a well defined star formation efficiency in their models which has an effect on the predicted number densities. The star formation efficiency is linked to the intensity of the LW background which sets the minimum mass where Pop III stars can form – therefore requiring additional simulations in order to accurately compare with the higher star formation efficiency.

If the star formation efficiency is indeed lower than $\epsilon \approx 0.1$, we see that spectroscopic detection of a Pop III galaxy is very improbable in the surveys included here. A way to improve the prospects for detection is therefore to consider even wider and/or deeper surveys. As stated, increasing the depth of the telescope has a larger impact than increasing the survey area – but going deeper unfortunately demands significantly higher exposure times.

3.2 The impact of the Pop III stellar IMF

As mentioned earlier, we use a fiducial lognormal IMF in the mass range 1 – $500 M_\odot$ for our predictions. The strength of the He II $\lambda 1640$ Å emission line is affected significantly by the IMF considered. This is due to the correlation between the abundance of high-mass stars ($\gtrsim 50 M_\odot$) and the amount of hard ionizing flux. A standard Salpeter IMF in the mass range ~ 1 – $100 M_\odot$ simply won't produce the required ionizing flux needed to detect the He II $\lambda 1640$ Å emission line without very strong magnification. As Fig. 2 reveals, the same goes for other Pop III IMFs which have suppressed high-mass tails, resulting in low He II $\lambda 1640$ Å fluxes and therefore poor prospects for detection. Similarly, the IMF impacts the continuum level at 1500 Å, which affects the detectability of the galaxies in photometric surveys. The results of Figs 1, 3, and 4 were extracted for the lognormal IMF with characteristic mass $m_c = 60 M_\odot$ and width $\sigma = 1$, and a mass range 1 – $500 M_\odot$. In the case that nature in reality provides us with a less top-heavy IMF, the requirements on the minimum number density increases by at least an order of magnitude for the other, less extreme models included here.

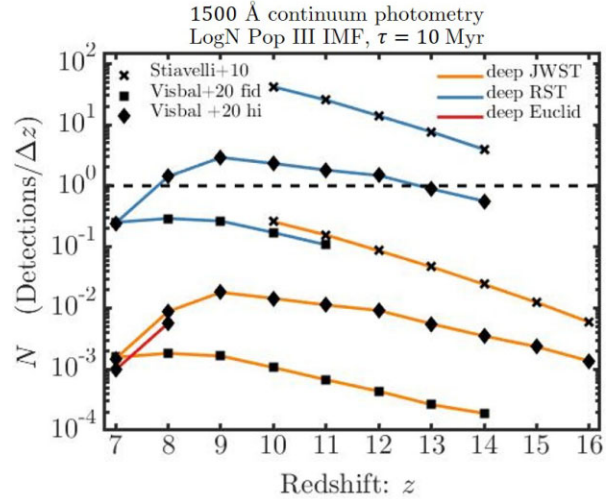


Figure 4. The expected number (N) of photometrically detected Pop III galaxies per redshift bin ($\Delta z = 1$) as a function of redshift for the different survey and simulation combinations. The redshift range applicable to the different surveys and simulations are restricted by the maximum allowed magnification $\mu_{\text{max}} = 1000$ as well as the limit at which the 1500 -Å continuum redshifts out of the surveys reddest filter. The restrictions can be notably seen for *Euclid* which is only applicable for the lowest redshift ranges combined with the Visbal et al. (2020) high-efficiency model due to magnifications limits. The deep *RST* survey is restricted to $z \lesssim 11$ due to magnification limits when combined with the fiducial Visbal et al. (2020) simulations, while for the other simulations, the 1500 -Å continuum redshifts out of the *RST* filters for $z \gtrsim 14$. The deep *JWST* survey is only restricted by magnification limitations when combined with the Visbal et al. (2020) fiducial simulations where it reaches $z \lesssim 14$. The different surveys are identified by the colour on the lines while the combination of that particular survey with a simulation is indicated by the respective black marker (cross, square or diamond). The dashed black line illustrates which survey and simulation combinations that are expected to detect at least one Pop III galaxy per redshift bin. As revealed, *RST* is the only telescope that is predicted to detect at least one Pop III galaxy per redshift bin over a considerable redshift range when combined with the Stiavelli & Trenti (2010) or the high-efficiency Visbal et al. (2020) simulations.

3.3 The impact of the Pop III galaxy age

For the constant star formation rate considered in this paper, the age of the Pop III burst is important early after star formation has initiated, as the He⁺-ionizing flux grows rapidly at the onset of star formation while stabilizing after ~ 2 Myr to remain more or less constant throughout the considered time-span. The actual variations in luminosity with age is only at the level of a few per cent in the range of ages 3–10 Myr – for the continuum level at 1500 -Å as well as the He II $\lambda 1640$ Å emission-line luminosity. Therefore, once the initial ~ 2 Myr burst of star formation has completed, the impact of the star-forming age does not significantly affect the detectability of Pop III galaxies. However, as the Pop III galaxy total stellar masses predicted by simulations provide the mass that has been converted into stars before the galaxy proceeds into chemically enriched star formation, the minimum required number density that we predict for a given Pop III galaxy total stellar mass is therefore specifically calculated for galaxies that form their stars over the 10 Myr period that we use as a fiducial age. Performing our calculations for the same total stellar mass but at lower star-forming age implies a higher SFR, which would render the galaxies brighter and thus more easily detectable. In order to investigate how this affects the detectability of Pop III galaxies, we also performed calculations at a lower star-

forming age. Looking at shorter star-forming ages also requires the included simulations to be scaled accordingly in order to not overestimate the number of pristine Pop III galaxies. The scaling of simulated number densities with age is linear while the effects of age on our calculated minimum required number densities scales quadratically (this can be seen by combining equations 1, 5, and 6). The slight variation of the luminosity with age for a constant SFR also enters quadratically into the minimum number density calculation, but as mentioned above, this effect is very small. For example, performing our calculations at a star-forming age of 3 Myr instead of 10 Myr lowers the minimum number density required to detect one object by ~ 1 order of magnitude while simultaneously lowering the number density predicted by simulations by 0.5 orders of magnitude – effectively reducing the gap between the simulations and our minimum required number density with ~ 0.5 orders of magnitude.

