On the Observability of Individual Population III Stars and Their Stellar-mass Black Hole Accretion Disks through Cluster Caustic Transits

Rogier A. Windhorst[1] , F. X. Timmes[1] , J. Stuart B. Wyithe[2] , Mehmet Alpaslan[3] , Stephen K. Andrews[4], Daniel Coe[5] ,
Jose M. Diego[6] , Mark Dijkstra[7], Simon P. Driver[4] , Patrick L. Kelly[8] , and Duho Kim[1]

[1] School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA; Rogier.Windhorst@asu.edu, Francis.Timmes@asu.edu
[2] University of Melbourne, Parkville, VIC 3010, Australia; SWyithe@physics.unimelb.edu.au
[3] New York University, Department of Physics, 726 Broadway, Room 1005, New York, NY 10003, USA
[4] The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
[5] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[6] IFCA, Instituto de Fisica de Cantabria (UC-CSIC), Avenida de Los Castros s/n, E-39005 Santander, Spain
[7] Institute of Theoretical Astrophysics, University of Oslo, NO-0315 Oslo, Norway
[8] University of California at Berkeley, Berkeley, CA 94720-3411, USA

Abstract

We summarize panchromatic Extragalactic Background Light data to place upper limits on the integrated near-infrared surface brightness (SB) that may come from Population III stars and possible accretion disks around their stellar-mass black holes (BHs) in the epoch of First Light, broadly taken from $z \simeq 7$–17. Theoretical predictions and recent near-infrared power spectra provide tighter constraints on their sky signal. We outline the physical properties of zero-metallicity Population III stars from MESA stellar evolution models through helium depletion and of BH accretion disks at $z \gtrsim 7$. We assume that second-generation non-zero-metallicity stars can form at higher multiplicity, so that BH accretion disks may be fed by Roche-lobe overflow from lower-mass companions. We use these near-infrared SB constraints to calculate the number of caustic transits behind lensing clusters that the *James Webb Space Telescope* and the next-generation ground-based telescopes may observe for both Population III stars and their BH accretion disks. Typical caustic magnifications can be $\mu \simeq 10^4$–10^5, with rise times of hours and decline times of $\lesssim 1$ year for cluster transverse velocities of $v_T \lesssim 1000 \, \mathrm{km \, s^{-1}}$. Microlensing by intracluster-medium objects can modify transit magnifications but lengthen visibility times. Depending on BH masses, accretion-disk radii, and feeding efficiencies, stellar-mass BH accretion-disk caustic transits could outnumber those from Population III stars. To observe Population III caustic transits directly may require monitoring 3–30 lensing clusters to AB $\lesssim 29$ mag over a decade.

Key words: accretion, accretion disks – galaxies: clusters: general – gravitational lensing: strong – infrared: diffuse background – stars: black holes – stars: Population III

1. Introduction

In this paper, we consider if the *James Webb Space Telescope* (*JWST*; Rieke et al. 2005; Gardner et al. 2006; Windhorst et al. 2008; Beichman et al. 2012) can observe First Light objects directly. *JWST*'s Near-InfraRed Camera (NIRCam) is expected to reach medium-deep to deep (AB $\simeq$ 28.5–29 mag) flux limits routinely, and in ultradeep surveys perhaps as faint as AB $\simeq$ 30–31 mag, once *JWST*'s on-orbit stray-light properties are mapped.

Unlensed Population III (Pop III) stars or their stellar-mass black hole (BH) accretion disks may have fluxes of AB $\simeq$ 35–43 mag at $z \simeq$ 7–25, and therefore are not directly detectable by *JWST*, not even via ordinary gravitational lensing targets (e.g., Rydberg et al. 2013), which typically have magnification factors of $\mu \simeq$ 10 or ~2.5 mag (e.g., Lotz et al. 2017). We use "μ" throughout to indicate the lensing magnification factor, and "SB" to indicate surface brightness.

However, cluster caustic transits, when a compact rest-frame UV source transits a caustic due to the cluster motion in the sky, or perhaps due to significant velocity substructure in the cluster, have great potential for magnifying such compact objects temporarily by factors of $\mu \simeq 10^3$–10^5 (e.g., Miralda-Escude 1991; Zackrisson et al. 2015; Diego et al. 2017; Kelly et al. 2017, 2018; Rodney et al. 2017). This could temporarily boost the brightness of a very compact object by

$\mu \simeq$ 7.5–12.5 mag, which may render it observable by *JWST*. If Pop III stars—and/or their resulting BH accretion disks—are numerous enough in the sky, it is therefore possible that individual Pop III stars or their BH accretion disks are temporarily lensed by foreground cluster caustics as the cluster transits across the background Pop III target. This could render an AB $\simeq$ 35–41.5 mag Pop III star at redshifts $z \simeq$ 7–17 temporarily visible to a medium-deep or deep (AB $\simeq$ 28.5–29 mag), well time-sequenced set of *JWST* observations.

The 2016 *Planck* results (Planck Collaboration et al. 2016a, 2016b, 2016d) reduced the polarization optical depth even further from earlier values—and reduced its errors—to $\tau \simeq 0.058 \pm 0.012$, thereby placing the redshift of reionization at approximately $z_{\mathrm{reion}} \simeq 7.8 \pm 0.9$ *if* it had occurred instantaneously. Sobral et al. (2015) discovered an object at $z \simeq 6.7$ with both a clear Lyα 1216 Å line and a possible He 1640 Å line, which may indicate a late, pristine stellar population dominated by very hot stars, possibly Pop III stars. That is, the Pop III star epoch may have ended around $z \simeq 7$ and could have started very early, at $z \gtrsim$ 20–40 (Trenti & Stiavelli 2009). Of course, at $z \gg 30$, the luminosity distance would be very large and render most Pop III stars fainter than $\gtrsim 43$ mag. In the hierarchical simulations of Sarmento et al. (2017, 2018), most of the early star formation (SF) occurs between $z \simeq 20$, when the star-forming population consists predominantly of pristine Pop III stars, and $z \simeq 7$, when the population is predominantly

polluted with metallicities of $Z \gtrsim 10^{-4} Z_\odot$. In this paper, we will therefore adopt a redshift range of $z \simeq 12 \pm 5$, where we may observe Pop III stars or their BH accretion disks directly with *JWST if* they are sufficiently strongly lensed during a cluster caustic transit. For brevity, we will take "Pop III" hereafter to include any objects at $z \gtrsim 7$ that may have been already (slightly) polluted by First Light objects.

To discuss the possibilities of cluster caustic transits by Pop III objects, we need to address four different main topics. In Section 2, we summarize constraints to the possible sky surface brightness (SB) from objects at $z \gtrsim 7$, which is the foremost constraint that we must first understand before we can predict a frequency of potential cluster caustic transits. In Section 3, we present the physical properties of Pop III stars from stellar evolution models with Hertzsprung–Russell (HR) diagrams through the hydrogen-depletion and helium-depletion stages and from these derive their mass–luminosity (ML) relation, their bolometric+IGM+K-corrections, and their relative contribution to the luminosity density in faint star-forming objects. In Section 4, we evaluate limits to the typical transverse velocities of massive lensing clusters, their typical caustic lengths, and the possible effects from microlensing, and estimate the cluster caustic transit times and rates for the Pop III star parameters from Section 3. In Section 5, we discuss the possible physical properties of Pop III stellar-mass BH accretion disks, and under what conditions these may be fed from early massive stellar binaries. In Section 6, we present estimates of the cluster caustic transit rates that may result from BH accretion disks. In Section 7, we discuss what a cluster caustic transit observing program for Pop III objects with *JWST* might look like. In Section 8, we summarize our conclusions.

Throughout, we use *Planck* cosmology (Planck Collaboration et al. 2016b): $H_0 = 66.9 \pm 0.9$ km s^{-1} Mpc^{-1}, matter density parameter $\Omega_m = 0.32 \pm 0.03$, and vacuum energy density $\Omega_\Lambda = 0.68 \pm 0.03$, resulting in a Hubble time of 13.8 Gyr. When quoting magnitudes, our fluxes are all in AB magnitudes (hereafter AB-mag), and our SB values are in AB-mag arcsec^{-2} (Oke & Gunn 1983), using $S_\nu = 10^{-0.40(AB - 8.90 \text{ mag})}$ in Jy.

2. Constraints to the Sky Surface Brightness from Objects at $z \gtrsim 7$

Before we can estimate the number of possible cluster caustic transits of Pop III objects, we must estimate the maximum possible contribution of Pop III stars and their stellar-mass BH accretion disks to the observed near-IR sky surface brightness. In Figure 1 and Sections 2.1–2.3, we therefore summarize the available data on the Extragalactic Background Light (EBL) that are directly relevant to our caustic transit calculations in Sections 4.4 and 6.2. Throughout, "EBL" will refer to the total Extragalactic Background Light, including any diffuse EBL component, while "iEBL" will refer to the integrated EBL extrapolated from the discrete galaxy counts.

2.1. Constraints from the Discrete Extragalactic Background Light

In Figure 1, the green unfilled squares at 2–3 μm indicate the Kelsall et al. (1998) COBE DIRBE sky-SB from zodiacal light, which is scattered sunlight. At 3–200 μm, these COBE DIRBE points are dominated by the $\sim$200 K thermal dust component in the zodiacal belt. Most of this dust is piled up in the asteroid

belt and is clearly a limiting factor for near–mid-IR observations, including for *JWST* observations at $\lambda \gtrsim 3.5$ μm. Figure 1 plots with solid green points the zodiacal foreground as measured from low-Earth orbit using the panchromatic *Hubble Space Telescope* (*HST*) Wide Field Camera 3 (WFC3) Early Release Science (ERS) observations of Windhorst et al. (2011, hereafter W11) and its precursor data from the Great Orbiting Observatories Deep Survey (GOODS) Advanced Camera for Surveys (ACS) data (Giavalisco et al. 2004). This includes the zodiacal sky measurements in the Hubble Ultra Deep Field (HUDF) by Hathi et al. (2008). The green dotted line is the solar energy spectrum (Kurucz 2005) normalized to these *HST* data.

All units in Figure 1 have been converted to νI_ν in units of nW m^{-2} sr^{-1}. For reference, 1.00 nW m^{-2} sr^{-1} corresponds to 28.41 mag arcsec^{-2} at 2.00 μm, which is indicated by the orange K-band SB-scale in AB-mag arcsec^{-2} on the right vertical axis of Figure 1. At other near-IR wavelengths, one can derive the SB-scale corresponding to the νI_ν scale on the left by adding $-2.5 \log (\lambda/2.0 \mu m)$ to the K-band scale on the right.

An important comment on the WFC3 ERS data of W11 is in order here. Figure 1 shows that 8 of the 10 ERS filters have sky-background measurements in line with the zodiacal foreground at those wavelengths. However, their bluest and reddest filters (WFC3/UVIS F225W and WFC3/IR F160W) have a sky level significantly in excess of the zodiacal foreground for this ecliptic latitude. This was expected for the F225W filter, as this bluest WFC3 filter was intentionally scheduled at the end of each available *HST* orbit, so that any Earthshine would add some sky level to the highly readnoise-limited UV images, since the zodiacal sky is darkest at the shortest *HST* wavelengths. Indeed, the resulting F225W background level was significantly higher than that expected from the zodiacal sky alone. In all other ERS filters, every possible effort was made to avoid the Earth's limb, but this was not fully successful for the WFC3/IR filter F160W, and its resulting sky background was $\sim$0.3 dex higher than expected, despite our scheduling attempts to avoid this. In the remaining eight ERS filters, the root mean square (rms) variation from the best-fit normalized solar energy spectrum is 10%–20%, illustrating that even in the case of requesting *HST* "LOW-SKY" observations—and going to great lengths in the Astronomer's Proposal Tool (APT) scheduling requests to make sure that the sky background remains close to the theoretical zodiacal minimum—some Earthshine may have nonetheless leaked into the low-Earth orbit observations.

The red dots in Figure 1 indicate the integrated EBL measurements derived from the panchromatic (0.1–500 μm) discrete galaxy counts from *GALEX*, *HST*, ground-based, *Spitzer*, *WISE*, and *Herschel* surveys, as summarized in Driver et al. (2016, hereafter D16), which incorporated the panchromatic *HST* galaxy counts at $\lambda \simeq 0.2$–2 μm to AB $\lesssim$ 29–30 mag discussed in W11. From 0.1–500 μm, the discrete galaxy counts converge well at almost all wavelengths, except for the less deep *Spitzer*/*WISE* galaxy counts at 8–12 μm, where the galaxy count extrapolation that yields the iEBL integral is $\sim$40% uncertain. Typically, the normalized differential galaxy counts in D16 reach a peak at AB $\simeq$ 19–25 mag, where most of their iEBL energy is contained. At all wavelengths except 8–12 μm, the normalized differential counts converge—with a slope flatter than 0.4 dex/mag—to a finite sky integral that results in a well-determined iEBL value
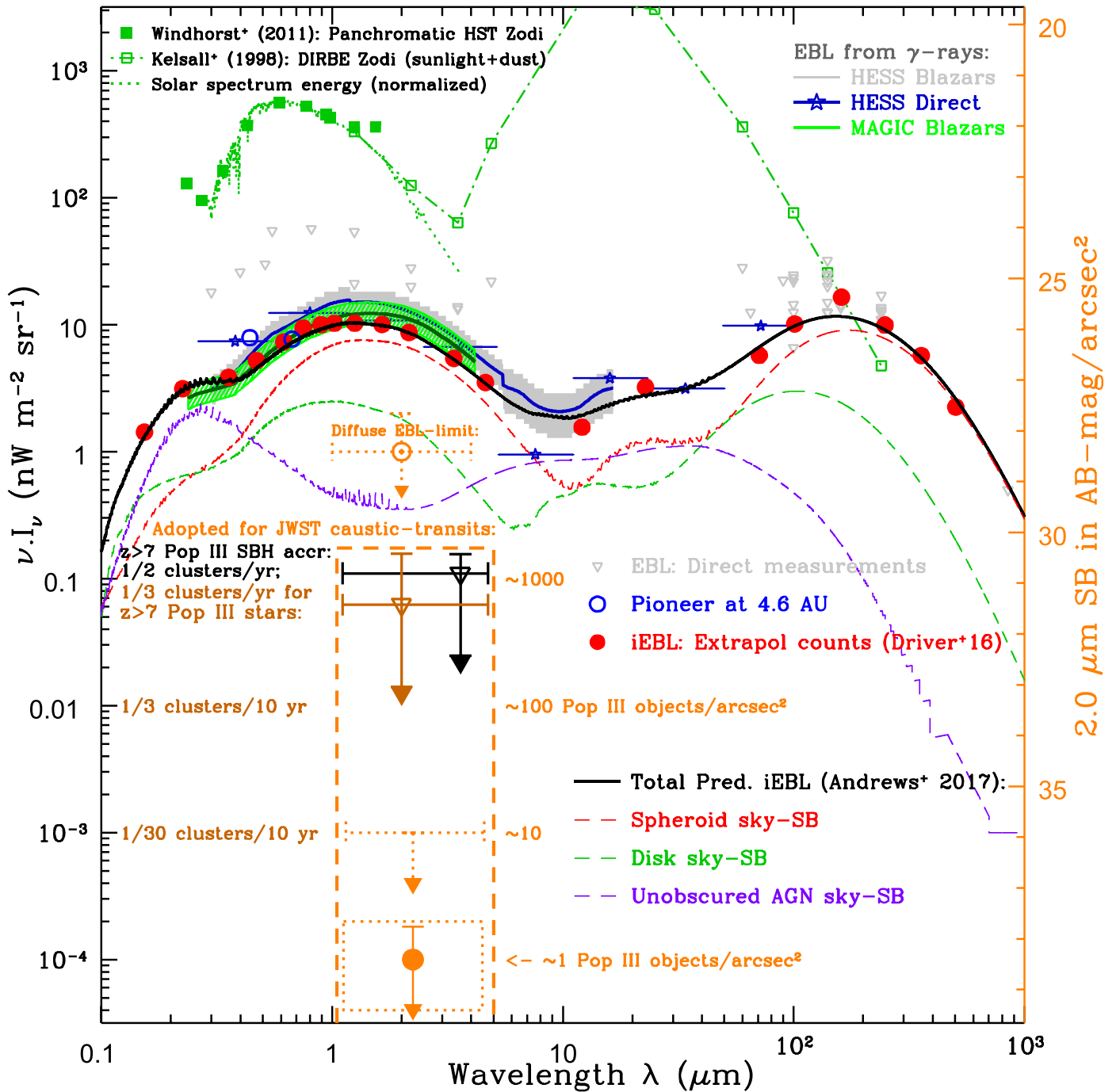
Figure 1. Summary of panchromatic backgrounds relevant to possible cluster caustic transits of Pop III stars and their stellar-mass black hole accretion disks. Green dotted–dashed lines with green unfilled squares indicate the scattered and thermal zodiacal foreground of Kelsall et al. (1998). Filled green squares indicate the panchromatic on-orbit zodiacal (labeled "Zodi") foreground values measured by *HST* (Windhorst et al. 2011). Light gray unfilled triangles indicate *direct* measurements of the Extragalactic Background Light from low-Earth orbit or L2 (for a review, see Dwek & Krennrich 2013). Blue unfilled circles indicate the direct *Pioneer* spacecraft EBL values measured beyond most of the zodiacal dust at 4.6 au. Red filled circles indicate the integrated and extrapolated (to AB $\gtrsim$ 30 mag) panchromatic galaxy counts (iEBL) of Driver et al. (2016 and references therein). The dashed red, green, and purple lines are iEBL model predictions for spheroids, disks, and unobscured AGNs, respectively (Andrews et al. 2017b). The solid black line is their total predicted iEBL. EBL constraints from H.E.S.S. γ-ray blazars are plotted as the light gray shaded region plus its dark-blue best fit, and for MAGIC blazars as the green shaded region with its dark-green best fit. The orange unfilled circle with dotted range is our "hard" upper limit for the diffuse 1–4 μm EBL, denoted as "Diffuse EBL-limit." The orange dashed box contains our adopted upper limits on the 1–4 μm near-IR sky-SB for Pop III stars at $z \gtrsim 7$ (dark orange) and for their stellar-mass BH accretion disks at $z \gtrsim 7$ (black). The possible range in SB from Pop III objects is indicated at the level of $\sim$1, 10, 100, and 1000 objects/arcsec². The filled orange circle indicates the approximate SB level of $\sim$1 Pop III star/arcsec². Cluster caustic transit rates that may be observed with *JWST* are listed in dark orange on the left for three SB levels, ranging from $\sim$1 caustic transit per three clusters per year to $\sim$1 per 30 clusters *if* monitored over 10 years. This is the lowest rate *JWST* could detect in a dedicated, large multiyear program. Details are given in Sections 2–7.

for discrete objects to within 10%–20%, including random errors, count extrapolation errors, and cosmic variance that were determined through Monte Carlo simulations. For clarity,

error bars are omitted from Figure 1, but these can be found in D16. The iEBL from discrete objects is thus well-determined to within $\lesssim$20% in general, as indicated by the small scatter in

the red dots in Figure 1 compared to the iEBL models of Andrews et al. (2017b).

The red, green, and purple dashed lines indicate the contributions that spheroids, disks, and unobscured active galactic nuclei (AGNs) at $z \lesssim 6$ may contribute to the EBL energy, following Andrews et al. (2017a, 2017b). Obscured AGNs in these models are incorporated into the spheroidal galaxies and not plotted separately. The contribution from unobscured AGNs to the discrete iEBL is uncertain, but at their median redshift of $z_{med} \simeq 2$, AGNs may produce enough total rest-frame UV radiation (at $\lambda_{rest} \gtrsim 912$–$1216$ Å) to contribute significantly to the observed near-UV background ($\lambda_{obs} \lesssim 0.4\,\mu$m). Even below $\lambda_{rest} \simeq 912$ Å, AGNs at $z \simeq 2$–3 may produce non-negligible LyC radiation (possibly made visible through outflows) to the reionizing budget at these redshifts (e.g., Madau & Haardt 2015; Smith et al. 2018). As we will discuss below, these discrete object iEBL measurements are directly relevant to the possible sky-SB contributed from unresolved objects, such as Pop III stars and their stellar-mass BH accretion disks at $z \gtrsim 7$.

The light gray unfilled downwards triangles in Figure 1 indicate the direct measurements or limits to the EBL, which are in general absolute measurements, and are summarized in detail in Dwek & Krennrich (2013) and D16. Most of these direct EBL estimates are a factor of 3–5× higher than the integrated and extrapolated discrete-objects counts (the iEBL), and about $\gtrsim 2\times$ higher in the far-IR, although the latter is in general agreement within the errors. Given that non-zodiacal foreground light may enter into the low-Earth orbit observations at the $\gtrsim 10\%$ level as discussed above, it is therefore possible that the true level of foreground (zodiacal+Earthshine and other straylight components) may have been under-subtracted in some of the direct EBL measurements.

2.2. Limits to the Diffuse Extragalactic Background Light

Here we summarize arguments that the diffuse EBL is likely smaller than the iEBL that comes from discrete objects, especially in the near-IR. This will help us derive our first constraints to the diffuse EBL that may be caused by Pop III stars and their stellar-mass BH accretion disks. Any real diffuse EBL could be due to faint Inter-galaxy Halo Light (IHL; Cooray et al. 2012), IntraCluster Light (ICL), or IntraGroup Light (IGL) not measured by Source Extractor-type algorithms (Bertin & Arnouts 1996) in discrete object surveys, or to truly diffuse, unresolved populations, such as Pop III stars and their BH accretion disks. Our reasoning that there may not be a large amount of near-IR diffuse light hidden is as follows:

(1) Independent diffuse EBL estimates at 0.3–20 μm come from γ-ray blazar spectra and how much these are distorted from their original power-law shape. When a γ-ray from the blazar hits an intervening EBL photon, this can result in pair production and energy loss in the power-law spectrum. This constrains the total EBL level that each of the low-redshift blazar γ-ray photons are exposed to (Dwek & Krennrich 2013; Lorentz et al. 2015). Figure 1 indicates the resulting EBL constraints as a gray shaded region+blue line and a green shaded region+dark-green line from the blazar surveys with the High Energy Stereoscopic System (H.E.S.S.; H.E.S.S. Collaboration et al. 2013, 2017) and the Major Atmospheric Gamma Imaging Cherenkov telescope (MAGIC; Ahnen et al. 2016), respectively. The MAGIC shaded region in Figure 1 is smaller than that of HESS, and closer to the red

iEBL points of D16. The extent to which the γ-ray blazar spectra deviate from their intrinsic power laws constrains the amplitude and shape of the foreground component of the EBL spectrum directly (Biteau & Williams 2015), which is completely independent from having to subtract the zodiacal foreground. Biteau & Williams (2015) found that the amount of diffuse EBL at $\lambda \simeq 1$–$5\,\mu$m is $\lesssim 1$–2 nW m^{-2} sr^{-1}. For a detailed discussion of these blazar data and their constraints on the EBL, we refer the reader to Dwek & Krennrich (2013) and D16. In short, the allowed amount of total 1–5 μm EBL from the γ-ray blazar spectral constraints is generally quite consistent with the integrated and extrapolated discrete galaxy counts (red dots in Figure 1) summarized in D16.

At 0.45–0.65 μm, the diffuse blazar EBL and the discrete iEBL measurements are—to within their errors—also consistent with the direct *Pioneer* spacecraft measurements (Matsuoka et al. 2011), which were made at a distance of 4.6 au from the Sun, i.e., well away from most of the zodiacal foreground brightness (blue unfilled circles in Figure 1). The direct *R*-band *Pioneer* EBL measurement was confirmed through the first measurement in a broader *R*-band with the Long Range Reconnaissance Imager instrument on board the *New Horizons* spacecraft on its way to Pluto at $\sim$7–16 au from the Sun (Zemcov et al. 2017), albeit with a larger error bar, which will improve as further *New Horizons* data are taken. At these very large distances from the Sun, the uncertainties due to the zodiacal foreground are much smaller than those from low-Earth orbit. Ground-based optical spectroscopy of dark clouds was done by Mattila et al. (2017) to remove the Diffuse Galactic Light (DGL), suggesting a diffuse EBL component at $\lambda \simeq 0.4$–$0.6\,\mu$m possibly as high as $\sim$4–6 nW m^{-2} sr^{-1}. The good correspondence at $\lambda \lesssim 1\,\mu$m among the iEBL from the discrete extrapolated galaxy counts (D16), the direct *Pioneer* and *New Horizons* *B+R*-band observations at 4.6–16 au, and the independent constraints from the H.E.S.S./MAGIC blazar γ-ray spectra in Figure 1, suggests that a low-redshift, truly diffuse EBL component at $\lambda \simeq 0.4$–$1\,\mu$m may not exceed the iEBL component itself, which is $\sim$4–10 nW m^{-2} sr^{-1}.

Despite uncertainties in the optical diffuse EBL, the 1–4 μm iEBL results are consistent with the blazar constraints on the diffuse EBL to within their errors. Below, we will therefore adopt an upper limit to the diffuse 1–4 μm EBL based on the difference between the γ-ray blazar constraints from H.E.S.S. + MAGIC and the integrated plus extrapolated galaxy counts of D16. If any diffuse 1–4 μm EBL were truly 3×–5× higher than what the red dots in Figure 1 indicate, such a high EBL level would have distorted the blazar spectra more than is observed in Figure 1. Comparing the H.E.S.S. and MAGIC blazar constraints to the EBL from the discrete galaxy counts in Figure 1 suggests that a diffuse 1–4 μm EBL component (Biteau & Williams 2015) may add $\sim$20% to the iEBL from discrete objects (D16).

(2) The deepest ground-based surveys with large telescopes do not detect an excessive amount of light in the outskirts of galaxies that have total fluxes of AB $\simeq$ 20–23 mag. It is precisely in this flux range where most of the iEBL is generated in the observed blue wavelength regime (see D16). For instance, Ashcraft et al. (2017) present 32 hr LBT *U*-band images sorted as a function of image FWHM value. The best 10% of their 320 images with the highest resolution (FWHM $\lesssim 0\rlap{.}''7$) reach AB $\lesssim$ 27.0 mag for point-source detection, while their best-depth 32 hr image has FWHM $\lesssim 1\rlap{.}''8$,

reaches AB $\lesssim$ 28.0 mag for point sources, and has a 1σ SB sensitivity of AB $\lesssim$ 32 mag arcsec^{-2}. Ashcraft et al. (2017) then compare the light profiles of 220 galaxies with total fluxes of AB $\simeq$ 20–23 mag in both their highest-resolution images and in their best-depth LBT U-band image, and find that no more than 5%–10% of the total galaxy flux is missing in the high-resolution images compared to the deeper low-resolution images. That is, at least in the U-band for galaxies AB $\simeq$ 20–23 mag—over which most of the iEBL is generated (see Section 2.1)—no more than 10% of the light appears to be hidden in the outskirts of these galaxies down to AB $\lesssim$ 32 mag arcsec^{-2}. The integrated and extrapolated U-band galaxy counts of W11, D16 and Ashcraft et al. (2017) are consistent with the H.E.S.S. and MAGIC blazar constraints at 0.36 μm, with little room to hide more than 10%–20% in diffuse EBL at 0.36 μm. Longer-wavelength studies of this depth have been done with the 10 meter Grand Canary Telescope (Trujillo & Fliri 2016), with similar results in the r-band.

(3) Combining the constraints from the previous two arguments, we derive the following limit to the diffuse 1–4 μm EBL: (1) a diffuse 1–4 μm EBL component can add $\sim$20% to the iEBL from discrete objects, and (2) no more than 10%–20% in diffuse EBL seems to be hidden in galaxy outskirts to AB $\lesssim$ 32 mag arcsec^{-2}. More could come from low-redshift ICL, but cluster galaxies comprise a small fraction of the total galaxy population. Some could come from IGL at low redshifts, since most galaxies reside in galaxy groups (Robotham et al. 2011). Where the IGL has been measured, it does not appear to dominate the total stellar light in galaxy groups (e.g., Robotham et al. 2011) or in galaxy clusters (e.g., Morishita et al. 2017; Griffiths et al. 2018 and references therein). In all, Figure 1 suggests that diffuse 1–4 μm EBL may well be as low as 20% of the discrete iEBL, or $\lesssim$1–2 nW m^{-2} sr^{-1} at 2 μm. We will use this level as a conservative or "hard" upper limit for any Pop III contribution to the near-IR EBL, as indicated by the orange circle with its dotted 1–4 μm range in Figure 1.

If the diffuse 1–4 μm EBL from Pop III stars or accretion disks at $z \gtrsim 7$ was much larger than our hard upper limit of $\lesssim$1–2 nW m^{-2} sr^{-1} at 2 μm, it would exceed the known components from unobscured AGNs and even galaxy disks (blue and green dashed lines in Figure 1) at $z \lesssim 6$, which would be unheard of at any other wavelength in the electromagnetic spectrum. That is, the diffuse 1–4 μm EBL from Pop III stars and/or their accretion disks is likely well below the level indicated by our hard upper limit at 1–2 nW m^{-2} sr^{-1} in Figure 1. In Section 4.4, we will estimate the Pop III caustic transit rate for a range of possible diffuse 1–4 μm EBL values, and estimate which SB levels may result in observable numbers of Pop III caustic transits during $JWST$'s lifetime.

2.3. Diffuse EBL Limits Adopted for Pop III Stars and Their Stellar-mass Black Hole Accretion Disks

Next, we adopt tighter constraints to the sky-SB from Pop III stars from recent theoretical and observational constraints, and from Pop III stellar-mass BH accretion disks using recent near-IR–X-ray power-spectrum results. We need both sky-SB constraints to estimate their cluster caustic transits in Sections 4.4 and 6.2, respectively.

The thermal brightness of the zodiacal belt rapidly increases at wavelengths $\lambda \gtrsim 4$ μm (Figure 1), and so in the calculations below we do not anticipate easily detecting Pop III caustic transits with $JWST$ at wavelengths longer than 4 μm. For Pop III objects at $z \gtrsim 7$, the wavelength range of interest is therefore $\lambda \simeq 1$–4 μm. The geometric average of this wavelength range is $\lambda = 2.0$ μm, which is also equal to the $JWST$ diffraction limit (Rieke et al. 2005). $JWST$ NIRCam will be most sensitive over the wavelength range of 2–3.5 μm, where the zodiacal sky from L2 is darkest (Figure 1 and W11).

2.3.1. Diffuse EBL Limits Adopted for Pop III Stars

Based on metallicity arguments, Madau & Silk (2005) provided a constraint suggesting that Pop III stars must contribute less than a few nW m^{-2} sr^{-1} to the (1–4 μm) InfraRed Background (IRB). This is consistent with our hard diffuse-EBL upper limit in Section 2.2. Cooray et al. (2012) provide a detailed Pop III model for reionization and estimate the Pop III flux to be $\lesssim$0.04 nW m^{-2} sr^{-1} (see their Figure 4), which we confirm below. Bovill (2016) suggests a Pop III star density of 0.1–10^3 stars per arcsec^2 between $z \simeq 10$–30, which we consider in more detail in Section 3.5. For their expected range in luminosities, Pop III stars could have an observed flux of AB $\simeq$ 35–41.5 mag over the redshift range of $z \simeq 7$–17 (see Section 3). As an example, if there existed $\sim$1000 Pop III stars of 100 $M_\odot$ each per arcsec^2, then their integrated 2.0 μm sky-SB would be $\gtrsim$33 mag arcsec^{-2} or $\lesssim$0.016 nW m^{-2} sr^{-1}, which is comparable to the Cooray et al. (2012) limit.

To confirm these numbers, we will estimate the average sky-SB from star-forming objects at $z \simeq 7$–8 from the actual HUDF data corrected for incompleteness. For our caustic transit calculations, we need to estimate the maximum possible SB from Pop III stars at $z \gtrsim 7$ to use as the most conservative upper limit. This needs to take into account that the steep faint end of the galaxy luminosity function (LF) at $z \gtrsim 7$ will contribute additional flux from unseen Pop III objects beyond the detection limit of the deepest HST and $JWST$ images, and an estimate of the maximum additional sky-SB from $z \simeq 9$ to $z \simeq 17$. We proceed with this calculation in three steps:

(a) *The average sky-SB from star-forming objects at $z \simeq 7$–8 from the actual HUDF data corrected for incompleteness*: we use the actual HUDF data at $z \simeq 7$–8 (Table A1 of Bouwens et al. 2015) to estimate the observed surface densities of star-forming objects at $z \simeq 7$ and $z \simeq 8$ to an average sky-SB. In the 4.7 arcmin^2 effective area of the WFC3/IR data, there are 56 dropout candidates detected at $z \simeq 7$ to the HUDF limit of AB $\simeq$ 30.0 mag, while there are 28 dropout candidates at $z \simeq 8$ to AB $\lesssim$ 30 mag. These can be directly converted to a total sky-SB, in this case from the objects detected to AB $\lesssim$ 30 mag. We need to correct these observed surface densities by about a factor of 1.8, since in the deepest HUDF WFC3/IR images, at least $\sim$45% of the detector pixels are covered by the wings of the foreground objects (Koekemoer et al. 2013). Our own insertion of artificial objects into the HUDF WFC3/IR images confirms this correction factor.

(b) *Maximum contribution from the steep faint end of the galaxy luminosity function down to the luminosity of single Pop III stars*: next, we correct this upper limit to the 2.0 μm sky-SB that comes from $z \simeq 7$–8 for the flux of objects that will have been missed below the current HUDF object detection limit of AB $\simeq$ 30 mag. At $z \simeq 7$ to $z \simeq 10$, the AB $\simeq$ 30 mag HUDF detection limits correspond to absolute magnitudes of $M_{AB} \simeq -17.5$ mag. According to the fits to the available galaxy LF data in Figures 5–6 of Finkelstein (2016), the faint-end slope of the galaxy LF at $z \gtrsim 7$ may become as steep as $\alpha \lesssim -2.0$ to -2.3,

while the characteristic Schechter luminosities (L^* or M^*) and space densities (Φ^*) may well continue to get fainter and decline at $z \gtrsim 7$, respectively. High-resolution hierarchical simulations of the faint-end galaxy LF-slope evolution with redshift (e.g., Morgan et al. 2015) suggested values of $\alpha \simeq -2.1$ from $z \simeq 11$ to $z \simeq 4$. This is about as steep as the Initial Mass Function (IMF) slope for more massive stars (Coulter et al. 2017), and would occur if the luminosity density is dominated by Pop III stars with $M \gtrsim 100 \, M_\odot$, for which $L \propto M$ approximately holds. We discuss this further in Sections 3.1 and 3.4.

We will adopt for simplicity in our extrapolation $\alpha \simeq -2.0$, so that each additional luminosity bin with objects at $z \gtrsim 7$ that are currently beyond the *HST* detection limit would contribute roughly equal amounts of energy to the sky-SB. Since the M^* values at $z \simeq 7$–8 in the best fits of Finkelstein (2016) are about $M^* \simeq -20.5$ mag, the sky-SB in the HUDF from objects that are currently resolved into galaxies comes effectively from a ~ 3 mag range in the observed LF. If we extrapolate this LF with a faint-end Schechter slope $\alpha = -2.0$ from $M_{AB} \simeq -17.5$ mag to $M_{AB} \simeq -7$ mag (i.e., the luminosity of a $20 \, M_\odot$-star; see Section 3), then the integrated $2.0 \, \mu m$ sky-SB will be $\sim 3\times$ brighter than the estimate from (a) alone.

Integrating the maximum SB that can come from Pop III stars at $z \gtrsim 7$ to $M_{AB} \simeq -7$ mag is meaningful and necessary, since at this luminosity a faint star-forming "object" would simply consist of a single unresolved Pop III star with $M \gtrsim 20 \, M_\odot$ and $M_{AB} \simeq -7$ mag, which is the faintest *JWST* could detect at $z \gtrsim 7$ during a favorable caustic transit (see Section 3). Given the homology relations in Section 3.1, the M–L relation for such massive stars becomes approximately $L \propto M$, so that the very faint-end slope of the object luminosity function may reflect the bright-end slope of the stellar-mass function at $z \gtrsim 8$.

