Title

Physically Motivated Sub-grid and Surrogate Models of Star Formation and Feedback in Eulerian Adaptive Mesh Cosmological Simulations

Permalink

https://escholarship.org/uc/item/7gq6x76z

Author

Wells, Azton Ian

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Physically Motivated Sub-grid and Surrogate Models of Star Formation and Feedback in Eulerian Adaptive Mesh Cosmological Simulations

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Physics

by

Azton Ian Wells

Committee in charge:

Professor Michael L. Norman, Chair
Professor Phillip Gill
Professor Michael Holst
Professor Dušan Kereš
Professor David Tytler

2022

The Dissertation of Azton Ian Wells is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

For my parents

For Jenny, whose steadfast support is an amazing source of comfort, inspiration, and motivation

For Liam and Desmond, its never too late to find the spark that drives you

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This work would not have been possible without the patience and adventurous spirit of Professor Michael Norman, who took on working with me despite my lack of astrophysics or computing knowledge, and then pursued the unknown in the application of deep learning to astrophysical simulations. I have immense gratitude to the `Enzo` and `YT` developer and user communities, which have been of constant help and enabled every step of this work. These are both collaborative open source codes representing efforts from many research groups, scientists, and students around the world.

Chapter 2, in full, is a formatted reprint of the material as it appears in Wells & Norman 2021, “Predicting Localized Primordial Star Formation with Deep Convolutional Neural Networks”. *The Astrophysical Journal Supplement Series*, 254:41. © American Astronomical society. Reproduced with permission. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a formatted reprint of the material as it appears in Wells & Norman 2022, “Connecting Primordial Star-forming Regions and Second-generation Star Formation in the Phoenix Simulations”. *The Astrophysical Journal*, 932:71. © American Astronomical society. Reproduced with permission. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is being prepared for submission for publication of the material as it may appear in Wells, A.I. & Norman, M.L. “The First Galaxies and the Effect of Heterogeneous Enrichment from Primordial Stars”. *The Astrophysical Journal* (2022). The dissertation author was the primary investigator and author of this work.

VITA

2002-2012	U.S. Air Force
2012	Associates of Science, Northwest Florida State College
2014-2015	Undergraduate Research Assistant, Dale Van Harlingen Laboratory
2015	Bachelor of Science, University of Illinois at Urbana-Champaign
2016-2022	Graduate Research Assistant, Laboratory for Computational Astrophysics
2022	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Hicks, Wells, Norman, Wise, Smith, & O'Shea (2021), "External Enrichment of Mini-halos by the First Supernovae", *The Astrophysical Journal*, 909:70

Wells & Norman (2021), "Predicting Localized Primordial Star Formation with Deep Convolutional Neural Networks". *The Astrophysical Journal Supplement Series*, 254:41

Wells & Norman (2022), "Connecting Primordial Star-forming Regions and Second-generation Star Formation in the Phoenix Simulations". *The Astrophysical Journal*, 932:71

Wells & Norman (2022), "The First Galaxies and the Effect of Heterogeneous Enrichment from Primordial Stars", *in preparation*

ABSTRACT OF THE DISSERTATION

Physically Motivated Sub-grid and Surrogate Models of Star Formation and Feedback in Eulerian Adaptive Mesh Cosmological Simulations

by

Azton Ian Wells

Doctor of Philosophy in Physics

University of California San Diego, 2022

Professor Michael L. Norman, Chair

Star formation and feedback in astrophysical simulations remains a longstanding challenge when attempting to model large regions of the universe. The small scales of star formation limit the computational volume of self-consistent simulations; however, many recent works have highlighted the importance of accurate star formation and feedback to match simulations with observations. Unfortunately, the first generation of stars is often neglected in modern simulations, despite the fact that these stars generate the first metal-rich regions that fuel all subsequent star formation. This work represents the first effort to model the effect of these primordial stars without the extreme resolution requirements of prior generation simulations. We find that specialized deep convolutional neural network architectures are competent at identifying primordial star formation in under-resolved simulations, predicting star forming regions that are matched well by resolved simulations. Based on our study of the *Phoenix Simulations*, we find that

primordial star forming regions have a large, but limited influence. We generate an interpretable linear regression model to predict the size of this region based on the number, masses, and ages of stars within the primordial population.