3.4 The redshift dependence

The dependence on redshift are two-fold in these calculations. First off, a more distant galaxy of a given mass (and therefore integrated luminosity) are simply dimmed according to the luminosity distance relation. Increasing the redshift therefore shifts the detection delimiter lines in Figs 1 and 2 to the right-hand side. Secondly, in our calculation of the comoving survey volume, the fixed redshift bin size of Δz implies that we are probing different comoving volumes at each redshift. Due to the slow increase of comoving distance with redshift, as we go to higher z , the bin size of Δz covers a smaller depth. The increase of transverse comoving length spanned by the telescope field of view does not compensate for this, implying that the comoving survey volume (V_c) decreases with redshift – which in turn increases n_{\min} – shifting the lines in Figs 1–2 upward. Typically, surveys will cover more than a single such Δz slice of spacetime which would imply that any of the surveys included here may observe Pop III galaxies at several redshifts provided that the number densities are high enough. This would in principle lower the minimum required number density at each redshift if we are content with detecting only one Pop III galaxy at any of the probed redshifts. This will only affect our results moderately, but the effect increases depending on how many individual redshift bins that are covered in the observations. Increasing/decreasing Δz results in larger/smaller volumes being probed at each redshift.

The redshift that is the most promising for detection depends on the star formation efficiency. Looking at Fig. 3 we see that for the high star formation efficiency of 0.1, Stiavelli & Trenti (2010) produce the required number density at redshift $z \approx 10$. In the more realistic scenario of lower star formation efficiencies, simulations are closest at, or around, the peak of Pop III star formation at $z \lesssim 10$, which motivates our choice of fiducial redshift for the other figures. Lower redshifts are, however, of great interest if the simulated Pop III number densities can remain at a sufficiently high level.

In Fig. 4, we display the expected number of photometric Pop III galaxy detections per redshift bin ($\Delta z = 1$) as a function of redshift. By applying the different surveys to the predicted number density from the simulations, we can estimate how many Pop III galaxies we expect to be detected. Once again, we assume the lognormal top-heavy IMF and 10 Myr star-forming age. Furthermore, we assume the star formation efficiency $\epsilon = 0.01$ for the Pop III galaxies forming in atomic cooling haloes predicted by Stiavelli & Trenti (2010), while the Visbal et al. (2020) simulations provide specific total stellar masses (as given by their high-efficiency and fiducial model). Notice that there are two different total stellar masses available for

each simulation – we use the lower as it is representative of the typical atomic cooling halo in which Pop III galaxies form. The Inayoshi et al. (2018) simulations have not been included due to their very low number densities – far from producing any detections. Furthermore, we have only plotted surveys which do not breach the maximum magnification limit and have also excluded surveys in which the 1500-Å continuum redshifts out of the telescopes available filters. As revealed by the figure, the only survey and simulation combination that predicts detections of Pop III galaxies across the whole considered redshift range, is the deep *RST* survey combined with the Stiavelli & Trenti (2010) simulation while the high-efficiency Visbal et al. (2020) simulations provide detection between redshifts $z \approx 8$ and 13 in the same survey.

Due to the nature of conducting high redshift surveys, where the telescopes scan a range of redshifts as one continuously reaches for fainter objects, we can estimate the expected number of detections across all redshifts where we have access to simulations. This reveals that the deep *JWST* survey would detect ~ 0.6 Pop III galaxies if combined with the Stiavelli & Trenti (2010) simulations (e.g. two times more survey area is required in order to produce approximately one detection). The deep *RST* survey produces a considerable number, about 92, of detections if combined with the Stiavelli & Trenti (2010) simulations, a total of ~ 12 if combined with Visbal et al. (2020) high-efficiency model or ~ 1 if combined with the fiducial Visbal et al. (2020) model. The wide *Euclid* photometric survey cannot provide detections without breaching the maximum magnification limit set in this paper.

4 DISCUSSION

Considering the results of this paper (based on the surveys in Table 1), the prospects for spectroscopic confirmation of the true Pop III nature of high-redshift galaxies using the He II $\lambda 1640$ emission line appear to remain grim. However, in this paper, we have set a stringent requirement on the kind of stellar population, namely that it is truly pristine, such that it contains absolutely no metals at all and simultaneously shows a strong He II $\lambda 1640$ line. When searching the skies in future high-redshift surveys, we will potentially find many cases of so-called hybrid Pop III galaxies, in which Pop III stars co-exist with metal-enriched stars (Sarmiento et al. 2018, 2019). These objects are still very interesting when it comes to understanding the formation of the first truly pristine galaxies as it will bring deeper insight into the times-scales related to metal pollution and the feedback effects present at the onset of star formation and put potential constraints on the masses of Pop III stars.