If we integrate down to the limit of a $1.5 \, M_\odot$ Pop III star luminosity of $M_{AB} \simeq +2$ mag at $z \gtrsim 7$ (see Section 3), then the maximum $2.0 \, \mu m$ sky-SB will be $\sim 5\times$ brighter than the estimate from (a). Since these are the coolest stars that can contribute to reionization (Section 3), we will use this multiplier to derive the most conservative upper limit to the $2.0 \, \mu m$ sky-SB that may come from $z \gtrsim 7$. The maximum $2.0 \, \mu m$ sky-SB we then obtain from the entire object LF to $M_{AB} = +2$ mag is 32.2 mag arcsec^{-2} at $z \simeq 7$ and 32.8 mag arcsec^{-2} at $z \simeq 8$.

(c) *Maximum contribution from the cosmic star formation history at $z \gtrsim 8$:* lastly, we need to correct these limits for the maximum contribution from the LF of star-forming objects at $z \simeq 9$–17 that is not yet accounted for. For this, we use a best fit of the cosmic star-formation history (SFH) data summarized by Madau & Dickinson (2014) and Finkelstein (2016). Equation (15) of Madau & Dickinson (2014) gives a best fit to the cosmic SFR data over the entire redshift range $0 \lesssim z \lesssim 8$:

$$\psi(z) = 0.015 \, \frac{(1+z)^{2.7}}{1 + [(1+z)/2.9]^{5.6}} \, M_\odot \, \text{yr}^{-1} \, \text{Mpc}^{-3}. \quad (1)$$

Their best fit has its peak in the cosmic SFR at $z \simeq 1.9$. The best-fit power-law slope for $z \gg 2$ is approximately $2.7 - 5.6 \simeq -2.9$, so that at $z \simeq 7$ the SFR is ~ 1.0 dex or 2.5 mag lower than at $z \simeq 1.9$. This decline is also seen in the more recent *HST* WFC3 data reviewed by Finkelstein (2015,

2016) and Madau & Fragos (2017), who find a slightly steeper decline of $\propto (1+z)^{-3.6}$ to $(1+z)^{-4.2}$ when only fitting the data for $z \gtrsim 2$. The difference in slope could be due to a truly steeper decline in the cosmic SFR at $z \gtrsim 8$, the smaller fitted redshift range used in these more recent papers, or larger incompleteness corrections for dropout samples at $z \gtrsim 7$, as discussed in (a).

To obtain the most conservative upper limit to the integrated sky-SB from $z = 7$ to $z = 17$, we will use the highest predicted SFR at $z \gtrsim 8$. Hence, we will use the extrapolation of Madau & Dickinson (2014) in Equation (1), since it is ~ 0.3 dex above the fits of Finkelstein (2016) and Madau & Fragos (2017) to the most recent WFC3 data at $z \simeq 8$–10. The extrapolation of Equation (1) is also consistent with the hierarchical model predictions of Sarmento et al. (2018) at $7 \lesssim z \lesssim 20$, which approximately match the Madau & Dickinson (2014) results at $z \simeq 7$–8. The extrapolation of Equation (1) thus yields the highest observed sky-SB that may come from star-forming objects at $z \gtrsim 8$, which is used in Section 4.4 to predict the highest level of caustic transits that could be seen. That is, if the true Pop III star sky-SB is lower than what we predict from Equation (1) here, then the caustic transit rates will be correspondingly smaller, as indicated in Figure 1 and discussed in Section 4.4.

With the Madau & Dickinson (2014) extrapolation of Equation (1), more than half of the sky-SB that comes from $7 \lesssim z \lesssim 17$ is already obtained from the redshift shell at $z \simeq 7$, while about 75% comes from the two redshift shells at $z \simeq 7$ and $z \simeq 8$ combined. Each redshift shell here is assumed to have a width of $\Delta z \simeq 1$. The contributions from the redshift shells at $z \gtrsim 12$ are negligibly small. Integrating the sky-SB produced by each redshift shell by Equation (1) from $z \simeq 7$–17 thus produces approximately $1.33\times$ the flux than that from the $z \simeq 7$–8 redshift shells alone, where we directly summed the observed sky-SB in (a).

This then results in a most conservative upper limit to the $2.0 \, \mu m$ sky-SB from star-forming objects at $7 \lesssim z \lesssim 17$ down to the luminosity of a single Pop III star. This upper limit to the $2.0 \, \mu m$ Pop III star sky-SB is $\gtrsim 31.4 \pm 0.6$ mag arcsec^{-2} or $\lesssim 0.06$ nW m^{-2} sr^{-1}, which is indicated by the dark orange unfilled triangle and its error and wavelength range in Figure 1.

2.3.2. Diffuse EBL Limits Adopted for Pop III Stellar-mass Black Hole Accretion Disks

Kashlinsky et al. (2012, 2015), Cappelluti et al. (2013), Helgason et al. (2016), and Mitchell-Wynne et al. (2016) provided estimates of the object-free IR-power spectrum. After carefully subtracting all objects in ultradeep *Spitzer* 3.6 and $4.5 \, \mu m$ images in the CANDELS GOODS-South field (Grogin et al. 2011; Koekemoer et al. 2011), these papers all found a consistent rather uniform power in the power spectrum on 100″– 1000″ scales with an rms (amplitude)^2 of $\lesssim 0.004$ nW^2 m^{-4} sr^{-2}, which is relatively flat on the angular scales where it is well-sampled, and is fairly similar between 3.6 and $4.5 \, \mu m$. Although it is possible that residual, very low-level detector systematics (Arendt et al. 2016) or DGL (Cooray et al. 2012) may have boosted this signal, the $3.5 \, \mu m$ power spectrum amplitude itself does provide an upper limit to the diffuse $3.5 \, \mu m$ sky-SB that may be generated by objects at $z \gtrsim 7$, as we will discuss below.

Cappelluti et al. (2013) cross-correlated the object-subtracted ultradeep *Spitzer* images with the deepest object-free 0.2–2 keV

Chandra images in the same CANDELS field, and found a similar power-spectrum signal on $\gtrsim 10''$ scales. Their power spectra when cross-correlated with the object-free soft (0.5–2 keV, or 1.2 keV in energy on average) *Chandra* images gave a stronger signal than when cross-correlated with the hard (2–4.5 keV or 4.5–7 keV) *Chandra* images. Cappelluti et al. (2017) fit the 0.3–7 keV energy spectrum of the X-ray background (XRB) with the redshifted X-ray spectra of known populations, and constrain the fraction of the XRB that can come from unresolved sources—possibly early black holes at $z \gtrsim 6$—to be $\lesssim 3\%$ of the peak in the supermassive black hole (SMBH) growth-rate curve at $z \simeq 1$–2.[9] Mitchell-Wynne et al. (2016) cross-correlated the object-free *Spitzer* 3.6 and 4.5 μm images with the deepest available object-free CANDELS *HST* ACS and WFC3 images at 0.6, 0.7, 0.85, 1.25, and 1.60 μm, and found no correlation with the *Spitzer* images, or even an anticorrelation, in these shorter *HST* wavelength filters.

This *Spitzer*–*Chandra* cross-correlation signal cannot be easily explained by DGL alone (Mitchell-Wynne et al. 2016). If this cross-correlation signal is real, the implication is that some of it may come from First Light objects at $z \gtrsim 7$. Some of this signal may come from an unresolved AGN or hard X-ray binary population in faint red bulge-dominated galaxies at lower redshifts (Cooray et al. 2012)—from objects both below the *Spitzer* and *Chandra* detection limits. But this signal has also been modeled with Primordial Black Holes (PBHs; Kashlinsky 2016), Direct Collapse Black Holes (DCBHs; Yue et al. 2013), or Obese Black Holes (OBHs; Natarajan et al. 2017) at $z \gtrsim 7$–8. If part of this 3.6–4.5 μm power-spectrum signal and the *Spitzer*–*Chandra* cross-correlation signal truly came from $z \gtrsim 7$, then it must have an X-ray component that is much hotter than 10 keV in the rest frame (i.e., $T \gtrsim 3 \times 10^7$ K).

Regardless of its correct explanation, the near-IR power-spectrum signal provides an upper limit to the 3–4 μm sky-SB that may come from Pop III BH accretion disks, the inner regions of which may reach X-ray temperatures, as we will discuss in Section 5.5.2. None of the evolutionary models for Pop III stars that we discuss in Section 3.1 reach temperatures much hotter than $T \simeq 10^5$ K, and so the redshifted spectral energy distribution (SED) of Pop III stars at $z \gtrsim 7$ alone cannot produce the *Spitzer*–*Chandra* cross-correlation signal.

Let us now consider the upper limit to the diffuse 3–4 μm sky-SB that may come from BH accretion disks at $z \gtrsim 7$. Since the possible 3.6 (and 4.5) μm sky-signal was derived from power spectra at $\theta \gtrsim 10''$–$1000''$ scales (Cappelluti et al. 2013; Kashlinsky et al. 2015; Mitchell-Wynne et al. 2016), we must first convert it to an upper limit to the actual *3.6 μm* signal in the sky. For this, we proceed as follows. The smallest angular scale $\theta \gtrsim 100''$ at which the 3.6 μm power-spectrum excess signal is seen corresponds to 4.3–5.2 Mpc physical scales at $z \simeq 7$–17 in our adopted cosmology with an average of 4.4 Mpc at $z \simeq 8$. (Note that the physical scale needs to be used in this argument, not the co-moving scale). As in Section 2.3.1, the redshift shell $7 \lesssim z \lesssim 8$ contains about half of the sky-SB that comes from $7 \lesssim z \lesssim 17$ if the source of this SB intrinsically declines as $\propto (1+z)^{-2.9}$, or more steeply. At $\lesssim 5$ Mpc scales, the overdensities $\Delta\rho/\rho$ are about unity at $z \simeq 0$ (Barkana & Loeb 2001). At redshift z, the physical overdensities $\Delta\rho/\rho$ would thus have been

$(1+z)\times$ lower, and so the signal amplitude itself (or the sky-SB of the signal) will scale with the fluctuation in the signal as $\rho \simeq (1+z)\Delta\rho$. That is, if a power spectrum that came from $z \gtrsim 7$ has an (amplitude)^2 at $100''$ scales of $\lesssim 0.004$ nW^2 m^{-4} sr^{-2}, then its linear flux amplitude must be less than $(1+z) \times \sqrt{(0.004)} \simeq (1+z) \times 0.06$ nW m^{-2} sr^{-1}, or $\lesssim 0.57$ nW m^{-2} sr^{-1}.

From their *Spitzer*–*Chandra* cross-correlation, Cappelluti et al. (2013) suggest that $\lesssim 20\%$ of the large-scale power of the cosmic infrared fluctuations is correlated with the spatial power spectrum of the X-ray fluctuations. Hence, we will here adopt that no more than 0.2×0.57 or 0.11 nW m^{-2} sr^{-1} of the 3.6 μm sky-SB may come from accreting sources at $z \gtrsim 7$. In Figure 1, we indicate this upper limit by the black unfilled triangle plus its error range in black. This limit is thus far only observationally constrained at 3.6 and 4.5 μm, but not yet at 2.0 μm, although deep *JWST* images of cirrus-free, low-extinction regions at the north ecliptic pole will provide sky-SB constraints at 2.0 μm as well (Jansen et al. 2017). At 3.6 μm, this current SB limit corresponds to $\gtrsim 30.2$ mag arcsec^{-2} following the wavelength-dependent conversion between nW m^{-2} sr^{-1} and AB-mag arcsec^{-2} in Section 2.1. Since the *Spitzer* power spectra and cross-correlation spectra with *Chandra* of Mitchell-Wynne et al. (2016) are fairly similar at both 3.6 and 4.5 μm in units of nW m^{-2} sr^{-1}, we will therefore adopt the equivalent sky-SB value of $\gtrsim 30.8$ mag arcsec^{-2} at 2.0 μm as the upper limit for BH caustic transit calculations, as indicated by the black upper limit in Figure 1.

In summary, Sections 2.3.1 and 2.3.2 yield rather similar upper limits to the 2.0 μm sky-SB that may come from Pop III stars or their stellar-mass BH accretion disks of $\gtrsim 31$ mag arcsec^{-2}. In what follows, we will therefore perform the caustic transit calculations assuming that the full 2.0 μm sky-SB signal of $\gtrsim 31$ mag arcsec^{-2} is either completely caused by Pop III stars (Section 4.4) or by their BH accretion disks (Section 6.2). For the plausible case where both Pop III stars and their BH accretion disks both contribute to the 2.0 μm sky-SB of $\gtrsim 31$ mag arcsec^{-2}, one could use a weighted sum of the caustic transit rates derived in Sections 4.4 and 6.2 for Pop III stars and their BH accretion disks, respectively. Where appropriate, we give size, lifetime, and obscuration arguments regarding the proportions of caustic transits of Pop III stars and their BH accretion disks that may be visible to *JWST* (Sections 3.1, 5.3, 6.2).

Note that for our caustic transit calculations, it does not matter whether the light that comes from $z \gtrsim 7$ exists in faint discrete objects that have already been detected down to the HUDF limit and contain Pop III stars and/or stellar-mass BH accretion disks, or whether this light is fully unresolved below the current HUDF object detection limit of AB $\simeq 30$ mag. Either way, the maximum 2.0 μm SB of ~ 31 mag arcsec^{-2} that can be produced at $z \gtrsim 7$ may be subject to cluster caustic transits.

3. Parameters Adopted for Pop III Stars

In this section, we present the physical properties of Pop III stars from stellar evolution models with HR diagrams through the hydrogen-depletion and helium-depletion stages, and from these derive their mass–luminosity relation, their bolometric +IGM+K-corrections, and their relative contribution to the luminosity density of a faint star-forming object.

---

[9] Throughout, "SMBH" indicates the rare supermassive black holes, while "BH" indicates the much more numerous stellar-mass black holes discussed in Sections 3, 5, and 6 of this paper.
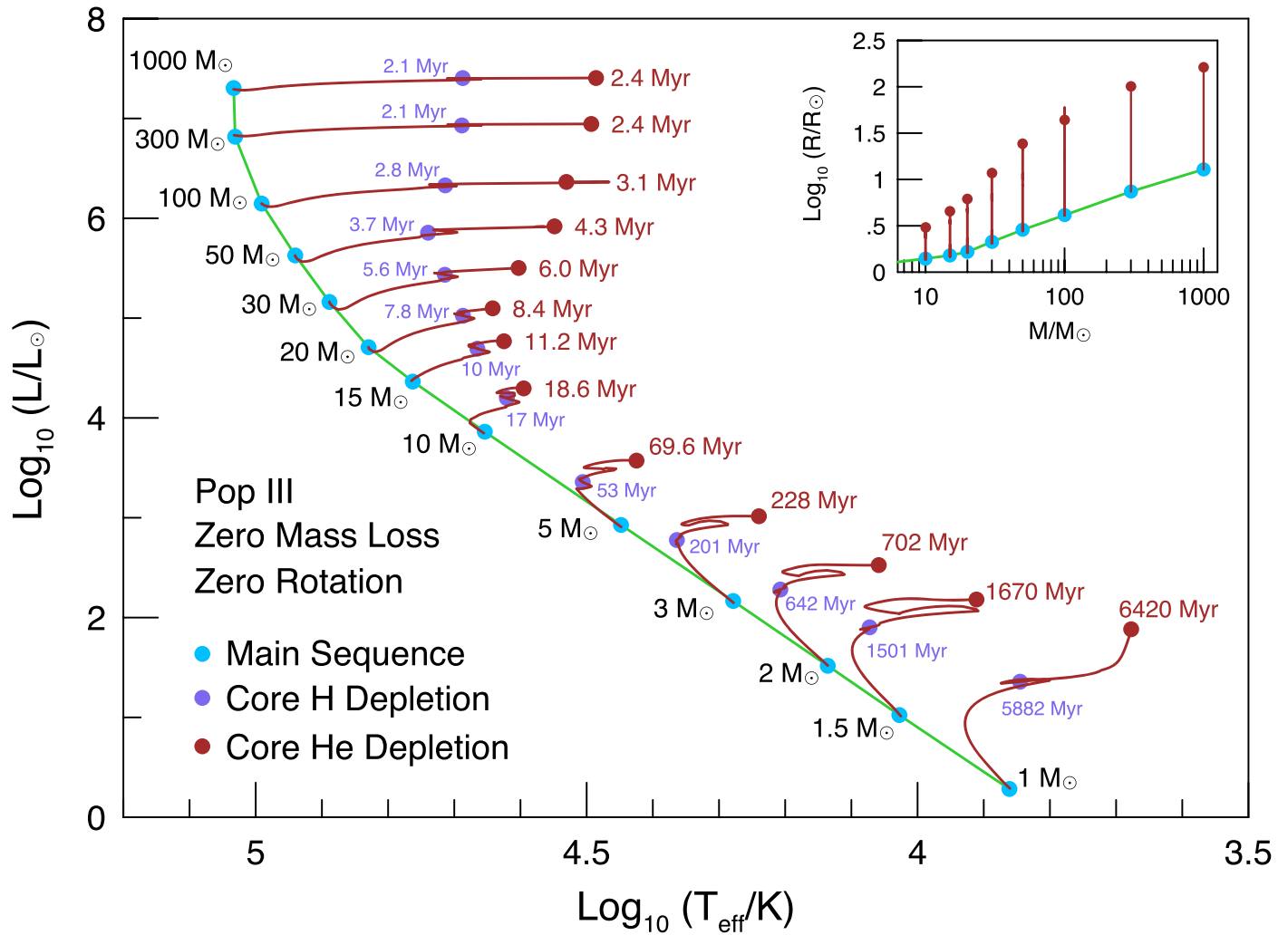
Figure 2. Loci of the zero-age main sequence in the HR diagram for non-rotating, zero mass loss, $Z = 0.00\ Z_\odot$ MESA models. Evolutionary tracks to core He depletion are shown, with the final model marked by filled circles and labeled by age. The actual model data are given in Tables 1–4. The inset plot shows the mass–radius evolution, with filled circles marking the location of ZAMS and core He depletion.

Simulations suggest that the fragmentation of primordial gas allows central concentrations with a range of stellar masses to form in a mini halo, depending on the dimensionality, spatial resolution, and local physics used in the simulations. For instance, Abel et al. (2002) presented a 3D hydrodynamical simulation to form the first stars, which resulted in a $100\,M_\odot$ star forming. In a higher-resolution simulation (Turk et al. 2009), a $50\,M_\odot$ clump breaks up into two cores, each with a forming star that likely will become a binary star. Radiation-hydrodynamic simulations of primordial clouds (Hosokawa et al. 2016) showed fragmentation into protostars with masses $M \simeq 10$–$1000\,M_\odot$, depending on the amount of UV feedback that was produced. Stacy et al. (2016) followed the formation of a mini halo with gas collapsing into central cores ranging from $20\,M_\odot$ to as many as ~30 stars with $M \lesssim 1\,M_\odot$.

Strong radiative feedback from the most massive stars may initially prevent lower-mass stars—and therefore binaries—from forming in a mini halo (e.g., Abel et al. 2002; Trenti & Stiavelli 2009). We discuss the low near-IR sky-SB that may result from this in Section 3.5. Given that more recent simulations resulted in the formation of lower-mass stars and

binaries, we will also allow for the possibility that slightly polluted lower-mass stars—and binaries—can form in the vicinity of previous more massive, zero-metallicity Pop III stars ($Z \lesssim 10^{-4}\,Z_\odot$; Sarmento et al. 2018). We discuss this in more detail in Sections 3.2 and 5.2. When the distinction is relevant, we refer to these slightly polluted stars as "Pop II.5." This paper will thus consider stars of zero or very low metallicity that cover the mass range of $1\,M_\odot \lesssim M \lesssim 1000\,M_\odot$.

3.1. Physical Parameters of Pop III Stars from MESA Models

We first need to outline the plausible physical parameter ranges for Pop III stars. Figure 2 shows the zero-age main sequence (ZAMS) in an HR diagram for stellar evolution models with $Z = 0.00\ Z_\odot$, and the inset shows their corresponding mass–radius relation. These non-rotating, zero-metallicity, zero mass-loss, single-star 1–$1000\,M_\odot$ models were calculated using the MESA software instrument (Paxton et al. 2011, 2013, 2015) with the same physical and numerical parameters as those in Farmer et al. (2015, 2016) and Fields et al. (2016). We also calculated MESA models for $Z = 10^{-8}\ Z_\odot$, and their results were very similar to $Z = 0.00\ Z_\odot$. This is because stars more massive than

Table 1
Adopted Pop III Star Physical Parameters from MESA Models[a]

Mass	Age Pre-MS	T_{eff}	$\log R$	$\log L_{bol}$	T_{eff}	$\log R$	$\log L_{bol}$	Age	T_{eff}	$\log R$	$\log L_{bol}$	Age	Time[b] AGB-MS
			at ZAMS			at Hydrogen Depletion				at Helium Depletion			
$(M_\odot)$	(Myr)	(K)	$(R_\odot)$	$(L_\odot)$	(K)	$(R_\odot)$	$(L_\odot)$	Myr	(K)	$(R_\odot)$	$(L_\odot)$	Myr	(Myr)
1.0	9.28	7.266e3	−0.0581	0.2825	6.999e3	0.5119	1.3576	5882	…[c]	…	…	6420	538
1.5	6.11	1.065e4	−0.0203	1.0227	1.181e4	0.3292	1.9015	1501	8.149e3	0.7913	2.1804	1670	169
2.0	3.02	1.367e4	0.0108	1.5177	1.611e4	0.2498	2.2815	642	1.145e4	0.6685	2.5249	702	60
3.0	1.38	1.899e4	0.0487	2.1654	2.311e4	0.1843	2.7770	201	1.736e4	0.5510	3.0138	228	27
5.0	0.56	2.805e4	0.0911	2.9274	3.206e4	0.1903	3.3581	53	2.658e4	0.4608	3.5732	70	17
10	0.23	4.508e4	0.1462	3.8618	4.174e4	0.3807	4.1972	17	3.938e4	0.4811	4.2968	19	1.6
15	0.13	5.789e4	0.1803	4.3647	4.624e4	0.5401	4.6937	10	4.215e4	0.6581	4.7691	11	0.8
20	0.09	6.754e4	0.2183	4.7082	4.864e4	0.6612	5.0240	7.8	4.386e4	0.7879	5.0975	8.4	0.6
30	0.05	7.737e4	0.3270	5.1619	5.180e4	0.8120	5.4347	5.6	4.006e4	1.0688	5.5016	6.0	0.5
50	0.03	8.713e4	0.4570	5.6283	5.490e4	0.9722	5.8562	3.7	3.536e4	1.3862	5.9200	4.3	0.5
100	0.02	9.796e4	0.6147	6.1470	5.173e4	1.2610	6.3303	2.8	3.392e4	1.6437	6.3627	3.1	0.3
300	0.02	1.074e5	0.8697	6.8172	4.882e4	1.6111	6.9301	2.1	3.165e4	2.0041	6.9631	2.4	0.3
1000	0.02	1.080e5	1.1090	7.3047	4.807e4	1.8740	7.4288	2.1	3.122e4	2.2119	7.3549	2.4	0.3

Notes.
[a] All physical Pop III star parameters were calculated using MESA models with zero initial metallicity ($Z = 0.00\ Z_\odot$), zero mass loss, zero rotation, and no stellar duplicity (i.e., no binaries/multiple stars). Pop III star parameters are listed with a sufficient number of significant digits to be able to integrate them assuming blackbody spectra, which is needed in Section 3.3.
[b] Ages are listed for the pre-main-sequence collapse (pre-MS), the core+shell hydrogen-burning phase (H-depletion), the core+shell helium-burning phase (He-depletion), and the total giant branch lifetime (i.e., the AGB–MS age difference). The latter provides an upper limit to the BH feeding times due to Roche-lobe overflow in non-zero-metallicity massive-star binaries, as discussed in Section 5.
[c] The $1.0\ M_\odot$, $Z = 0.00\ Z_\odot$ model did not ignite helium and may thus turn directly into a helium white dwarf, so no AGB parameters are listed here. (Its AGB ages and K-corrections in Section 3.3 are those of the $Z = 10^{-8}\ Z_\odot$ model, which did end in a white dwarf).

$\sim 2\ M_\odot$ make enough of their own carbon in their cores to run the CNO cycle appropriate for their mass. In other words, there is a floor metallicity, which—if not provided by the star's birth composition—will be made by the star itself, and convective episodes may bring part of these self-made metals to the stellar photosphere. Here, we adopt the set of $Z = 0.00\ Z_\odot$ MESA models for Pop III stars and will discuss the possible effects of metallicity in more detail below and in Section 5.2.

There may be model-dependent variations in the MS ages, depending on the age definitions and on the chemical mixing algorithms used (e.g., convection, overshooting, etc). For details, we refer to Paxton et al. (2011, 2013, 2015). In our MESA models, the ZAMS by definition starts when the nuclear luminosity reaches 90% of the total stellar luminosity. The Terminal-Age Main Sequence (TAMS) is defined when the central hydrogen mass fraction drops to below 10^{-6} of the star's core mass, which is when the ZAMS ends. At this point, shell hydrogen burning dominates the energy production, and can be taken as the "beginning" of the "Giant Branch." Core helium depletion is defined as the stage in the star's evolution when the fraction of ^{4}He drops below 10^{-6} of the core mass of the star. These definitions are more precise than the common use of "Red Giant Branch" or "Asymptotic Giant Branch," but for the sake of brevity, we will henceforth refer to these latter stages as the "RGB" and "AGB," respectively. The MS age adopted here is defined as the time between the start of the ZAMS and the start of the RGB (core-hydrogen depletion), while the "Giant Branch" (GB) age is defined between the start of the RGB and the end of the AGB, when the star has run out of ^{4}He.

Stars more massive than $\gtrsim 100\ M_\odot$ are radiation-pressure dominated. For CNO burning, constant electron scattering, and radiative transport, the ZAMS "homology" relations for massive stars (Hoyle & Lyttleton 1942; Faulkner 1967; Pagel

& Portinari 1998; Bromm et al. 2001; Portinari et al. 2010) are

$$\left(\frac{R}{10\ R_\odot}\right) \simeq \left(\frac{Z}{10^{-7}\ Z_\odot}\right)^{1/11}\left(\frac{M}{400\ M_\odot}\right)^{5/11},$$

$$\left(\frac{T_{eff}}{10^5\ \mathrm{K}}\right) \simeq \left(\frac{Z}{10^{-7}\ Z_\odot}\right)^{-1/20}\left(\frac{M}{100\ M_\odot}\right)^{1/40},$$

$$\left(\frac{L}{L_\odot}\right) \simeq \left(\frac{L_{edd}}{L_\odot}\right) \simeq 10^5\left(\frac{M}{M_\odot}\right), \qquad (2)$$

and they approximate the trends shown by our detailed MESA models for $Z \lesssim 10^{-8}\ Z_\odot$ (Figure 2 and Table 1).

Metallicity affects the evolution of single stars in four distinct ways: it sets their initial abundance, and it impacts their energy generation, opacity, and their mass-loss mechanism. For low-metallicity stars, these homology relations approximate the metallicity dependence of their radii and luminosities. For metallicities $Z \lesssim 10^{-4}\ Z_\odot$, Equation (2) suggest that the ZAMS tracks in Figure 2 will shift systematically by a factor of ~ 0.7 toward lower effective temperatures almost independently of mass, while the ZAMS luminosities would be nearly independent of the metallicity. The ZAMS radii will correspondingly shift by a factor of ~ 2 toward larger values. Given the other much larger uncertainties in our Pop III star caustic transit calculations in Section 4.4, we will adopt the physical parameter values of the zero-metallicity Pop III stars (or $Z = 0.00\ Z_\odot$) from our MESA modeling runs.

Equation (2) suggests that the bolometric luminosities of zero-metallicity Pop III stars—as modeled in our MESA runs—are to first order directly proportional to their ZAMS mass, while the mass–radius and mass–T_{eff} relations have much shallower slopes. All three parameters in Equation (2) need to

be carefully traced as a function of ZAMS mass for our caustic transit calculations in Section 4.4.

A closer inspection of Figure 2 suggests that the third line of Equation (2) is only approximately correct for stars with $M \gtrsim 100\,M_\odot$. Over the mass range of $1 \lesssim M \lesssim 1000\,M_\odot$, the bolometric ZAMS luminosities of Pop III stars in Figure 2 scale to a better approximation with ZAMS mass M as

$$
\begin{aligned}
L &\simeq L_{100}\,(M/100\,M_\odot)^{1.16}, & (100 \lesssim M \lesssim 1000\,M_\odot), \\
&\simeq L_{100}\,(M/100\,M_\odot)^{2.06}, & (20 \lesssim M \lesssim 100\,M_\odot), \\
&\simeq L_{20}\,(M/20\,M_\odot)^{3.20}, & (1 \lesssim M \lesssim 20\,M_\odot),
\end{aligned}
\tag{3}
$$

where L_{100} and L_{20} are the luminosities of a $100\,M_\odot$ and a $20\,M_\odot$ star, respectively. The first two segments in this equation are rescaled to the parameters of a $100\,M_\odot$ star, and the third to a $20\,M_\odot$ star. The first segment is the nearly linear mass–luminosity relation for the most massive ($M \gtrsim 100\,M_\odot$) Pop III stars in Equation (2), the second segment is a good approximation for Pop III stars in the intermediate-mass range ($20 \lesssim M \lesssim 100\,M_\odot$), and the third segment is the M/L relation for Pop III stars with $M \simeq 1$–$20\,M_\odot$, which has a slope of ~ 3.2, similar to the slope of the M/L relation for lower-mass stars in our own Galaxy. Our caustic transit calculations in Section 4.4 are dependent on stellar luminosity, and so we will propagate the segmented M/L relation of Equation (3) into the relevant equations (Equations (19)–(30)) in Section 4.4.

Here we discuss in more detail the MESA Pop III star physical parameters that are needed to estimate their resulting caustic transit rates at $z \gtrsim 7$:

Masses: The mass range for Pop III stars that have luminosities bright enough for caustic transit detection by *JWST* at $z \gtrsim 7$ is 30–$1000\,M_\odot$ (see Table 1 and Section 4.4). This corresponds to a logarithmic mass range of $M \simeq 175\,M_\odot \pm 0.75$ dex, with a corresponding bolometric absolute magnitude range of $M_{\rm AB} \simeq -10.8 \pm 2.5$ mag. As discussed in Section 5.1, LIGO has detected several BHs at the lower end of this mass range, some of which may be the remnants of Pop III stars. BH accretion disks in stellar binaries are therefore considered in Sections 5–6 as possible additional sources of caustic transits at $z \gtrsim 7$. When computing physical quantities below, we use a solar mass of 1.989×10^{30} kg (Mamajek et al. 2015).

Temperatures: Effective temperatures were determined by integrating a finely zoned MESA photospheric model inward to an optical depth of $\tau = 1$. This radial location becomes the effective radius $R_{\rm eff}$ at which the effective temperature is $T_{\rm eff}$. Zero-metallicity Pop III stars have ZAMS photospheric temperatures ranging from $T_{\rm eff} \simeq 7300$–108,000 K for masses in the range of $M \simeq 1$–$1000\,M_\odot$, as summarized in Table 1. During the RGB stage, Pop III stars with $M \simeq 1$–$1000\,M_\odot$ have lower temperatures ranging from 7000–55,000 K, while during the AGB stage, their temperatures are even lower, ranging from 6300–44,000 K (see Figure 2). For both post-MS stages, the photospheric temperatures change nonmonotonically with mass. This will affect their bolometric, IGM, and K-corrections in a nonlinear way as a function of mass and stellar evolution stage (see Section 3.3). When calculating their bolometric corrections using blackbody curves below, we use as reference a solar photospheric temperature of 5772 K (Mamajek et al. 2015; Prša et al. 2016).

Radii: Figure 2 and Table 1 show that zero-metallicity Pop III stars have ZAMS effective radii ranging from $R_{\rm eff} \simeq 0.9$–13 $R_\odot$. These are in line with previous predictions (e.g., Woosley et al. 2002; Hirschi 2007; Ohkubo et al. 2009; Yusof et al. 2013). In 2015, the IAU adopted—for stellar normalization purposes—a value of the solar radius of $1.00\,R_\odot \equiv 695,700$ km (Mamajek et al. 2015; Prša et al. 2016), which was guided by recent space-based measurements (e.g., Emilio et al. 2012). Hence, the Pop III star ZAMS radii in Table 1 range from $R_{\rm Pop\ III} = 6.05 \times 10^8 - 8.97 \times 10^9$ m, which are the values we use for Pop III star caustic transit rate predictions in Section 4.4. The Pop III star radii are at most between $1.3\times$–$5.8\times$ larger during their RGB phase, and $2.3\times$–$14\times$ larger during their AGB phase (see Table 1).

Luminosities: The MESA models shown in Figure 2 yield the bolometric absolute magnitudes of Pop III stars as a function of their mass and for different stellar evolution ages. Zero-metallicity Pop III stars have ZAMS luminosities ranging from $L_{\rm bol} \simeq 1.9\,L_\odot$–$2.0 \times 10^7\ L_\odot$ for masses $M \simeq 1$–$1000\,M_\odot$, respectively. During the RGB stage, their luminosities range from $L_{\rm bol} \simeq 23\,L_\odot$–$2.7 \times 10^7\,L_\odot$, while during the AGB stage, they are $L_{\rm bol} \simeq 40\,L_\odot$–$2.4 \times 10^7\,L_\odot$. These are the full bolometric stellar luminosities as predicted by the MESA code.

Several stellar atmosphere calculations have suggested that zero-metallicity Pop III stars in the ZAMS, RGB, and AGB stages can be approximated as blackbody emitters (e.g., Bromm et al. 2001) due to the lack of atomic absorption features or line blanketing in their spectra. For our calculations below, we therefore approximate the Pop III stars with blackbody curves of the same photospheric temperatures and radii from the MESA models summarized in Tables 2–4. We integrated these blackbody curves in I_ν from hard X-ray to radio wavelengths, and use their listed stellar radii (in kilometers using $R_\odot$ above) to predict the theoretical luminosities integrated under the full *Planck* curve for stars of that size. We use these results to convert their bolometric luminosities to observed apparent magnitudes in *JWST*'s near-IR filters, which requires the distance modulus (DM) as a function of redshift, their bolometric corrections, corrections for IGM transmission, and their K-corrections (see Section 3.3).

To normalize these calculations, we use the bolometric luminosity of the Sun, $L_\odot = 3.828 \times 10^{26}$ W, which by definition is produced by a blackbody with the effective temperature ($T_{\rm eff} = 5772$ K) and radius ($R_\odot = 6.957 \times 10^8$ m) adopted for the Sun (Mamajek et al. 2015). This corresponds to an absolute bolometric AB-magnitude of the Sun, which is by definition $M_{\rm bol}(\odot) \equiv +4.74$ mag (Bessell et al. 1998; Casagrande et al. 2006). Hence, all our Pop III star absolute magnitudes in Tables 2–4 are normalized to this $M_{\rm bol}$ value of the Sun.

This worked well for all Pop III stars in our MESA runs, except for the $1.0\,M_\odot$ zero-metallicity AGB model, whose MESA predictions were 2% lower than a blackbody curve at its specified $T_{\rm eff}$. Its MESA model did not ignite helium, so the star may turn directly into a helium white dwarf. Hence, no AGB parameters are listed in Table 1 for a $1.0\,M_\odot$ zero-metallicity AGB star. To permit caustic calculations for a $1.0\,M_\odot$ AGB star in any case, Table 4 lists the parameters for the $Z = 10^{-8}\,Z_\odot$ MESA model, which did result in a white dwarf.