Finally, we combine the prior works to predict primordial star formation and feedback in cosmological simulations and compare the new framework to literature-standard simulations that employ a metallicity floor. We analyze the impact of heterogeneous metal enrichment by studying the protogalaxies ($10^6 \lesssim M_v/M_\odot \lesssim 10^8$) and their stellar populations. We find that ignoring metallicity requirements for enriched star formation results in a up to 30% excess in stellar mass created. Further, using a metallicity floor causes an early underproduction of stars before $z = 21$ that reverses to overproduction by $z = 18$, creating $\sim 20\%$ excess stellar mass and 8.6% excess in protogalaxy number by $z = 14.95$. Heterogeneous metallicity conditions greatly increase the range of halo observables, e.g., stellar metallicity, stellar mass, and absolute magnitude. The increased range leads to better agreement with observations of ultra-faint dwarf galaxies when compared to metallicity-floor simulations. `StarNet` generates protogalaxies with $M_* \lesssim 10^3 M_\odot$, so it may additionally model low-luminosity protogalaxies more effectively than a metallicity floor criterion.

Chapter 1

General Introduction

Despite the rise of petascale computing and the recent dawn of exascale computer systems, astrophysical simulations continue to push the limits of the most advanced supercomputers. While dark matter (DM) or hydrodynamic DM + gas simulations can now simulate massive volumes (e.g., Vogelsberger et al., 2014; Emberson et al., 2019), the inclusion of more complete physical models, such as resolved star formation and feedback (SFF) processes, places severe restrictions on the size of simulations (Hopkins et al., 2018; Smith et al., 2015). This, of course, is detrimental to simulations that would aim to model large swathes of the observable universe. Further complicating the matter, advances in observational astronomy combine modern telescopes with data analysis techniques to produce very precise results. The astrophysical and cosmological simulation communities can no longer rely on DM only simulations to reproduce observables, as the lower-order effects of baryons and galactic processes are measurable in observations. Further, decent models of SFF contribute to deducing unobservables, e.g., the primordial stellar initial mass function (IMF) (Welsh et al., 2019; Cooke et al., 2017; Welsh et al., 2020; Vogelsberger et al., 2020).

Much to the chagrin of every astrophysicist, modelling large, statistically significant and representative volumes with accurate models of SFF is currently beyond the computing capabilities of even next-generation exascale systems. In our collective surrender to technological limitations, many compromises are made in the name of modelling stellar and galactic systems "well enough." Some involve using galaxy particles, where the star formation rate (SFR) is tuned to observable galaxies, e.g., the mass-loading factor of galactic winds is an user-defined parameter requiring tuning to observations (Oppenheimer & Davé, 2006; Vogelsberger et al., 2013; Vogelsberger et al., 2014). Although matching observations with simulations is important, tuning parameters to match observables in the modern universe offers no assurance that the parameter is correct for earlier times, particularly those times before reionization that severely lack observational data.

In other cases, star clusters are modelled instead of single stars (Hopkins et al., 2018; Wise et al., 2012; Wise et al., 2012). In these cases, special care must be paid to how the star cluster interacts with its environment through, e.g., stellar winds, supernovae, and radiation processes (e.g., Rosdahl et al., 2016).

Largely ignored in most simulations, there is another initial condition besides the DM and baryonic density perturbations that sets the stage for later star formation and is necessary to begin the formation of the first galaxies: metals from Population III stars. These first stars form from primordial gas consisting of only H, He, and H_2 (for a review, see Bromm, 2013). They are also thought to be massive, and while their IMF is not well known, modern evidence (Welsh et al., 2019, 2020; Cooke et al., 2017) indicates that the IMF almost certainly includes masses that would have led to supernovae of various types (Heger & Woosley, 2002; Nomoto et al., 2006; Woosley & Heger, 2015). These first Population III supernovas (SNe) were the first source of metals–those elements heavier than helium–which are efficient coolants that enable protostellar cloud collapse to begin in ways that resemble modern star formation (Larson, 2003).

Modelling these first stars directly and generating the metallicity initial conditions of a simulation *in situ* is extremely costly. The molecular H_2 clouds that contribute to the formation of Population III stars measure only tens of parsecs with proto-stellar cores on the scale of astronomical units (Abel et al., 2000; Bromm et al., 2002), whereas statistically useful regions of the universe, i.e., regions much larger than the Lagrangian region of the Milky Way, measure volumes by gigaparsecs. Therefore, when attempting precision modelling of the Population III era with associated SFF, one must choose whether to simulate the Population III SFF explicitly or to adopt a simplification scheme to compensate for the lack thereof, such as adopting a metallicity floor that assumes a set metallicity for all points in space. The course of this dissertation began at recognizing the limitation of current simulation paradigms by studying the first stars in Wise et al. 2012 and Xu et al. 2016. These works explicitly model the Population III era, but due to high resolution and ray-tracing radiation algorithms, could only achieve $z \sim 7$ in relatively small $1-300$ Mpc^3 volumes. We additionally examined modern works that ignore the Population III era, and note that some disagreement could be caused by the metallicity floor simplification (e.g., Regan et al. 2017; Wheeler et al. 2019). What is needed is a method that can generate a heterogeneous metallicity initial condition that reflects the actual formation and impact of Population III stars without the severe resolution and timestep requirements of modeling primordial SFF explicitly.