Furthermore, in the estimations of the gravitational lensing probabilities used in this paper, we based our calculations on surveys performed along random sightlines. This generally provides probabilities that are much lower than what one would expect when targeting known galaxy clusters acting as lenses. In these so called cluster lenses one has enhanced probabilities of achieving the very highest magnifications which may very well be the cases which turn out to provide us with the first spectroscopic detection of a truly pristine Pop III galaxy (see section below). When utilizing gravitational lensing as a necessary condition for detection one must also keep in mind that there are complicating factors that may compromise observations due to the presence of so called blends – i.e. the contamination of flux from other objects, e.g. galaxies at other redshifts along the line of sight, for instance, galaxies in a galaxy cluster responsible for the lensing of the Pop III object. This is phenomena that may, in some cases, render observations that meet all the other required criteria for detection useless.

4.1 Comparison to cluster lensing surveys

In this paper, we have focused on the utility of blank fields (randomly located in the sky) that due to their large survey areas could be able to detect a considerable number of gravitationally lensed Pop III galaxies. The large survey areas considered in this paper stands in stark contrast to the very narrow fields relevant when targeting specifically selected cluster lenses that provide high probabilities for achieving very high magnifications. We can make a rough estimate on the capability of individual cluster lenses for the detection of Pop III galaxies compared to blank fields by taking the magnification data assembled in Zackrisson et al. (2012) that provide the area on the sky over which high ($\mu > 10$) magnifications are attained when targeting individual cluster lenses. Given the brightness of the Pop III galaxies considered in this paper, and the detection limit of the telescope in question, we can calculate the minimum required magnification and therefore also the area over which this magnification is realized. Applying this area to the minimum required number density of Pop III galaxies that we calculated for our randomly oriented wide-field surveys, we can make order of magnitude estimates on whether a cluster lens or a wide-field survey performs better. We consider the approved early release science program by Treu et al. (2017) and the guaranteed observer time program by Windhorst et al. (2017) which target several known cluster lenses with *JWST* at deep/medium sensitivity (detection limits around $m_{AB} \sim 28.5$ –29 in photometry, and $\sim 3.5 \times 10^{-19}$ erg s $^{-1}$ cm $^{-2}$ in spectroscopy) as representative surveys to compare our calculations with.

For spectroscopy on the He II $\lambda 1640$ Å emission line, we found that our wide-field surveys would require a minimum number density of ~ 26 cMpc $^{-3}$ in order to detect a single Pop III galaxy of mass $4.4 \times 10^4 M_{\odot}$ (the total stellar mass of an atomic cooling halo with $\epsilon = 0.01$), at $z = 10$. In order to detect such a Pop III galaxy with a sensitivity of 3.5×10^{-19} erg s $^{-1}$ cm $^{-2}$, one requires $\mu \sim 178$. This magnification is attained, on average, over an area of ~ 0.082 arcmin 2 in a typical cluster lens. This area maps to a comoving volume of ~ 0.79 cMpc 3 in the source frame of potential Pop III galaxies at $z = 10$. Applying this volume to the number density of ~ 26 cMpc $^{-3}$ reveals that a single cluster lens is about 20 times better at spectroscopically detecting Pop III galaxies than the wide-field surveys considered in this paper. Taking the number density of Pop III galaxies with age ≤ 10 Myr that is predicted by simulations, given by ~ 0.23 cMpc $^{-3}$ at $z = 10$ for the Stiavelli & Trenti (2010) model, one finds that a total of ~ 6 cluster lenses has to be targeted in order to detect one Pop III galaxy with spectroscopy. When looking at photometry, on the other hand, the scenario is different. The survey with the best performance for photometric detection of the continuum at 1500 Å that we found in this paper was the wide *RST* survey (at $z = 10$) which required a minimum number density of ~ 0.004 cMpc $^{-3}$. Performing a similar calculation for photometry shows that a cluster lens observation operating at a depth of $m_{AB} \sim 28.5$ (requiring $\mu \sim 172$ over 0.085 arcmin 2) would detect a factor of several hundreds fewer Pop III galaxies than the wide *RST* survey. Fig. 4 reveals that the wide *RST* survey would detect ~ 40 Pop III galaxies at $z = 10$, given the number densities predicted by Stiavelli & Trenti (2010). This can be compared to cluster lensing which would rather need to target ~ 5 cluster lenses in order to photometrically detect one single Pop III galaxy at $z = 10$, for the given depth of the telescope. Pushing the telescope depth to $m_{AB} \sim 29.5$, we crudely estimate that a single cluster lens could detect at least one Pop III galaxy.

These estimates showcase the utility of wide photometric surveys which manages to reach the depth required to detect Pop III galaxies with the help of gravitational lensing while simultaneously covering

very large areas on the sky. The benefit of targeting a cluster lens, where the probabilities for high magnifications are better, is greatly counteracted by the large relative difference between the areas covered by a wide-field survey and a cluster lensing survey – therefore resulting in better detection prospects for the wide-field surveys discussed in this paper. For spectroscopy, which inherently requires very deep exposures in order to detect indicative emission lines in high- z Pop III galaxies, the benefit of frequently achieving strong magnifications in cluster lenses results in very advantageous detection prospects – at a relatively small cost in survey area.

4.2 Comparison to previous studies

Previous work has showed that we are not limited by the areal number density of minihalos (the potential hosts of Pop III galaxies) in the field of view of typical telescopes (e.g. Schauer et al. 2020) – which continue to show that an *ultimately large telescope* reaching AB magnitudes of ~ 39 would be able to detect even the very first Pop III stars, without gravitational lensing.