Ages: Pre-MS ages were estimated from the collapse time of a gas cloud of the specified mass—using atomic and molecular H cooling—to the onset of ZAMS stage, and are not added to

Table 2
Implied ZAMS Pop III Star Observational Parameters Relevant to Caustic Transit Calculations

Mass[a] ZAMS ($M_\odot$)	$T_{\rm eff}$[b] (K)	Radius[c] at ZAMS ($R_\odot$)	$L_{\rm bol}$[d] ($L_\odot$)	$M_{\rm bol}$[e] (AB)	Bolo+IGM+K-corr[f] $z = 7$ (AB-mag)	$z = 12$ (AB-mag)	$z = 17$ (AB-mag)	ZAMS $m_{\rm UV}$[g] $z = 7$ (AB-mag)	$z = 12$ (AB-mag)	$z = 17$ (AB-mag)	$t_{\rm rise}$[h] caust (hr)	Transit[i] Rate (/cl/yr)
1.0	7.266e3	0.87	1.92	+4.03	+4.44	+3.13	+2.61	57.71	57.74	58.07	0.17	8×10^5
1.5	1.065e4	0.95	10.5	+2.18	+1.45	+0.42	−0.06	52.87	53.18	53.55	0.18	1.1×10^4
2.0	1.367e4	1.03	32.9	+0.95	+0.30	−0.59	−1.06	50.49	50.93	51.31	0.20	1.5×10^3
3.0	1.899e4	1.12	146.	−0.67	−0.51	−1.26	−1.72	48.06	48.64	49.03	0.22	182
5.0	2.805e4	1.23	846.	−2.58	−0.70	−1.35	−1.80	45.96	46.65	47.04	0.24	29.1
10	4.508e4	1.40	7.28e3	−4.91	−0.22	−0.79	−1.23	44.10	44.88	45.27	0.27	5.70
15	5.789e4	1.51	2.32e4	−6.17	+0.23	−0.30	−0.75	43.30	44.10	44.50	0.29	2.78
20	6.754e4	1.65	5.11e4	−7.03	+0.56	+0.04	−0.40	42.77	43.59	43.99	0.32	1.74
30	7.737e4	2.12	1.45e5	−8.16	+0.88	+0.36	−0.08	41.95	42.78	43.17	0.41?	0.82?
50	8.713e4	2.86	4.25e5	−9.33	+1.17	+0.66	+0.22	41.08	41.91	42.31	0.55*	0.37*
100	9.796e4	4.12	1.40e6	−10.63	+1.47	+0.96	+0.52	40.08	40.91	41.31	0.80*	0.15*
300	1.074e5	7.41	6.56e6	−12.30	+1.71	+1.21	+0.77	38.64	39.48	39.88	1.43*	0.039*
1000	1.080e5	12.9	2.02e7	−13.52	+1.72	+1.22	+0.78	37.44	38.28	38.68	2.48*	0.013*

Notes.

[a] Stellar mass in $M_\odot$. The physical parameters listed are for Pop III ZAMS stars in Table 1, as modeled with the MESA code for zero initial metallicity, zero mass loss, no rotation, and no stellar duplicity.

[b] Pop III star photospheric temperature $T_{\rm eff}$ in Kelvin.

[c] Pop III star radius $R_{\rm eff}$ at $T_{\rm eff}$ in $R_\odot$.

[d] Pop III star bolometric luminosity $L_{\rm bol}$ in $L_\odot$.

[e] Pop III star bolometric absolute magnitude $M_{\rm bol}$ in AB-mag.

[f] Combined bolometric + IGM + K-corrections to Pop III star $L_{\rm bol}$ at $z = 7$, $z = 12$, and $z = 17$, respectively, from Section 3.3.

[g] Pop III star apparent rest-frame-UV AB-magnitudes at $z = 7$, $z = 12$, and $z = 17$ in 2016 *Planck* cosmology (Planck Collaboration et al. 2016b) using the NIRCam filters that sample the rest-frame UV 1500 Å, assuming K-corrections as in columns 6–8 and no dust (for a discussion of dust, see Section 6.1). Distance moduli used are DM = 49.24, 50.58, and 51.42 mag at $z = 7$, $z = 12$, and $z = 17$, respectively.

[h] Upper limits to the caustic transit rise time $t_{\rm rise}$ (in hours) as estimated in Section 4.4. Asterisks indicate Pop III star ZAMS masses $M \gtrsim 50\,M_\odot$, for which cluster caustic transit events are possibly observable to the detection limits of *JWST* medium-deep to deep survey epochs reaching AB $\lesssim 28.5$–29 mag, assuming that caustic transit magnifications of $\mu \simeq 10^3$–10^5 can elevate Pop III stars with AB $\lesssim 35$–41.5 mag temporarily above these *JWST* detection limits. Stars with $M \lesssim 30\,M_\odot$ (labeled "?") likely remain undetectable even through caustic transits.

[i] The cluster caustic transit rate for Pop III stars (number of events per cluster per year) at $z = 12$ as estimated in Section 4.4 by directly using Equation (27). Appendices C–D summarize the uncertainties relevant to caustic transits.

the other ages below. Table 1 lists the MESA ages for the (estimated) pre-main sequence (pre-MS) collapse time, the core hydrogen-burning phase (H depletion), the shell hydrogen- and core+shell helium-burning phases (He depletion), and for the total time spent on the giant branches. That is, $(\tau_{\rm AGB}-\tau_{\rm MS})$ estimates the lifetime of the Red Giant Branch, Hot Horizontal Branch, and Asymptotic Giant Branch together (RGB+HHB +AGB). Note that for the most massive Pop III stars, HHB (core He burning) is of very short duration, and the stars essentially quickly transit from shell H burning to shell He burning in one smooth, nearly horizontal giant branch toward cooler $T_{\rm eff}$ values. We will refer to the combined RGB+HHB +AGB phases as the giant branch ("GB"). For all Pop III stars in Table 1, the time between core H depletion and core He depletion is about 8%–14% of the time between ZAMS and core H -depletion, with an average of ∼12%. For our caustic transit calculations in Section 4.4, we will take the approximate duration of the Pop III RGB and AGB stages each to be about 6% of their ZAMS lifetime.

For Pop III stars, we will use the post-MS lifetimes in Section 5 to estimate the maximum time that a lower-mass He-burning star may be feeding the accretion disk around a BH that was left over from a more massive Pop III companion star at $z \gtrsim 7$. This assumes no major mass exchange during the prior stellar evolution stages, i.e., we assume that stars in multiple systems evolve in isolation during the ZAMS stage following the MESA models.

For the most massive Pop III stars, their MS lifetime τ scales roughly as mass/luminosity. Since for the highest masses ($M \gtrsim 100\,M_\odot$) luminosities are directly proportional to their ZAMS mass (see Equation (2)), the MESA models yield MS ages of 5.6–2.1 Myr that are only weakly dependent on the ZAMS mass for the mass range of 30–1000 $M_\odot$, respectively (see Table 1). The shortest MS lifetime possible is ∼2.07 Myr, which happens when the star is radiating at the Eddington luminosity, and so its age becomes nearly independent of mass and only a function of fundamental constants. The MESA models in Table 1 indeed approach this MS age limit for $M \simeq 300$–1000 $M_\odot$ to within the modeling uncertainties.

In summary, Tables 1–4 show that Pop III stars in the mass range of $M \simeq 30$–1000 $M_\odot$ have ZAMS photospheric temperatures of 77,000–108,000 K, bolometric luminosities of $L_{\rm bol} \simeq 10^{5.2}$–$10^{7.3}\,L_\odot$, stellar radii of $R_{\rm MS} \simeq 2$–13 $R_\odot$, and main-sequence (MS) lifetimes of $\tau_{\rm MS} \simeq 2.1$–5.6 Myr. They may therefore be bright enough for occasional caustic transit detections by *JWST*, which is summarized in columns 9–13 of Tables 2–4, as calculated in Section 4.4.

As discussed in Sections 2.3 and 4.4, we only use upper limits to the integrated 1–4 μm sky-SB to estimate the maximum Pop III object caustic transit rate. Hence, the actual Pop III star lifetimes do not directly enter these calculations. However, plausible differences in Pop III star GB lifetimes as a function of ZAMS mass are relevant when estimating limits to the caustic transit rates from Pop III stellar-mass BH accretion

Table 3
Implied Red Giant Branch Pop III Star Observational Parameters Relevant to Caustic Transit Calculations

Mass[a] GB ($M_\odot$)	$T_{\rm eff}$[b] (K)	Radius[c] at Hydrogen Depletion ($R_\odot$)	$L_{\rm bol}$[d] ($L_\odot$)	$M_{\rm bol}$[e] (AB)	Bolo+IGM+K-corr[f] $z = 7$ (AB-mag)	$z = 12$ (AB-mag)	$z = 17$ (AB-mag)	Giant Branch $m_{\rm UV}$[g] $z = 7$ (AB-mag)	$z = 12$ (AB-mag)	$z = 17$ (AB-mag)	$t_{\rm rise}$[h] Caust (hr)	Transit[i] Rate (/cl/yr)
1.0	6.999e3	3.25	22.8	+1.35	+4.83	+3.48	+2.96	55.42	55.41	55.73	0.63	9×10^4
1.5	1.181e4	2.13	79.7	−0.01	+0.91	−0.06	−0.53	50.13	50.51	50.88	0.41	1.0×10^3
2.0	1.611e4	1.78	191.	−0.96	−0.19	−1.01	−1.47	48.08	48.60	48.99	0.34	175
3.0	2.311e4	1.53	598.	−2.20	−0.69	−1.39	−1.84	46.35	46.99	47.38	0.30	39.8
5.0	3.206e4	1.55	2.28e3	−3.66	−0.63	−1.25	−1.70	44.95	45.67	46.07	0.30	11.8
10	4.174e4	2.40	1.57e4	−5.75	−0.34	−0.92	−1.36	43.15	43.91	44.31	0.46	2.33
15	4.624e4	3.47	4.94e4	−6.99	−0.18	−0.74	−1.19	42.06	42.84	43.24	0.67?	0.87?
20	4.864e4	4.58	1.06e5	−7.82	−0.10	−0.65	−1.09	41.32	42.11	42.51	0.88*	0.44*
30	5.180e4	6.49	2.72e5	−8.85	+0.02	−0.53	−0.97	40.41	41.20	41.60	1.25*	0.19*
50	5.490e4	9.38	7.18e5	−9.90	+0.13	−0.42	−0.86	39.47	40.26	40.66	1.81*	0.081*
100	5.173e4	18.2	2.14e6	−11.09	+0.02	−0.53	−0.98	38.17	38.96	39.36	3.52*	0.024*
300	4.882e4	40.8	8.51e6	−12.59	−0.09	−0.65	−1.09	36.57	37.35	37.75	7.88*	0.006*
1000	4.807e4	74.8	2.68e7	−13.83	−0.12	−0.67	−1.12	35.29	36.07	36.47	14.44*	0.002*

Note.
[a–i] are as in Table 2. All parameters in this table are for Pop III stars at hydrogen depletion (RGB).

disks compared to those from Pop III stars. This is discussed in Sections 5–6, in which we consider the case where Pop III BH feeding may be done by lower-mass companion stars as soon as these can form in slightly polluted environments.

3.2. Multiplicity of Massive Stars

The effect of binaries and stars of higher multiplicities is a complex subject that can have a significant effect on population synthesis models of galaxies (e.g., Zhang et al. 2010; Conroy 2013; Stanway et al. 2016). For our present purposes, we must address the fact that the multiplicity factor MF is nearly unity for O-stars, at least in the local universe (e.g., Duchêne & Kraus 2013):

$$\mathrm{MF} = \frac{(B + T + Q)}{(S + B + T + Q)} \simeq 1. \qquad (4)$$

Here, S is the number of single stars in a coeval stellar population, B the number of binary stars, T the number of triples, and Q the number of quads, etc., implying that one gets essentially a factor $\gtrsim 2$ increase in the luminosity from binary, triple, and quad stars together.

3.2.1. Multiplicity—Low-mass End

At least 30% of all lower-mass stars in our own Galaxy occur in binaries (Kiminki & Kobulnicky 2012; Sana et al. 2012; Duchêne & Kraus 2013; Mayer et al. 2017), but at $z \gtrsim 7$, this fraction is unknown. The exact ratio of Pop III stars to BH accretion disks that are present will depend on the Pop III IMF slope, which is also unknown (e.g., Greif et al. 2011; Susa et al. 2014; Guszejnov et al. 2016; Ishiyama et al. 2016). We consider possible effects from the IMF slope in Sections 3.4, 4.4, and 5.

Table 1 shows that the pre-MS lifetimes of Pop III stars with $M \lesssim 1.5$–$2\,M_\odot$ would be $\gtrsim 3$–6 Myr, and thus generally exceed the $\lesssim 4$ Myr He-depletion age of 50–1000 $M_\odot$ stars. Hence, for coeval stellar populations with a large number of stars and a sufficiently flat mass function (i.e., $dN/dM \propto M^{-\alpha}$ with $\alpha \simeq 2$), a significant number of $\gtrsim 50\,M_\odot$ Pop III stars may be present that will have polluted the surrounding ISM with

their AGB mass loss—and supernovae in the right mass range —before stars with $M \lesssim 1.5$–$2\,M_\odot$ can have finished forming via their Hayashi tracks may be present. Hence, most early low-mass stars with $M \lesssim 1.5$–$2\,M_\odot$ may already have been polluted by coeval or precursor massive Pop III stars ($M \gtrsim 50\,M_\odot$), unless these low-mass stars formed in very isolated environments well away from the massive Pop III stars. The formation of lower-mass Pop III stars may also have been prevented by the very strong UV radiation field of nearby more massive Pop III stars, as discussed in Section 3.5.

The very first stars likely did not form until $z \lesssim 35$ (age $\lesssim 79$ Myr), or they would have left small-scale imprints on the Cosmic Microwave Background (CMB). Even if $\lesssim 1.5$–$2\,M_\odot$ Pop III stars had formed as early as $z \simeq 35$, Table 1 shows that their MS ages will be $\gtrsim 640$ Myr, so these low mass Pop III stars would not reach the giant branch until well below $z \lesssim 7.3$ (cosmic age 720 Myr) when reionization has essentially completed (Section 1). Therefore, Pop III stars with $M \lesssim 1.5$–$2\,M_\odot$—if they did manage to form as part of binary systems—could not fill their Roche lobes at $z \gtrsim 7$, and would not be relevant to Pop III BH accretion disk feeding at $z \gtrsim 7$ if stellar binaries were their progenitors. In Section 5, we will therefore not consider Pop III stars with masses $M \lesssim 2\,M_\odot$. With $T_{\rm eff} \simeq 10^4$ K (Table 1), low-mass Pop III stars could, however, contribute some reionizing flux. Their fractional contribution to the UV luminosity density depends on the value of the Pop III IMF slope (see Section 3.4).

3.2.2. Multiplicity—High-mass End

Under the assumption that (slightly polluted) massive stars at $z \gtrsim 7$ may occur in binary or multiple systems, then, for a Salpeter (1955) slope or flatter IMF, stars with $M \gtrsim 30\,M_\odot$ may have a lower-mass companion with $M \lesssim 30\,M_\odot$. The last column of Table 1 shows that these lower-mass companion stars with $M \gtrsim 2\,M_\odot$ will be in their RGB–AGB stage for $\tau_{\rm GB} \lesssim 30$–$60$ Myr, i.e., generally much longer than the plausible ages of a massive Pop III star in the binary. They could thus be feeding the BH that was left over from the more massive Pop III star after 2.4–6 Myr. As long as the more massive star—during its short GB lifetime ($\tau_{\rm GB}$)—

Table 4
Implied AGB Pop III Star Observational Parameters Relevant to Caustic Transit Calculations

Mass[a] AGB ($M_\odot$)	T_{eff}[b] (K)	Radius[c] at Helium Depletion ($R_\odot$)	L_{bol}[d] ($L_\odot$)	M_{bol}[e] (AB)	Bolo+IGM+K-corr[f] $z = 7$ (AB-mag)	$z = 12$ (AB-mag)	$z = 17$	AGB m_{UV}[g] $z = 7$ (AB-mag)	$z = 12$ (AB-mag)	$z = 17$	t_{rise}[h] Caust (hr)	Transit[i] Rate (/cl/yr)
1.0	6.312e3[j]	5.23[j]	39.8[j]	+0.74	+6.01	+4.57	+4.03	55.99	55.89	56.19	1.01	1.4×10^5
1.5	8.149e3	6.18	151	−0.71	+3.36	+2.14	+1.64	51.89	52.01	52.35	1.19	4.0×10^3
2.0	1.145e4	4.66	335	−1.57	+1.06	+0.07	−0.40	48.73	49.08	49.45	0.90	273
3.0	1.736e4	3.56	1.03e3	−2.79	−0.36	−1.15	−1.60	46.09	46.64	47.03	0.69	28.9
5.0	2.658e4	2.89	3.74e3	−4.19	−0.72	−1.38	−1.82	44.33	45.01	45.41	0.56	6.43
10	3.938e4	3.03	1.98e4	−6.00	−0.42	−1.00	−1.45	42.82	43.57	43.97	0.58	1.71
15	4.215e4	4.55	5.88e4	−7.18	−0.33	−0.90	−1.34	41.73	42.50	42.89	0.88?	0.64?
20	4.386e4	6.14	1.25e5	−8.00	−0.27	−0.84	−1.28	40.97	41.74	42.14	1.19*	0.32*
30	4.006e4	11.7	3.17e5	−9.01	−0.40	−0.98	−1.42	39.83	40.59	40.98	2.26*	0.11*
50	3.536e4	24.3	8.32e5	−10.06	−0.55	−1.15	−1.59	38.63	39.37	39.77	4.70*	0.036*
100	3.392e4	44.0	2.31e6	−11.17	−0.59	−1.19	−1.64	37.49	38.22	38.61	8.50*	0.012*
300	3.165e4	101.	9.19e6	−12.67	−0.64	−1.26	−1.71	35.93	36.65	37.04	19.49*	0.003*
1000	3.122e4	163.	2.26e7	−13.65	−0.65	−1.28	−1.72	34.94	35.66	36.05	31.45*	0.001*

Notes.
[a]–[i] are as in Table 2. All parameters in this table are for Pop III stars at helium depletion (AGB).
[j] The 1.0 $M_\odot$, $Z = 0.00\ Z_\odot$ MESA model did not ignite helium and may thus become a helium white dwarf. To complete our calculations, its AGB parameters listed are for the $Z = 10^{-8}\ Z_\odot$ model, which did end in a white dwarf.

does not transfer the majority of its mass to the less massive companion star, the resulting accretion timescale around the BHs that were leftover from $M \gtrsim 30\ M_\odot$ stars (Section 5.3) would be mainly driven by the much longer τ_{GB} of the less massive star when it leaves the MS and fills its Roche lobe.

A reasonable upper limit for the BH-feeding timescale by a lower-mass star in a binary filling its Roche lobe is thus the $\lesssim 60$ Myr GB age of an $M \gtrsim 2\ M_\odot$ star. Table 1 suggests that the lower limit on the timescale of BH accretion disk feeding from the more massive companion stars is $\lesssim 0.3$ Myr. In Section 5.3, we will therefore assume a lifetime range for BH accretion disk feeding of 0.3–60 Myr. That is, Pop III BH accretion may last up to $\lesssim 10\times$ longer than that of the Pop III stars themselves, but it could also be $\sim 10\times$ shorter. The *Spitzer–Chandra* power-spectrum results may already hint at a BH contribution, as discussed in Section 2.3.2. The MESA stellar evolution physics of zero-metallicity stars summarized in Table 1 thus provides a theoretical framework that allows massive Pop III stars to leave stellar-mass BHs with accretion disks that may feed up to $\sim 10\times$ longer than these massive Pop III stars live themselves. We will discuss the implications of this in Sections 5–6.

3.2.3. Massive Star Multiplicity at Low Redshifts

O stars in nearby surveys show significant multiplicity ($\gtrsim 80\%$) and have a rather flat mass-ratio distribution:

$$q \equiv M_{sec}/M_{pri}, \quad (5)$$

where M_{pri} is the more massive star (Duchêne & Kraus 2013). In theory, the q-value distribution can be as steep as the Salpeter (1955) mass function slope, i.e.,

$$N(q) \propto q^{-2.35}. \quad (6)$$

The observed q-distribution of nearby O-stars seems to have a slope much flatter than the IMF slope (Duchêne & Kraus 2013). Since nearby surveys of double/multiple stars may suffer from flux bias, very faint low-mass stellar companions around more massive stars are harder to find. In Sections 5–6, we will

therefore assume that slightly polluted early massive stars have a mass ratio no steeper than the IMF slope, if they already occur in binaries.

Figure 2 of Duchêne & Kraus (2013) suggests that the majority of nearby binary OB stars have typical orbital periods in the range of ~ 10–130 days and typical orbital separations between 0.067–0.51 au or $14 \lesssim D \lesssim 110\ R_\odot$. Larger and smaller separations can occur as well, some as small as a few $R_\odot$. Each of their Roche lobes will be about half that in effective radius, or $7 \lesssim R \lesssim 55\ R_\odot$. Following the MESA models of Tables 1–4, it is therefore possible that if Pop III or II.5 stars exist in binaries at $z \gtrsim 7$ as in OB binary stars today, their lower-mass companion stars with $M \gtrsim 2\ M_\odot$ will fill their Roche lobes in the GB stage at $z \gtrsim 7$ with sizes $2 \lesssim R \lesssim 160\ R_\odot$ (see column 3 of Tables 3–4). This would feed the BH remnant from the more massive Pop III star during their AGB–MS lifetime, which could last up to $\lesssim 60$ Myr (Section 3.2.1). Their BH UV accretion disk half-light radii (r_{hl}) are estimated in Section 5.5 to be in the range of $1 \lesssim r_{hl} \lesssim 30\ R_\odot$, and so in general will fit inside these Roche lobes when the companion star in the binary reaches the AGB stage.

Since we do not include Pop III star multiplicity in the MESA models, this will render the Pop III caustic transit rates of Section 4.4 more conservative. This is because caustic transit detections may be $\gtrsim 0.75$ mag brighter for binary stars—and possibly multipeaked in their detailed time sequence—than we estimate for single Pop III stars in Section 4.4. One example of a multiple caustic transit event has been suggested for a possible massive binary star at $z \simeq 1.5$ (Kelly et al. 2018). Future work that includes the evolution of massive binaries with possible mass exchange will need to model the evolutionary tracks of zero- or very low-metallicity stars, and how this may affect the BH feeding timescales.

3.3. Bolometric Corrections after IGM Transmission and K-corrections

The luminosities and absolute magnitudes of Pop III stars summarized in Section 3.1 and Tables 1–4 were calculated by

the MESA code in bolometric solar units without making bolometric corrections, corrections for IGM transmission, or K-corrections. We thus need to correct the theoretical Pop III star luminosities for these effects to predict the m_{AB} values observed in the *JWST* filters at a given redshift. The exact bolometric and K-corrections cannot be computed until the actual object redshifts have been estimated from the eight-band *JWST* NIRCam photometry and/or measured with *JWST* NIRSpec or NIRISS spectra. For our current photometric predictions, we will therefore proceed as outlined below.

3.3.1. Pop III Star Bolometric+IGM Corrections

We use zero-metallicity blackbody spectra for Pop III stars with ZAMS, RGB, and AGB T_{eff} values and radii from Tables 2–4 to estimate the bolometric correction (BC) and their IGM corrections as follows. Pop III stars of $30 \lesssim M \lesssim 1000\,M_\odot$ with photospheric temperatures of $T = 77{,}400$–$108{,}000$ K (Table 1) have SEDs with rest-frame wavelength peaks in I_ν around $\lambda_{max} \simeq 620$–444 Å, respectively.

The IGM at $z \gtrsim 7$ is opaque for rest-frame $\lambda \lesssim 1216$ Å due to the significant fraction of neutral hydrogen in the immediate foreground of each First Light object (Haardt & Madau 2012). Hence, for our caustic transit calculations in Sections 4.4 and 6.2, it only matters that the Pop III star or its stellar-mass BH accretion disk is UV-bright down to the 1216 Å Lyα limit, since the opaque neutral hydrogen forest at $z \gtrsim 7$ will certainly block the shorter hard-UV wavelengths from both the stars and BH accretion disks. Therefore, although many Pop III objects may have unreddened UV SED β slopes (Calzetti et al. 1994) much steeper than those corresponding to $T \simeq 30{,}000$ K, in what follows we will only consider their luminosities and fluxes in the rest-frame UV continuum at 1216–2000 Å. At $z \simeq 7$–17, this is the only rest-frame wavelength range that *JWST* will sample to AB $\simeq 28$–30 mag over its most sensitive NIRCam wavelength range of 1–4 μm.

To predict their fluxes as observed in the *JWST* 1–5 μm filters, we integrated the assumed Pop III blackbody spectra in I_ν and computed the fraction of flux longwards of $\lambda = 1216$ Å compared to the total bolometric flux integrated from $\lambda = 0$ to $+\infty$. The SEDs of Pop III stars with $M \simeq 30$–$1000\,M_\odot$ as observed through the *JWST* NIRCam filters are predicted to be about 9.5×–23× fainter than their bolometric model luminosities from column 5 of Tables 2–4. After accounting for the drop in IGM transmission to 0% at $z \gtrsim 7$, the actual BC of Pop III stars would thus make them about 2.4–3.4 mag fainter in absolute magnitude and in the m_{AB} magnitude to be observed with *JWST*. The average BC that we need to apply is thus about +2.9 mag. This is much less than the formal BC of ~ 6 mag implied for a $T \simeq 10^5$ K star (Flower 1996). This is because the Pop III SED below Lyα 1216 Å is completely blocked by the IGM at $z \gtrsim 7$, and so does not enter the BC. Note that we derive the BC with the opposite sign as defined in Flower (1996), since we need to go from theoretical bolometric values to the predicted values to be observed with *JWST*.

3.3.2. Pop III Star K-corrections

For each fixed *JWST* filter, we need to apply the K-correction to the Pop III star flux at $\lambda \gtrsim 1216$ Å that makes it through the IGM. Hogg (1999) and Hogg et al. (2002 and references

therein) define the K-correction in I_ν units as follows:

$$K = -2.5 \log\left[(1 + z)L_{(1+z)\nu}/L_{\nu_e}\right] \quad \text{(mag)}. \quad (7)$$

This includes the effects of bandpass shifting due to the object's redshift and the change of the object's rest-frame SED with frequency or wavelength. The factor of $(1+z)$ accounts for the fact that the flux and luminosity are not bolometric, but are flux densities per unit frequency (Hogg et al. 2002). Due to the complete IGM absorption for $\lambda \lesssim 1216$ Å at $z \simeq 7$–17, our specific K-term needs to correct the flux predicted in the bluest available *JWST* filter that is completely longwards of redshifted Lyα for these effects. Hence, the luminosity ratio in Equation (7) only needs to account for the brighter flux shortward of this filter that still makes it through the IGM.

At $z = 7$, the bluest available *JWST* filter that is above the Lyman-break (F115W or short J-band) samples rest-frame $\lambda_{rest} \simeq 1440$ Å, while at $z = 12$ the F200W filter (short K-band) samples $\lambda_{rest} \simeq 1645$ Å, and at $z = 17$ the F277W filter samples $\lambda_{rest} \simeq 1540$ Å. Since the most massive Pop III ZAMS stars have nearly uniform temperatures of $T \simeq 10^5$ K (Table 1), all 1–5 μm *JWST* filters sample the Rayleigh–Jeans tail of their SEDs at $z \simeq 7$–17, so that $I_\nu \propto \nu^2$. The peak in their rest-frame (I_ν) SEDs thus occurs only a factor of 2–3.5 below the central rest-frame wavelength sampled by these filters. We then compute this K-correction from the flux bluewards of each filter down to the 1216 Å IGM transmission cutoff. Using Equation (7), the K-correction follows from these wavelength ratios as

$$K \simeq -2.5 \log\left[(1 + z)(1216/\lambda_c)^2\right] \text{(mag)}. \quad (8)$$

Here, λ_c is the central rest-frame wavelength of the bluest *JWST* filter used for object detection at each redshift listed in Tables 2–4. The wavelength ratio $(\lambda_{Ly\alpha}/\lambda_c)^2$ reflects the fact that we can only sample the $I_\nu \propto \nu^2$ tail of the very blue SEDs longwards of Lyα at $z \gtrsim 7$. For the extremely blue Pop III stars, this total K-term amounts to about -1.9 mag brighter at $z = 7$, -2.5 mag brighter at $z = 12$, and -2.9 mag brighter at $z = 17$. The K-correction gets somewhat brighter at the higher redshifts, because the $(1+z)$ factor dominates the term from the wavelength or frequency ratio. For all longer wavelength *JWST* filters, the K-corrections can be computed similarly.

3.3.3. The Combined Bolometric+IGM+K-corrections

The combined bolometric+IGM+K-corrections (hereafter "BIK"-corrections) were calculated in two independent ways. First, all integrations were done in the observed frame while folding their blackbody SEDs with the *JWST* NIRCam filters and integrating the bolometric+IGM corrections as described in Section 3.3.1, and the K-corrections as in Section 3.3.2. Second, the calculations were done in the rest frame after folding the blackbody SEDs with the appropriately blueshifted *JWST* NIRCam filter curves, and integrating between each filters' rest-frame FWHM wavelength cutoffs. The NIRCam interference filters were designed to resemble block functions (Rieke et al. 2005), so this approximation is valid. Both methods gave similar m_{AB} flux results to within $\lesssim 0.05$–0.20 mag, with the second method producing m_{AB} fluxes that are on average only 0.09 mag brighter than the first method. The m_{AB} values listed in Tables 2–4 are therefore the slightly brighter values from the second method, since our
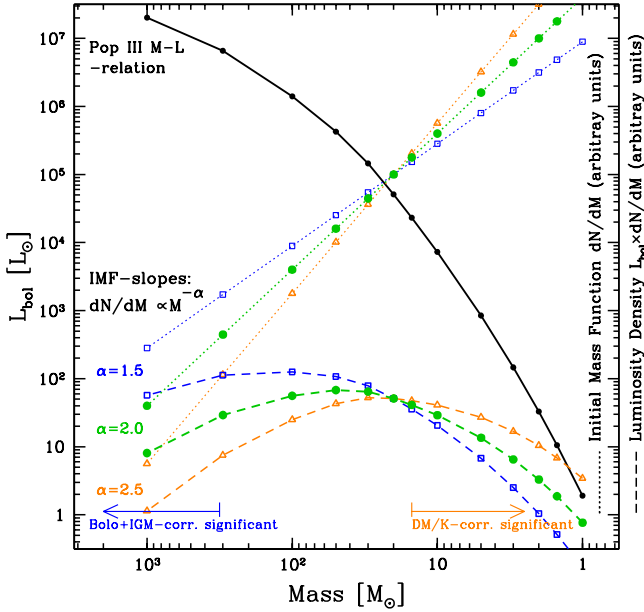
Figure 3. Luminosity density (dashed curves) for early star-forming objects inferred from the ZAMS Pop III mass–luminosity relation (solid black line) from Table 1 in Section 3.1. The ZAMS Pop III M/L relation is folded with three different IMF slopes (dotted lines), ranging from $\alpha = 1.5$ (top heavy; blue), $\alpha = 2.0$ (normal; green), and $\alpha = 2.5$ (steep IMF; orange). For a Pop III IMF slope of $\alpha = 2.0$, the luminosity density peaks at around $30\,M_\odot$, while most of the population's luminosity density is produced between 10–$100\,M_\odot$.

caustic transit calculations use in all cases the upper limits in predicted fluxes.

These combined BIK-corrections need to be added to the M_{AB} values from Tables 2–4 to yield the predicted m_{AB} values that *JWST* would observe:

$$m_{AB}(z) = M_{AB} + DM(z) + (BC + IGM + K)(z), \qquad (9)$$

where the distance moduli DM at redshift z are listed in the footnotes of Table 2. To first order, the bolometric+IGM and K-corrections are comparable in magnitude, but are opposite in sign for the Pop III stars in Table 1. The combined BIK-corrections are, therefore, in general, modest, but can, on average, result in making objects ~ 1–2 mag fainter than what the intrinsic bolometric luminosities from the MESA models would yield.

We list the combined corrections for $z = 7$, $z = 12$, and $z = 17$ in columns 6–8 of Tables 2–4 for the ZAMS, RGB, and AGB, respectively. Columns 9–11 list the resulting apparent rest-frame UV AB-magnitudes for these redshifts that result from the absolute bolometric magnitudes in column 5 and these combined corrections, respectively. These are directly used in the calculation of our caustic transit rise times and rates in Section 4.4, which are listed in columns 12–13, respectively.

For the lowest-mass stars that have $T_{eff} \lesssim 10^4$ K in Tables 2–4, the combined BIK-corrections are significantly positive, since for such cool stars most of their blackbody-like SED is redshifted to well longwards of the bluest *JWST* filter that samples above the 1216 Å break at the anticipated object redshift. These cool, low-mass Pop III stars are thus predicted to be always very faint (AB $\gtrsim 50$ mag), and permanently out of reach for *JWST* caustic transits. At the intermediate temperatures of Pop III RGB and AGB stars of $T_{eff} \simeq 50{,}000$ $-30{,}000$ K, respectively (see Figure 2 and Section 3.1), the combined BIK-corrections are in general negative but no

brighter than -1 to -2 mag (Tables 3–4), because the peak of their blackbody SED falls between the Lyα 1216 Å IGM cutoff and the rest-frame wavelength range covered by the bluest NIRCam filter that *JWST* will use to detect the Pop III object at $z \simeq 7$–17, so that the full benefit of the K-correction for a very blue SED is achieved. For much hotter ZAMS Pop III star temperatures of $T_{eff} \simeq 10^5$ K, the BIK-corrections are generally positive but no dimmer than $+1$ to $+2$ mag (Table 2), because most of the energy in their blackbody SED now falls well below the Lyα 1216 Å IGM cutoff. Hence, the combined BIK-corrections are more advantageous for detecting the cooler Pop III RGB and AGB phases than for the hotter Pop III ZAMS stars. This can be seen in columns 6–8 of Tables 2–4 and will be folded into the caustic transit rate calculations in Section 4.4.

3.4. Luminosity Density from the Mass–Luminosity Relation and Initial Mass Function

The ZAMS Pop III mass–luminosity relation discussed in Section 3.1 and Table 2 has important implications for the mass range that dominates the luminosity density of a faint star-forming object at $z \gtrsim 7$. This is indicated in Figure 3, where the ZAMS M/L relation from Table 2 is indicated by the solid black line:

$$L_{bol} \propto M^\lambda, \qquad (10)$$

with the mass-dependent slope λ from Table 2, as approximated by the segmented power laws in Equation (3).

In our caustic transit calculations of Section 4.4, we assume that small early star-forming objects exist at $z \gtrsim 7$. These will be mostly fainter than the *HST* or *JWST* detection limits, and contribute a total 1–4 μm sky-SB whose upper limits were discussed in Section 2.3.1. In this context, it is necessary to consider which stars will dominate the luminosity density of these faint star-forming objects, which is defined as the number of stars per unit area on the sky. We consider their IMF to be a power law,

$$dN/dM \propto M^{-\alpha}, \qquad (11)$$

with three different IMF slopes in Figure 3 (dotted curves), ranging from "top heavy" ($\alpha = 1.5$; blue), "intermediate" ($\alpha = 2.0$; green), and "steep" ($\alpha = 2.5$; orange), that bracket a range of plausible IMFs (dotted curves; e.g., Scalo 1986; Bastian et al. 2010; Coulter et al. 2017). The ZAMS Pop III M/L relation is folded with these three IMF slopes in Figure 3 to yield the luminosity density:

$$L_{bol}(dN/dM) \propto M^\lambda M^{-\alpha} \propto M^\delta, \qquad (12)$$

where $\delta = (\lambda - \alpha)$ is the slope of the luminosity density versus mass relation. For an IMF slope $\alpha \simeq 2.0$ and the mass-dependent slope of the M/L relation in Equation (3), we infer a strongly positive slope of the luminosity density versus mass relation (dashed green line): $\delta \simeq +1.2$ for $M \simeq 1$–$20\,M_\odot$, a nearly zero slope ($\delta \simeq 0.1$) for $M \simeq 20$–$100\,M_\odot$, and a negative slope ($\delta \simeq -0.8$) for $M \simeq 100$–$1000\,M_\odot$.