Thesis Outline

The remainder of this dissertation will address the core topic of SFF in astrophysical simulations via the following format. The remainder of the introduction outlines common tools used throughout this work, including `Enzo` and a short primer on deep learning methods. The first element of our endeavor to develop a surrogate model of SFF is presented in Chapter 2. There, we study deep convolutional neural networks (DCNNs) and their application to predicting localized star formation in simulations, culminating in a model to identify which grid cells of a simulation are have the requisite conditions to host primordial star formation. Chapter 3 presents the *Phoenix Simulations* that were used to develop training data for our surrogate models. With the *Phoenix*, we also develop an interpretable linear regression model of the feedback region from primordial stars and present new findings regarding the formation of Population II.1 star clusters and subsequent other early enriched star formation (Population II.2) cluster formation. Chapter 4 presents multiple efforts: first, creating a physically motivated, resolution intelligent, SFF routine, the **S**cale-intelligent **T**erminal-momentum **A**lgorithm for **R**ealistic **S**tellar **S**ources (`STARSS`) to simulate SFF from enriched stars. Second, using `STARSS` with the combined models of Chapters 2 and 3 to create a DCNN-based surrogate model of the effect of Population III stars. The resulting `StarNet` simulations are used to study the impact of heterogeneous metallicity initial conditions on protogalaxies and mini-halos beyond their Population III star-forming phase.

The codebase to develop and use `StarFind` and the StarNetRuntime developed in Chapters 2 and 4 is composed of approximately 40,000 lines of Python[1]. To ease an interested reader into the code, descriptions of these frameworks along with helpful guides are available in the appendices. There, I also include initial conditions and parameter files to replicate the simulations in this work.

---

[1] `StarFind` development: https://github.com/azton/StarNet
StarNetRuntime: https://github.com/azton/StarNetRuntime

1.1 `Enzo` Simulations

Simulations in this body of work were performed with `Enzo` (Bryan et al., 2014; Brummel-Smith et al., 2019a). `Enzo` is a public, community maintained Eulerian grid-based hydrodynamic simulation code with adaptive mesh refinement (AMR). It has been used for broad ranging studies including high-resolution, high-redshift galaxy formation (Wise & Abel, 2011; Wise et al., 2012; Wise et al., 2012), single dwarf galaxies (Emerick et al., 2019), cosmological re-ionization era zoom-in simulations (Xu et al., 2016), and solar mass resolution simulations of the first primordial stars (Smith et al., 2015). In this section, we will outline the functionality and methods that `Enzo` uses.

Although many hydrodynamic methods have been implemented in `Enzo` since its inception, this project exclusively uses the piecewise parabolic method (PPM) (Colella & Woodward, 1984). Each timestep in `Enzo` solves the Eulerian equations of ideal hydrodynamics including gravity in a comoving coordinate system. `Enzo` solves conservation of mass

$$\frac{\partial \rho}{\partial t} + \frac{1}{a}\nabla \cdot (\rho \mathbf{v}) = 0 \tag{1.1}$$

with scale factor a, density ρ, and velocity $\boldsymbol{v}$; conservation of momentum

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \frac{1}{a}\nabla \cdot \left(\rho \mathbf{v}\mathbf{v} + \mathbf{I}p^*\right) = -\frac{\dot{a}}{a}\rho \mathbf{v} - \frac{1}{a}\rho \nabla \phi \tag{1.2}$$

with isotropic pressure p^*, identity matrix $\boldsymbol{I}$, and gravitational potential ϕ; conservation of total (kinetic + thermal) energy

$$\frac{\partial E}{\partial t} + \frac{1}{a}\nabla \cdot \left[(E + p^*)\mathbf{v}\right] = -\frac{\dot{a}}{a}\left(2E\right) - \frac{\rho}{a}\mathbf{v} \cdot \nabla \phi - \Lambda + \Gamma \tag{1.3}$$