Liu & Bromm (2020) calculate the number of Pop III galaxies per 10 arcmin 2 with total stellar mass $M_{\star} = 10^3 M_{\odot}$ as a function of redshift for a variety of metal-mixing models. At $z = 10$, they predict numbers of such Pop III galaxies as large as 4×10^3 – 2×10^4 per 10 arcmin 2 , ranging from their pessimistic to their optimistic model. In our calculations, including the effects of gravitational lensing, we find that the minimum number of Pop III galaxies per 10 arcmin 2 required for \sim one detection of an $M_{\star} = 10^3 M_{\odot}$ object with the best-performing surveys – the wide and deep *RST* photometric surveys – is given by $\sim 3.7 \times 10^4$ per 10 arcmin 2 . This once again points out the importance of investigating whether high-efficiency star formation really is ruled out. For example, considering Pop III galaxies with total stellar masses a factor of 10 higher, i.e. $M_{\star} = 10^4 M_{\odot}$ puts us at a required minimum of $\sim 3.7 \times 10^2$ Pop III galaxies per 10 arcmin 2 , at $z = 10$, for the wide and deep *RST* survey. Even more massive Pop III galaxies – as high as $M_{\star} \approx 8 \times 10^4$ – $2.5 \times 10^6 M_{\odot}$ – have been suggested by Yajima & Khochfar (2017) at redshift $z = 7$. Whether the number density of such high-mass objects is sufficiently high remains unknown.

Zackrisson et al. (2011b) show that 100-h exposures at $z = 10$ with *JWST* should be able to detect Pop III galaxies of masses $M_{\star} \sim 5 \times 10^5 M_{\odot}$, having formed stars for 10 Myr, without the use of gravitational lensing. In our surveys we require similar masses for exposures of the same depth. Looking at the range $z = 7$ –15, Zackrisson et al. (2012) showed that *JWST* imaging reaching ~ 29.7 in AB magnitude could find ≈ 1 Pop III galaxy in one single NIRCam field of view when targeting well selected cluster lenses, in this case MACS J0717.5+3745, which provides higher probabilities for larger magnifications. These predictions were based on Pop III galaxies forming in atomic cooling haloes with $\epsilon = 0.001$ and a top-heavy IMF with zero Lyman continuum leakage. As a comparison, using the randomly oriented wide or deep *RST* surveys presented in this paper, we calculate that in order to find one Pop III galaxy at $z = 10$ – that is, only in one redshift bin $\Delta z = 1$ – Stiavelli & Trenti (2010) and Visbal et al. (2020) predict number densities that are almost exactly at the requirements for photometric detection when using the star formation efficiency $\epsilon = 0.001$ – the exact number of detections depends on whether the high or low-mass data points are chosen in the simulations, but the difference is quite modest. Since we only consider one redshift bin at $z = 10$, the number of detected Pop III galaxies would be increased by a factor of a few when integrated over the whole range of redshifts $z = 7$ –15 – which would increase the number of detections by a few. Furthermore, large sky

surveys conducted by e.g. *RST* will most likely also include strong-lensing clusters – but as shown, even without deliberately targeting cluster lenses, the wide and deep *RST* photometric surveys require number densities that are more or less identical to the predictions by simulations at $z = 10$, even for $\epsilon = 0.001$. As we discussed earlier, cluster lenses would offer a better alternative for spectroscopic detection of Pop III galaxies while photometry is arguably much more effective in a wide-field survey like the ones planned with *RST*.

Zackrisson et al. (2015) explored a scenario similar to the one in this paper by looking at the possibility of finding Pop III galaxies in randomly selected fields with magnifications up to $\mu \sim 1000$. They find that given an areal number density of ≈ 300 objects arcmin $^{-2}$, a 100 deg 2 survey with a limiting magnitude of $m_{AB} = 28$ would detect at least one Pop III galaxy of mass $\sim 1.5 \times 10^4 M_\odot$ at $z = 7-13$ with magnifications $\mu \gtrsim 300$. In our predictions, we find that the wide and deep *RST* survey, covering 2000 deg 2 at $m_{AB} = 27$ and 40 deg 2 at $m_{AB} = 29$, respectively, have similar prospects for detection of a few Pop III galaxies without breaching the magnification limit ($\mu_{max} = 1000$) that we set for this paper. Once again, integration of the whole redshift range will increase the expected number of galaxies that will be detected by a factor of a few.

4.3 The impact of variations in $\mathcal{L}_{\text{He II } 1640}$ and $\mathcal{L}_{\text{UV } 1500}$

Other important factors to take into consideration are the expected deviations from the emission-line fluxes and UV continuum predicted in, e.g. Raiter et al. (2010). Combining equations (1), (5), and (6) suggests that the minimum required number density has an inverse square dependence on the He II $\lambda 1640$ Å emission-line luminosity. This is valid for the range $\mu > 10$ in which the slope of the magnification function is $P(> \mu) \propto \mu^{-2}$. Therefore, a reasonable factor of ≈ 3 variation in the luminosity of the He II $\lambda 1640$ Å emission line will shift the required number density by a factor of ≈ 9 for a given Pop III galaxy total stellar mass, therefore shifting the coloured lines in the figures by an order of magnitude. Similarly, in the region of lower magnification $\mu \approx 1-10$, the magnification probability function has a steeper μ dependence and therefore introduces even larger variations which could be relevant for objects at lower redshift. Referring to Fig. 1(a), we see that an order of magnitude decrease in the minimum required number density cannot close that gap to the Stiavelli & Trenti (2010) simulations – neither does it manage to bridge the gap to the simulations by Inayoshi et al. (2018), Visbal et al. (2020). Looking at Fig. 3, we see that across the redshift span presented, an order of magnitude decrease in the minimum required number density would render the Stiavelli & Trenti (2010) simulations detectable for $\epsilon = 0.1$ in the range $z \approx 10-13$. In order for the fiducial, Visbal et al. (2020) simulations to be made detectable (achieved by lowering the green, $\epsilon = 0.001$, line down to the green line with black square markers), one would require He II $\lambda 1640$ Å emission-line luminosities that are a factor of ≈ 100 higher than predicted in the Raiter et al. (2010) stellar population synthesis models – such variations should be regarded as highly unlikely.