For an IMF slope of $\alpha \simeq 2.0$, most of the Pop III ZAMS bolometric energy from faint star-forming objects at $z \gtrsim 7$ is thus produced by stars with masses between 10–$100\,M_\odot$, with a somewhat smaller contribution from stars with $M \simeq 100$–$1000\,M_\odot$, and a much smaller contribution from $M \simeq 1$–$10\,M_\odot$, which is compounded by the significant K-correction for the lowest-mass stars (Section 3.3). For an

IMF slope of $\alpha \simeq 2.0$, the Pop III luminosity density peaks at around $30\,M_\odot$ with a broad plateau (green dashed curve in Figure 3). These are the Pop III stars with the most advantageous bolometric+IGM+K-correction values (Tables 2–4) and are within reach for *JWST*, assuming caustic transits can occur as described in Section 4.4.

For a top-heavy IMF, most of the luminosity density would be produced by stars with $M \simeq 100\,M_\odot$ (blue dashed lines in Figure 3), while for a very steep Salpeter-like IMF, most energy is still produced by stars as massive as $M \simeq 20\,M_\odot$ (orange dashed lines). Hence, irrespective of any reasonable IMF slope at $z \gtrsim 7$, the Pop III star mass–luminosity relation implies that the highest near-IR sky-SB will be produced by stars with $20 \lesssim M \lesssim 100\,M_\odot$. These are precisely the stars that are most likely to become visible during caustic transits, as discussed in Section 4.4.

3.5. Estimating the Surface Brightness from Massive Pop III Stars

Mas-Ribas et al. (2016) give the number of Lyman–Werner (LW) photons produced by ZAMS Pop III stars as a function of their mass. A $300\,M_\odot$ Pop III star emits $\dot{N} = 3.1 \times 10^{49}$ photons s^{-1} in the LW band, which spans the energy range of $E = 11.2$–13.6 eV. From these numbers, we find a flux of $m_{\mathrm{LW}} \simeq 38.6$ mag at $z \simeq 12$. The ZAMS Pop III stars from the MESA runs in Tables 1–2 are about 0.3 mag brighter in their bolometric M_{AB} magnitude.

We estimate the surface density of Pop III stars using the models of Sarmento et al. (2018) in which the Pop III star formation rate density reaches $\sim 10^{-3.5}\,M_\odot$ yr^{-1} Mpc^{-3} at $z = 12$ (see their Figure 2). If we assume that these consist of $M = 300\,M_\odot$ stars that each live $t \simeq 2 \times 10^6$ yr (Schaerer 2002, see Table 1 here), then the total number of Pop III stars per arcsec^2 is

$$\approx 0.03 \left(\frac{\mathrm{SFR}_{\mathrm{Pop\,III}}}{10^{-3.5} M_\odot/\mathrm{yr}} \right) \left(\frac{M_*}{300\,M_\odot} \right)^{-1} \frac{\mathrm{stars}}{('')^2 \Delta z}. \quad (13)$$

If each of these stars has $m_{\mathrm{AB}} \gtrsim 38.6$ mag (see Table 2), then the total surface brightness in Pop III stars could be fainter than ~ 36.0 mag arcsec^{-2}.

Theoretically, it appears unlikely that Pop III stars alone can fully account for an IR background with SB $\simeq 31$ mag arcsec^{-2}. In order to reach SB $\simeq 31$ mag arcsec^{-2}, we need $\sim 10^3$ massive Pop III stars per arcsec^2, each with AB $\simeq 38.5$ mag. Most of these must be at lower redshift ($z \lesssim 12$), because of the strong redshift dependence of dS/dz, as discussed in Section 2.3.1. This would require $\gtrsim 100$ massive Pop III stars per arcsec^2 at $z \simeq 12$, with a weak z-dependence. This is much larger than the numbers we calculate above, although close to the "no LW" case of Trenti & Stiavelli (2009).

Strong LW radiation from massive Pop III stars can significantly suppress the formation of subsequent lower-mass stars in their immediate environment, resulting in a possibly much dimmer total Pop III sky-SB of ~ 36 mag arcsec^{-2}. This SB-level is also indicated in orange in Figure 1. In Sections 4.5 and 7–8, we will therefore consider a range of Pop III near-IR SB-levels and the scope of observing programs needed for *JWST* to detect their resulting caustic transits in each case.

4. Estimates of Cluster Caustic Transits for Pop III stars

The question that we address in this section is: under what conditions could *JWST* detect the individually lensed Pop III stars of Section 3 at very high magnification, identified as a sudden onset of an AB $\simeq 28.5$ mag point source, corresponding to a caustic crossing in which two additional critical images[10] are formed?

4.1. Selection of Lensing Clusters for JWST Caustic Transit Observations of Pop III Objects

To address the caustic transit rate and duration for Pop III stars, we first need to evaluate the plausible limits on the transverse velocities of massive lensing clusters, their typical caustic lengths, and the possible effects from microlensing. From this, we will in Section 4.4 estimate the cluster caustic transit rates for the Pop III stars of Section 3.

A Pop III caustic transit observing program with *JWST* would likely select the best lensing clusters that also have matching deep *HST* images in previous epochs—including WFC3 IR data—such as the Hubble Frontier Field clusters (HFF; e.g., Kawamata et al. 2016; Acebron et al. 2017; Lagattuta et al. 2017; Lotz et al. 2017; Mahler et al. 2018) or the CLASH clusters (e.g., Postman et al. 2012; Rydberg et al. 2015).

The HFF clusters were chosen to have the capability for significant lensing magnification. Many are highly elongated and could well have significant internal velocities between cluster subcomponents and/or a significant space velocity compared to the nearby cosmic web, as discussed in Section 4.2 and Appendix A. Indeed, in two of these clusters (MACS J0416–2403 and MACS J1149.5+2223) to date, possible caustic transits have been identified at lower redshifts ($z \simeq 1.0$–1.5; Diego et al. 2017; Kelly et al. 2018; Rodney et al. 2017).

A *JWST* lensing cluster program should select the best lensing clusters with redshifts $0.3 \lesssim z \lesssim 0.5$. This is because of a combination of the following two factors. First, the SED of the 5–8 Gyr old stellar population in these clusters at $z \simeq 0.4$ peaks at $\lambda \simeq 1.6\,\mu$m in the rest frame (e.g., Kim et al. 2017). This includes the SED of the substantial ICL that is present in massive virialized clusters. Also, the zodiacal foreground in *JWST*'s second Lagrange point (L2) orbit strongly declines between 1–3.5 μm (see Figure 1), so the best *JWST* sensitivity per unit time is obtained in the observed wavelength range of 2–3.5 μm. This is the critical wavelength range for detecting First Light objects at $z \lesssim 17$. Hence, ideally, the redshift of lensing clusters to be observed with *JWST* should be kept at $z \lesssim 0.5$. This is so that the rest-frame peak SED of the cluster galaxies and the ICL does not redshift as much into the most sensitive 2.5–3.5 μm NIRCam filters and thus compromise the ability to make First Light object detections, including Pop III caustic transits.

Higher-redshift clusters will of course have lower ICL and cluster galaxy brightness, because of the stronger cosmological $(1+z)^4$ SB dimming. They are also less massive by selection, and may therefore not always be the most optimal gravitational lenses. Because of their younger ages, they may also be less virialized. There exist exceptional clusters at $z \gtrsim 0.5$, of

---

[10] The critical images are two formally infinite magnification images of a point source, which form on the critical curve for a source at the location of the caustic line.
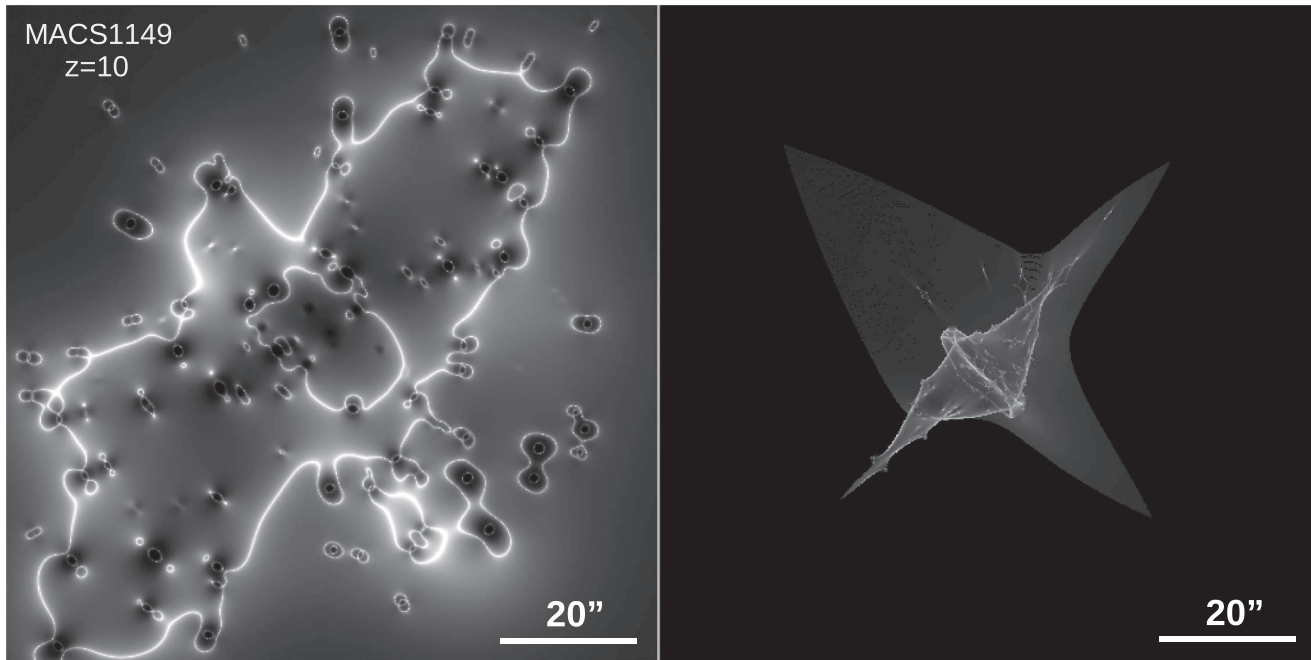
Figure 4. Left panel: Example of the lensing magnification map for galaxy cluster MACS J1149.5+2223 at $z \simeq 0.4$ and a background source at $z = 10$ (e.g., Lotz et al. 2017 and references therein). Light from the cluster galaxies is not shown to avoid overcrowding, but can be found in these papers. The white areas mark the critical curves, where the maximum lensing magnification is observed from this cluster for a background source with half-light radius $r_{hl} \lesssim 0''\!.5$ at $z = 10$. The lightest regions have the highest magnification ($\mu \gtrsim 10$–20), while the darkest regions are areas of low magnification ($\mu \simeq 1$ or even $\mu \lesssim 1$) around the cluster member galaxies. Right panel: Example of the caustic map produced by the cluster mass model for a background source at $z = 10$. This is the location where a point source at $z = 10$ produces maximum magnification. The total length of the cluster caustics can be as large as $L \simeq 100''$, which we adopt as upper limit to the typical caustic length in our caustic transits calculations.

course, that could be used for lensing studies with *JWST*, such as, e.g., El Gordo at $z \simeq 0.87$ (Zitrin et al. 2013).

In Section 4.5, we suggest that 3–30 clusters need to be monitored by *JWST* for several years for Pop III caustic transit studies. In the end, practical arguments, such as available *HST* images—especially at shorter wavelengths ($\lambda \lesssim 0.6\ \mu$m) than those that *JWST* can observe—the quality of available lensing models, ancillary data such as available ground-based spectra and X-ray images, and the ability to schedule *JWST* observations for at least half a year during each *JWST* cycle will likely determine which cluster sample is best suited for Pop III caustic transit observations with *JWST*.

4.2. Maximum Plausible Transverse Velocity of Lensing Clusters

In this section, we consider the possible maximum transverse (or tangential) velocity, v_T, of a massive cluster, which has visible substructure in both its measured redshift/velocity distribution as well as in its spatial extent on the sky (see Figure 4 and Appendix A).

4.2.1. General Limits to Transverse Velocities of Nearby Clusters

First, we will summarize what typical space velocities are seen for clusters locally to get an idea what v_T values to expect for massive clusters at $z \gtrsim 0.3$. The best-determined CMB dipole value is $3364.5 \pm 2.0\ \mu$K (Fixsen et al. 1996; Hinshaw et al. 2009; Planck Collaboration et al. 2016c). Compared to the best-determined current CMB temperature values of $T = 2.72548 \pm 0.00057$ (Fixsen 2009; Planck Collaboration et al. 2016c), this corresponds to the average velocity of the solar system of $370.1 \pm 0.2\ \mathrm{km\,s}^{-1}$ in the direction of

$l^{\mathrm{II}} = 264°\!.00 \pm 0°\!.03$ and $\quad b^{\mathrm{II}} = 48°\!.24 \pm 0°\!.20$ in Galactic coordinates (Hinshaw et al. 2009; Planck Collaboration et al. 2014, 2016a).

The Local Group is falling into the Virgo Cluster at $\sim$250 km s^{-1}(e.g., Dressler 1991). More recent studies suggest that the bulk velocity of Virgo plus the Local Group toward the CMB is $631 \pm 20\ \mathrm{km\,s}^{-1}$ (e.g., Hoffman et al. 2015, 2017; Watkins & Feldman 2015a, 2015b). This bulk motion may be as much due to gravitational pull from a Great Attractor in the direction of the Shapley overdensity as well as a push from large local underdensities or voids roughly in the opposite direction (e.g., Hoffman et al. 2017). That is, the net space velocity of the solar system moving with the Local Group and the Virgo Cluster toward the Great Attractor would correspond to a one-dimensional velocity of $\sim$631/$\sqrt{3} \simeq 364\ \mathrm{km\,s}^{-1}$ when viewed from a random direction. This would be close to its v_T value when viewed from a random point in space. To calculate the net v_T value below that includes the solar system motion, we will use the actual solar system velocity of $370\ \mathrm{km\,s}^{-1}$ toward the CMB from the Planck Collaboration et al. (2016a) results above, which is similar in value.

From a large sample of clusters, Bahcall & Oh (1996) suggest a $\lesssim$5% probability of finding clusters with one-dimensional peculiar velocities greater than $600\ \mathrm{km\,s}^{-1}$, while the one-dimensional cluster peculiar velocity ranges between 300–$600\ \mathrm{km\,s}^{-1}$ (rms) for various cosmological models, in line with the bulk velocity implied for the Virgo cluster above. It thus seems reasonable for nearby massive clusters to have transverse space velocities of $300 \lesssim v_T \lesssim 600\ \mathrm{km\,s}^{-1}$.

These velocities are substantially smaller than the relative velocity of the subcluster components seen in the Bullet cluster at $z \simeq 0.296$ (e.g., Tucker et al. 1998; Clowe et al. 2006). The

relative transverse velocity of the two Bullet subcluster components may be as high as $3000 \lesssim v_T \lesssim 4500$ km s^{-1} (Molnar et al. 2013). Based on cosmological simulations with the largest possible volumes, Thompson & Nagamine (2012) and Watson et al. (2014) emphasize that the probability of merging subclusters with masses exceeding $10^{14} M_\odot$ and velocities this high are rare, except perhaps in non-standard models (Angus & McGaugh 2008). Based on hydrodynamical models, Springel & Farrar (2007) suggest a more modest transverse velocity for the Bullet cluster subcomponents of ~ 2700 km s^{-1}. The question then arises: what are reasonable values for v_T for the massive ($M \gtrsim 10^{15} M_\odot$) and best lensing clusters at $0.3 \lesssim z \lesssim 0.5$ to be selected for observations of the First Light epoch with *JWST*?

4.2.2. Maximum Transverse Velocities Adopted for Lensing Clusters at $0.3 \lesssim z \lesssim 0.5$

Throughout, we take (V_T, s) to mean the net transverse velocity, accounting for the transverse motion of both the observer, the lens, and the source planes. The effective transverse velocity V_T, s when observing a source at $z_s \simeq 7$–17 lensed by a cluster at $z_d \simeq 0.4$ is computed as follows, starting with the sum of the relevant velocity vectors scaled with the appropriate angular diameter distances, using Equation B9 of Kayser et al. (1986):

$$\boldsymbol{V}_T, s = \boldsymbol{v}_s/(1 + z_s) + \boldsymbol{v}_{\rm obs}(D_{ds}/D_d)/(1 + z_d) - \boldsymbol{v}_T(D_s/D_d)/(1 + z_d). \quad (14)$$

Here, the first term due to the source motion at $z \simeq 7$–17 is negligible at $v_s \lesssim 30/(1 + z_s) \lesssim 2$–4 km s^{-1} (Barkana & Loeb 2002), using the expected small velocity dispersion of the low-mass halos they likely reside in at high redshifts. The unknown bulk motion of the halo at $z \gtrsim 7$ could increase this to several $100/(1+z)$ km s^{-1}, or $\lesssim 40$ km s^{-1}. In either case, this first term is much smaller than the last two terms. The second term is due to the velocity of $v_{\rm obs} \simeq 370$ km s^{-1} of the solar system (moving in the Galaxy and the Local Group, and with the Virgo cluster) toward the CMB from Section 4.2.1—modulated by the Earth's motion around the Sun of ~ 30 km s^{-1}–and scales with the ratio of the angular diameter distance between the deflector to source, D_{ds}, and the angular diameter distance to the cluster deflector, D_d. The third term v_T is due to the transverse cluster motion itself, and scales with the angular diameter distance ratio of the source to the deflector, D_s/D_d. We ignore here the intrinsic velocities of any microlenses in the lens plane, since these are demagnified in the source plane by large factors, as shown by Kayser et al. (1986) in the high-magnification regime of interest.

To assess the transverse velocity v_T for lensing clusters at $0.3 \lesssim z \lesssim 0.5$, we perturbed the observed redshift distribution of three promising HFF clusters with a random space velocity and determine how much of the projected space velocity can be added to the transverse direction, before the other projected component added to the line-of-sight velocities disturbs the observed cluster redshift distribution too much. Details of this simulation are given in Appendix A and its figure. These show that adding space velocities with projected transverse components much larger than $v_T \simeq 1000$ km s^{-1} imply projected components of this space velocity added along the line of sight that are not consistent with the available redshift data in the cluster core, although somewhat smaller values are certainly allowed. This results in an upper limit of $v_T \lesssim 1000$ km s^{-1} for the maximum transverse velocity of these clusters at $0.3 \lesssim z \lesssim 0.5$ in the plane of the sky. For some substructures in each cluster, the v_T values may well be as high as 1000 km s^{-1}. Since we cannot currently distinguish if the whole cluster or only several of its subclumps are moving transversely at $v_T \lesssim 1000$ km s^{-1}, the integral caustic length that we defined in Section 4.3 and use in Section 4.4 to calculate the caustic transit rate is therefore also an upper limit.

For the HFF cluster subsample discussed in Appendix A at least, it seems appropriate to use $v_T \lesssim 1000$ km s^{-1} for the maximum transverse velocity of these clusters in the plane of the sky, as projected from their space velocity. Kelly et al. (2018) find that for the cluster MACS J0416–2403—behind which their caustic-transiting star at $z \simeq 1.5$ was identified—the transverse velocity is about 1000 km s^{-1}. From large N-body simulations, Watson et al. (2014) find that most pairwise halo velocities are $\lesssim 3000$ km s^{-1} with a median of ~ 1000 km s^{-1}.

The maximum effect for caustic transits in Equation (14) is obtained when the solar system velocity vector, $\boldsymbol{v}_{\rm obs}$, toward the CMB as projected on the sky at the cluster location is exactly anti-aligned with the transverse cluster velocity vector, $\boldsymbol{v}_T$, which is captured by the minus sign in Equation (14). If both lens and observer are going in the same direction, one would obtain the smallest value for this velocity difference. The actual transverse vector sum will be different for each cluster by an unknown amount, depending on how the unknown direction of the transverse vector of each cluster, $\boldsymbol{v}_T$, aligns with the velocity vector, $\boldsymbol{v}_{\rm obs}$, of the solar system toward the CMB. Because both projected transverse velocities are vectors, the typical expected value of the velocities is the rms, not the velocity difference. Hence, we add both in quadrature, so that for $z_d \simeq 0.4$ and $z_s \simeq 12$ we obtain

$$|V_T, s| \simeq \sqrt{[370\sqrt{2/3} \times 0.40]^2 + [1000 \times 0.48]^2} \quad (15)$$

using the velocities discussed above and the appropriate angular diameter distance ratios for our adopted cosmology. This amounts to $V_T, s \simeq 495$ km s^{-1}. When exactly anti-aligned, the two velocity components would just add without the transverse projection factor $\sqrt{2/3}$ of the solar system velocity at the location of the lensing cluster, so that $V_T, s \lesssim 502$ km s^{-1}. Given the significant differences in the allowed v_T values between the three HFF clusters discussed in Appendix A, we will adopt for our caustic transit calculations in Section 4.4 an upper limit of $V_T, s \lesssim 1000$ km s^{-1}.

4.3. Estimates of Cluster Caustic Lengths for Pop III Caustic Transit Calculations

Gravitational lensing modeling will result in lensing maps in the plane of the cluster. The clusters selected for *JWST* are assumed to be in the redshift range $0.3 \lesssim z \lesssim 0.5$, following the arguments of Section 4.1. An example of a lensing map is shown in Figure 4(a) for sources at $z = 10$ behind the HFF cluster MACS J1149.5+2223 at $z \simeq 0.4$ (e.g., Lotz et al. 2017). Detailed lensing maps of the HFF clusters have been made by, e.g., Jauzac et al. (2014, 2015), Lam et al. (2014), Diego et al. (2015a, 2015b, 2016a, 2016b), and Caminha et al. (2017). An example of a caustic map for a source at $z = 10$ behind MACS J1149.5+2223 is given in Figure 4(b). These

caustic maps are similar for Pop III sources at $7 \lesssim z \lesssim 17$, which may be observable to JWST via caustic transits.

Making lensing maps is a complex process that introduces its own uncertainties, which depend, e.g., on the detailed input cluster mass distribution, the number and redshift distribution of available sources with multiple images through which the lensing model gets refined, the point spread function (PSF), quality, and the depth of the images at *HST* resolution used to reconstruct the lensed sources, and on a number of other factors such as the actual amount of cluster substructure present. For details, we refer to, e.g., Meneghetti et al. (2017), where errors in the reconstruction of HFF-like clusters are discussed based on simulated data.

In essence, the more numerous the input redshifts and the better the input imaging data are, the more reliable the lensing model will become. Ideally, the entire gravitational field of the cluster with all its substructure, dwarf galaxies, detailed ICL distribution, and stellar microlenses would have been modeled. For exact caustic transit modeling of a known source at lower redshift (e.g., Diego et al. 2017; Kelly et al. 2018), detailed lensing modeling is necessary, since a caustic transit may have been observed at one location, where the local gravitational lensing model then exactly matters for the correct interpretation of the observed data. This is, e.g., the case when a known background galaxy provides the stellar object that transits the cluster caustic at a specific location.

For the present work, detailed lensing models are not required, since Pop III stars at $z \gtrsim 7$ may be present everywhere at average sky-SB levels no brighter than the upper limits adopted in Section 2.3, where we calculated that most of their diffuse flux will come from objects that are well beyond the *HST* and *JWST* point-source detection limits. For the current caustic transit calculations, we will assume that the integrated near-IR sky-SB of these "unresolved" objects at $z \gtrsim 7$ is rather uniform (see Section 2.3.1). That is, we do not need to know the exact lensing model at each location along the caustic, since Pop III caustic transits at $z \gtrsim 7$ can happen anywhere at unpredictable locations along the caustic lines in the cluster.

Instead, we need the general properties of the caustics to estimate the rate of transits and an estimate of the maximum magnification around the caustic as a function of source angular size. We are interested in the statistics/probability of seeing these rare events. Hence, the typical global properties of the lenses, such as the area in the background plane with magnifications above a given threshold or the statistical presence of microlenses that can modify these properties in the high-magnification regime, are the only quantities that are relevant here.

For the present goal of order-of-magnitude estimates of Pop III object caustic transits at $z \gtrsim 7$, we will thus assume average caustic lengths L and geometry. Line integration of the lensing models in clusters like Figure 4(b) shows that their typical total caustic length is $L_{\mathrm{caust}} \lesssim 100''$, which we will use as an upper limit in the caustic transit calculations in Sections 4.4 and 6.2.

4.3.1. Details of Lensing Magnification Near the Cluster Caustics

Although usually represented in the source plane for convenience, caustics actually form in the observer plane, and it is the relative motion with respect to these caustics that produce the peaks in the observed light curves (see Section 4.2.2). For most clusters, caustics tend to adopt a diamond-like shape aligned in the same direction as the main symmetry axis of the ellipsoid that encloses the cluster (see Figures 4(a) and (b)). Since these three planes are uncorrelated, the vector of the relative transverse velocity can point in any direction with respect to the caustic pattern. As discussed in Section 4.2.2, we should consider all possible directions when estimating caustic transit rates. In detail, the velocity of the caustics is complex, because the shape of the network changes in addition to the transverse movement. Given the other larger uncertainties in cluster geometry, v_T, and L_{caust}, we will henceforth ignore the projection effects from the angle, i, between the cluster's unknown main velocity vector and the main direction of the caustic at each location, which will average out to $\langle \sin(i) \rangle \simeq 1/2$.

When a background star crosses a cluster caustic, it can be magnified by a factor of up to $\mu \simeq 10^5 - 10^6$ for a short period of time (few weeks to months; see Section 4.4), depending on the strength of the caustic and the stellar radius. This magnification can thus boost the apparent brightness of the star by $\sim 12.5 - 15$ mag. Possible modifications from microlensing are discussed in Section 4.3.2 and Appendices B–B.2. Fainter Pop III stars with AB $\simeq 41 - 43$ mag could then be observed with *JWST* to AB $\lesssim 28.5 - 29$ mag during one of these caustic crossing events. At larger distances from the caustic ($\simeq 1$ pc), the magnification is more moderate ($\mu \simeq 10^3$), and only the brightest stars with AB $\lesssim 36$ mag could be observed via a caustic transit, but they could remain visible for many years because they remain visible farther away from the caustic. Microlensing can reduce these magnifications and spread the microlensed events over a larger area still, which lengthens their visibility in time. This is discussed further in Appendices B–B.2.

For the more ubiquitous fold caustics, the magnification near a caustic varies with the distance to the caustic, d, as

$$\mu = B_o / \sqrt{d}, \qquad (16)$$

where B_o is a constant that depends on the derivatives of the gravitational potential. For clusters like the HFFs, B_o is normally in the range 10–20, while d is expressed in arcseconds (see, e.g., Miralda-Escude 1991; Diego et al. 2017, for a detailed discussion). Hence, for a background Pop III star at $z \gtrsim 7$, magnifications of order $\mu \simeq 10^3$ can be attained once the background star is $\simeq 1$ pc away from the caustic (or $d \simeq 0.''001$). For an HFF-like cluster with $L_{\mathrm{caust}} \simeq 100''$ at $z \simeq 12$, this implies that an area of ~ 0.1 arcsec^2 in the source plane can magnify background stars by more than a factor of $\mu \simeq 10^3$, so that any star brighter than AB $\simeq 36$ mag can be lensed to above the detection limit of *JWST* and produce a double image separated by less than $0.''5$ around the critical curve. When we get closer to the critical curve, the double-lensed image that would appear on each side of the critical curve will be unresolved at *JWST*'s near-IR resolution of $\sim 0.''08$ FWHM if the separation between the two images is smaller than ~ 25 milliarcsec. At these separations, the total magnification would be $\mu \simeq 10^4$, and any star brighter than AB $\simeq 38$ mag could be lensed to above the detection limit of *JWST*. This corresponds to an area of $\simeq 3 \times 10^{-4}$ arcsec^2 in the source plane.

At even smaller distances to the true caustic, fainter and smaller stars would become visible, but the probability of magnifying a star in this narrower region would be smaller. Clearly, there is a trade-off between the luminosity function

slope of the background stars and the probability of being magnified above a certain value, which is given by the area in the source plane, $A(>\mu)$, that has a magnification larger than μ. This magnification area seems to follow a power law for the more ubiquitous fold caustics, $A(>\mu) = B_o/\mu^2$, where B_o will vary somewhat from cluster to cluster.

Owing to this scaling of the area with μ^2, it is easy to see (Kelly et al. 2018) that the optimal trade-off between the luminosity function and $A(>\mu)$ happens when the stellar LF slope is close to $\alpha \simeq -2$, where $dN/dL \propto L^\alpha$ is the luminosity function of the background stars. One possible complication when observing a lensed bright star at moderate magnifications (i.e., a star brighter than AB $\simeq 36$ mag and with $\mu \lesssim 10^3$) is that the timescale for the flux variation could be very long ($\gtrsim$ hundreds of years). This makes it a challenging task to distinguish between a lensed Pop III star and a larger unlensed substructure, such as a globular cluster in the background galaxy. Spectroscopy of brighter caustic transits will be necessary to help reveal their nature, as discussed in Section 7. Microlensing fluctuations will likely make the light curve of a caustic transit more variable and spread over a longer period of time, as discussed in Section 4.3.2 and Appendices B–B.2.

4.3.2. Possible Role of Microlenses during Caustic Transits

For completeness, we will also consider here the case where the caustics are disrupted by microlenses, as discussed in Kelly et al. (2018) and Diego et al. (2017). A way to distinguish a Pop III star from a small background substructure would be through microlensing by low-mass stars and stellar remnants in the intracluster medium (e.g., Lewis et al. 2000). First, the timescale for microlensing would be on the order of days to weeks instead of years. Second, microlensing by star-like objects in the lens plane would affect only very compact objects in the background, such as Pop III stars and their stellar-mass BH accretion disks. Larger objects magnified by factors of $\mu \simeq 10^3$ would be much larger than the Einstein radius of the microlenses, resulting in microlensing being irrelevant for such objects. Third, microlensing events would take place around the critical curve, at separations of $\sim 0.''1$ on either side of the critical curve, instead of just at the location of the critical curve. Finally, if microlenses are ubiquitous in the lens plane, a single bright star in the background can be responsible for multiple peaks, all of them with exactly the same spectrum, which would increase the rate of observed caustic transit events. These would have to be monitored over the long term (see Section 7.3) and modeled in detail for proper interpretation.

As discussed in Diego et al. (2017), the relatively high magnifications near the critical curve of a cluster amplify not only the background object but also the lensing distortion produced by otherwise negligible microlenses from the intracluster medium. In the magnification regime of $\mu \simeq 10^3$ (about 1 pc from the caustic at z $\simeq 12$), a microlens with $M \simeq 1\,M_\odot$ behaves as a microlens with an effective mass of hundreds of $M_\odot$. These large effective masses can magnify a Pop III star by $\mu \gtrsim 10^4$, instead of by the expected factor of $\mu \simeq 10^3$ that would occur without microlenses. This translates into a temporary boost on timescales of days to weeks of ~ 3 mag with respect to the case without microlenses, depending on the microlens mass and the radius of the Pop III star. Multiple microlens magnification events can occur for many years before or after the Pop III star aligns with the cluster caustic, thereby greatly increasing the chance of seeing these stars. On the other hand, a large number of microlenses

around the critical curve may disrupt the caustic in such a way that extreme magnifications of order $\mu \simeq 10^5$ are no longer possible. Hence, only the brightest Pop III stars may be observed this way, thereby reducing the pool of background objects that can be observed by *JWST*. For more details on caustic transits in the presence of microlensing, we refer the reader to the discussion in Diego et al. (2017).

To compute the caustic transit rates in this paper, we consider two cases: with and without microlenses. The case without microlenses is more straightforward, since it involves only the properties of the caustic, the sky-SB of background Pop III stars, and the relative velocity between the caustic and the stars. As mentioned above, all Pop III stars brighter than AB $\simeq 35$–41.5 mag could potentially be observed during a caustic crossing ($\mu_{\rm max} \lesssim 10^3$–$10^5$). The caustic transit rate for this particular case is discussed in Section 4.4. The case *with* microlenses renders similar results, but is more uncertain since it depends on the actual IMF of faint microlensing stars in the foreground cluster ICL. We refer the reader to Appendices B–B.2 for further details of possible Pop III object caustic transits in the presence of microlenses.

4.4. Implied Estimates of Cluster Caustic Transits for Pop III Stars Without Microlensing

The question that we will address in this section is: if a fraction of the diffuse near-IR background is generated by Pop III stars—with a conservative upper limit to their near-IR sky-SB of $\gtrsim 31$ mag arcsec^{-2} (Section 2.3)—then what is the probability that *JWST* will catch one of these Pop III stars being lensed by a cluster caustic transit?

For our calculations, we start with the premise that this maximum 1–4 μm sky-SB is made up of ZAMS Pop III stars with AB $\gtrsim 37.5$ mag at $z \gtrsim 7$ (Table 2). During their RGB and AGB stages, these Pop III stars may be as "bright" as AB $\gtrsim 35$ mag at $z \gtrsim 7$ (Tables 3–4). Pop III stars in the mass range of $30 \lesssim M \lesssim 1000\,M_\odot$ are the most likely to be detected by *JWST* at $z \gtrsim 7$ at AB $\lesssim 28.5$–29 mag if the caustic magnifications reach $\mu \gtrsim 10^4$–10^5. We will assume that the geometrical optics approximation still holds in this very small source regime. To reference our calculations following Equation (3), we define the apparent magnitude of a $100\,M_\odot$ star with luminosity L_{100} at $z = 12$ as m_{100}, and that of a $20\,M_\odot$ star with luminosity L_{20} at $z = 12$ as m_{20}. The AB-magnitudes at other redshifts scale with the DM and BIK-corrections in Tables 2–4.

Our caustic transit calculations depend on stellar luminosity, which depends on the ZAMS mass following the Pop III star mass–luminosity relation of Equation (3). To generalize our caustic transit calculations in the relevant equations below, we will propagate the three different mass-dependent power-law slopes in Equation (3) over the entire Pop III ZAMS mass range of $1 \lesssim M \lesssim 1000\,M_\odot$. For $10 \lesssim M \lesssim 1000\,M_\odot$, the ZAMS radii scale as

$$R = R_{100}(M/100\,M_\odot)^{0.45},\qquad(17)$$

following the first line of Equation 2 in Section 3.1.

Given a population of Pop III stars with luminosity L, the number density $N(L)$ required to make up a surface brightness of AB $\simeq 31$ mag arcsec^{-2} follows from

$$(m_{100} - 31) = 2.5\log_{10}(N(L) \times L_{100}),\ {\rm or:}$$
$$N(L) \times L_{100} = 10^{(m_{100}-31)/2.5}.\qquad(18)$$

Given the segmented M/L relation in Equation (3), we can generalize this as a function of mass M as follows:

$$N(M) = 10^{\frac{(m_{100} - 2.5 \log (L/L_{100}) - 31)}{2.5}}$$
$$= \left(\frac{L}{L_{100}}\right)^{-1} 10^{\frac{(m_{100} - 31)}{2.5}}$$
$$\simeq \left(\frac{M}{100}\right)^{-1.16} 10^{\frac{(m_{100} - 31)}{2.5}}, \text{ for } M \gtrsim 100 \, M_\odot,$$
$$\simeq \left(\frac{M}{100}\right)^{-2.06} 10^{\frac{(m_{100} - 31)}{2.5}}, \text{ for } 20 \lesssim M \lesssim 100 \, M_\odot,$$
$$\simeq 2.33 \left(\frac{M}{20}\right)^{-3.20} 10^{\frac{(m_{20} - 31)}{2.5}}, \text{ for } M \lesssim 20 \, M_\odot,$$

$$(19)$$

in units of arcsec^{-2}, while all masses are in $M_\odot$. The extra constant in the last line of Equations (19)–(30) reflects the change in normalization at $100 \, M_\odot$ in the first two mass ranges to $20 \, M_\odot$ in the last mass range.