with energy E, cooling rate Λ, and heating rate Γ; gravitational potential ϕ as

$$\nabla^2 \phi = \frac{4\pi G}{a}\left(\sum \rho - \langle \rho \rangle\right) \tag{1.4}$$

with total (dark matter + baryonic) density $\sum \rho$ and cosmic mean density $\langle \rho \rangle$; metal from stars and supernovae are evolved according to

$$\frac{\partial \rho_z}{\partial t} + \frac{1}{a}\nabla \cdot (\rho_z \mathbf{v}) = \dot{\rho}_{\mathbf{z,sf}} \tag{1.5}$$

with metals from stellar sources $\rho_{z,sf}$. The final equation to close the set is the equation of state

$$e = \frac{p}{\gamma - 1} \tag{1.6}$$

with ideal gas energy e, pressure p, and ratio of specific heats γ. `Enzo` additionally ensures mass conservation of chemical species and their reactions for species i and comoving density n_i by evolving

$$\frac{\partial n_i}{\partial t} + \frac{1}{a} \cdot (n_i \mathbf{v}) = \sum_j k_{ij}(T) n_i n_j + \sum_j \Gamma_j^{\mathrm{ph}} n_j \tag{1.7}$$

where $k_{ij}(T)$ are temperature dependent two-body rate coefficients, and Γ_j^{ph} are destruction and creation coefficients due to photoionization or photodissociation. The simulations in this body of work include both a redshift dependent uniform Lyman-Werner (LW) H_2 dissociating radiation background as well as ionizing radiation from point stellar sources. The radiative transfer equation to evolve these fields in comoving coordinates (Gnedin & Ostriker, 1997) is given by

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \frac{a_{em}}{a}\hat{n} \cdot \nabla I_\nu - \frac{H}{c}\left(\nu \frac{\partial I_\nu}{\partial \nu} - 3 I_\nu\right) = -\kappa_\nu I_\nu + j_\nu. \tag{1.8}$$

$I_\nu \equiv I(\nu, \boldsymbol{x}, \Omega, t)$ is the specific intensity of radiation with dimensions of energy per time (t) per solid angle (Ω) per frequency (ν). The second term, (a_{em}/a) accounts for cosmic expansion since radiation emission, the third term describes dilution of radiation and cosmic expansion, and the right-hand side describes absorption (κ_ν) and emission (j_ν) from all sources. The photoionization and photoheating coefficients (Wise & Abel, 2011) are given by

$$k_{\mathrm{ph}} = \frac{P(1 - \mathrm{e}^{\tau})}{n_{\mathrm{abs}} V_{\mathrm{cell}} \mathrm{d}t_P} \tag{1.9}$$

$$\Gamma_{\mathrm{ph}} = \kappa(E_{\mathrm{ph}} - E_i) \tag{1.10}$$

where in Equation 1.9, $P, \tau, V_{\mathrm{cell}}, n_{\mathrm{abs}}$ are the photon number, optical depth, cell volume, and number density of the absorbing medium respectively, and in 1.10, E_{ph} and E_i are the energies of the photon and ionization. The radiation transport from point sources is handled by the `Enzo+Moray` ray-tracing method (Wise & Abel, 2011). Chemistry is evolved by `Enzo` using Cloudy lookup tables (Smith et al., 2008; Ferland et al., 2017) and can include cooling rates that consider density, electron density, metal density, and temperature. Dark matter dynamics in `Enzo` are calculated using the particle-mesh method (Hockney & Eastwood, 1988; Bryan et al., 2014). In this method, the gravitational potential, ϕ, is calculated via fast fourier transforms. The resulting accelerations from the potential are then interpolated back to the dark matter particles, and used to determine hydrodynamic states for that time step.