Another important factor that would impact the He II $\lambda 1640$ Å emission-line luminosity is the escape fraction of He $^+$ -ionizing photons. The escape fraction of ionizing radiation has proven to be very important for the strength of HI emission lines (e.g. Wise et al. 2014). However, the extent to which He $^+$ -ionizing photons ($h\nu > 54.4$ eV) escape in the chemically pristine environments relevant for the birth Pop III galaxies, are not necessarily as severe as for metal

enriched environments. Johnson et al. (2009) showed that only newly forming galaxies with quite high total stellar mass ($2.5 \times 10^4 M_\odot$) combined with a very top-heavy IMF (in that case a stellar population of 250 stars, all with mass $100 M_\odot$) produces a significant escape fraction (~ 0.8) of He $^+$ -ionizing photons. In the case of lower total stellar mass ($2.5 \times 10^3 M_\odot$) and/or less top-heavy IMF, the escape fraction is entirely negligible – which highlights the importance of the star formation efficiency and Pop III IMF. The simulations included in our paper spans a range of total stellar masses which is highly dependant on our choice of star formation efficiency. The Visbal et al. (2020) models have quite low star formation efficiency built into their models, and would therefore avoid the problem with a high escape fraction, while our fiducial choice for star formation efficiency in the Stiavelli & Trenti (2010) and Inayoshi et al. (2018) models produce Pop III galaxies with masses as high as $\sim 10^4 M_\odot$ which could therefore stand the risk of suffering from a large He $^+$ ionizing escape fraction – this would decrease the strength of the He II $\lambda 1640$ Å emission line and subsequently require higher Pop III galaxy number densities. The top-heavy IMF used in our paper has a characteristic mass of $m_c = 60 M_\odot$, $\sigma = 1$, and an allowed mass range of $1-500 M_\odot$. For a total stellar mass of $2.5 \times 10^4 M_\odot$ this would produce ~ 80 stars with mass above $100 M_\odot$ – which may result in a considerable He II $\lambda 1640$ Å escape fraction according to Johnson et al. (2009).

Similarly, variations in the continuum level at 1500 Å introduces the same order of magnitude changes in the minimum number density required for photometric detection – as in the case for spectroscopic detection of the He II $\lambda 1640$ Å emission line. Looking at Fig. 1(b), we see that an order of magnitude downward shift of the coloured lines would suggest that even the galaxies simulated in the fiducial Visbal et al. (2020) models could be detectable by wide/deep *RST* photometric surveys – whereas the number of densities predicted by Inayoshi et al. (2018) continues to remain at a level much too low even for photometric detection.

4.4 Simulation predictions versus reality

The accomplishment of this paper is unfortunately to have shown that we should not expect to spectroscopically confirm Pop III galaxies in future planned surveys targeting large and randomly selected areas in the sky (random in the sense that the areas are not specifically targeted based on large magnifications from gravitational lensing). This, however, hinges on the predictions of contemporary simulations regarding the number density of atomic cooling haloes and the expected total stellar mass of the Pop III galaxies embedded in them – and our analysis helps to separate the detectable and undetectable parts of the parameter space. This hopefully compels the scientific community to further explore the possibilities of achieving high star formation efficiencies without making significant sacrifices with the number density of Pop III galaxies. At this moment, we cannot say with certainty that such high-efficiency models are ruled out and should therefore consider them seriously.

A candidate Pop III galaxy was recently reported by Vanzella et al. (2020), showing strong Ly α emission at $z \approx 6.6$, which may suggest that the lower star formation efficiencies provided by simulations are actually sufficient for spectroscopic detection of the Ly α emission lines. Applying the computational machinery used in this paper with the Ly α luminosity predicted in Raiter et al. (2010) for our fiducial lognormal top-heavy IMF ($m_c = 60 M_\odot$, with mass range $1-500 M_\odot$), assuming 50 per cent transmission of Ly α through the intergalactic medium, at a source redshift of $z = 7$ (the lowest

redshift that we have included simulations for) – our model shows that Visbal et al. (2020) predict number densities that are high enough to detect the Ly α emission line in ~ 1 Pop III galaxy of total stellar mass $\approx 4 \times 10^4 - 2 \times 10^5 M_\odot$ using the very deep *JWST* survey. This neither defers nor confirms whether Vanzella et al. (2020) actually found a true Pop III galaxy. Rather, we see that finding lensed Ly α emitting Pop III galaxies close to the end of reionization should statistically be quite plausible, assuming that the transmission of Ly α through the intergalactic medium is sufficient.