To be observed with *JWST* in a single epoch at a flux limit of AB $\lesssim 28.5$ mag, a star of mass M would need a lensing magnification of

$$\mu(M) = 10^{\frac{(m_{100} - 2.5 \log (L/L_{100}) - 28.5)}{2.5}}$$
$$= \left(\frac{L}{L_{100}}\right)^{-1} 10^{\frac{(m_{100} - 28.5)}{2.5}}$$
$$\simeq \left(\frac{M}{100}\right)^{-1.16} 10^{\frac{(m_{100} - 28.5)}{2.5}}, \text{ for } M \gtrsim 100 \, M_\odot,$$
$$\simeq \left(\frac{M}{100}\right)^{-2.06} 10^{\frac{(m_{100} - 28.5)}{2.5}}, \text{ for } 20 \lesssim M \lesssim 100 \, M_\odot,$$
$$\simeq 2.33 \left(\frac{M}{20}\right)^{-3.20} 10^{\frac{(m_{20} - 28.5)}{2.5}}, \text{ for } M \lesssim 20 \, M_\odot.$$

$$(20)$$

Now the typical magnification at a distance $d_\mu \simeq 1''$ from a caustic is $\mu(1'') = 10$ (see Section 4.3 and Figure 4(b)), using the conservative lower value of $B_o \simeq 10$ in Section 4.3.1, so that

$$\mu \simeq 10 \frac{1}{\sqrt{d_\mu}}.$$

$$(21)$$

The angular distance (in arcsec) from the true caustic maximum corresponding to a magnification μ is then

$$d_\mu(M) = 100 \, \mu^{-2}$$
$$= 100 \left(\frac{L}{L_{100}}\right)^2 10^{\frac{-(m_{100} - 28.5)}{1.25}} \text{ arcsec}$$
$$\simeq 100 \left(\frac{M}{100}\right)^{2.32} 10^{\frac{-(m_{100} - 28.5)}{1.25}}, \text{ for } M \gtrsim 100 \, M_\odot,$$
$$\simeq 100 \left(\frac{M}{100}\right)^{4.12} 10^{\frac{-(m_{100} - 28.5)}{1.25}}, \, 20 \lesssim M \lesssim 100 \, M_\odot,$$
$$\simeq 18.4 \left(\frac{M}{20}\right)^{6.4} 10^{\frac{-(m_{20} - 28.5)}{1.25}}, \, 1 \lesssim M \lesssim 20 \, M_\odot.$$

$$(22)$$

To estimate the relevant timescales, we need to know the crossing time for a caustic passing over a distance d_μ, and the crossing time for a typical radius of a ZAMS Pop III star with $M \simeq 30$–$1000 \, M_\odot$. Our adopted cosmology yields 3740 pc per arcsec at $z \simeq 12$. Using the 2012 IAU value for one astronomical unit (au) of 149.6×10^9 m (Prša et al. 2016), $1.0 \, R_\odot$ (or 695, 700 km; Section 3.1) then corresponds to $\sim 6.03 \times 10^{-12}$ arcsec at $z \simeq 12$. Hence, the ZAMS Pop III stars in Table 1 are $\sim 5.2 \times 10^{-12}$–7.78×10^{-11} arcsec across at $z \simeq 12$, and at most between $\sim 1.3\times$–$14\times$ larger during their RGB–AGB phases, which together last $\sim 8\times$ shorter than the ZAMS (Section 3.1). Pop III RGB–AGB star caustic transits will thus be more rare, although according to Tables 3–4 also ~ 1.5–2.5 mag brighter than those of ZAMS Pop III stars (Table 2). To obtain lower limits to the caustic transit rise times and upper limits to their caustic transit rates, we will therefore use the Pop III star ZAMS parameters in Tables 1–2.

The upper limit of V_T, $s \lesssim 1000 \, \text{km s}^{-1}$ (Section 4.2 and Appendix A) corresponds to an angular speed of $d\theta/dt \simeq 1.83 \times 10^{-7}$ arcsec/yr for a galaxy cluster at $z \simeq 0.4$. At this redshift, there are ~ 5590 pc/arcsec in *Planck* cosmology. Using Equation (17), the crossing times for the Pop III star radius $R(M)$ across the $\mu > 10^4$ magnification region $\theta_\mu(M)$—needed to make stars of mass M detectable to *JWST*—leads then to a mass-dependent Pop III star caustic transit timescale of

$$t_R(M) = \frac{R}{v_{T,s}} = \frac{R}{R_{100}} \frac{R_{100}}{v_{T,s}} \simeq \frac{R_{100}}{v_{T,s}} \left(\frac{M}{100}\right)^{0.45}, \quad (23)$$

and:

$$\theta_\mu(M) = \left(\frac{d\theta}{dt}\right)^{-1} \frac{100''}{\mu^2(M)}$$
$$\simeq \frac{100}{d\theta/dt} \left(\frac{M}{100}\right)^{2.32} 10^{\frac{-(m_{100} - 28.5)}{1.25}}, \text{ for } M \gtrsim 100 \, M_\odot,$$
$$\simeq \frac{100}{d\theta/dt} \left(\frac{M}{100}\right)^{4.12} 10^{\frac{-(m_{100} - 28.5)}{1.25}}, \, 20 \lesssim M \lesssim 100 \, M_\odot$$
$$\simeq \frac{18.4}{d\theta/dt} \left(\frac{M}{20}\right)^{6.4} 10^{\frac{-(m_{20} - 28.5)}{1.25}}, \, \text{for } M \lesssim 20 \, M_\odot. \quad (24)$$

This implies that the brightening time—defined as the time for the magnification to go from zero to its maximum value—for a Pop III star is very short (~ 0.5–3 hr) when it transits the caustic starting at the "highest-magnification edge." The star would then stay bright for several months to a year, with brightness decaying as $1/\sqrt{t - t_o}$, where $(t - t_o)$ is the time since the stellar disk started the caustic crossing at time t_o.

Also, this entire process is reversible, so one could witness a very slow rise of an object's flux as $1/\sqrt{t_f - t}$ when it moves toward the caustic starting from the low-magnification end, followed by an abrupt disappearance once it crosses the caustic at some future time t_f. We will discuss below and in Section 7 how *JWST* may detect each of these cases.

To calculate how often we expect such a brightening, we assume that the cluster has a length $L_{\text{caust}} \simeq 100''$ of caustic (see Section 4.3). To calculate the rate at which Pop III lensing occurs, we need the area crossed by the caustics per unit time.

This change in area is given by

$$\frac{dA}{dt} = L_{caust} \times v_T = 100 \times 2.4 \times 10^{-14} \text{ arcsec}^2 / s$$

$$\simeq 1.6 \times 10^{-5} \left(\frac{L_{caust}}{100''}\right)\left(\frac{v_T}{1000 \text{ km s}^{-1}}\right)('')^2 \text{ yr}^{-1}. \quad (25)$$

The surface areas referred to here are all in the source plane, and so there is no depletion correction for magnification. The number of events therefore follows from the surface density of Pop III stars N, yielding

$$\frac{dN_{lens}}{dt} = N(M) \times \frac{dA}{dt}$$

$$= \frac{dA}{dt}\left(\frac{L}{L_{100}}\right)^{-1} 10^{\frac{(m_{100}-31)}{2.5}} \text{ yr}^{-1}. \quad (26)$$

To quantify the values for N, t_μ, and $\frac{dN_{lens}}{dt}$, we base our numbers on the discussion of the physical parameters of Pop III stars in Section 3.1. From Table 2, the luminosity of a $100\,M_\odot$ star is $\sim 1.40 \times 10^6\,L_\odot$, giving an absolute magnitude of -10.63 AB-mag, using $M \equiv +4.74$ mag for the absolute magnitude of the Sun (see Section 3.1). Including the bolometric+IGM+K-corrections of Table 2, the corresponding apparent magnitude m_{100} at $z = 12$ is then $m_{100} = 40.91$ mag, assuming no extinction. For a discussion of dust, see Section 6.1. From Table 1, the radius of a $100\,M_\odot$ star is $R_{100} = 4.12\,R_\odot$. We then find the following caustic crossing rate for lensed Pop III stars:

$$\frac{dN_{lens}}{dt} = N(M) \times \frac{dA}{dt}$$

$$= 0.064 \frac{L_{caust}}{100''} \frac{v_T}{1000}\left(\frac{M}{100}\right)^{-1.16}, \text{ for } M \gtrsim 100\,M_\odot,$$

$$= 0.064 \frac{L_{caust}}{100''} \frac{v_T}{1000}\left(\frac{M}{100}\right)^{-2.06}, \; 20 \lesssim M \lesssim 100\,M_\odot,$$

$$= 1.76 \frac{L_{caust}}{100''} \frac{v_T}{1000}\left(\frac{M}{20}\right)^{-3.20}, \; 1 \lesssim M \lesssim 20\,M_\odot \quad (27)$$

per year. The duration of a brightening time is then

$$t_R(M) = 9.1 \times 10^{-5}\left(\frac{M}{100}\right)^{0.45} \text{yr}, \quad (28)$$

which for a $100\,M_\odot$ ZAMS star is about 0.80 hr for $v_T \lesssim 1000$ km s^{-1}. The range in rise times in Table 2 is ~ 0.4–2.5 hr for $M \simeq 30$–$1000\,M_\odot$ Pop III stars.

The time spent above the detection limit is

$$t_\mu(M) \simeq 0.4\left(\frac{v_T}{1000}\right)^{-1}\left(\frac{M}{100}\right)^{2.32}, \text{ for } M \gtrsim 100\,M_\odot,$$

$$\simeq 0.4\left(\frac{v_T}{1000}\right)^{-1}\left(\frac{M}{100}\right)^{4.12}, \text{ for } 20 \lesssim M \lesssim 100\,M_\odot,$$

$$\simeq 5.3 \times 10^{-4}\left(\frac{v_T}{1000}\right)^{-1}\left(\frac{M}{20}\right)^{6.4}, \text{ for } M \lesssim 20\,M_\odot \quad (29)$$

in units of years. This assumes that after the flux has peaked upon caustic crossing, the flux declines as $1/\sqrt{t}$ following Equation (21), assuming constant velocity v_T.

The number of lensed Pop III stars visible at a given time is then

$$N_{lens} = t_\mu(M)\frac{dN_{lens}}{dt}$$

$$\simeq 0.026 \frac{L_{caust}}{100''}\left(\frac{M}{100}\right)^{1.16}, \text{ for } M \gtrsim 100\,M_\odot,$$

$$\simeq 0.026 \frac{L_{caust}}{100''}\left(\frac{M}{100}\right)^{2.06}, \text{ for } 20 \lesssim M \lesssim 100\,M_\odot,$$

$$\simeq 9.4 \times 10^{-4}\frac{L_{caust}}{100''}\left(\frac{M}{20}\right)^{3.20}, \; 1 \lesssim M \lesssim 20\,M_\odot. \quad (30)$$

Note that while $t_\mu(M)$ and $\frac{dN_{lens}}{dt}$ are sensitive to the transverse velocity v_T, the visible number of events is not. Equations (27)–(30) contain the key relations of this paper for calculating Pop III star caustic transits.

For an IR background of $\gtrsim 31$ mag arcsec^{-2} (Section 2.3) made up of AB $\simeq 41$ mag Pop III stars with $M \simeq 100\,M_\odot$, we estimate that one lensing event can be observed above a flux limit of AB $\simeq 28.5$ mag per cluster per ~ 2.7 years, or one event when monitoring ~ 3 clusters during a year. Because these events should stay detectable at $\mu > \mu(M = 100\,M_\odot)$ for $t_\mu \simeq 0.4$ years, this implies that ~ 0.15 such lensed Pop III sources per cluster would be observed above the flux limit at any given time.

These results are sensitive to the luminosity of the Pop III stars. For example, let us instead try the extreme case where the Pop III stars are $1000\,M_\odot$, and so have AB $\simeq 38.3$ mag at $z \simeq 12$ rather than AB $\simeq 40.9$ mag. This implies that the source needs to be magnified less, and so it can be observed while farther from the caustic, with a visible time of $t_\mu = 40$ yr. However, the rate is lower with $dN_{lens}/dt \simeq 0.013$ per year per cluster, giving one event per cluster per 75 years. In this case of brighter, more massive stars, we find that ~ 0.5 events per cluster would be visible at any given time.

Thus, for $100\,M_\odot$ Pop III stars, about six clusters observed twice about six months apart would make the likelihood of observing a lensed Pop III star of order unity, while for more massive stars, detecting a new lensing event (with a time baseline limited to 1 year) would require the observation of a larger number of clusters in proportion to the mass M. For lower-mass stars, fewer clusters would need to be observed, as long as they can appear magnified above the detection thresholds of Tables 2–4.

The observed rate of events will thus also depend on the mass function of Pop III stars. A mass-function weighted average over Equation (27) is

$$\left\langle\frac{dN_{lens}}{dt}\right\rangle = 0.064\left(\frac{L_{caust}}{100''}\right)\left(\frac{v_T}{1000 \text{ km s}^{-1}}\right)$$

$$\times \int_{M_{min}}^{M_{max}} dM\left(\frac{M}{100}\right)^{-1}\frac{dP}{dM}dM \text{ yr}^{-1}. \quad (31)$$

Here, dP/dM is the normalized instantaneous mass function of Pop III stars, i.e., the population of stars available to be lensed, not the entire IMF. We assume a power-law mass function $dP/dM \propto M^{-\alpha}$ with slope $|\alpha| > 1$ in the range

$M_{\min} < M < M_{\max}$, leading to

$$\left\langle \frac{dN_{\text{lens}}}{dt} \right\rangle = 0.064 \left(\frac{L_{\text{caust}}}{100''} \right) \left(\frac{v_T}{1000 \text{ km s}^{-1}} \right) \left(\frac{M_{\min}}{100} \right)^{-1}$$

$$\times \left(\frac{\alpha - 1}{\alpha} \right) \left(\frac{1 - (M_{\min}/M_{\max})^{\alpha}}{1 - (M_{\min}/M_{\max})^{\alpha - 1}} \right) \text{yr}^{-1}. \quad (32)$$

For $M_{\max} \gg M_{\min}$, the last term is close to unity, while for steep mass functions with $\alpha \gg 1$ the integral converges to the value of $\langle \frac{dN_{\text{lens}}}{dt} \rangle \rightarrow \frac{dN_{\text{lens}}}{dt}(M_{\min})$. The choice of the mass function slope $\alpha \simeq 2.0$ was discussed in Sections 2.3.1 and 3.4, and is used below.

We chose the lower-mass boundary here at $30 M_{\odot}$ since such stars may be visible through caustic transits to *JWST* (Tables 2–4), and such stars may produce BH leftovers in the mass range already observed by LIGO at $M \gtrsim 14 M_{\odot}$ (e.g., Abbott et al. 2016e, see also Section 5.1). Lower Pop III stellar masses would render the stars too faint to be reasonably observed at AB $\lesssim 28.5$ mag through caustic transits in a single *JWST* epoch (Table 2), except for perhaps RGB–AGB Pop III stars with $M \gtrsim 15$–$20 M_{\odot}$. The latter may be visible because their *K*-corrections are more advantageous (Tables 3–4) than for the much hotter ZAMS Pop III stars (Section 3.3.3).

For the lowest Pop III star mass considered here ($M_{\min} = 30 M_{\odot}$), its physical parameters given in Section 3.1, and adopting an IMF slope of $\alpha \simeq 2.0$, we get the following upper limits for $L \simeq 100''$ and $V_T, s \lesssim 1000 \text{ km s}^{-1}$:

$$\left\langle \frac{dN_{\text{lens}}}{dt} \right\rangle \leqslant 0.064 \left(\frac{30}{100} \right)^{-1} \left(\frac{1.0}{2.0} \right) \text{yr}^{-1}. \quad (33)$$

For a ZAMS Pop III star mass function slope of $\alpha \simeq 2$, the weights for each mass bin in Table 2 are very similar at 0.23–0.17 following Equation (32).

The resulting total transit rates for stars with $M \gtrsim 30 M_{\odot}$ that are in principle observable with *JWST* across the caustics are then predicted to be $dN_{\text{lens}}/dt \lesssim 0.30$ events per cluster per year. These caustic transits that may be visible to *JWST* are marked with an asterisk in columns 12–13 of Tables 2–4.

To this we need to add the caustic transit rates expected for the RGB from Table 3 and the AGB from Table 4. These must be weighted with their approximate lifetimes compared to the ZAMS, which are ~6% of the ZAMS lifetime (Section 3.1) for each of the RGB and AGB phases detectable by *JWST*. This amounts to an additional 0.01 transits per cluster for each of the RGB and AGB phases. Hence, the weighted total number of caustic transits for ZAMS, RGB and AGB Pop III stars in Tables 2–4 are ~0.32 per cluster per year.

4.5. Observing Strategies for JWST to Detect Pop III Stars via Caustic Transits

From Section 4.4, it follows that in order to see one Pop III star caustic transit per year at the top of the Pop III star mass function ($M \gtrsim 15$–$30 M_{\odot}$), one would need to observe about three clusters at least two times per year about six months apart in one to two successive *JWST* cycles. Observing more often when scheduling allows for clusters at higher zodiacal latitude would, of course, be preferred. The first exposure pair would be needed to identify a potential Pop III star caustic transit event, and the last pair is needed to monitor its expected decay on a timescale of less than one year. Imaging in all eight broadband NIRCam filters is essential to identify the high-redshift dropout nature of a potential caustic transit event and to identify foreground interloping events, which will be more numerous but interesting in their own right. For the brighter caustic transit events, follow-up spectroscopy should be attempted to confirm the nature of the transit, as described in Section 7.

The caustic transit rate of 0.32 per cluster per year is indicated by the orange upper limit in Figure 1. If the actual 2.0 μm SB of Pop III stars is dimmer than ~31 mag arcsec^{-2}, then their caustic transit rate would be correspondingly lower. This is indicated in Figure 1 for SB levels (in light orange) that are $10\times$, $100\times$, and $1000\times$ dimmer than ~31.0 mag arcsec^{-2}, with the corresponding caustic transit rates indicated in dark orange. The minimum number of caustic transits *JWST* could reasonably see—in a large monitoring program spread over many years—is a Pop III SB of ~36 mag arcsec^{-2}, which would require monitoring 30 clusters at least twice every year over 10 years. Such a large *JWST* observing program could reach the level of ~10 Pop III objects per arcsec^2 and would need to be a dedicated multiyear community effort.

To reach levels of only a few Pop III objects per arcsec^2 (SB $\gtrsim 37$ mag arcsec^{-2} in Figure 1) through *JWST* caustic transits would either require observing either ~100 clusters per year for 10 years—prohibitive in terms of *JWST* time—or the existence of stellar-mass BH accretion disks that are feeding much longer than massive Pop III stars live on average. This is discussed in Section 5.3.

Appendix C discusses the uncertainty estimates in the main parameters that determine the caustic transit rates and rise times of Pop III stars at $z \gtrsim 7$. The combined uncertainty in their caustic transit rates follows from the multiplicative sources of error in Equations (19), (25), and (26). These are the adopted effective caustic length L_{caust} (with ~0.3 dex uncertainty), the cluster transverse velocity v_T (~0.3 dex), the Pop III stellar luminosity L at $z \gtrsim 7$ (~0.2 dex), the uncertainty from the presence of microlensing in the ICL ($\gtrsim 0.5$ dex), and the uncertainty in the 1–4 μm sky-SB from Pop III stars ($\gtrsim 0.5$ dex). Further details are given in Appendix C.

These five main parameters that determine the Pop III star caustic transit rates are independent. Therefore, the combined uncertainty in the Pop III star caustic transit rates follows from taking these factors in quadrature and is estimated to be at least 0.7 dex, which is indicated by the vertical (dark orange) error range in Figure 1. For this reason, a *JWST* survey to find caustic transits at $z \gtrsim 7$ will need to be prepared to cover at least this factor of 5 uncertainty in Pop III star caustic transit rates. Since these uncertainty factors can be larger, *JWST* may need to observe at least 3–30 clusters per year during the first couple years of its lifetime. Such a survey would need to be maintained until a sufficient number of Pop III star caustic transits have been detected, at which point the actual Pop III star caustic transit rate can be better estimated, and the survey strategy updated accordingly.

If Pop III stars at $z \gtrsim 7$ are weakly clustered, their SB may be fairly uniform compared to the size of the caustics (see Section 4.3 and Figure 4(b)). Therefore, one could instead monitor fewer clusters for a correspondingly longer period of time. That is, for the anticipated 5–10 year lifetime of *JWST*

Table 5
Pop III Stellar Mass black hole Accretion Disk Parameters Adopted for Caustic Transit Calculations

Mass[a] ZAMS ($M_\odot$)	$M_{\rm compact}$[b] BH ($M_\odot$)	R_s[c] BH (km)	Radius[d] of the UV Accretion Disk ($R_\odot$)	$L_{\rm bol}$[e] ($L_\odot$)	$M_{\rm bol}$[f] AB-mag	Bolo+IGM+K-corr[g] $z=7$ (AB-mag)	$z=12$	$z=17$	$m_{\rm AB}$ Limits at[h] $z=7$ (AB-mag)	$z=12$	$z=17$	$t_{\rm rise}$[i] ($z=12$) (hr)	Transit[j] Rate (/cl/yr)
BH accretion-disk bolometric luminosities and UV half-light radii scaling from microlensed quasars (Blackburne et al. 2011)													
30	$\sim$5.0 BH	15	1.4	$\lesssim 4.2 \times 10^4$	$\gtrsim -6.8$	-0.6	-1.4	-1.7	$\gtrsim 41.8$	$\gtrsim 42.4$	$\gtrsim 42.9$	0.27?	$\gtrsim 0.58$?
50	$\sim$24 BH	72	3.0	$\lesssim 2.0 \times 10^5$	$\gtrsim -8.5$	-0.4	-1.2	-1.5	$\gtrsim 40.3$	$\gtrsim 40.9$	$\gtrsim 41.4$	0.58*	$\gtrsim 0.15$*
100	$\sim$65 BH	195	4.9	$\lesssim 5.4 \times 10^5$	$\gtrsim -9.6$	-0.2	-0.9	-1.3	$\gtrsim 39.4$	$\gtrsim 40.0$	$\gtrsim 40.5$	0.95*	$\gtrsim 0.06$*
300	$\sim$230 BH	690	9.2	$\lesssim 1.9 \times 10^6$	$\gtrsim -11.0$	-0.2	-1.0	-1.3	$\gtrsim 38.1$	$\gtrsim 38.6$	$\gtrsim 39.2$	1.8*	$\gtrsim 0.02$*
1000	$\sim$720 BH	2160	16.3	$\lesssim 6.0 \times 10^6$	$\gtrsim -12.2$	-0.2	-0.9	-1.3	$\gtrsim 36.8$	$\gtrsim 37.5$	$\gtrsim 37.9$	3.2*	$\gtrsim 0.01$*
BH accretion-disk bolometric luminosities and UV half-light radii estimated from multi-color thin-disk model													
30	$\sim$5.0 BH	15	1.9	$\lesssim 3.1 \times 10^4$	$\gtrsim -6.5$	-0.6	-1.4	-1.7	$\gtrsim 42.1$	$\gtrsim 42.8$	$\gtrsim 43.2$	0.37?	$\gtrsim 0.84$?
50	$\sim$24 BH	72	4.5	$\lesssim 1.8 \times 10^5$	$\gtrsim -8.4$	-0.4	-1.2	-1.5	$\gtrsim 40.4$	$\gtrsim 41.1$	$\gtrsim 41.5$	0.87*	$\gtrsim 0.18$*
100	$\sim$65 BH	195	7.8	$\lesssim 5.9 \times 10^5$	$\gtrsim -9.7$	-0.2	-0.9	-1.3	$\gtrsim 39.3$	$\gtrsim 40.0$	$\gtrsim 40.4$	1.51*	$\gtrsim 0.06$*
300	$\sim$230 BH	690	15.8	$\lesssim 2.0 \times 10^6$	$\gtrsim -11.0$	-0.2	-1.0	-1.3	$\gtrsim 38.0$	$\gtrsim 38.6$	$\gtrsim 39.1$	3.1*	$\gtrsim 0.02$*
1000	$\sim$720 BH	2160	29.8	$\lesssim 6.6 \times 10^6$	$\gtrsim -12.3$	-0.2	-0.9	-1.3	$\gtrsim 36.7$	$\gtrsim 37.4$	$\gtrsim 37.8$	5.8*	$\gtrsim 0.01$*

Notes.
[a] Pop III ZAMS stellar mass in $M_\odot$ from Table 1.
[b] Resulting Pop III stellar BH mass in $M_\odot$, following Woosley et al. (2002). Note that for Pop III stellar masses of $100 \lesssim M \lesssim 200\ M_\odot$ there are likely no BH leftovers (see Section 5.1), which the weighting in Section 6.2 includes.
[c] Resulting Pop III BH Schwarzschild radius R_s in kilometers.
[d] Adopted Pop III BH rest-frame UV-accretion disk half-light radius $r_{\rm hl}$ in $R_\odot$. The top tier of BH UV accretion radii (and bolometric luminosities) is inferred by scaling from observed microlensed quasars (Blackburne et al. 2011), the bottom tier was estimated from the multicolor thin-disk model discussed in Section 5.5.2. For a standard multicolor accretion disk around a BH of mass M, we get about $R_{\rm UV} \lesssim 40{,}000\ R_s$.
[e] Adopted Pop III BH accretion disk bolometric luminosity in $L_\odot$. The quoted luminosities and resulting rest-frame UV magnitudes are upper limits, since they assume that the BH accretion disk is constantly feeding at the stated luminosities for maximum lifetimes discussed in Section 5.3–5.4. Therefore, the resulting caustic BH accretion disk transit rates in column 14 are lower limits.
[f] Resulting Pop III BH accretion disk absolute bolometric AB-magnitude $M_{\rm bol}$.
[g] Combined bolometric+IGM+K-correction to the Pop III star $M_{\rm bol}$ at $z=7$, $z=12$, and $z=17$, respectively, calculated as in Sections 3.3 and 5.5.2.
[h] BH accretion disk apparent AB-magnitudes at $z=7$, $z=12$, and $z=17$ in 2016 *Planck* cosmology (Planck Collaboration et al. 2016b) using the NIRCam filters that sample rest-frame UV 1500 Å, assuming K-corrections as in columns 7–9 and no dust (see Section 6.1). Distance moduli used are DM $=49.24$, 50.58, and 51.42 mag at $z=7$, $z=12$, and $z=17$, respectively.
[i] Pop III BH accretion disk caustic transit rise time $t_{\rm rise}$ at $z=12$ as estimated in Section 4.4. Asterisks (*) indicate BH masses $M \gtrsim 24$–$65\ M_\odot$. For their accretion disks, caustic transit events are possibly observable to the detection limits of *JWST* medium-deep to deep survey epochs reaching AB $\lesssim 28.5$–29 mag, assuming caustic transit magnifications of $\mu \simeq 10^4$–10^5 can elevate Pop III stellar-mass BH accretion disks with AB $\lesssim 41.5$ mag temporarily above these *JWST* detection limits. Details are in Section 6.2.
[j] The cluster caustic transit rate of stellar-mass BH UV accretion disks as estimated in Sections 4.4 and 6.2, but directly applying Equations (19) and (25), rather than the general expression in Equation (27), which is only valid for the M/L relation in Equation (3) and Table 2 for Pop III stars.

(see Section 7), one could instead monitor a number of well-understood lensing clusters at high zodiacal latitude every few months during *JWST*'s lifetime. Any of these possibilities would constitute a minimum observing program to potentially identify Pop III star caustic transits during the lifetime of *JWST*. The program could then be adjusted after the number of caustic transits at $z \gtrsim 7$ is known after the first couple of years when monitoring a number of clusters. The presence of microlensing will likely also require observing the clusters more frequently to catch caustic transits at shorter timescales, as discussed in Section 4.3.2 and Appendix B.

5. Parameters Adopted for Pop III Star Black Hole Accretion Disks

The question that we address in this section is: under what conditions could *JWST* see the UV accretion disks of Pop III stellar-mass black holes lensed individually through cluster caustic transits at very high magnification? To address this, we first need to discuss the plausible range in physical properties of Pop III stellar-mass BH accretion disks at $z \gtrsim 7$, and under what conditions these may be fed from early massive stellar binaries for the expected range in IMF slope (Section 3.4) and metallicity evolution (Section 5.2). We refer the reader to recent work on PBHs (Kohri et al. 2014), DCBHs (Yue et al. 2013), or OBHs (Natarajan et al. 2017) for other direct BH feeding mechanisms. Their surface densities and accretion rates are uncertain, but if these could be estimated from theory, one could use the same formalism as in Sections 4.4 and 6.2 here to estimate their caustic transit rates.

5.1. Range in Pop III Stellar Black Hole Masses

The mass of the final Pop III star end product is more nuanced than just the BH mass. For example, theoretical models predict that stars in the general mass range of $100\ M_\odot \lesssim M \lesssim 200\ M_\odot$ do not lose much mass and that they may undergo an $e^+ - e^-$ pair-creation instability (Barkat et al. 1967; Fraley 1968; Wheeler 1977; Sugimoto & Nomoto 1980; Bond et al. 1984; Fryer et al. 2001; Woosley et al. 2002; Kozyreva et al. 2017; Woosley 2017). Such stars may undergo thermonuclear explosions that completely disrupt the star without forming a stellar-mass BH and eject a large amount of iron-group elements, especially ^{56}Ni (e.g., Smith et al. 2007;

Kozyreva & Blinnikov 2015). Theoretical models predict that stars with $260\,M_\odot \lesssim M \lesssim 5 \times 10^5\,M_\odot$ enter the pair-instability region but are too massive to be disrupted. They undergo standard core collapse and form intermediate-mass black holes (IMBHs; Fryer et al. 2001; Chatzopoulos et al. 2013; Belczynski et al. 2016). Figure 12 of Woosley et al. (2002) offers a map of the Pop III initial–final mass relation for massive stars from stellar evolution theory. For the Pop III ZAMS mass range in our MESA models, we adopt similar end products. Their end-product mass and the BH Schwarzschild radii R_s, which are used in our caustic transit calculations for stellar-mass BH accretion disks, are listed in Table 5.

In this context, we briefly consider possible constraints from the recent LIGO detections on stellar-mass BHs at $z \lesssim 0.1$ (Abbott et al. 2016a, 2016c). These are very plausibly examples of merging BH pairs with $M \simeq 29\text{–}36\,M_\odot$, $14\text{–}21\,M_\odot$, and $19\text{–}31\,M_\odot$, respectively, about 1–3 Gyr ago (Abbott et al. 2016b, 2016d, 2016e, 2017a). de Mink & Mandel (2016) suggest that these BHs are possibly left over from later (Pop II) starbursts about 5–12 Gyr before $z \simeq 0.1$, with a median age of ~ 7 Gyr for these mass pairs, which in 2016 *Planck* cosmology corresponds to a range in their formation redshift of $z_f \simeq 0.7\text{–}10$ with a median of $z_f \simeq 1.1$. If true, such BHs may not have had significant accretion rates since their progenitor-star supernovae (SNe) went off 5–12 Gyr before their binary merger produced gravitational waves at their detection distance of $z \lesssim 0.1$. In each LIGO case, a pair of massive stars of somewhat unequal masses formed, and so their evolutionary scenarios may have resulted in accretion onto the black holes left by the more massive parent stars with $M \gtrsim 30\text{–}80\,M_\odot$ after it produced an SN.

Another issue that we need to consider in this section is the lowest ZAMS mass that can with some fidelity produce a BH, also for Pop III stars at $z \gtrsim 7$. This is a very active topic of research where different groups are getting different results (Sukhbold & Woosley 2014, 2016; I. Petermann & F. X. Timmes 2018, private communication). Depending on the models used (1D, 2D, or 3D, with or without rotation), the "compactness" of the end product is rather uncertain in the mass range of $10 \lesssim M \lesssim 30\,M_\odot$. Rotation and binary interaction can produce different initial–final mass landscapes (e.g., Yoon et al. 2008). For $10 \lesssim M \lesssim 30\,M_\odot$, not all models get a clean explosion. On the other hand, at $M \gtrsim 30\,M_\odot$ nature can produce SNe with BH remnants, since LIGO has already seen $14\text{–}36\,M_\odot$ BHs at $z \lesssim 0.1\text{–}0.2$. Hence, Pop III stars with $10 \lesssim M \lesssim 30\,M_\odot$ may yield BHs, while for $M \gtrsim 30\,M_\odot$ they most likely do.

For our calculations of Pop III BH accretion disk caustic transits, we will assume that Pop III stars with $M \gtrsim 30\,M_\odot$—with the exception of the mass range of $100 \lesssim M \lesssim 200\,M_\odot$—can and will produce BHs of roughly 15%–70% of the ZAMS Pop III stellar mass, or $M \simeq 5\text{–}720\,M_\odot$ (deduced from Figure 12 of Woosley et al. 2002; see column 2 of Table 5 here). A full treatment of the evolution of Pop III binary or multiple stars, their end products, and their impact on Pop III BH accretion disks is beyond the scope of this study, and needs to be the focus of more detailed modeling in future work.

The actual resulting BH masses themselves are not as relevant for our caustic transit calculations. It only matters that such BHs exist—and for $M \gtrsim 14\,M_\odot$ LIGO has clearly shown that they do—*and* that they accrete while producing a

sufficiently high UV luminosity to be detected by *JWST* during a caustic transit. Any accretion (Frank et al. 2002) would have to be maintained for $\gtrsim 0.1$ yr in the rest frame at $z \gtrsim 7$ (i.e., ~ 1 yr in the observed frame) with L_{bol} values $\gtrsim 10^5\,L_\odot$ (Tables 2–4) to be possibly seen transiting across a cluster caustic by *JWST* and decay for about a year or less above the *JWST* detection threshold (Section 4.4).

5.2. Evolution in Metallicity and Massive Star Duplicity

Since we do not know the duplicity nor the separation distribution of Pop III stars, nor of the first polluted O-stars in mini halos, we need to consider a range of possibilities. Pop III BHs with $5 \lesssim M \lesssim 720\,M_\odot$ may accrete more steadily via Roche-lobe overflow from a (slightly polluted) Pop II.5 companion star of lower ZAMS mass, as discussed in Section 3.2.

The second scenario is much more common for O stars nearby, given their very high multiplicity, but may not be common for Pop III stars at $z \gtrsim 7$. Trenti & Stiavelli (2009) suggest that as soon as a massive Pop III star first forms in a mini halo, its powerful Lyman–Werner UV radiation field may prevent lower-mass Pop III stars from forming in its immediate surroundings. Self-shielding by very dense surrounding hydrogen gas against this UV radiation may allow some neighboring lower-mass Pop III stars to still form. Trenti & Stiavelli (2009) therefore also discuss mini halos that may have more than one Pop III star. In their models, Pop III stars generally start forming at $z \lesssim 30\text{–}40$ (cosmic ages $\gtrsim 99\text{–}65$ Myr, respectively), followed by slightly polluted Pop II.5 stars that quickly ramps up at $z \lesssim 28\text{–}35$ (cosmic ages $\gtrsim 109\text{–}79$ Myr, respectively), or about $\sim 10\text{–}15$ Myr later in cosmic time. Sarmento et al. (2018) present hydrodynamical simulations that narrow the Pop III star redshift range from $z \simeq 20$ to $z \simeq 7$. In their models, pristine Pop III stars are still the dominant population at $z \simeq 20$, while at $z \simeq 7$, slightly polluted ($Z \lesssim 10^{-4}\,Z_\odot$) "Pop II.5" stars outnumber Pop III stars by $\sim 10\text{:}1$. In other words, Pop III stars may have polluted their surroundings quickly enough that within 10–15 Myr, many lower-mass stars that formed in their neighborhood already have somewhat non-zero metallicities.