4.5 Other spectroscopic probes

Having focused on the He II $\lambda 1640$ emission line as the best observational signature for detection of pristine Pop III galaxies, we have perhaps overlooked the utility of other emission lines that are useful in the search for Pop III galaxies. We mention the Ly α line as a very prominent emission line that could differentiate a Pop III galaxy from enriched stellar populations. This, however, restricts us to redshifts after reionization has completed due to the absorption of Ly α in the intergalactic medium. This does not render the line useless since we also note from Fig. 3 that lower redshifts may indeed be very promising in the search for Pop III galaxies. As shown in, for example, Liu & Bromm (2020), small pockets of pristine gas is expected to remain after reionizations ends, enabling the study of Ly α emission from Pop III stars in some rare regions at lower redshift. Furthermore, starting with *JWST* we will finally be able to probe the rest-frame optical at the redshift frontier, such as the Balmer emission lines – providing us with yet another useful tool in the search for Pop III. Being unaffected by the neutral intergalactic medium, the H α emission line will be within the wavelength range of *JWST* up to $z \approx 7$ using NIRSpec, while the H β emission line remains within range up to $z \approx 10$. These emission lines will generally be stronger than the He II $\lambda 1640$ emission line, and will therefore be more easily detected at high redshift. For our fiducial IMF, star-forming age ($\tau \approx 10$ Myr) at zero metallicity, we have that the ratio of hydrogen-ionizing photons to He $^+$ ionizing photons is given by $\frac{Q_{\text{H}}}{Q_{\text{He}^+}} = 24.6$ for case B recombination (Schaerer 2002; Raiter et al. 2010) – suggesting that the H α and H β luminosities are ≈ 5.3 and 1.9 times higher than the He II $\lambda 1640$ luminosity respectively – however, these ratios will vary with IMF, age, etc. In the best circumstances (strongest He II lines), we may find $\frac{Q_{\text{H}}}{Q_{\text{He}^+}} \sim 10$, leading to H α and H β luminosities a factor of ~ 2.2 stronger and a factor of ~ 0.77 weaker, respectively, when compared to the He II $\lambda 1640$ luminosity. Furthermore, the He II $\lambda 4686$ emission line can be used as a probe for a highly ionizing stellar population, but this line is a factor of ~ 8 weaker than the He II $\lambda 1640$ emission line.

If very strong Balmer line equivalent widths are detected together with a complete lack of metal lines, one will have a good case for a Pop III galaxy or a galaxy with extremely metal-poor chemical composition. Inoue (2011) showed that extremely metal-poor galaxies may be identified using constraints on the rest frame equivalent width (EW) of the Ly α ; (EW(Ly α) > 460 Å), He II $\lambda 1640$; (EW(He II) > 1 Å), O III; (EW(O III) < 20 Å), H α ; (EW(H α) > 3200 Å) and H β ; (EW(H β) > 540 Å) emission lines – where the limits are stringent enough to avoid confusion with observations of very young (~ 1 Myr) galaxies of higher metallicity. The O[III] $\lambda 5007$ emission line would serve as the best indicator of metals when performing spectroscopy on selected targets. Inoue (2011) continues to show that a ratio O[III]/H β < 0.1 is sufficient to identify a galaxy as metal-free or extremely metal poor – which can be confirmed for SFR $\sim 3 M_\odot$ galaxies up to $z \sim 8$ in deep exposures with NIRSpec on the *JWST*.

4.6 Spectroscopic follow-up

Photometrically detected objects showing indicative Pop III colour signatures may in principle be targeted for follow-up spectroscopy – rendering the He II $\lambda 1640$ emission-line detectable for sufficiently deep exposures. As shown in Grisdale et al. (2021), Pop III galaxies with total stellar masses of $\geq 10^4 M_\odot$, a top-heavy IMF and a compact stellar distribution would have a detectable He II $\lambda 1640$ emission line for redshifts $z = 4-10$ with the HARMONI spectrograph on the *ELT*. In our paper, we find that the sensitivity of the *ELT* ($F_{\text{limit}} \sim 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$) is sufficient to allow spectroscopic follow-up on Pop III galaxies identified with deep *RST* photometry across the whole relevant redshift range – restricted only by the maximum magnification which limits us to $z \lesssim 11$ for the fiducial Visbal et al. (2020) while only placing a restriction on redshift for the high-efficiency model and the Stiavelli & Trenti (2010) model due to redshifting out of the *RST* filters (see Fig. 4).

5 CONCLUSIONS

Here, we summarize the key conclusions and take-home messages from this paper:

(i) Simulations predict number densities that are likely too low for spectroscopic detection of gravitationally lensed Pop III galaxies in wide-area surveys using *JWST*, *RST*, and *Euclid*. Allowing magnifications $\mu > 1000$, we find that all included telescopes provide similar prospects for spectroscopic detection. However, due to limitations in the highest applicable magnification when observing extended objects like galaxies, *JWST* is the telescope that is best poised for spectroscopic detections of the He II $\lambda 1640$ emission line. Photometric surveys will however, very likely, pick up lensed Pop III galaxies if star formation efficiencies are sufficiently high ($\epsilon \gtrsim 0.005$) but this is model-dependent since not all simulations predict number densities that are sufficient for detection. When it comes to photometric detection of the UV continuum at 1500 Å, the wide and deep *RST* survey performs better than all other telescopes and surveys – requiring Pop III galaxy number densities that are ~ 2 orders of magnitude lower than required for detection in planned photometric *JWST* surveys and ~ 1.5 orders of magnitude lower than required for the wide *Euclid* survey (see e.g. Fig. 1). For the wide *Euclid* survey, we find minimum required number densities that are sufficient for photometric detections based on the Stiavelli & Trenti (2010) simulations for $\epsilon = 0.01$, although while generally breaching the maximum allowed magnification set in this paper for all included redshifts.

(ii) The star formation efficiency of Pop III star formation is paramount to the ability of detecting Pop III galaxies in gravitationally lensed fields due to its correlation with the total stellar mass formed, and therefore the brightness of Pop III galaxies. The poor prospects for spectroscopic detection of the He II $\lambda 1640$ emission line, using our fiducial IMF and star-forming age (τ), can only be remedied if star formation efficiencies would reach $\epsilon \gtrsim 0.1$ while at the same time keeping the Pop III galaxy number density unchanged (see Fig. 3).