Comparing the estimated pre-MS lifetimes (τ_{preMS}) to the MESA ZAMS–AGB lifetimes in Table 1, we found in Section 3.1 that Pop III stars with $M \simeq 20\text{–}1000\,M_\odot$ live short enough ($\lesssim 8$ Myr) that they may have polluted the material from which stars with $M \simeq 1\text{–}1.5\,M_\odot$ formed at $z \gtrsim 7$, since their pre-MS lifetimes are longer than 6–9 Myr. Hence, it is possible that most early low-mass stars ($M \simeq 1\text{–}1.5\,M_\odot$) may have been polluted by massive Pop III stars as early as $z \lesssim 20$, and certainly at lower redshifts down to $z \simeq 7$. This then also means that it is possible that very low-metallicity Pop II.5 stars may have formed at $z \lesssim 20$ in the vicinity of Pop III stars, perhaps some close enough to form binaries or multiple-star systems with those Pop III stars. In any case, the first polluted O stars likely also appeared at $z \lesssim 20$. For the latter, the duplicity fraction may have quickly increased from the very low duplicity values expected for true zero-metallicity Pop III stars—with lower-mass companions not forming due to their significant LW radiation (Trenti & Stiavelli 2009)—to the much higher duplicity fraction seen in O stars today (see Section 3.2).

The metallicity evolution of stellar populations is not well-known at high redshifts ($z \gtrsim 4$; Maiolino et al. 2008; Kim et al.

2017). Trenti & Stiavelli (2007) and Sarmento et al. (2018) suggest that mini halos and the IGM can get quickly enriched (to $Z \simeq 10^{-4}$ $Z_\odot$) by a progenitor Pop III SN. The hydrodynamical simulations of Sarmento et al. (2018) use Adaptive Mesh Refinement (AMR) to sample the mass range of $M \simeq 10^{5.5}$–$10^8 M_\odot$ over the redshift range of $z \simeq 8$ to $z \simeq 16$, where their predicted metallicities range from $Z \simeq 0.1$ $Z_\odot$ at $M \simeq 10^8 M_\odot$ to $Z \simeq 0.003$ $Z_\odot$ at $M \simeq 10^{5.5} M_\odot$. Over this mass and redshift range, their mass–metallicity relation has a slope of $\Delta \log(Z/Z_\odot)/\log(M/M_\odot) \simeq 0.5$–0.6. At masses below $M \sim 10^{5.5} M_\odot$, their AMR simulations have insufficient mass resolution, but if the mass–metallicity relation were to continue with this slope to single stellar masses as low as $M \lesssim 10^3 M_\odot$, then the non-pristine stars at $z \gtrsim 7$ could indeed have metallicities as high as $Z \simeq 10^{-3.5} Z_\odot$. Madau & Fragos (2017) suggest that at $z \simeq 7$–10, the metallicity of massive ($M \sim 10^8$ $M_\odot$) star-forming objects may be as high as 0.03–0.1 $Z_\odot$. For the low-mass environments in which slightly polluted Pop II.5 stars form, a metallicity of $Z \gtrsim 10^{-4} Z_\odot$ is thus plausible.

Recent observation (e.g., Badenes et al. 2018) has shown that metal-poor ($Z \lesssim 0.3$ $Z_\odot$) stars have a multiplicity fraction $2\times$–$3\times$ higher than metal-rich ($Z \sim Z_\odot$) stars. Theoretical work on star formation (e.g., Machida et al. 2009) suggested a higher binary frequency in lower-metallicity gas, and that a majority of stars are born as members of binary/multiple systems for $Z \lesssim 10^{-4}$ $Z_\odot$. Hence, for non-zero metallicities, at least the binary fraction increases with decreasing metallicity. Physically, this occurs because metal-line cooling becomes significant above a threshold of $Z \gtrsim 10^{-4} Z_\odot$, which decreases the fragmentation of the gas clouds that form the stars. We do not know if this trend continues to hold for truly zero-metallicity Pop III stars at $z \gtrsim 7$, but it seems possible that any non-zero-metallicity massive star will form and evolve in an environment with a significant binary fraction (see, e.g., Adams et al. 2006; Adams 2010 for a discussion).

What matters for the current work is that, while some massive stars with zero or very low metallicity may still exist at $z \simeq 7$, at the same time, a sufficient fraction of polluted stars ($Z \gtrsim 10^{-4}$ $Z_\odot$) already exists at $z \simeq 7$–17. The latter are critical, since they likely formed with a significant fraction of binaries and so play an essential role in BH accretion disk feeding via Roche-lobe overflow during its post-MS evolution.

Mass transfer is not currently considered in the MESA code. Future work needs to include detailed star formation scenarios with full metallicity evolution in the ISM at $z \gtrsim 7$, their subsequent evolutionary tracks at very low metallicities, and include scenarios of binary evolution that incorporate mass transfer and address how mass transfer affects the BH-feeding timescales.

5.3. Range in Stellar-mass Black Hole Accretion Lifetimes

We will consider here that any BHs left over after a massive Pop III star's death may accrete from a surrounding lower-mass, low-metallicity star filling its Roche lobe during its post-MS evolution, causing a UV-bright accretion disk. The accretion timescales onto these BHs in stellar binaries are not well-known, but may have plausibly lasted as long as the GB lifetimes of the less massive star in a binary when it fills its Roche lobe. Following the Pop III MS lifetimes from Section 3.1 and Table 1, this can happen within $\lesssim 12\%$ of

their MS ages, or within 0.3–60 Myr after the first SN of the more massive star in the pair has occurred. The question then arises: how often can this scenario have happened for Pop III stars at $z \gtrsim 7$, whose stellar-mass BH remnants would still be around today as leftovers from the First Light epoch?

If a fraction $(1 - \epsilon)$ of the matter is accreted at the Eddington rate, where ϵ denotes the radiative efficiency, then the mass of the BH will increase exponentially with a characteristic timescale of $t_E = 4\pi G \mu m_p/(\sigma_e c \epsilon) \simeq 45$ Myr. If all remnants of Pop III stars accreted at the Eddington rate for $\gtrsim 10^8$ yr, then this would increase the BH mass by orders of magnitude, which would increase the mass density of BHs to values that are excluded by constraints on the present-day mass density of BHs (see, e.g., Tanaka et al. 2012). Steady BH feeding from accretion disks for $\gtrsim 10^8$ yr would have likely given rise to BHs that will grow catastrophically to $\gg 10^2 M_\odot$, and may quickly produce massive BHs with $M \gtrsim 10^3$–$10^5 M_\odot$ or more and become Ultra-Luminous X-ray sources (ULXs). Although UV-bright accretion disks around such massive BHs would be easier to detect with *JWST* during caustic transits (see Sections 4.4 and 6.2), they will likely also be much rarer.

For Pop III stellar-mass BH accretion disks, we will therefore consider lifetimes of at least $\gtrsim 0.3$ Myr from the massive binary argument in Section 3.2.2. In Section 5.5, we will assume that the BH accretion disks are constantly feeding at the luminosities predicted for maximum lifetimes of $\lesssim 60$ Myr, during which the lowest-mass ($M \gtrsim 2.0 M_\odot$) companion AGB stars would fill their Roche lobes before reionization is complete at $z \simeq 7$ (Section 3.2.1). Given the uncertain accretion times, the BH accretion disk UV luminosities derived in Section 5.5 are upper limits, so their caustic transit rates in Section 6.2 are lower limits.

5.4. Efficiency of Massive Pop III Star Black Hole Accretion Disks

Following the arguments of Section 3.5, if $N \lesssim 10^3$ massive Pop III stars per arcsec^2 contribute to the near-IR sky-SB of AB $\gtrsim 31$ mag arcsec^{-2} at 2.0 μm, then a large fraction (f_{BH}) of them will leave behind BHs. Accretion onto these BHs will give rise to additional flux in the IRB. The ratio of the Pop III to Pop III remnant contribution can be estimated from

$$\frac{S_{\text{Pop III}}}{S_{\text{BH}}} = \frac{1}{f_{\text{BH}}} \frac{t_{\text{PopIII}}}{t_{\text{acc}}} \frac{L_{\text{Pop III}}}{L_{\text{BH}}}. \qquad (34)$$

Figure 2 shows that a $300 M_\odot$ Pop III star has a luminosity of $L_{\text{bol}} \simeq 2.5 \times 10^{40}$ erg s^{-1}, which agrees quite closely with the Eddington luminosity associated with an almost equal-mass BH, which is $\sim 10^{40}$ erg s^{-1} following Equation (34). If we assume that the fraction f_{BH} of Pop III stars that collapses into BHs produces a BH of $\sim 15\%$–70% of the original stellar ZAMS mass, then we expect $L_{\text{Pop III}} \simeq L_{\text{BH}}$ at least at early times. We then obtain

$$\frac{S_{\text{Pop III}}}{S_{\text{BH}}} = \frac{1}{f_{\text{BH}}} \frac{t_{\text{PopIII}}}{t_{\text{acc}}}. \qquad (35)$$

The efficiency of gas accretion onto stellar-mass BHs formed by Pop III stars is discussed by Milosavljević et al. (2009). It is possible that radiative feedback seriously limits the efficiency of gas accretion. Time-averaged Eddington ratios of $\sim 1\%$ have

been reported by, e.g., Park & Ricotti (2012), although this ratio could be smaller. If accretion occurs during the typical ~0.3–60 Myr adopted for early massive binaries, then these accretion times are less than 1%–10% of the available Hubble time at $z \gtrsim 7$. Hence, BHs may have been feeding with a duration of $\lesssim$1%–10% of the total available time.

If we take into account that the mass of the BH grows with time, then it is plausible that $(S_{\text{Pop III}}/S_{\text{BH}}) \lesssim 1$, i.e., the remnants of Pop III stars may contribute more to the near-IR sky-SB than the Pop III stars themselves. If Pop III remnants form the seeds for SMBHs, including the rare $M_{\text{BH}} \simeq 10^9 M_\odot$ black holes that are seen in quasars at $z \gtrsim 6$ (Willott et al. 2003; Jiang et al. 2007; Kurk et al. 2007), then at least a small fraction of them must accrete at practically the Eddington rate with a duty cycle of ~100% (e.g., Willott et al. 2010). If a small fraction of the remnants accrete so efficiently, then it is not unexpected that a much larger fraction will accrete with duty cycles intermediate between 1% and 100%. Depending on how large the fraction of more slowly accreting BHs is, this population could contribute significantly more to the near-IR sky-SB and to caustic transits than the Pop III stars themselves.

5.5. Stellar-mass Black Hole Accretion Disk Radii and Luminosities

Pop III stars with masses $M \simeq 30$–$1000 M_\odot$ can leave BHs behind with $M \simeq 5$–$720 M_\odot$ (Section 5.1), except for the mass range around 100–200 $M_\odot$, where they seem to produce no BHs (Woosley 2017). The Schwarzschild radii of these Pop III BHs will thus be in the range of $R_s \simeq 15$–2200 km, as listed in column 3 of Table 5. Using this range of BH masses and Schwarzschild radii, this section summarizes available constraints on the resulting sizes and luminosities of stellar-mass BH accretion disks. Since these parameters are more uncertain than those of Pop III stars, we will estimate them in two independent ways to permit a consistency check. The resulting UV accretion disk radii, bolometric luminosities, and corresponding M_{AB} magnitudes are listed in columns 3–6 of Table 5, which are described for both methods in the next two subsections.

5.5.1. Estimates by Scaling from Observed Microlensed Quasar Results

A first estimate of R_{accr} and L_{accr} can be made from observed microlensing results on strongly lensed quasars at $z \simeq 1$–2 by Blackburne et al. (2011). These authors present accretion disk sizes, temperatures, and luminosities from their quasar images that were monitored extensively with ground-based telescopes and through *Chandra* X-ray fluxes. Their Equation (2) gives a simple relationship among accretion disk half-light radius r_{hl} or R_{accr}), the quasar SMBH mass, and the observed wavelength, which in their case is the observed optical that samples rest-frame ~2500 Å:

$$R_{\text{accr}} \simeq 1.68 \times 10^{14} \left(\frac{M_{\text{BH}}}{10^9 M_\odot} \right)^{2/3} \left(\frac{\lambda}{\mu\text{m}} \right)^{4/3} m. \quad (36)$$

We rescale this for the *JWST* NIRCam near-IR filters F115W–F277W, which sample Pop III objects at $z \simeq 7$–17 approximately in the rest-frame UV at $\lambda \simeq 1500$ Å. From their multicolor microlensing photometry, Blackburne et al. (2011) derive SMBH masses of order $(0.04$–$2) \times 10^9 M_\odot$ and

bolometric luminosities in the range of $L_{\text{bol}} \simeq (0.1$–4) $\times 10^{46}$ erg s^{-1}. Their Table 8 suggests that for all 12 quasars, the bolometric accretion disk luminosity scales with the SMBH mass approximately as

$$L_{\text{bol}} \simeq 3.2 \times 10^{46} \left(\frac{M_{\text{BH}}}{10^9 M_\odot} \right) \text{erg s}^{-1}. \quad (37)$$

For a solar luminosity of 3.828×10^{33} erg s^{-1}, this corresponds to quasar accretion disk luminosities of ~$(0.3$–10) $\times 10^{12} L_\odot$.

These are remarkable direct constraints to quasar rest-frame UV accretion disk sizes and their luminosities. We do not know if we can scale these values down from their observed mass range to our range of BH masses of ~5–720 $M_\odot$ adopted in Table 5. Blackburne et al. (2011) suggest from the data over their mass range that their half-light radii scale as

$$r_{\text{hl}} \propto M_{\text{BH}}{}^\rho, \quad (38)$$

with a best fit of $\rho \simeq 0.27 \pm 0.17$. This is flatter than the $\rho = 2/3$ slope implied by the multicolor accretion disk theory in Equation (36). If we scale our UV accretion disk radii down with $\rho \simeq 0.27$ from their SMBH mass range, then we obtain very large radii ($R_{\text{UV}} \gtrsim 10^3 R_\odot$) and luminosities for Pop III stellar-mass BH accretion disks. This suggests that the flat ρ slope derived from their quasar sample may not hold down to Pop III BH masses, as may be caused by the strong dependence on BH mass of the tidal forces around each BH. We therefore adopt a slope between these values of $\rho \simeq 0.5$, which is consistent with the Blackburne et al. (2011) value within their errors and still provides a good fit to their data given the small dynamic range in M_{BH} in their sample. In Section 5.5.2, we suggest that $\rho \simeq 0.5$ produces more consistent overall results for the multicolor thin-disk accretion model. When we scale the Blackburne et al. (2011) UV accretion-disk radii down with $\rho \simeq 0.5$, we obtain the BH UV half-light radii listed in the top tier of Table 5. These are in the range of $R_{\text{UV}} \simeq 1$–16 $R_\odot$ for $M_{\text{BH}} \simeq 5$–720 $M_\odot$. The bolometric luminosities listed in the top tier of Table 5 were scaled down directly with Equation (37) and are in the range of 4×10^4–$6 \times 10^6 L_\odot$ for $M_{\text{BH}} \simeq 5$–720 $M_\odot$, respectively.

5.5.2. Estimates from Multicolor Accretion Disk Theory

In this section, we compare the stellar-mass BH accretion disk sizes and luminosities as scaled down from the quasar observations in Section 5.5.1 to theoretical estimates.

In the simplest form, accretion disks around BHs are assumed to be "multicolor" thin disks, which consist of a series of concentric shells, each of which emit blackbody radiation characterized by its radially dependent temperature (e.g., Shakura & Sunyaev 1973; Remillard & McClintock 2006; Blackburne et al. 2011). In the rest-frame UV–optical (at $\nu_{\text{Ly}\alpha} \lesssim 2.466 \times 10^{15}$ Hz), the spectrum of the accreting BH is dominated by the thermal disk component. In the very inner part of the accretion disk, other radiation mechanisms will likely produce significant X-ray emission, such as synchrotron radiation in the presence of strong central magnetic fields, inverse Compton radiation, or thermal bremsstrahlung (Shakura & Sunyaev 1973, 1976). Only the harder part of this redshifted X-ray emission will make it past the IGM and potentially be detected by *Chandra*, but what matters for any

JWST detection is the amount of associated rest-frame UV emission that makes it past the IGM at $\lambda \gtrsim 1216$ Å. In the multicolor thin accretion disk model, the temperature increases with radius as

$$T \propto r^{-3/4}. \tag{39}$$

Using Equation (38), gas in the innermost stable orbit at $R \simeq 3R_s$ has a maximum temperature of about

$$T_{max} \simeq 10 \left(\frac{M_{BH}}{100\,M_\odot} \right)^{-\tau} \text{keV}. \tag{40}$$

Standard thin-disk accretion theory suggests a slope of $\tau = 1/4$, but since we adopted $\rho \simeq 1/2$ in Equation (38), we need to use $\tau = 3/8$ here to maintain consistency with Equation (39). The multicolor accretion disk models predict similar UV half-light radii R_{UV} for either slope τ, since the largest SED differences occur well below rest-frame 1216 Å, and this part of the SED does not make it past the IGM at $z \gtrsim 7$.

We will assume here that the maximum temperature of the inner accretion disk in Equation (40) for a $100\,M_\odot$ BH needs to be at least 10 keV, or $T_{max} \simeq 3.87 \times 10^7$ K. This is so that their hard X-ray photons can make it past the neutral hydrogen at $z \gtrsim 7$ (Haardt & Madau 2012), and when redshifted from $z \simeq 7$–17, still be in principle observable in the *Chandra* soft X-ray band, which covers 0.5–2.0 keV. This argument is based on the following: if part of the *Spitzer–Chandra* cross-correlation power-spectrum signal (Cappelluti et al. 2013; Mitchell-Wynne et al. 2016) came from redshift $z \gtrsim 7$, then the sources that cause it must be both *Spitzer* 3–4 μm and *Chandra* X-ray sources, as discussed in Section 2.3.2. Both of these papers discussed PBHs as possible candidates for the *Spitzer–Chandra* cross-correlation signal. As discussed in Section 3.1, Pop III stars alone cannot cause this *Spitzer–Chandra* cross-correlation signal, since they reach only a maximum temperature of $T_{eff} \lesssim 10^5$ K.

As we move out in radius, the temperature drops as in Equation (39). We assume that each concentric radius interval in the multicolor accretion disk emits as a blackbody with its own temperature. The largest radius that will contribute to the UV emission is the one where the blackbody curve peaks in the UV longwards of Lyα. For our 1500 Å rest-frame UV reference, this largest ring needs to have a temperature of $T \gtrsim 3.2 \times 10^4$ K. This suggests that we need to go out in radius where the temperature is a factor of $\gtrsim 1200\times$ lower than in the inner ring. Hence, we need to integrate out to $r \simeq 13,000\,R_{min}$, where $R_{min} \simeq 3\,R_s$ is the radius of the innermost stable orbit around the BH. Plugging in the numbers above then yields $R_{accr}(UV) \simeq 1.6 \times 10^7$ km for the maximum radius of the UV-emitting region, or $\sim 17\,R_\odot$ for $M_{BH} = 100\,M_\odot$. Integration of the actual multicolor thin-disk light profiles for a $100\,M_\odot$ BH yields a half-light radius r_{hl} that is about $1.7\times$ smaller than this, as shown below.

We use the multicolor accretion disk model in Equations (39)–(40) for the Pop III BH mass range of 5–720 $M_\odot$ in Table 5 to predict their UV half-light radii r_{hl} and their bolometric and UV luminosities. Their UV half-light radii r_{hl} are then simply integrated from the part of the radially dependent UV accretion disk SED that makes it past the IGM at $z \gtrsim 7$. These results are listed in the bottom tier of Table 5 and shows UV half-light radii in the range of $R_{UV} \simeq 2$–30 $R_\odot$.

At these r_{hl} values, the multicolor accretion disks have an effective temperature of $T_{eff} \simeq 47,500$–48,000 K for $M \simeq 5$–720 $M_\odot$. Bolometric+IGM+K-corrections were applied to the bolometric luminosities in Table 5, as for Pop III stars in Section 3.3. For multicolor accretion disks with $T_{eff} \simeq 47,700$ K, these combined BIK-corrections are –0.3, –1.1, and –1.5 mag at $z = 7$, $z = 12$, and $z = 17$, respectively. These are comparable to the values in Tables 3–4 for Pop III RGB and AGB stars with monochromatic blackbody disks of similar temperatures. The bolometric luminosities predicted for the multicolor thin accretion disks are in the range of 3×10^4–$7 \times 10^6\,L_\odot$ for $M_{BH} \simeq 5$–720 $M_\odot$, and are listed in the bottom tier of Table 5.

To check our multicolor accretion disk models for consistency, we first verified that they reproduce the Blackburne et al. (2011) UV half-light radii obtained for accretion disks of $z \simeq 1$–2 quasars from Equation (36). Second, we apply our multicolor accretion disk model to $M \simeq 10^9\,M_\odot$ SMBHs known to be present in quasars at $z \gtrsim 6$ (e.g., Willott et al. 2003; Jiang et al. 2007; Kurk et al. 2007). The above equations imply a UV accretion-disk diameter of $2R_{UV} \simeq 2 \times 10^5\,R_\odot$ for a $10^9\,M_\odot$ SMBH, which is ~ 1000 au or 0.005 pc across, corresponding to a light travel time of $\lesssim 5$ days in the rest frame. This is comparable to the accretion-disk sizes of QSOs inferred from variability studies, where the somewhat larger broad-line region can be light-days–weeks across (e.g., Kozłowski et al. 2010; Butler & Bloom 2011). Our multicolor thin accretion disk model also predicts the unobscured rest-frame UV luminosity for the rare quasars with a $10^9\,M_\odot$ SMBH at $z \gtrsim 6$, which is $M_{UV} \simeq -27$ AB-mag (Fan et al. 2001, 2003). Hence, their BIK-corrected near-IR fluxes are predicted to be $m_{AB} \gtrsim 21$ mag at $z \gtrsim 7$, which is comparable to what is observed for the highest-redshift quasars (e.g., Mortlock et al. 2011). This extrapolation to QSOs at $z \gtrsim 6$ then justifies the slightly modified choices of $\rho \simeq 1/2$ (instead of 2/3) and $\tau \simeq 3/8$ (instead of 1/4) above.

In summary, Table 5 shows that both estimates of the R_{UV} and L_{bol} of Pop III stellar-mass BH accretion disks in Section 5.5.1 and in this section are within a factor of two or less. We will therefore adopt the two scaling methods in the equations above and assume that the resulting range of properties in Table 5 captures the properties of Pop III BH UV accretion disks sufficiently well to make an order-of-magnitude estimate of the cluster caustic transit rates for Pop III BH accretion disks.

Given the unknown accretion efficiencies compared to Eddington, or the unknown accretion lifetimes compared to the maximum accretion lifetimes possible (Section 3.1), the values in Table 5 are upper limits to the Pop III BH UV accretion disk luminosities. That is, the luminosities and resulting M_{bol} and m_{AB} values in Table 5 assume that BH accretion disks radiate at steady-state levels inferred by the multicolor accretion-disk model for maximum lifetimes as discussed in Section 5.3.

In conclusion, the inner stellar-mass BH accretion disks may be significantly hotter than the typical $T \simeq 10^5$ K temperatures of Pop III stars, plausibly reaching X-ray temperatures at the innermost radii, and reaching $\sim 30,000$ K at the outermost radii. Their UV-bright accretion disks—if unobscured by surrounding dust—have SEDs that can make it in part through the neutral IGM at $z \gtrsim 7$ with UV radii $\lesssim 40,000\,R_s$. Their rest-frame UV radii are $R_{UV} \simeq 1$–30 $R_\odot$, and their UV luminosities are at most 3×10^4–$7 \times 10^6\,L_\odot$ for $M_{BH} \simeq 5$–720 $M_\odot$,

respectively. Pop III stellar-mass BH accretion disk radii may thus be similar to, or somewhat larger than, the 1–13 $R_\odot$ radii of the ZAMS Pop III stars in Tables 1–2, but no larger than the Pop III RGB or AGB star radii in Tables 3–4. They would fit well within the $\sim$7–55 $R_\odot$ Roche lobe sizes seen in massive binaries discussed in Section 3.2.2, and so are eligible for feeding from a less massive RGB/AGB star in the binary that is filling its Roche lobe. This assumes that subsequent generations of (slightly) polluted massive stars at $z \gtrsim 7$ already have high-enough metallicity to form binaries. The predicted stellar-mass BH accretion disk UV radii and maximum luminosities are similar to those of Pop III RGB–AGB stars in the 10–300 $M_\odot$ range. We use this to estimate the BH accretion disk caustic transit time and rates in Section 6.2.

5.6. White Dwarfs and Neutron Stars from Low-mass Pop III Stars

For completeness, we will briefly consider here the potential impacts of White Dwarfs (WDs) that likely result from low-mass stars ($M \lesssim 5\,M_\odot$) at $z \gtrsim 7$ (see, e.g., the $Z = 5 \times 10^{-3}$ $Z_\odot$ sample of Romero et al. 2015), and of Neutron Stars (NSs) that likely result from Pop III stars at ZAMS masses $M \lesssim 20\,M_\odot$, since both will be far more common than Pop III stellar-mass BHs (see Table 5 and the IMF slopes in Figure 3).

Table 1 implies that NSs would not appear until 8–70 Myr after their progenitor stars with $5 \lesssim M \lesssim 20\,M_\odot$ form, while WDs would appear at least $\gtrsim$230 Myr after their progenitor stars with $M \lesssim 5\,M_\odot$ form at $z \simeq 7$–17. If the first stars form at $z \simeq 35$–40, then the first NSs would thus appear soon thereafter, but the first WDs would not appear until $z \lesssim 14$. In either case, the first NS or WD mergers at $z \gtrsim 7$ would have only 500–700 Myr to occur. Hence, we will not consider NS–NS mergers such as those recently found by LIGO (Abbott et al. 2017b, 2017c) and identified by ground-based follow-up campaigns (e.g., Chornock et al. 2017; Cowperthwaite et al. 2017) in a nearby galaxy, nor potential NS–WD or WD–WD mergers, as these are far more rare than regular accretion onto either a WD or an NS.

The duration of regular accretion onto WDs or NS before they undergo a nuclear explosion on their surfaces depends mainly on their accretion rates. These in turn depend on the binary separation, masses of the two components, evolutionary state of the companion, and the nature of the explosion. For WDs (novae and super-soft X-ray sources), the recurrence timescales are $\sim$20 to $\sim$10,000 years (Shara et al. 1986; Cannizzo et al. 1988; Wolf et al. 2013; Henze et al. 2015; Shafter et al. 2015; Shafter 2017), and are likely too rare to average out to a flux that could be detected during a cluster caustic transit. For NSs (X-ray bursters), the recurrence timescales can be hours to weeks (Tanaka & Shibazaki 1996; Watts 2012). Their luminosities when averaged over $\gtrsim$0.1 years at $z \gtrsim 7$ would determine if such objects could be seen via caustic transits by *JWST*. In all cases, their surface layers explode, after which they may resume accretion and may approach their previous steady-state luminosity. A proper description of WD and NS accretion will thus not only require the multicolor thin-disk models that we use for BH accretion disks in Section 5.5.2, but also a quantitative modeling of these episodic nuclear detonation events, which is beyond the scope of the current paper. Future work will need to consider if accretion onto Pop III NS or WDs can be steady enough and luminous enough to be a source of caustic transits that are potentially observable by *JWST*.

6. Estimates of Caustic Transits for Pop III Star Black Hole Accretion Disks

In this section, we discuss the possible effects of dust produced by Pop III stars, and then present our estimates of the cluster caustic transit rates resulting from stellar-mass BH accretion disks as described in Section 5.

6.1. Possible Effects from Dust Generated by Pop III Stars

True zero-metallicity massive stars, with all modeling investigations to date, have significantly reduced mass loss. The normal driver of massive-star winds—radiation pressure from scattering off metals, is not present. Alternatives such as rotational mixing, some dredge-up scenario to bring core material to the photosphere, or (epsilon- and kappa-) pulsation mechanisms, are too weak to cause much mass loss (Castor et al. 1975; Götberg et al. 2017; Renzo et al. 2017). Thus, zero-metallicity massive stars may not be shrouded by dusty circumstellar material. For metallicities of $Z \lesssim 10^{-4}$ $Z_\odot$ (or even $\lesssim 10^{-5}$ $Z_\odot$), the winds (hence dust) will be at levels more common for massive stars seen nearby, although still significantly reduced.

If Pop III stars—or the slightly polluted Pop II.5 stars—did manage to produce stellar winds during their main sequence and Blue–Red Supergiant (BSG-RSG/AGB) phases, this could have deposited dust into the surrounding medium. When a fraction of Pop III stars goes off as Pair Instability SuperNovae (PISNe), they would deposit additional metals into their immediate surroundings. Dust formation in the circumstellar material of initially zero-metallicity Pop III stars could thus have added a non-trivial extinction/reddening factor, especially in their late stellar evolution and subsequent BH accretion disk stages. Hence, we should consider possible cases where either Pop III stars or their stellar-mass BH accretion disks are significantly reddened by dust, or both.

For non-rotating Pop III stars, this dust could be distributed rather uniformly and obscure most of the Pop III stars and their BH accretion disks, but for rotating stars, the situation may be quite different. We do know that Gamma-Ray Bursters (GRBs) are quite visible from γ-ray to radio waves when viewed from the right direction. The same is true for unobscured versus obscured AGNs—much of their visibility is viewing-angle dependent with respect to the dust torus. We therefore must consider that at least a fraction of Pop III stars with significant stellar rotation produced BH accretion disks that are visible under certain viewing angles and produced an equal amount of UV-continuum radiation as the Pop III stars themselves, or perhaps more. Evolving rotating Pop III star models, dust production, and their likely non-uniform dust-expulsion mechanism are currently too uncertain to take into account in the model calculations and will require more detailed numerical modeling in future work.

If both Pop III stars and their stellar-mass BH accretion disks were fully unobscured, then the average $\gtrsim 2 \times 10^6$ yr MS lifetime of Pop III stars (Section 3.1) and maximum BH accretion disk lifetimes—as visible in the rest-frame UV—of $\lesssim$60 Myr would determine their visible ratio. Some fraction of Pop III BH accretion disks may not be fully obscured, as would be implied by the *Spitzer–Chandra* power spectrum results in Section 2.3.2, if some of this signal came from $z \gtrsim 7$ (Cappelluti et al. 2013, 2017). In reality, nature may have well-produced some observable combination of obscured and

unobscured Pop III stars and their BH accretion disks, as it has for the iEBL from spheroids, disks, and unobscured AGNs at lower redshifts in Figure 1. For that reason, we allowed the maximum sky-SB of $\gtrsim 31.0$ mag arcsec^{-2} of Section 2.3 to be either fully caused by Pop III stars or by their BH accretion disks, or by a combination of the two not exceeding this SB level. *JWST* may be able to distinguish between the two through chromatic effects of caustic transits, as discussed in Section 7.2.

6.2. Implied Estimates of Cluster Caustic Transits for Pop III Star Black Hole Accretion Disks without Microlensing

In this section, we present estimates of the cluster caustic transit rates resulting from stellar-mass BH accretion disks as described in Section 5. To first order, for Pop III stellar-mass BHs, the same principles apply as above, so unless stated otherwise, we use the equations in Section 4.4.

As discussed in Section 5.5.2, the expected BH accretion disk radii are similar to, or somewhat larger than, the 1–13 $R_\odot$ radii of the ZAMS Pop III stars in Tables 1–2, but no larger than Pop III RGB or AGB star radii in Tables 3–4. The maximum BH accretion disk luminosities are in general similar to those of Pop III RGB–AGB stars in the 10–100 $M_\odot$ range, or $\sim 10^4$–$10^7 L_\odot$.

Pop III stellar-mass BH accretion disks—when lensed through cluster caustic transits—thus also have rise times of order one to several hours. For their similar luminosities, the decline times will then be also of the order of a year, as discussed in Section 4.4. These, together with their transit rates predicted for the Pop III BH radii and luminosities in Section 5.5, are listed in Table 5.

For a Pop III ZAMS mass function slope of $\alpha \simeq 2$, the weights for each of the mass bins for BH accretion disks in Table 5 are very similar, following Equation (32). The resulting total transit rates for Pop III stellar-mass BH accretion disks with $M_{\rm BH} \gtrsim 24$–720 $M_\odot$ that are in principle observable with *JWST* to AB $\lesssim 28.5$–29 mag across the caustics are predicted to be $\gtrsim 0.18$ per cluster per year for the top tier in Table 5, and $\gtrsim 0.24$ per cluster per year for the bottom tier, respectively.

Because the luminosities and the resulting $M_{\rm bol}$ and $m_{\rm AB}$ values in Table 5 are upper limits (Section 5.5), the inferred BH accretion-disk transit rates are lower limits, as indicated in Table 5. That is, if the actual accretion efficiencies were $10\times$ lower, then the BH accretion luminosities would be ~ 2.5 mag fainter, and the caustic transit rates could be several times higher. This is because there would be $10\times$ as many faint objects per mass bin that make up the near-IR SB adopted in Section 2.3 that contribute to caustic transits above the detection limit, but there would also be fewer mass bins contributing above the *JWST* detection limit.

The limits to the caustic transit rates of stellar-mass BH accretion disks of ~ 0.2 per cluster per year are similar to caustic transit rate of ~ 0.32 per cluster per year obtained for Pop III ZAMS+RGB+AGB stars (Section 4.4). For BHs, they could be several times higher, depending on their actual accretion efficiency.

We briefly discuss this in the context of the lifetime differences between Pop III stars and their stellar-mass BH accretion disks that could affect the mix of caustic transits *JWST* may observe. In Table 5, BHs with UV accretion disks bright enough to be detected by *JWST* during caustic transits have $M_{\rm BH} \simeq 24$–720 $M_\odot$ and AB $\simeq 37$–42 mag. Pop III stars with

$30 \lesssim M \lesssim 1000 M_\odot$ that produce BHs have ZAMS ages of 5.6–2.1 Myr (Table 1) with an average of ~ 3 Myr. Pop III stars of masses $M \simeq 2$–20 $M_\odot$ live considerably longer than these during their AGB stage, where they could fill their Roche lobes for up to 0.6–60 Myr, with an IMF-weighted average GB age of ~ 6 Myr. Hence, during their AGB stage 2–20 $M_\odot$ stars could feed the BH that is leftover from a 30–1000 $M_\odot$ star for a maximum duration that is significantly longer than the ZAMS lifetime of the massive Pop III star that produced this BH.

In summary, depending on how steady and efficient BH feeding by a lower-mass AGB star in its Roche lobe is, stellar-mass BH accretion disks may be about as likely as Pop III stars at $z \gtrsim 7$ to cause cluster caustic transits that could be observed by *JWST*, and possibly more likely. Stellar-mass BH accretion disks with an SB $\simeq 31.0$ mag arcsec^{-2} (or ~ 1 nW m^{-2} sr^{-1}) could produce about one caustic transit per five clusters per year, and perhaps as many as one event per two clusters per year. As for the Pop III stars in Section 4.4, a dedicated *JWST* program that monitors three clusters per year for a number of years could possibly detect several caustic transits for stellar-mass BH accretion disks. If their SB were to be as dim as ~ 36.0 mag arcsec^{-2}, which corresponds to ~ 10 Pop III BH accretion disks per arcsec^2 (see Figure 1), then 30 clusters would have to be monitored for up to 10 years to detect any caustic transits from BH accretion disks at $z \gtrsim 7$.

Appendix D discusses the uncertainty estimates in the main parameters that determine the caustic transit rates and rise times of Pop III stellar-mass BH accretion disks. As in Section 4.5 and Appendix C, the combined uncertainty in their caustic transit rates follows from the adopted effective caustic length $L_{\rm caust}$ (with ~ 0.3 dex uncertainty), the cluster transverse velocity v_T (~ 0.3 dex), and the uncertainty from the presence of microlensing in the ICL ($\gtrsim 0.5$ dex). The uncertainty in the Pop III stellar-mass BH accretion disk luminosity L is larger than for Pop III stars. This is due to their uncertain accretion efficiency or accretion duration, which we assume is uncertain by at least 0.5 dex, as discussed in Section 5.4 and Appendix D. On the other hand, the uncertainty in the 1–4 μm sky-SB from Pop III stellar-mass BH accretion disks may be smaller than that of Pop III stars, since the *Spitzer–Chandra* power spectrum results (Section 2.3.2) hint at a possible contribution from (stellar-mass) BHs at $z \gtrsim 7$. As discussed in Appendix D, the error in their power spectrum signal is estimated at $\gtrsim 0.15$ dex.

Since these parameters are independent, the combined uncertainty for the caustic transit rates of stellar-mass BHs is thus at least ~ 0.7 dex, but for somewhat different reasons than for Pop III stars. This is indicated by the vertical black error range in Figure 1. Given the *Spitzer–Chandra* power-spectrum signal discussed in Section 2.3.2 and Appendix D, and the possibility that their non-steady-state luminosities may increase their caustic transit rates for a given near-IR sky-SB, as discussed in Section 5.4, the caustic transit rates for stellar-mass BH accretion disks may be closer to the upper value indicated in black in Figure 1. Within the uncertainties detailed in Appendices C–D, it is thus possible that stellar-mass BH accretion disks at $z \gtrsim 7$ may outshine the sky-SB from Pop III stars in the observed near-IR, and that they may produce correspondingly more caustic transits. Only a long-term, dedicated observing program may be able to tell the difference between these two possible sources of caustic transits at $z \gtrsim 7$, as discussed in Section 7.

In conclusion, Pop III star rotation, the way dust is produced and expelled during and after the Pop III star evolutionary sequence, the massive star binary fraction, and the subsequent stellar-mass BH accretion timescales and accretion efficiency may well in the end determine which of the two has the best chance to be detected by *JWST* via cluster caustic transits.

7. Possible Observing Programs to Detect Pop III Caustic Transits

JWST's lifetime requirement is five years, and its lifetime goal is 10 years (Gardner et al. 2006). *JWST*'s actual mass is currently about 200 kg under its allotted 6500 kg launch mass, so its propellant tank has been completely filled, enabling a maximum possible lifetime of 11–14 years with proper angular momentum management if no hardware components and their spares fail before that time. *JWST* carries a number of *HST*/SM4 heritage parts, and the *HST*/WFC3 hardware is operating just fine three years past its design lifetime. Hence, contemplating a compelling time-domain science case for a *JWST* mission with a 5–10 year baseline is possible.

7.1. Characteristics of a JWST Survey to Find Pop III Caustic Transits at $z \gtrsim 7$

To observe caustic transits from First Light objects, a dedicated *JWST* observing program will be required of at least several and up to 30 clusters for a duration of 1–10 years (see Figure 1). Depending on their exact contribution to the diffuse 1–4 μm sky-SB ($\lesssim$0.01–0.1 nW m^{-2} sr^{-1}), such a *JWST* observing program to detect individual Pop III stars and/or their stellar-mass BH accretion disks at $z \gtrsim 7$ may well require monitoring—in the optimistic case that most of the NIR power-spectrum signal comes from $z \gtrsim 7$—a few suitable galaxy clusters during a year. In the most pessimistic case where there exist really only a few Pop III objects per square arcsecond and/or that most of them are shrouded by dust, *JWST* may need to monitor 30 clusters twice a year for a good fraction of *JWST*'s 5–10 year lifetime to detect a few Pop III caustic transits. All of these cluster observations would require coeval images in four NIRCam filter pairs and/or four NIRISS filters to constrain the spectral signature and redshift of a Pop III caustic transit candidate. These would appear as $z \gtrsim 7$ dropout candidates that vary with time, either increasing rapidly and then slowly fading, or vice versa. Their rise time are of the order of hours, while their fading times are a good fraction of a year.

Both cases pose interesting challenges to *JWST* IR-array data reduction techniques: great care must be taken that a sudden increase in magnified object flux during a caustic transit does not get rejected as an artifact or as a cosmic ray in the series of images taken that day. Also, care must be taken that a slow increase in magnified flux of an object that approaches the caustic from the other side does not get misinterpreted as a slowly variable faint Galactic brown dwarf star or a weak variable AGN (e.g., Cohen et al. 2006). The nature of such "reverse transits" may therefore not be obvious when first identified observationally by *JWST*.

Could Pop III caustic transits cause a real difference in the luminosity function at $z \gtrsim 7$ in the field (e.g., Bouwens et al. 2017) compared to clusters (Livermore et al. 2017)? If in the most optimistic case several Pop III objects at $z \gtrsim 7$ were always seen transiting a cluster caustic in any given year, then this could artificially boost the number of $z \gtrsim 7$ objects seen behind clusters. This may not be obvious if such Pop III objects resided in small star-forming objects that are well below the *HST* or *JWST* detection limits, so one would not know in advance to expect caustic transits at these locations. This is unlike the caustic transiting objects detected by Kelly et al. (2018) and Rodney et al. (2017), where there was a known faint galaxy at a given location on the cluster caustic. Although such caustic transit detections of Pop III objects at $z \gtrsim 7$ behind clusters could be real, they may need additional lensing magnification corrections in order to represent the unlensed background universe at $z \gtrsim 7$, and so could affect the derived steady-state LF. Detailed *JWST* studies of high-quality LFs at $z \gtrsim 7$ that are well-sampled behind different clusters may reveal cluster-to-cluster differences in caustic properties. Cosmic variance of the $z \gtrsim 7$ population will also require averaging over a significant number of lines of sights, by observing a number of clusters with *JWST*.

Microlensing by faint stars in the lensing cluster ICL may decrease the magnifications from $\sim$10^4–10^5 to $\gtrsim$10^3, but greatly lengthen the visibility time of the caustic transit, where a transiting microlensed object may be visible for many decades or longer. We outlined an observing strategy that *JWST* may use to observe these objects. To minimize the effects from microlensing in the modeling of caustic transits, one could also target some compact galaxy clusters at $0.3 \lesssim z \lesssim 0.5$ that have a smaller fraction of ICL compared to their total galaxy light at the $z \gtrsim 7$ lensing contours, but that —due to their compactness—have excellent lensing properties (e.g., Griffiths et al. 2018).

7.2. Possible Spectral Differences Between Pop III Star and Stellar-mass BH Caustic Transits

A dedicated multiband *JWST* monitoring program of well-studied lensing clusters may be able to detect the chromatic differences expected between caustic transits of stellar-mass BH accretion disks and those of Pop III stars, perhaps including spectroscopic confirmation.

The one significant difference between Pop III stellar-mass BH accretion disks and Pop III stars is likely the presence of a hard X-ray component that contributes very significantly at the inner accretion disk radii and that will also have a significant energy tail longwards of Lyα 1216 Å. No such X-ray component would exist for the Pop III stars themselves, since, ignoring their limb darkening and any starspots, their stellar photospheres have nearly uniform temperatures of $T \simeq 10^5$ K (Section 3.1). Hence, Pop III stars will not show significant chromatic behavior that may be traced during a caustic transit, but BH accretion disks could show such chromaticity if they were detected close to the actual caustic transit.

Any differences in the effective UV radii between BH accretion disks and Pop III stars are important, since the maximum magnification obtained during a caustic transit increases strongly for objects with smaller effective UV radii (Equation (21) in Section 4.4). A very hot BH accretion disk crossing a caustic could have much larger magnifications at its smallest intrinsic X-ray–UV-bright radii, since these radii contribute a larger fraction of the total energy longwards of Lyα than they do for Pop III stars. Since the maximum magnification scales as $1/\sqrt{R_{\mathrm{UV}}}$ (Equations (16) and (21)), their smallest UV-bright radii (Section 5.5) could undergo a maximum magnification, μ_{max}, that is considerably larger

during a caustic crossing, which could boost their observed rates accordingly compared to Pop III stars.

Specifically, the inner (bluer) part of the BH accretion disk would be magnified much more than its outer (redder) part. The ratio in magnifications should follow $\sqrt{(r_{out}/r_{in})}$, where $r_{out} > r_{in}$ are the largest and smallest BH UV-accretion disk radii discussed in Section 5.5.2, respectively. This will result in chromaticity due to lensing, where the shape of the BH light curve peaks during a caustic crossing would depend more strongly on rest-frame UV wavelength, unlike that of the Pop III stars. For *JWST*, there could be a $\sim$1 dex difference in magnification between the bluer and redder filters for an object undergoing a caustic transit at $z \gtrsim 7$. If a caustic transit maximum is observed almost simultaneously in different *JWST* filters, we could then constrain the BH mass using Equation (38), assuming its scaling holds with slope $\rho \simeq 1/2$ to stellar BH masses (see Section 5.5). This would be an indirect way of confirming that part of the light observed from a $z \gtrsim 7$ object undergoing a caustic transit originates in accretion disks around stellar-mass BHs.

Pop III stars may also be detected or confirmed by *JWST* in other ways. For instance, Macpherson et al. (2013) consider the prospect of finding a Pop III hypernova "in flagrante," and suggest a detection rate of 2.78×10^{-6} per *JWST* field of view (FOV) and a probability of 37% that *JWST* will serendipitously image an afterglow during its lifetime. What *JWST* truly will find from the Pop III epoch may include these and other unexpected surprises. It is therefore critical that *JWST* First Light surveys are well-designed to optimize the possible detection of Pop III objects directly.

7.3. Role of Next-generation Ground-based Optical–Near-IR Telescopes in Caustic Transits

JWST will be able to detect and monitor caustic transits during its 5–10 year lifetime. It is therefore useful to consider which other facilities can observe caustic transits on longer timescales. *JWST*'s unique advantage is its very dark zodiacal sky in L2 (AB $\gtrsim$ 23–24 mag arcsec^{-2} at $\lambda \simeq$ 2.0–3.5 μm; Figure 1), and its stable PSF over a relative wide FOV (Rieke et al. 2005, 2$\farcm$2 $\times$ 4$\farcm$4). Together with its 25 m^2 collecting area, *JWST* should be able to reach AB $\gtrsim$ 28.5 mag routinely (Windhorst et al. 2008). The next generation 25–40 m ground-based telescopes—the European Extremely Large Telescope (E-ELT), the Giant Magellan Telescope (GMT), and the Thirty Meter Telescope (TMT)[11]—will have a much larger collecting area and narrower PSFs when using Multi-Conjugate (laser-assisted) Adaptive Optics, although perhaps not as stable as *JWST*'s PSFs, and they will have lower Strehl ratios. They will also have a 1–2 μm sky foreground that is $\gtrsim$7 mag brighter than *JWST*'s in L2. As a consequence, the next-generation ground-based telescopes may be able to reach AB $\lesssim$ 29 mag in integrations of hours at 1–2 μm, but—given their adaptive optics—only over a smaller FOV ($\lesssim$20$''$ $\times$ 20$''$– 1$'$ $\times$ 1$'$). Ground-based telescopes will have reduced sensitivity at wavelengths $\lambda \gtrsim$ 2–2.2 μm because of the strongly increasing thermal foreground. For that reason, *JWST* will be able to better address any chromatic differences between caustic transits of Pop III stars and their stellar-mass BH accretion disks (Section 7.2), especially those at $z \gtrsim 12$ that require

several very sensitive filters at $\lambda \gtrsim 2$ μm, where ground-based telescopes cannot reach AB $\sim$ 29 mag due to the much brighter thermal foreground.

Confirming spectra of caustic transits by Pop III stars or their stellar-mass BH accretion disks could be taken with the *JWST* NIRISS and NIRSpec spectrographs, and also with the next-generation near-IR spectrographs on the ELT, GMT, and TMT telescopes. Of particular interest would be detecting the 1640 Å He line, which is expected to be present in the ionized regions around Pop III stars, or their BH accretion disks, with $T \gtrsim 10^5$ K (Schaerer 2002; Sobral et al. 2015).

We do not need to catch a caustic transit event at the precise moment of crossing the caustic. It may be sufficient if a Pop III star is seen $\lesssim$10 yr before or after a caustic crossing, when the typical magnification may well be of order 10^4, which can make a Pop III star with AB $\simeq$ 38 mag visible to *JWST*. In five years' time, the observed flux would increase (or decrease) steadily by a factor of $\sim\sqrt{2}$, which could be identified as a star heading toward (or away from) a caustic. Perhaps the caustic transit of such stars will not be observed during *JWST*'s lifetime, but the next-generation ground-based telescopes will be able to continue to monitor such stars for a much longer period, when a given star appears to be heading toward a caustic in several years' time.

In summary, the next generation ground-based telescopes can monitor at 1–2 μm over a much longer period than *JWST*— individual Pop III caustic transits that *JWST* will have detected at 1–4 μm during its lifetime, and also discover new ones on timescales longer than *JWST*'s lifetime. This capability would be particularly useful to follow up on caustic transits that may be affected by microlensing, and so may stretch out over many decades. Because of its much wider 1–4.5 μm wavelength range over which it can reach AB $\simeq$ 29 mag, *JWST* will be essential to distinguish between possible chromatic differences between Pop III stars and BH caustic transits.

8. Summary and Conclusions

The following are the main conclusions of our paper:

(1) The panchromatic (0.1–500 μm) discrete galaxy counts (Driver et al. 2016) converge well at almost all wavelengths, resulting in iEBL values from discrete objects that are well-determined (to within 20%) and similar to those obtained at 1–4 μm from γ-ray blazar spectral distortions. Therefore, limits to the diffuse 1–4 μm EBL are likely below $\lesssim$1–2 nW m^{-2} sr^{-1}, which we consider as "hard" upper limits for any Pop III contribution to the EBL.

(2) Based on recent near-IR (Kashlinsky et al. 2012, 2015; Mitchell-Wynne et al. 2016) and near-IR–X-ray power-spectrum (Cappelluti et al. 2013) results and theoretical estimates, we adopt tighter constraints to the sky-SB from Pop III BH accretion disks of $\lesssim$0.11 nW m^{-2} sr^{-1} (i.e., sky-SB $\gtrsim$ 31 AB-mag arcsec^{-2} at 2.0 μm). From observational and theoretical considerations of the cosmic SFH, we adopt similar upper limits to the 2.0 μm SB for Pop III stars themselves.

(3) These adopted near-IR Pop III sky-SB values lead to a predicted rate of $\lesssim$0.32 Pop III star caustic transits per cluster per year that may be observable with *JWST* to AB $\lesssim$ 28.5 mag, with rise times of less than a few hours and decay timescales of less than a year, or vice versa, depending on from which direction the Pop III object approaches the caustic: starting at the "sharp edge" of the caustic, or starting at the other side that declines smoothly as $1/\sqrt{d}$. Microlensing by intracluster

---

[11] http://www.eso.org/sci/facilities/eelt/, http://www.gmto.org/resources/, and http://www.tmt.org/.

medium objects can reduce transit magnifications but lengthen visibility times.

(4) For Pop III stellar-mass BH accretion disks and their anticipated accretion times of 0.3–60 Myr, we suggest cluster caustic transit rates that are similar to those of Pop III stars, amounting to $\gtrsim 0.2$ Pop III BH accretion disk caustic transits per massive cluster per year. The BH feeding timescales compared to the Pop III star lifetimes—and the amount and distribution of self-produced dust around the Pop III stars and their subsequent BH accretion disks—will determine which one of the two compact UV sources will yield the most frequent cause of cluster caustic transits that could be observed by *JWST* or the next-generation 25–40 m ground-based telescopes.

(5) In the case where the actual caustic transit rates from Pop III stars or their stellar-mass BH accretion disks are much lower than our suggested predictions, the actual detection rate—or upper limits thereto—by *JWST* over its 5–10 year lifetime will significantly constrain the Pop III objects that our universe contains. Any firmly detected Pop III caustic transit would be one of the most exciting First Light discoveries with *JWST*.

If no Pop III caustic transits are seen with *JWST* by monitoring ∼30 clusters over 5–10 years, despite a long dedicated campaign, then the SB of Pop III stars and their stellar-mass BH accretion disks may truly be fainter than SB $\gtrsim$ 36–37 mag arcsec^{-2} at 2.0 μm. In other words, the true Pop III star density would be very low indeed, with only a few Pop III stars per arcsec^2 in the 1–4 μm sky. Although not as exciting as a number of significant caustic transit detections at $z \gtrsim 7$, such a null experiment would be interesting in itself, as it would significantly constrain the sky-SB of Pop III objects at $z \gtrsim 7$ that may contribute to the diffuse EBL. Either way, the experiment would allow *JWST* to directly constrain the First Light epoch.

In summary, unlensed Pop III stars or their stellar-mass BH accretion disks may have fluxes of AB $\simeq$ 35–41.5 mag at $z \simeq$ 7–17, and so will not be directly detectable by *JWST*. However, cluster caustic transits with magnifications of $\mu \simeq 10^4$–10^5 may well render them temporarily detectable to *JWST* in medium-deep to deep observations (AB $\lesssim$ 28.5–29 mag) on timescales of months to a year, with rise times less than a few hours. Deep and well time-sequenced observations of the best-lensing clusters carried out throughout *JWST*'s lifetime would fulfill its promise to the US Congress and citizens as NASA's "First Light" telescope.

Appendix A
Perturbing the Cluster Velocity Distribution to Constrain the Maximum v_T Value

To test the maximum values of v_T for galaxy clusters likely to be observed by the community for caustic transit observations, we examined the available redshift space distribution of the galaxies in three well-studied HFF clusters. For a circularized cluster in virial equilibrium, the central distribution of cluster galaxies in redshift space (i.e., projected distance, R_{proj}, from the cluster center as a function of line-of-sight velocity v_{los}) is expected to resemble a "trumpet" (Diaferio 1999; Alpaslan et al. 2012). This is clearly visible in the left-hand panels of Figure 5, where we show the redshift space distribution of galaxies in the HFF clusters Abell 2744, MACS J0416–2403, and MACS J1149.5+2223. Redshift information for Abell 2744 was taken from Owers et al. (2011), and for MACS J0416–2403 plus MACS J1149.5 +2223 from Ebeling et al. (2014). In the middle panels of Figure 5, we display the observed velocity distribution of v_{los} from the left-hand panels as thick black lines. We expect the distribution of the v_{los} for the central cluster to peak around 0 km s^{-1} with respect to the cluster redshift. But all three HFF clusters are embedded in significant large-scale structures in velocity space, with typical separations between different structures along the line of sight between ∼500 and ∼2000 km s^{-1}. MACS J0416–2403 is composed of two merging clusters, with a small difference in redshift between them. The black lines suggest that the central core of all three clusters appears to have significant substructure in velocity space, especially for MACS J0416–2403 and MACS J1149.5 +2223, whose central cores are significantly non-Gaussian in their redshift distribution $N(v)$. Each line of sight may have a number of virialized substructures—including the main cluster itself—each with approximately a Gaussian velocity distribution $N(v)$. One could think of the velocity distribution in the central core of the black lines as some combination of virialized Gaussians with a broader non-virialized component of galaxies, a fraction of which are falling into the main cluster.

The question then arises: by how much can we perturb the space velocity of the cluster itself before the v_{los} distribution is noticeably changed from the observed redshift distribution,
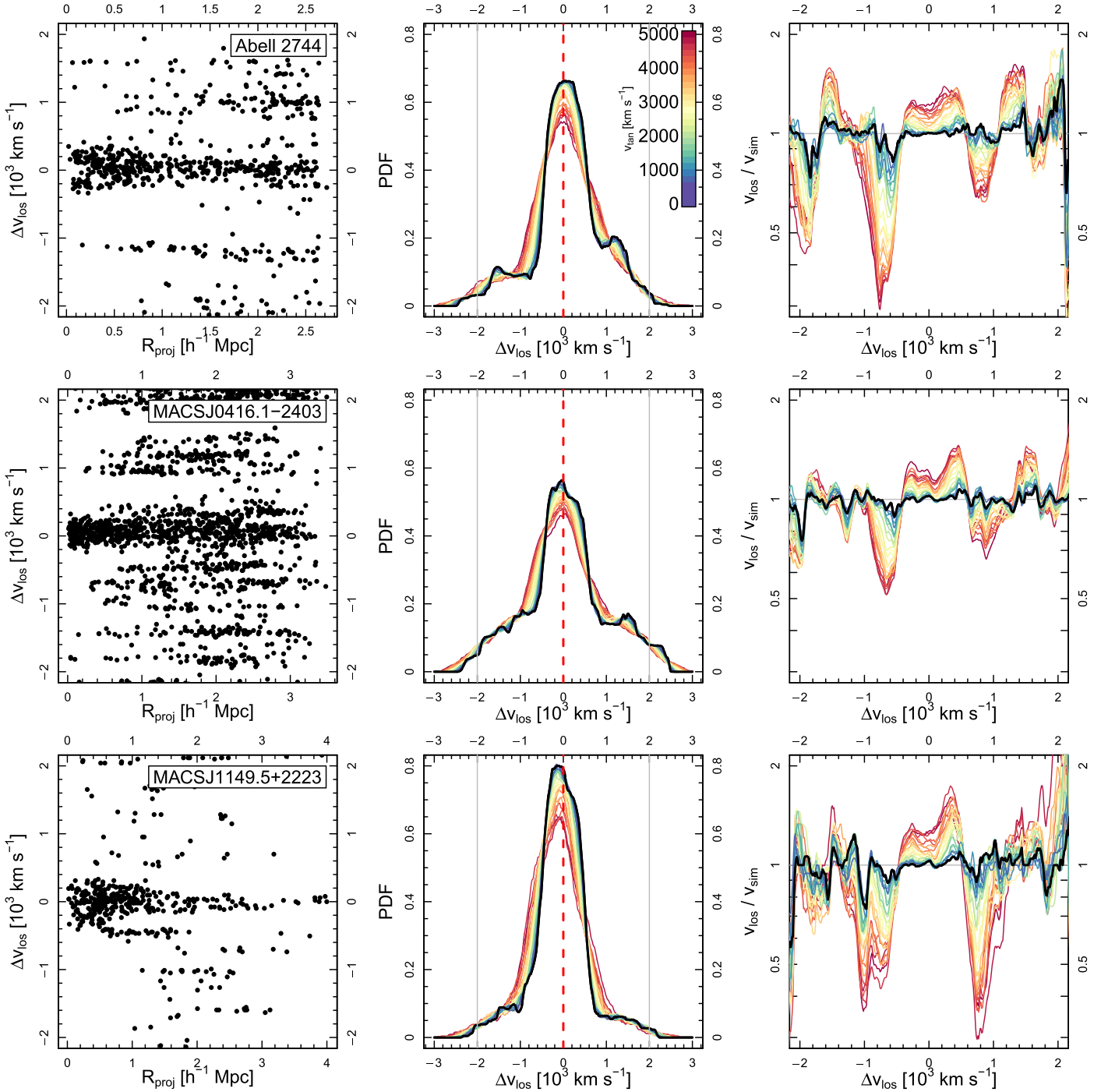
Figure 5. Redshift space distribution of galaxies in three lensing clusters (from the HFF program) suitable for lensing of First Light objects and caustic transit studies at $z \gtrsim 7$ with *JWST*. The left-hand panels show the projected radial distance from the cluster center of each galaxy as a function of its line-of-sight velocity. The middle panels show the distribution of line-of-sight velocities for each cluster, where the thick black lines indicate the real velocity data from the left-hand panels. The colored lines show the modified line-of-sight velocity distributions, after a random space velocity has been added to the whole cluster, affecting both its v_{los} and its v_T vectors. The resulting added transverse velocity increments range between $v_T \simeq 500$ to 5000 km s^{-1}, as indicated by the colorscale bar in the top middle panel. The right-hand panels display the residuals from the middle panels for better visibility between the models. We only detect a significant deviation from the observed v_{los} distribution when $v_T \gtrsim 1700$–2300 km s^{-1}. We adopt $v_T \lesssim 1000$ km s^{-1} as an upper limit to the transverse velocity of these clusters in the plane of the sky when calculating the possible caustic transit rate of First Light objects that may be seen by *JWST* at $z \gtrsim 7$. At lower v_T values, differences between the perturbed $N(v)$ model and the actual redshift data cannot be distinguished.

which then also provides a limit to the maximum transverse velocity component that can be randomly added? To implement this, we take the line-of-sight velocities of all galaxies within the central cluster itself and perturb them with a space velocity vector that is aligned at 45° with respect to the line of sight. We ensure that each galaxy is perturbed by a similar value to mimic the effect of a true space velocity on the observed redshift distribution of the central cluster. Galaxies whose projected distance is between the cluster center and the median projected distance for that cluster are perturbed by a vector whose magnitude is 500 km s^{-1} less than galaxies at the outskirts of the cluster and whose projected distance is greater

than that of the median. This simulates the effects of a differential velocity disturbance, since the galaxies closer to the cluster potential well likely experience a velocity perturbation resulting from the cluster space velocity that is smaller in magnitude compared to the local velocity dispersion. The magnitude of the perturbing velocity was drawn from a normal distribution centered on the number given in the colored legend of Figure 5, with a standard deviation of 500 km s^{-1}. The angle was drawn from a uniform distribution ranging from 40° to 50°. Finally, we extract the y-component of this resulting velocity vector, and assign that to be the new line-of-sight velocity for each galaxy. The colored lines in the middle and right panels of Figure 5 show these modified line-of-sight velocity distributions after the space velocity has been added to the cluster and decomposed into the v_{los} and the v_T vector components for each galaxy. The added transverse velocity increments range between $v_T \simeq 500$ to 5000 km s^{-1}, as indicated by the color bar.

In the right-hand panels of Figure 5, we display the residuals from the middle panels for better visibility between the models. We only detect a significant deviation from the measured v_{los} distribution if we increase the space velocity such that the components added to the transverse velocity exceed $v_T \gtrsim 1000$ km s^{-1}, where the excess becomes clearly visible in Figure 5 at $v_T \gtrsim 2000$ km s^{-1} (green–red curves). A simple normalized χ^2 estimate for each fit shows that the reduced χ^2 v_{los} values start to exceed unity when space velocities have been added with transverse velocity components considerably higher than 1000 km s^{-1}. For Abell 2744, the reduced χ^2 values exceed unity at $v_T \gtrsim 1900$ km s^{-1}, for MACS J0416–2403 at $v_T \gtrsim 2300$ km s^{-1}, and for MACS J1149.5 +2223 at $v_T \gtrsim 1715$ km s^{-1}. We obtain similar results when we add space velocities to each galaxy that are more randomized in angle, or when we add more randomized values of the space velocity to each of the cluster subclumps.

In conclusion, Figure 5 thus shows that adding space velocities with projected transverse components much larger than $v_T \simeq 1000$ km s^{-1} imply projected components of this space velocity added along the line of sight that are not consistent with the available redshift data in the cluster core. We will thus adopt an upper limit of $v_T \lesssim 1000$ km s^{-1} for the maximum transverse velocity of these clusters at $0.3 \lesssim z \lesssim 0.5$ in the plane of the sky when calculating the possible Pop III caustic transit rates of First Light objects that may be seen by *JWST* at $z \gtrsim 7$. For some substructures in each cluster, the v_T values may well be as high as 1000 km s^{-1}, or perhaps somewhat higher.

Appendix B
Caustic Transit Rates in the Presence of Microlensing

B.1. The General Case

When microlenses are present, the rate of caustic transit events is sensitive to the mass function of low-mass stars in the cluster ICL (e.g., Miralda-Escude 1991). If brighter Pop III stars are more common, relatively modest microlensing peaks with $\mu_{max} \simeq 5 \times 10^3$ can momentarily amplify a bright star (AB $\lesssim 37$ mag) to above the detection limit of *JWST*. The rate of events will then be dominated by the microlens peaks of the brightest stars, and the rate of events will be proportional to the optical depth of microlenses. Significantly fainter stars may be sufficiently magnified only if the disruption of the cluster

caustic by microlenses is moderate. In this case, most events will be produced by relatively faint Pop III stars crossing the lightly disrupted caustic, which can have a maximum magnification of $\mu_{max} \simeq 10^5$ (see Section 4). The rate of events will be dominated in this case by the anticipated, much more numerous fainter stars (AB $\simeq 40$–42 mag) when crossing the caustic, and will be proportional to the surface mass density of stars in the ICL. A rough estimate of the expected rate of events, R, in the presence of microlenses can be obtained based on the predictions from Kelly et al. (2018) and Diego et al. (2017), who modeled the HFF clusters with $v_T \simeq 1000$ km s^{-1}:

$$R(\text{yr}^{-1}) = A(>\mu)\rho_* \, r(\Sigma, v_T) = \frac{B_o}{\mu^2} \, \rho_* \, r(\Sigma, v_T). \quad (41)$$

Here, $A(>\mu)$ is the area in the source plane above a given magnification μ, which scales as B_o/μ^2 in the presence of microlenses and at high magnification. Also, ρ_* is the surface mass density of Pop III stars above $z = 7$, so that $A(>\mu) \cdot \rho_*$ is the number of Pop III stars undergoing microlensing at a particular moment. Last, $r(\Sigma, v_T)$ is the rate of microlens caustic (hereafter "microcaustic") events a moving object in the background ($z > 7$) would encounter if the surface mass density of microlenses is Σ, and the background object is moving with a transverse velocity v_T with respect to the network of micro-caustics. Kelly et al. (2018) estimated that $r(\Sigma, v_T)$ is of order 0.1 yr^{-1} for an event like Icarus. Using the equations of Diego et al. (2017) and assuming the total length of the caustics to be $L \simeq 100''$, we estimate that $B_o \simeq 1.8 \times 10^{-4}$ arcsec^2, or A ($\mu \gtrsim 3 \times 10^3$) $\simeq 0.002$ arcsec^2 for an HFF-like cluster in the presence of microlenses. The value $\mu \simeq 3 \times 10^3$ is adopted to select regions in the source plane associated with microlensing peaks that will reach $\mu_{max} \gtrsim 10^4$. Then, if there are $\rho_* \simeq$ 100 Pop III objects/arcsec^2 brighter than AB $\simeq 38$ mag, we would expect for each HFF-like cluster $R(\text{yr}^{-1}) \simeq 0.2 \times r$ $(\Sigma, v_T) \simeq 0.02$ caustic transits yr^{-1} if we extrapolate the results of Kelly et al. (2018) to the entire caustic region, i.e., one event when monitoring five HFF-like clusters for 10 years.

We note that both the estimates with microlensing in this appendix and those in Section 4.3 based on the adopted transverse velocity are in good agreement. These numbers should be compared with the expected caustic transit rate if we assume there are no microlenses. In this case, the magnification would be described by $\mu \simeq 20/\sqrt{d}$, which implies $d = 1.6 \times 10^{-5}$ arcsec for $\mu \simeq 10^4$. Note that the above expression would give $\mu = (20/\sqrt{1.6 \times 10^{-5}}) \simeq 5000$, but the total magnification would be $\mu = 10^4$ when we account for the double image produced in the image plane. If the perimeter of the caustic region is $L \simeq 100''$, then the area over which a magnification larger than $\mu \simeq 10^4$ can be attained is $\simeq 0.0016$ arcsec^2. This number is comparable to the value estimated above when microlenses are included (~ 0.002 arcsec^2).

The similarity of results obtained with and without microlenses could have been anticipated from basic principles. Owing to flux conservation, the number of photons collected after integrating for a long period (tens to hundreds of years) should be the same independent of the presence (and number) of microlenses. The distribution of microlenses determines how this magnification is redistributed. A lens plane without microlenses results in large magnifications concentrated in a unique narrow region around the caustic (i.e., a single very

bright peak; see Section 4.4), while a lens plane populated with microlenses will break apart the single caustic into multiple (smaller) microcaustics. Thus, the rate of high magnification events ($\mu_{max} \gtrsim 10^4$) would be similar whether or not there are microlenses, but in the case without microlenses, we would see a single very bright peak when the Pop III star crosses the caustic, while in the case with microlenses, we would see many (smaller) peaks hundreds of years before (or after) the star crosses the position of the cluster caustic (e.g., Diego et al. 2017). Extreme magnification ($\mu_{max} \simeq 10^6$) can be attained only when microlenses are not included, and may only occur for the lower-mass BH accretion disks whose inner X-ray–UV bright core radii may be much smaller than those of Pop III stars (Section 5.5.2), so that all large magnifications are concentrated in a single peak around the cluster caustic. When microlenses are included, the caustic region is expanded in size, as shown in Diego et al. (2017), so there is a higher probability for a star in the source plane to align with a microcaustic. However, the magnified peaks will be correspondingly fainter, so only the rarer, brighter Pop III stars or the brighter stellar-mass BH accretion disks may produce caustic transits that can be observed by *JWST*.

B.2. The Case of Relatively Bright Sources

If a background source at $z \gtrsim 7$ were to be relatively "bright" (AB $\lesssim 37$ mag for the unlensed source), virtually all microlensing peaks—with magnifications of about one to several thousand (Kelly et al. 2018)—can be observed if the star is sufficiently close to the critical curve (i.e., $\mu \gtrsim 10^3$). The number of events in that case will be approximately equal to the number of microcaustics that the background source encounters as it moves across the web of microcaustics. In this section, we therefore present an estimate of the case when *JWST* may observe when the background object is relatively bright. Instead of computing the probability of an event based on the area above a given magnification as in Appendix B.1, we can simply estimate the number of times a microcaustic is crossed, since a rare but very bright Pop III star (AB $\lesssim 35$–37.5 mag; see Table 2–4) may be directly observed, if the microlensing magnification from the cluster ICL at its location is at least $\mu \simeq 10^3$. In this case, even modest microlenses with subsolar masses can produce changes in flux of 0.5 magnitudes or more. Having a bright star undergoing such frequent encounters with microcaustics is possible, as discussed by, e.g., Kelly et al. (2018). These variations in flux may be observed with *JWST* when the star crosses the network of microlensing caustics.

The probability of having a microlensing event at a given distance, θ, from the critical curve is given by the effective optical depth to microlensing, which is defined as the fractional area at that location that is being affected by microlensing. Using the model in Diego et al. (2017), the effective optical depth from microlenses is given by

$$\tau(\theta) \simeq 21 \times 10^{-2} \, \Sigma(M_\odot \, \mathrm{pc}^{-2}) / \theta(''). \quad (42)$$

For sources at high redshift, the critical curve moves to distances of order $1'$ from the center of the Brightest Cluster Galaxy (BCG), where the impact of microlenses is expected to be small but still not necessarily negligible. At these angular distances, Diego et al. (2017) estimated for MACS1149 that the surface mass density of microlenses decreases by

approximately two orders of magnitude with respect to the one estimated at the position of Icarus. This corresponds to a mass surface density of $\Sigma \simeq 0.1 \, M_\odot \, \mathrm{pc}^{-2}$. Hence, if we restrict our analysis to the region where the effective optical depth of microlenses reaches the saturation level (i.e., $\tau \simeq 1$), then this implies an angular distance of $\theta \simeq 20$ milliarcsec (mas). That is, the two counter-images of the lensed background object would appear on either side of the critical curve and be separated by ~ 40 mas. At these separations, both counter-images would form a single unresolved—or at best a slightly resolved—object in the *JWST* mosaics. At 20 mas distance from the critical curve, the model in Diego et al. (2017) predicts that the magnification from the cluster is approximately $\mu \simeq 5000$, so that any star brighter than AB $\simeq 38$ mag transiting the caustic at this location could be detected by *JWST* to AB $\lesssim 29$ mag.