(iii) We conclude that the top-heavy lognormal IMF with characteristic mass $m_c = 60 M_\odot$ and mass range $1-500 M_\odot$ provide the best circumstances for spectroscopic detection among the IMFs included in this paper, but it still requires number densities about two orders of magnitude higher than the predictions from simulations. The less top-heavy IMFs included in this paper require number densities an additional ~ 1.5 orders of magnitude higher for spectroscopy

(see Fig. 2) and ~ 1 orders of magnitude higher for photometry (Section 3.2).

(iv) We find that during the first 10 Myr after the onset of star formation, the required number densities for spectroscopic and/or photometric detection varies very little, only at the level of a few per cent based on the luminosity of the galaxies. Looking instead at a shorter burst of star formation, e.g. 3 Myr, does, however, imply a brighter galaxy (higher SFR) given that the same total stellar mass is to be formed. This improves the detection prospects by lowering the minimum required number densities by about 1 order of magnitude while simultaneously lowering the number of pristine Pop III galaxies predicted by simulations by 0.5 orders of magnitude due to the shorter observational time window. This effectively reduces the gap between the simulations and our minimum required number density with ~ 0.5 orders of magnitude.

(v) In this paper, simulations of Pop III galaxies come closest to the minimum required number densities for detection at $z \approx 10$ (see Fig. 3 and Section 3.4). Whether our required minimum number density will coincide with the number densities predicted by simulations at lower redshift is not yet clear and therefore remains a topic of future research.

(vi) Considering variations from the spectral synthesis model used in this paper (Raiter et al. 2010), we find that a factor of ≈ 3 increase in the predicted He II $\lambda 1640$ emission-line luminosity corresponds to an order of magnitude decrease in the number density required for spectroscopic detection. At, e.g. $z = 10$, this would not manage to render the Stiavelli & Trenti (2010) simulations detectable, assuming $\epsilon \approx 0.01$, and cannot bridge the gap to the simulations by Visbal et al. (2020) for reasonable variations in the emission-line luminosity (see Section 4.3).

(vii) We make order of magnitude estimates which argues that cluster-lensing surveys perform better than wide-field surveys for spectroscopic detection of Pop III galaxies. However, we argue that for photometric detection, the wide fields surveys perform considerably better. For every Pop III galaxy detected in a wide-field survey, we estimate that a single cluster lens would detect a factor of $\gtrsim 10$ more galaxies with spectroscopy. In the case of photometry, one would instead require observations of several hundreds of cluster lenses for every galaxy detected in a wide-field survey in order to be competitive.

(viii) Gravitationally lensed Pop III galaxies detected in the photometric deep *RST* survey would be sufficiently bright for spectroscopic follow-up on the He II $\lambda 1640$ emission line using e.g. the *ELT*.

ACKNOWLEDGEMENTS

We thank the anonymous referee for comments that improved the scientific content of the paper. AV, EZ, and MS acknowledge funding from the Swedish National Space Agency. EV acknowledges support from NSF grant AST-2009309. MS acknowledges support from the Swedish Collegium for Advanced Study. EZ also acknowledges a sabbatical fellowship from AI4Research at Uppsala University.

DATA AVAILABILITY

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

REFERENCES

Akeson R. et al., 2019, preprint ([arXiv:1902.05569](https://arxiv.org/abs/1902.05569))
Barkana R., Haimes Z., Ostriker J. P., 2001, *ApJ*, 558, 482

- Bouwens R. J. et al., 2011, *ApJ*, 737, 90
Bromm V., 2013, *Rep. Prog. Phys.*, 76, 112901
Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
Chandrasekhar S., 1964, *ApJ*, 140, 417
Eisenstein D. J., Ferruit P., Rieke M. J., 2017, NIRC*am*-NIRSpec galaxy assembly survey - GOODS-N, JWST Proposal. Cycle 1
Feltre A. et al., 2020, *A&A*, 641, A118
Ferruit P., 2017, NIRSpec WIDE MOS Survey - COSMOS, JWST Proposal. Cycle 1
Finkelstein S. L. et al., 2017, The Cosmic Evolution Early Release Science (CEERS) Survey, JWST Proposal ID 1345. Cycle 0 Early Release Science
Finkelstein S. L. et al., 2021, The Webb Deep Extragalactic Exploratory Public (WDEEP) Survey: Feedback in Low-Mass Galaxies from Cosmic Dawn to Dusk, JWST Proposal. Cycle 1
Grisdale K., Thatte N., Devriendt J., Pereira-Santaella M., Slyz A., Kimm T., Dubois Y., Yi S. K., 2021, *MNRAS*, 501, 5517
Haemmerlé L., Woods T. E., Klessen R. S., Heger A., Whalen D. J., 2018, *MNRAS*, 474, 2757
Heger A., Woosley S. E., 2002, *ApJ*, 567, 532
Hilbert S., White S. D. M., Hartlap J., Schneider P., 2008, *MNRAS*, 386, 1845
Hirano S., Bromm V., 2018, *MNRAS*, 476, 3964
Hirano S., Yoshida N., Sakurai Y., Fujii M. S., 2018, *ApJ*, 855, 17
Hosokawa T., Hirano S., Kuiper R., Yorke H. W., Omukai K., Yoshida N., 2016, *ApJ*, 824, 119
Inayoshi K., Li M., Haimes Z., 2018, *MNRAS*, 479, 4017
Inoue A. K., 2010, *MNRAS*, 401, 1325
Inoue A. K., 2011, *MNRAS*, 415, 2920
Jaacks J., Thompson R., Finkelstein S. L., Bromm V., 2018, *MNRAS*, 475, 4396
Johnson J. L., Aykatalp A., 2019, *ApJ*, 879, 18
Johnson J. L., Greif T. H., Bromm V., 2008, *MNRAS*, 388, 26
Johnson J. L., Greif T. H., Bromm V., Klessen R. S., Ippolito J., 2009, *MNRAS*, 399, 37
Kulkarni M., Visbal E., Bryan G. L., 2019, *ApJ*, 882, 178
Labbe I. et al., 2021, UNCOVER: Ultra-deep NIRC*am* and NIRSpec Observations Before the Epoch of Reionization, JWST Proposal. Cycle 1
Latif M. A., Khochfar S., Whalen D., 2020, *ApJ*, 892, L4
Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
Liu B., Bromm V., 2020, *MNRAS* 497 2839–2854
Magg M., Hartwig T., Agarwal B., Frebel A., Glover S. C. O., Griffen B. F., Klessen R. S., 2018, *MNRAS*, 473, 5308
Marchetti L., Serjeant S., Vaccari M., 2017, *MNRAS*, 470, 5007
Mas-Ribas L., Dijkstra M., Forero-Romero J. E., 2016, *ApJ*, 833, 65
Nanayakkara T. et al., 2019, *A&A*, 624, A89
Oesch P. et al., 2021, FRESCO: The First Reionization Epoch Spectroscopic COmplete Survey, JWST Proposal. Cycle 1
Oesch P. A. et al., 2012, *ApJ*, 759, 135
Peacock J. A., 1982, *MNRAS*, 199, 987
Pello R., Schaerer D., 2003, Rev. Mex. Astron. Astrofis. Conf. Ser., 16, 225
Planck Collaboration XVI, 2014, *A&A*, 571, A16
Raider A., Schaerer D., Fosbury R. A. E., 2010, *A&A*, 523, A64
Regan J. A., Wise J. H., Woods T. E., Downes T. P., O'Shea B. W., Norman M. L., 2020, *Open J. Astrophys.*, 3, 15
Rieke M. et al., 2019, Bull. Am. Astron. Soc., 51, 45
Robertson A., Smith G. P., Massey R., Eke V., Jauzac M., Bianconi M., Ryczanowski D., 2020, *MNRAS*, 495, 3727
Rydberg C.-E., Zackrisson E., Lundqvist P., Scott P., 2013, *MNRAS*, 429, 3658
Rydberg C.-E., Whalen D. J., Maturi M., Collett T., Carrasco M., Magg M., Klessen R. S., 2020, *MNRAS*, 491, 2447
Sarmiento R., Scannapieco E., Cohen S., 2018, *ApJ*, 854, 75
Sarmiento R., Scannapieco E., Côté B., 2019, *ApJ*, 871, 206
Saxena A. et al., 2020, *A&A*, 636, A47
Schaerer D., 2002, *A&A*, 382, 28
Schaerer D., 2003, *A&A*, 397, 527
Schauer A. T. P., Glover S. C. O., Klessen R. S., Clark P., 2020, *MNRAS*, 507, 1775

- Skinner D., Wise J. H., 2020, *MNRAS*, 492, 4386
- Stacy A., Bromm V., Loeb A., 2011, *MNRAS*, 413, 543
- Stacy A., Greif T. H., Klessen R. S., Bromm V., Loeb A., 2013, *MNRAS*, 431, 1470
- Stiavelli M., Trenti M., 2010, *ApJ*, 716, L190
- Surace M. et al., 2018, *ApJ*, 869, L39
- Surace M., Zackrisson E., Whalen D. J., Hartwig T., Glover S. C. O., Woods T. E., Heger A., Glover S. C. O., 2019, *MNRAS*, 488, 3995
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Trenti M., Stiavelli M., Shull J. M., 2009, *ApJ*, 700, 1672
- Treu T. L. et al., 2017, Through the Looking GLASS: A JWST Exploration of Galaxy Formation and Evolution from Cosmic Dawn to Present Day, JWST Proposal ID 1324. Cycle 0 Early Release Science
- Tumlinson J., 2006, *ApJ*, 641, 1
- Vanzella E. et al., 2020, *MNRAS*, 494, L81
- Vikaews A., Zackrisson E., Binggeli C., 2020, *MNRAS*, 492, 1706
- Visbal E., Bryan G. L., Haiman Z., 2017, *MNRAS*, 469, 1456
- Visbal E., Bryan G. L., Haiman Z., 2020, *ApJ*, 897, 95
- Whalen D. J., Surace M., Bernhardt C., Zackrisson E., Pacucci F., Ziegler B., Hirschmann M., 2020, *ApJ*, 897, L16
- Windhorst R. A. et al., 2017, JWST Medium-Deep Fields - Windhorst IDS GTO Program, JWST Proposal. Cycle 1
- Windhorst R. A. et al., 2018, *ApJS*, 234, 41
- Wise J. H., Cen R., 2009, *ApJ*, 693, 984
- Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, *MNRAS*, 442, 2560
- Yajima H., Khochfar S., 2017, *MNRAS*, 467, L51
- Yoon S. C., Dierks A., Langer N., 2012, *A&A*, 542, A113
- Zackrisson E. et al., 2012, *MNRAS*, 427, 2212
- Zackrisson E., Inoue A. K., Rydberg C.-E., Duval F., 2011a, *MNRAS*, 418, L104
- Zackrisson E., Rydberg C.-E., Schaerer D., Östlin G., Tuli M., 2011b, *ApJ*, 740, 13
- Zackrisson E., González J., Eriksson S., Asadi S., Safranek-Shrader C., Trenti M., Inoue A. K., 2015, *MNRAS*, 449, 3057

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