In principle, we could see twice the caustic transit rate in this case, since this unresolved image would contain fluctuations form both sides of the critical curve (i.e., from its positive and negative parity). In reality, the rate on the side with negative parity is expected to be a factor $\sqrt{2}$ times smaller than the rate on the side with positive parity. This difference in rate can be obtained from Oguri et al. (2018), who estimated that the maximum dimension (or cross-section C_S) of the microcaustic on the side with positive parity is $C_S(M) = \theta_e(M)\sqrt{\mu_t}/\mu_r$, while its shape is that of a stretched diamond. The Einstein radius of the microlens, θ_e, depends on the mass of the microlens and the angular-diameter distances from the observer to the lens, from the lens to the background object, and from the observer to the background object. On the side with negative parity, this extension is smaller by a factor of $\sqrt{8}$, but the caustic is divided into two semi-diamond shapes, so that a star crossing the microcaustic would cross caustic lines four times instead of twice. Hence, the effective length (or cross-section) of the caustic on the side with negative parity is smaller by a factor of $\sqrt{8}/2 = \sqrt{2}$.

With the above ingredients, it is possible to estimate the expected number of caustic crossings for relatively bright sources (AB $\lesssim 37$ mag), which are now the microcaustics formed by the local microlenses. Each microcaustic is shaped as a diamond, or a double semi-diamond, on the sides with positive and negative parity, respectively (for details, see, e.g., Diego et al. 2017; Oguri et al. 2018). For simplicity, we assume that the microcaustic crossing events take place within the region where the saturation level to lensing is reached (i.e., the 40 mas region surrounding the critical curve mentioned above). In the regime where the saturation level to lensing has not been reached, one can still use the relation between the position in the source plane, β, and the position in the image plane, θ, given by standard lensing theory, $\beta = \theta^2/C$, where the constant C depends on the lens strength (i.e., the gradient of the lensing potential), and both β and θ are given with respect to the caustic and critical curves, respectively. For a cluster like MACS 1149, Diego et al. (2017) estimated $C \simeq 68''$. For this particular value of C, one obtains $\beta \simeq 6$ microarcsec, or 2.5×10^{-2} pc at $z \simeq 10$. The distance traveled by a moving background star with respect to the caustic network is $d(v) = 1 \times 10^{-4}(v_T/1000 \, \mathrm{km \, s}^{-1})$ pc yr^{-1}. During this time, the star may encounter multiple microcaustics, depending on the surface density of microlenses. For simplicity, we assume that the background source is moving perpendicular to

the maximum extension of the diamond-shaped macrocaustics. This is a reasonable assumption, since the microlens caustics are typically stretched by very large factors, so to first order they can be approximated by straight parallel lines. Since each microcaustic has a cross-section $C_S(M)$ that scales with $\theta_e(M)$ (see Oguri et al. 2018 or the expressions above), the yearly rate of intersections with a microcaustic of mass M is then given by

$$r(M) = 2(1 + \sqrt{2}) \, . \, n(M)\mu \, . \, d(v) \, . \, 2C_S(M). \qquad (43)$$

Here, the first term, $2(1 + \sqrt{2})$, accounts for the events produced on either side of the critical curve and the fact that a microcaustic is crossed twice (four times for the side with negative parity). The second term, $n(M)\mu$, is the number density of microlenses with mass M in the lens plane, $n(M)$, which is increased by a factor μ in the source plane. The third term, $d(v)$, is the distance traveled by the background object in one year. The last term, $C_S(M)$, is the cross-section of a microcaustic of mass M. For realistic distributions of $n(M)$, one should integrate $r(M)$ to compute the caustic transit rate, but for our purposes we adopt the simple scenario where all microlenses have similar masses of $M \simeq 1 M_\odot$. In that case, we get $n(M) = \Sigma(M) = 0.1$ pc^{-2} and $r(m) = 4.7 \times 10^{-7}\mu_t^{3/2}$ per year. Here we assumed that $v_T \lesssim 1000$ km s^{-1}, $z_{lens} = 0.5$, and $z_{source} = 10$, for which one obtains $\theta_e = 2.3 \times 10^{-6}$ arcsec, or 0.0097 pc at $z = 10$. Also, $\mu = \mu_t\mu_r$, so that the dependency with μ_r cancels out and the rate depends only on $\mu_t^{3/2}$. To estimate the value of μ_t, we adopt the model in Diego et al. (2017), according to which $\mu_t = \mu/5$ and $\mu \simeq 100/\theta$, where θ is in arcseconds. At the point where the effective optical depth of microlenses reaches the saturation level, $\theta = 0\rlap{.}''02$, we get $\mu_t \simeq 1000$, and the caustic transit rate becomes $r(M) \simeq 0.015$ per year.

The above rate is the expected rate per year for one background star intersecting $n(M) = \Sigma \simeq 0.1$ pc^{-2} stars per pc^2 with mass $M \simeq 1 M_\odot$. During the time the star is moving across the saturation region and toward the main caustic of the cluster, it will intersect many microcaustics until it reaches the main caustic, after which it fades away forever from our vantage point. We ignore the equally likely case where the star approaches from the main caustic from the other direction, which is discussed in Sections 4.4 and 7.1, but is observationally much harder to recognize. We can then estimate the time it takes to cross the saturation region as 0.0097 (pc)/10^{-4} (pc/year) = 97 years, so in this time the background star would cross 1.5 microcaustics before reaching the main cluster caustic. The final boosting factor is expected to be modest in the regions of the critical curve with a very small density of microlenses, which would amount to $\sim$2.5 caustic crossings instead of just one.

Finally, we note that the approximations made above assumed a very conservative low density of microlenses, about two orders of magnitude smaller than in the outskirts of the BCG region. There may be certain regions along the critical curve where the density of microlenses increases very significantly, for instance near an area with a larger fraction of ICL or near a cluster member galaxy. If in these areas the number density of microlenses, $n(M)$, increases substantially, the rate would increase by a similar amount. Assuming one could estimate the luminosity function of Pop III stars at $z \gtrsim 7$ in the future from a long-term monitoring program of cluster

caustic transits, one would expect their LF to show a significant excess at the highest luminosities, as a consequence of the caustic/microcaustic crossings boosting their observed luminosities the most. This would be akin to the lensing tail observed in the bright end of the high-redshift submillimeter galaxy luminosity function, such as, for instance, has been seen for the lensed *Herschel* sample (e.g., Negrello et al. 2017).

Appendix C
Uncertainty Estimates for Caustic Transit Rates of Pop III Stars at $z \gtrsim 7$

Here we estimate the uncertainties in the caustic transit rates and rise times of Pop III stars at $z \gtrsim 7$. The combined uncertainty in their caustic transit rates follows from the multiplicative sources of error in Equations (19), (25), and (26). These are the adopted effective caustic length L_{caust} (in arcsec), the cluster transverse velocity v_T (in km s^{-1}), the stellar luminosity L (in L_{100}), and the 1–4 μm sky-SB from Pop III stars (in mag arcsec^{-2}).

The error on the effective cluster caustic length L_{caust} is estimated to be $\sim$0.3 dex (Figure 4(b)), which incorporates the measurement errors in tracing L along the caustics, and the differences in caustic lengths between current lensing models (see Section 4.3). The error in the cluster transverse velocity v_T is estimated to be at least 0.3 dex, following the discussion in Section 4.2.2 and Appendix A. This includes the uncertainty in the v_T values as constrained in Figure 5 and their v_T values as projected onto the plane of the sky that assumed an average foreshortening of $\langle \sin(i) \rangle \simeq 1/2$, as discussed in Section 4.3.1.

Choi et al. (2016) predicted the stellar luminosities, radii, and T_{eff} values over a wide range of masses ($0.2 \lesssim M \lesssim 30 M_\odot$) and metallicities ($-4.0 \lesssim Z/Z_\odot \lesssim +0.5$) using the same MESA models as in Section 3.1, and compared these to data from a large number of detached eclipsing binaries in our Galaxy. They find that the stellar luminosities predicted by the MESA models in general follow the detached eclipsing binary data to within 0.2 dex. Although not anchored yet in data for (nearly) zero-metallicity stars with $M \gtrsim 30 M_\odot$ as needed for Pop III stars at $z \gtrsim 7$, we will adopt the 0.2 dex error in the predicted luminosities and radii for lower-mass stars to be representative of more massive, (nearly) zero-metallicity Pop III stars at $z \gtrsim 7$. Future work will need to Monte Carlo model the shape of the error distribution in these parameters (e.g., Fields et al. 2018).

The uncertainties in the caustic rise times follow for each mass from the two parameters in Equation (23): R_{100} and v_T. Choi et al. (2016) also find that the predicted stellar radii follow the detached eclipsing binary data to within $\sim$0.2 dex. This is less than the 0.3 dex uncertainty in the v_T values. Since both parameters are independent, the resulting uncertainties in the caustic rise times are thus $\lesssim$0.4 dex.

For Pop III binary or multiple stars, the situation may be more complex, as discussed in Sections 3.2 and 5.2, but can be approximated as follows. Unless a lot of mass exchange happens continuously in Pop III binaries, to first order, a binary—which will generally be of unequal masses (see Equation (6))—consists of two stellar photospheres with luminosities and radii that are determined by their ZAMS mass (Section 3.1). Without mass exchange, their radii will be in the range 2–13 $R_\odot$ during the ZAMS stage (see Table 2), while their typical binary separations are expected to be $\sim$10–100 $R_\odot$ (Section 3.2.3). The more common lower-mass Pop III stars with $M \lesssim 20 M_\odot$ in a binary

are simply too faint to be seen during caustic transits (Tables 2–4). However, when both Pop III stars in a binary have masses $M \gtrsim 30$–$50\,M_{\odot}$, caustic transits of Pop III binaries will, to first order, consist of multiple peaks, each with a transit rise time of less than a few hours as specified in Equation (28), while for constant v_T these events will thus be separated in time by hours to days. Therefore, to first order, Pop III star multiplicity does not affect the calculated caustic transit rates, other than producing two successive caustic transits that are likely separated by hours to days, such as potentially already observed at $z \simeq 1.5$ by Kelly et al. (2018). *JWST* epochs observed hours to days apart may thus observe multiple caustic transits for Pop III binary stars and identify each by its different SED colors, as discussed in Section 7. In any case, massive Pop III binary stars will likely lead to a double caustic transit, and as long as both caustic transit events are observationally recognized as coming from stars with different radii (rise times) and SEDs, they will not lead to a significant overcounting of Pop III binary star caustic transits.

Fast-rotating massive stars will evolve more toward the blue than their non-rotating counterparts, since their mass loss is not driven by the classic line-driven winds (Castor et al. 1975), but by wave transport, which is not incorporated in our current models. The effects of this on the radii and luminosities of massive stars will need to be addressed in future work.

The presence of microlensing in the foreground cluster ICL will require adjusting these calculations, as discussed in Section 4.3.2 and Appendices B–B.2. To first order, microlensing may reduce the magnification of each caustic transit event from values of $\mu \gtrsim 10^4$ to several thousands (Diego et al. 2017), but multiply the number of caustic transit events seen per unit time accordingly, while preserving the total lensed flux during the crossing of all (micro-)caustics by this object. Therefore, depending on the IMF slope, as discussed in Section 3.4 and Equation (32), the rarer, but more massive Pop III stars with $M \gtrsim 100$–$50\,M_{\odot}$ may be magnified more often than without microlensing at the expense of the more common, lower-mass Pop III stars ($M \lesssim 50\,M_{\odot}$), which may now become invisible more often due to their smaller (microlensed) magnification. Because the same star could be seen microlensed several times over a decade or longer (Diego et al. 2017), this could lead to overcounting of the caustic crossings, unless the observations allow us to recognize that these caustic crossings all came from a star with the same radius (rise time) and SED. Given this possible overcounting from microlensing, we take the uncertainty in the caustic transit rates induced by microlensing to be at least 0.5 dex.

The last major uncertainty in the caustic transit calculations is the value of the 1–4 μm sky-SB that comes from Pop III stars at $z \gtrsim 7$, which according to the discussion in Section 2.3.1 is $\gtrsim 31.4$ mag arcsec^{-2}. Given that recent hierarchical models yield values of the SFR at $z \simeq 7$–10 (Sarmento et al. 2018) that are within a factor of 2–3 from the Madau & Dickinson (2014) SFR at $z \simeq 7$–8 (Equation (1)), we will adopt the uncertainty in the 1–4 μm Pop III star sky-SB to be $\sim$0.3–0.5 dex. It is not likely that the true 1–4 μm sky-SB is much higher than this amount by many times this uncertainty, because the fitted values to the SFH data at $7 \lesssim z \lesssim 10$ by Madau & Dickinson (2014), Madau & Fragos (2017), and Finkelstein (2016) do not permit this, and because at least $\gtrsim 75\%$ of the near-IR sky-SB that comes from $7 \lesssim z \lesssim 17$ is already produced in the redshift bins at $z \simeq 7$–8 (Section 2.3.1). But it is possible that the 1–4 μm sky-SB from Pop III stars is significantly lower by

factors of 10–100 or more, as discussed in Sections 3.5 and 4.5, and as indicated by the (light orange) range in sky-SB levels in Figure 1 that may come from Pop III stars. The resulting Pop III star caustic transit rates discussed in Section 4.4 can therefore be regarded as upper limits, and the consequences of this for the *JWST* observing strategy are discussed in Sections 4.5 and 7.

Appendix D
Uncertainty Estimates for Caustic Transit Rates of Stellar-mass BH Accretion Disks at $z \gtrsim 7$

Here we estimate the uncertainties in the caustic transit rates of Pop III stellar-mass BH accretion disks at $z \gtrsim 7$. We will follow the same reasoning as in Appendix C, with some important differences. The error estimates for the total caustic length L_{caust} and the cluster transverse motion v_T are $\sim$0.3 dex for each, as discussed in Appendix C. The combined uncertainty from overcounting due to microlensing also remains at 0.5 dex (Appendix C).

The main differences with uncertainties in the caustic transit rates of Pop III stars are twofold. First, the uncertainty in the adopted 3–4 μm sky-SB from stellar-mass BH accretion disks is significant, like it is for Pop III stars, but unlike that of Pop III stars, it is not necessarily an upper limit. Following the discussion in Section 2.3.2, the IR and IR–X-ray power-spectrum results observed in the object-free *Spitzer* and *Spitzer–Chandra* images, respectively (Kashlinsky et al. 2012; Cappelluti et al. 2013; Mitchell-Wynne et al. 2016) have an (amplitude)^2 that is at least $\sim$0.3 dex uncertain between these papers. As discussed in Sections 2.3.2 and 5.5.1, the *Spitzer–Chandra* power-spectrum results hint at a component caused by (stellar-mass) BHs, since Pop III stars simply do not get hot enough to cause this signal. Since this power spectrum did not come from discrete objects seen down to either the *Spitzer* or *Chandra* detection limits, it is possible that a significant fraction of the near-IR sky-SB of $\gtrsim 31$ mag arcsec^{-2} that we derived in Section 2.3.2 comes from (stellar-mass) BH accretion disks at $z \gtrsim 7$. We therefore adopt the uncertainty in the near-IR sky-SB signal itself for stellar-mass BH accretion disks to be at least half this, or $\gtrsim 0.15$ dex, following the derivation in Section 2.3.2.

Second, unlike that of Pop III stars, the uncertainty in the predicted luminosities of stellar-mass BH accretion disks is no longer smaller than the uncertainty in the other parameters. The two methods of Sections 5.5.1 and 5.5.2 predicted their luminosities consistently (top and bottom tiers of Table 5), but this assumed that these BHs were always accreting. The largest uncertainty in L comes from their accretion efficiency, or accretion duration, as discussed in Section 5.4, which we assume is uncertain by at least 0.5 dex. This could reduce their luminosities from the steady-state values that we adopted in Section 5.5, and so increase their caustic transit rates for a given sky-SB, as discussed in Section 6.2.

ORCID iDs

Rogier A. Windhorst ⓘ https://orcid.org/0000-0001-8156-6281
F. X. Timmes ⓘ https://orcid.org/0000-0002-0474-159X
J. Stuart B. Wyithe ⓘ https://orcid.org/0000-0001-7956-9758
Mehmet Alpaslan ⓘ https://orcid.org/0000-0003-0321-1033
Daniel Coe ⓘ https://orcid.org/0000-0001-7410-7669
Jose M. Diego ⓘ https://orcid.org/0000-0001-9065-3926

Simon P. Driver ⓘ https://orcid.org/0000-0001-9491-7327
Patrick L. Kelly ⓘ https://orcid.org/0000-0003-3142-997X
Duho Kim ⓘ https://orcid.org/0000-0001-5120-0158

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, PhRvL, 116, 061102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016b, PhRvL, 116, 241103
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016c, ApJL, 818, L22
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016d, ApJL, 832, L21
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016e, ApJL, 833, L1
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, PhRvL, 118, 221101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, PhRvL, 119, 161101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, ApJL, 848, L13
Abel, T., Bryan, G. L., & Norman, M. L. 2002, Sci, 295, 93
Acebron, A., Jullo, E., Limousin, M., et al. 2017, MNRAS, 470, 1809
Adams, F. C. 2010, ARA&A, 48, 47
Adams, F. C., Proszkow, E. M., Fatuzzo, M., & Myers, P. C. 2006, ApJ, 641, 504
Ahnen, M. L., Ansoldi, S., Antonelli, L. A., et al. 2016, A&A, 590, A24
Alpaslan, M., Robotham, A. S. G., Driver, S., et al. 2012, MNRAS, 426, 2832
Andrews, S. K., Driver, S. P., Davies, L. J. M., et al. 2017a, MNRAS, 464, 1569
Andrews, S. K., Driver, S. P., Davies, L. J. M., et al. 2017b, MNRAS, 470, 1342
Angus, G. W., & McGaugh, S. S. 2008, MNRAS, 383, 417
Arendt, R. G., Kashlinsky, A., Moseley, S. H., & Mather, J. 2016, ApJ, 824, 26
Ashcraft, T. A., Windhorst, R. A., Jansen, R. A., et al. 2017, PASP, submitted (arXiv:1703.09874)
Badenes, C., Mazzola, C., Thompson, T. A., et al. 2018, ApJ, in press (arXiv:1711.00660)
Bahcall, N. A., & Oh, S. P. 1996, ApJL, 462, L49
Barkana, R., & Loeb, A. 2001, PhR, 349, 125
Barkana, R., & Loeb, A. 2002, ApJ, 578, 1
Barkat, Z., Rakavy, G., & Sack, N. 1967, PhRvL, 18, 379
Bastian, N., Covey, K. R., & Meyer, M. R. 2010, ARA&A, 48, 339
Beichman, C. A., Rieke, M., Eisenstein, D., et al. 2012, Proc. SPIE, 8442, 84422N
Belczynski, K., Heger, A., Gladysz, W., et al. 2016, A&A, 594, A97
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bessell, M. S., Castelli, F., & Plez, B. 1998, A&A, 333, 231
Biteau, J., & Williams, D. A. 2015, ApJ, 812, 60
Blackburne, J. A., Pooley, D., Rappaport, S., & Schechter, P. L. 2011, ApJ, 729, 34
Bond, J. R., Arnett, W. D., & Carr, B. J. 1984, ApJ, 280, 825
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2017, ApJ, 843, 41
Bovill, M. S. 2016, Presentation at the October 2016 Montreal JWST Workshop, http://craq-astro.ca/jwst2016/agenda_en.php/
Bromm, V., Kudritzki, R. P., & Loeb, A. 2001, ApJ, 552, 464
Butler, N. R., & Bloom, J. S. 2011, AJ, 141, 93
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Caminha, G. B., Grillo, C., Rosati, P., et al. 2017, A&A, 600, A90
Cannizzo, J. K., Shafter, A. W., & Wheeler, J. C. 1988, ApJ, 333, 227
Cappelluti, N., Kashlinsky, A., Arendt, R. G., et al. 2013, ApJ, 769, 68
Cappelluti, N., Li, Y., Ricarte, A., et al. 2017, ApJ, 837, 19
Casagrande, L., Portinari, L., & Flynn, C. 2006, MNRAS, 373, 13
Castor, J. I., Abbott, D. C., & Klein, R. I. 1975, ApJ, 195, 157
Chatzopoulos, E., Wheeler, J. C., & Couch, S. M. 2013, ApJ, 776, 129
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Chornock, R., Berger, E., Kasen, D., et al. 2017, ApJL, 848, L19
Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJL, 648, L109
Cohen, S. H., Ryan, R. E., Jr., Straughn, A. N., et al. 2006, ApJ, 639, 731
Conroy, C. 2013, ARA&A, 51, 393
Cooray, A., Gong, Y., Smidt, J., & Santos, M. G. 2012, ApJ, 756, 92
Coulter, D. A., Lehmer, B. D., Eufrasio, R. T., et al. 2017, ApJ, 835, 183
Cowperthwaite, P. S., Berger, E., Villar, V. A., et al. 2017, ApJL, 848, L17
de Mink, S. E., & Mandel, I. 2016, MNRAS, 460, 3545
Diaferio, A. 1999, MNRAS, 309, 610
Diego, J. M., Broadhurst, T., Chen, C., et al. 2016a, MNRAS, 456, 356
Diego, J. M., Broadhurst, T., Molnar, S. M., Lam, D., & Lim, J. 2015a, MNRAS, 447, 3130
Diego, J. M., Broadhurst, T., Wong, J., et al. 2016b, MNRAS, 459, 3447
Diego, J. M., Broadhurst, T., Zitrin, A., et al. 2015b, MNRAS, 451, 3920

Diego, J. M., Kaiser, N., Broadhurst, T., et al. 2017, ApJ, submitted (arXiv:1706.10281)
Dressler, A. 1991, Natur, 350, 391
Driver, S. P., Andrews, S. K., Davies, L. J., et al. 2016, ApJ, 827, 108 (D16)
Duchêne, G., & Kraus, A. 2013, ARA&A, 51, 269
Dwek, E., & Krennrich, F. 2013, APh, 43, 112
Ebeling, H., Ma, C.-J., & Barrett, E. 2014, ApJS, 211, 21
Emilio, M., Kuhn, J. R., Bush, R. I., & Scholl, I. F. 2012, ApJ, 750, 135
Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, AJ, 122, 2833
Fan, X., Strauss, M. A., Schneider, D. P., et al. 2003, AJ, 125, 1649
Farmer, R., Fields, C. E., Petermann, I., et al. 2016, ApJS, 227, 22
Farmer, R., Fields, C. E., & Timmes, F. X. 2015, ApJ, 807, 184
Faulkner, J. 1967, ApJ, 147, 617
Fields, C. E., Farmer, R., Petermann, I., Iliadis, C., & Timmes, F. X. 2016, ApJ, 823, 46
Fields, C. E., Timmes, F. X., Farmer, R., et al. 2018, ApJ, in press (arXiv:1712.06057)
Finkelstein, S. L. 2016, PASA, 33, e037
Finkelstein, S. L., Ryan, R. E., Jr., Papovich, C., et al. 2015, ApJ, 810, 71
Fixsen, D. J. 2009, ApJ, 707, 916
Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, ApJ, 473, 576
Flower, P. J. 1996, ApJ, 469, 355
Fraley, G. S. 1968, Ap&SS, 2, 96
Frank, J., King, A., & Raine, D. J. 2002, Accretion Power in Astrophysics (Cambridge: Cambridge Univ. Press), 398
Fryer, C. L., Woosley, S. E., & Heger, A. 2001, ApJ, 550, 372
Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, SSRv, 123, 485
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJL, 600, L93
Götberg, Y., de Mink, S. E., & Groh, J. H. 2017, A&A, 608, A11
Greif, T. H., Springel, V., White, S. D. M., et al. 2011, ApJ, 737, 75
Griffiths, A., Conselice, C. J., Alpaslan, M., et al. 2018, MNRAS, 475, 2853
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Guszejnov, D., Krumholz, M. R., & Hopkins, P. F. 2016, MNRAS, 458, 673
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Hathi, N. P., Jansen, R. A., Windhorst, R. A., et al. 2008, AJ, 135, 156
Helgason, K., Ricotti, M., Kashlinsky, A., & Bromm, V. 2016, MNRAS, 455, 282
Henze, M., Ness, J.-U., Darnley, M. J., et al. 2015, A&A, 580, A46
H.E.S.S. Collaboration, Abdalla, H., Abramowski, A., et al. 2017, A&A, 606, A59
HESS Collaboration, Abramowski, A., Acero, F., et al. 2013, A&A, 550, A4
Hinshaw, G., Weiland, J. L., Hill, R. S., et al. 2009, ApJS, 180, 225
Hirschi, R. 2007, A&A, 461, 571
Hoffman, Y., Courtois, H. M., & Tully, R. B. 2015, MNRAS, 449, 4494
Hoffman, Y., Pomarède, D., Tully, R. B., & Courtois, H. M. 2017, NatAs, 1, 0036
Hogg, D. W. 1999, arXiv:astro-ph/9905116
Hogg, D. W., Baldry, I. K., Blanton, M. R., & Eisenstein, D. J. 2002, arXiv:astro-ph/0210394
Hosokawa, T., Hirano, S., Kuiper, R., et al. 2016, ApJ, 824, 119
Hoyle, F., & Lyttleton, R. A. 1942, MNRAS, 102, 177
Hunter, J. D. 2007, CSE, 9, 90
Ishiyama, T., Sudo, K., Yokoi, S., et al. 2016, ApJ, 826, 9
Jansen, R. A. & Webb Medium Deep Fields IDS GTO team 2017, in American Astronomical Society Meeting Abstracts 230, #216.02
Jauzac, M., Clément, B., Limousin, M., et al. 2014, MNRAS, 443, 1549
Jauzac, M., Richard, J., Jullo, E., et al. 2015, MNRAS, 452, 1437
Jiang, L., Fan, X., Vestergaard, M., et al. 2007, AJ, 134, 1150
Kashlinsky, A. 2016, ApJL, 823, L25
Kashlinsky, A., Arendt, R. G., Ashby, M. L. N., et al. 2012, ApJ, 753, 63
Kashlinsky, A., Mather, J. C., Helgason, K., et al. 2015, ApJ, 804, 99
Kawamata, R., Oguri, M., Ishigaki, M., Shimasaku, K., & Ouchi, M. 2016, ApJ, 819, 114
Kayser, R., Refsdal, S., & Stabell, R. 1986, A&A, 166, 36
Kelly, P. L., Diego, J. M., Nonino, M., et al. 2017, ATel, 10005
Kelly, P. L., Diego, J. M., Rodney, S., et al. 2018, NatAs, in press (arXiv:1706.10279)
Kelsall, T., Weiland, J. L., Franz, B. A., et al. 1998, ApJ, 508, 44
Kim, D., Jansen, R. A., & Windhorst, R. A. 2017, ApJ, 804, 28
Kiminki, D. C., & Kobulnicky, H. A. 2012, ApJ, 751, 4
Koekemoer, A. M., Ellis, R. S., McLure, R. J., et al. 2013, ApJS, 209, 3
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kohri, K., Nakama, T., & Suyama, T. 2014, PhRvD, 90, 083514
Kozłowski, S., Kochanek, C. S., Udalski, A., et al. 2010, ApJ, 708, 927

Kozyreva, A., & Blinnikov, S. 2015, MNRAS, 454, 4357
Kozyreva, A., Gilmer, M., Hirschi, R., et al. 2017, MNRAS, 464, 2854
Kurk, J. D., Walter, F., Fan, X., et al. 2007, ApJ, 669, 32
Kurucz, R. L. 2005, MSAIS, 8, 189
Lagattuta, D. J., Richard, J., Clément, B., et al. 2017, MNRAS, 469, 3946
Lam, D., Broadhurst, T., Diego, J. M., et al. 2014, ApJ, 797, 98
Lewis, G. F., Ibata, R. A., & Wyithe, J. S. B. 2000, ApJL, 542, L9
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, ApJ, 835, 113
Lorentz, M., Brun, P., & Sanchez, D. 2015, Proc. 34th ICRC (The Hague), 34, 777
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97
Machida, M. N., Omukai, K., Matsumoto, T., & Inutsuka, S.-I. 2009, MNRAS, 399, 1255
Macpherson, D., Coward, D. M., & Zadnik, M. G. 2013, ApJ, 779, 73
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Madau, P., & Fragos, T. 2017, ApJ, 840, 39
Madau, P., & Haardt, F. 2015, ApJL, 813, L8
Madau, P., & Silk, J. 2005, MNRAS, 359, L37
Mahler, G., Richard, J., Clément, B., et al. 2018, MNRAS, 473, 663
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Mamajek, E. E., Prsa, A., Torres, G., et al. 2015, arXiv:1510.07674
Mas-Ribas, L., Dijkstra, M., & Forero-Romero, J. E. 2016, ApJ, 833, 65
Matsuoka, Y., Ienaka, N., Kawara, K., & Oyabu, S. 2011, ApJ, 736, 119
Mattila, K., Väisänen, P., Lehtinen, K., von Appen-Schnur, G., & Leinert, C. 2017, MNRAS, 470, 2152
Mayer, P., Harmanec, P., Chini, R., et al. 2017, A&A, 600, A33
Meneghetti, M., Natarajan, P., Coe, D., et al. 2017, MNRAS, 472, 3177
Milosavljević, M., Bromm, V., Couch, S. M., & Oh, S. P. 2009, ApJ, 698, 766
Miralda-Escude, J. 1991, ApJ, 379, 94
Mitchell-Wynne, K., Cooray, A., Xue, Y., et al. 2016, ApJ, 832, 104
Molnar, S. M., Broadhurst, T., Umetsu, K., et al. 2013, ApJ, 774, 70
Morgan, R. J., Windhorst, R. A., Scannapieco, E., & Thacker, R. J. 2015, PASP, 127, 803
Morishita, T., Abramson, L. E., Treu, T., et al. 2017, ApJ, 846, 139
Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Natur, 474, 616
Natarajan, P., Chadayammuri, U., Jauzac, M., et al. 2017, MNRAS, 468, 1962
Negrello, M., Amber, S., Amvrosiadis, A., et al. 2017, MNRAS, 465, 3558
Oguri, M., Diego, J. M., Kaiser, N., Kelly, P. L., & Broadhurst, T. 2018, PhRvD, 97, 023518
Ohkubo, T., Nomoto, K., Umeda, H., Yoshida, N., & Tsuruta, S. 2009, ApJ, 706, 1184
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Owers, M. S., Randall, S. W., Nulsen, P. E. J., et al. 2011, ApJ, 728, 27
Pagel, B. E. J., & Portinari, L. 1998, MNRAS, 298, 747
Park, K., & Ricotti, M. 2012, ApJ, 747, 9
Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3
Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4
Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15
Planck Collaboration, Adam, R., Aghanim, N., et al. 2016a, A&A, 596, A108
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, A&A, 594, A13
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016c, A&A, 594, A2
Planck Collaboration, Aghanim, N., Armitage-Caplan, C., et al. 2014, A&A, 571, A27
Planck Collaboration, Aghanim, N., Ashdown, M., et al. 2016d, A&A, 596, A107
Portinari, L., Casagrande, L., & Flynn, C. 2010, MNRAS, 406, 1570
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Prša, A., Harmanec, P., Torres, G., et al. 2016, AJ, 152, 41
Remillard, R. A., & McClintock, J. E. 2006, ARA&A, 44, 49
Renzo, M., Ott, C. D., Shore, S. N., & de Mink, S. E. 2017, A&A, 603, A118
Rieke, M. J., Kelly, D., & Horner, S. 2005, Proc. SPIE, 5904, 1

Robotham, A. S. G., Norberg, P., Driver, S. P., et al. 2011, MNRAS, 416, 2640
Rodney, S. A., Balestra, I., Bradac, M., et al. 2017, NatAs, in press (arXiv:1707.02434)
Romero, A. D., Campos, F., & Kepler, S. O. 2015, MNRAS, 450, 3708
Rydberg, C.-E., Zackrisson, E., Lundqvist, P., & Scott, P. 2013, MNRAS, 429, 3658
Rydberg, C.-E., Zackrisson, E., Zitrin, A., et al. 2015, ApJ, 804, 13
Salpeter, E. E. 1955, ApJ, 121, 161
Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Sci, 337, 444
Sarmento, R., Scannapieco, E., & Cohen, S. 2018, ApJ, in press (arXiv:1710.09878)
Sarmento, R., Scannapieco, E., & Pan, L. 2017, ApJ, 834, 23
Scalo, J. M. 1986, FCPh, 11, 1
Schaerer, D. 2002, A&A, 382, 28
Shafter, A. W. 2017, ApJ, 834, 196
Shafter, A. W., Henze, M., Rector, T. A., et al. 2015, ApJS, 216, 34
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Shakura, N. I., & Sunyaev, R. A. 1976, MNRAS, 175, 613
Shara, M. M., Livio, M., Moffat, A. F. J., & Orio, M. 1986, ApJ, 311, 163
Smith, B. M., Windhorst, R. A., Jansen, R. A., et al. 2018, ApJ, 853, 191
Smith, N., Li, W., Foley, R. J., et al. 2007, ApJ, 666, 1116
Sobral, D., Matthee, J., Darvish, B., et al. 2015, ApJ, 808, 139
Springel, V., & Farrar, G. R. 2007, MNRAS, 380, 911
Stacy, A., Bromm, V., & Lee, A. T. 2016, MNRAS, 462, 1307
Stanway, E. R., Eldridge, J. J., & Becker, G. D. 2016, MNRAS, 456, 485
Sugimoto, D., & Nomoto, K. 1980, SSRv, 25, 155
Sukhbold, T., & Woosley, S. E. 2014, ApJ, 783, 10
Sukhbold, T., & Woosley, S. E. 2016, ApJL, 820, L38
Susa, H., Hasegawa, K., & Tominaga, N. 2014, ApJ, 792, 32
Tanaka, T., Perna, R., & Haiman, Z. 2012, MNRAS, 425, 2974
Tanaka, Y., & Shibazaki, N. 1996, ARA&A, 34, 607
Thompson, R., & Nagamine, K. 2012, MNRAS, 419, 3560
Trenti, M., & Stiavelli, M. 2007, ApJ, 667, 38
Trenti, M., & Stiavelli, M. 2009, ApJ, 694, 879
Trujillo, I., & Fliri, J. 2016, ApJ, 823, 123
Tucker, W., Blanco, P., Rappaport, S., et al. 1998, ApJL, 496, L5
Turk, M. J., Abel, T., & O'Shea, B. 2009, Sci, 325, 601
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Watkins, R., & Feldman, H. A. 2015a, MNRAS, 447, 132
Watkins, R., & Feldman, H. A. 2015b, MNRAS, 450, 1868
Watson, W. A., Iliev, I. T., Diego, J. M., et al. 2014, MNRAS, 437, 3776
Watts, A. L. 2012, ARA&A, 50, 609
Wheeler, J. C. 1977, Ap&SS, 50, 125
Willott, C. J., Albert, L., Arzoumanian, D., et al. 2010, AJ, 140, 546
Willott, C. J., McLure, R. J., & Jarvis, M. J. 2003, ApJL, 587, L15
Windhorst, R. A., Cohen, S. H., Hathi, N. P., et al. 2011, ApJS, 193, 27 (W11)
Windhorst, R. A., Hathi, N. P., Cohen, S. H., et al. 2008, AdSpR, 41, 1965
Wolf, W. M., Bildsten, L., Brooks, J., & Paxton, B. 2013, ApJ, 777, 136
Woosley, S. E. 2017, ApJ, 836, 244
Woosley, S. E., Heger, A., & Weaver, T. A. 2002, RvMP, 74, 1015
Yoon, S.-C., Cantiello, M., & Langer, N. 2008, in AIP Conf. Ser. 990, First Stars III, ed. B. W. O'Shea & A. Heger (Melville, NY: AIP), 225
Yue, B., Ferrara, A., Salvaterra, R., Xu, Y., & Chen, X. 2013, MNRAS, 433, 1556
Yusof, N., Hirschi, R., Meynet, G., et al. 2013, MNRAS, 433, 1114
Zackrisson, E., González, J., Eriksson, S., et al. 2015, MNRAS, 449, 3057
Zemcov, M., Immel, P., Nguyen, C., et al. 2017, NatCo, 8, 15003
Zhang, F., Han, Z., Li, L., Guo, J., & Zhang, Y. 2010, Ap&SS, 329, 249
Zitrin, A., Menanteau, F., Hughes, J. P., et al. 2013, ApJL, 770, L15