1.2 Star Formation Algorithms

1.2.1 Primordial Stars

Population III stars are the first generation of stars (see Bromm & Larson, 2004, for a review). They are distinct from early enriched stars (Population II) or modern star formation (Population I) in that they form from primordial gas that matches the composition of early universe. Because of this, there are no appreciable quantities of metals[2], and the collapsing gas that forms the star is cooled via the rotational, vibrational and electronic transitions of H_2 (Bromm, 2013). Since H_2 is a less effective coolant than metals (Draine, 2011), e.g., oxygen or carbon, there is some question how the collapse of a gaseous cloud would proceed. In particular, how much the cloud fragments will determine the mass of the stars that can form from that cloud. Although this is still an open area of research, modern work suggests that the Population III IMF is much more top-heavy than the Population I/II IMF (Stacy et al., 2010, 2016; Ishigaki et al., 2018; Park et al., 2021; Latif et al., 2022): the average mass of Population III stars is thought to be 10-20 $M_\odot$. In addition, Population III stars have a potentially very large upper mass cutoff, with some studies focusing on $M_* \lesssim 300$ $M_\odot$ (Heger & Woosley, 2002), and others examining the possibility that they are the source of supermassive black hole (SMBH) seeds having masses $M_* \gtrsim 10^5$ $M_\odot$ (Schleicher et al., 2013). The less massive, $M_* \lesssim 300$ $M_\odot$ option is used in `Enzo`, and leads to multiple endpoints of the stellar lifetime. While lower-mass 10-40 $M_\odot$ Population III stars generate Type II SN or hypernova (HN), massive stars with $140 < M_*/M_\odot < 260$ result in pair-instability supernova (PISN) (Heger et al., 2003). Other stellar masses collapse as collisionless black holes. The primordial SNe serve as the source for metal for the next generation of stars, and determines the metallicity distribution function (MDF) of those second-generation stars. In `Enzo`, Population III stars are modeled as single stars, not clusters: each particle has an assigned mass, luminosity, metal yield, and explosion energy. The specifics of Population III SFF is detailed in Chapter 3; here, it is simply noted that the star formation criteria are resolution dependent, predominantly because it requires a specific H_2 fraction and

---

[2] Elements heavier than He

baryon number density for star formation to proceed. More importantly, the SN feedback from primordial stars is strongly resolution dependent. The foundation of the feedback model is to resolve the Sedov-Taylor (ST) blastwave of the supernova remnant (SNR) (Sedov, 1946; Taylor, 1950) and deposit the energy of the remnant in thermal form, neglecting all kinetic energy or momenta that may exist in the remnant at that time. Crucially, deposition of energy in thermal form has strict resolution requirements, as low resolution results in diffusing hot gas instead of generating the SNR shockwave that is desired (Katz, 1992; Abadi et al., 2003; Slyz et al., 2005; Kimm & Cen, 2014) (the "overcooling" problem). The strict resolution requirements of primordial SFF is a primary motivator of the work to develop a surrogate method to generate a heterogeneous metallicity field from Population III stars.

1.2.2 Enriched Star Clusters

Star formation and feedback of Population II stars is modeled in `Enzo` by particles that represent a star cluster. Their formation criteria is similar to Population III stars, except require finite, user-defined metallicity, and do not require H_2. The mass of the cluster is derived from the computational grid as 7% of the gas mass with number density $n_b > 1000\ \mathrm{cm}^{-3}$. This is a resolution dependent criterion, as lower resolution meshes will struggle to have substantial gas mass at that density. Similar to the Population III SN feedback, Population II SN are modelled by depositing energy in thermal form in a sphere that approximates the ST solution. This, of course, has the same pitfall as thermal energy deposition in the Population III case; however, the problem is exacerbated for Population II stars because the energy of deposition is much lower and therefore even more susceptible to overcooling. The clusters also model their feedback via continuous injection, i.e., the rate of energy, mass and metals returned to the grid is fixed and does not represent individual SNe events. This method is known to be problematic at low and middling resolution (e.g., Hopkins et al., 2018). The shortcomings of the Population II model is the primary motivator for development of the physically motivated `STARSS` algorithm, detailed in Chapter 4.

1.3 Methods: Machine Learning and Deep Learning

Machine learning (ML) and deep learning (DL) form a broad category of statistical models to represent underlying data distributions. The *raison d'être* for these unique forms of statistical model is to have parameters that are not explicitly chosen, but found by iterative optimization, which is of particular use when the underlying data distribution is too complex to be represented by only a handful of parameters. This section will provide a very brief overview of the basics of DL; extremely good and comprehensive sources are available freely on the web (e.g., http://d2l.ai/). ML and DL are entirely built on the presumption that predictions can be made by combining input features with a weight-factor; the output of the model, $\hat{Y}$, is given by the multiplication of input values(X) and weights (w): $\hat{Y} = X \times w + b$ with a bias, b. If we consider that X is not a scalar quantity, but, for example, a data point with multiple features, then the rows of $\boldsymbol{X}$ represent individual datum and each column is an individual feature. We can then recast the equation for $\hat{\boldsymbol{Y}}$ in vector form as

$$\hat{\boldsymbol{Y}} = \boldsymbol{X} \cdot \boldsymbol{w} + b. \tag{1.11}$$

Equation 1.11 results in only linear behaviour. To create non-linearity, we include an *activation function* so that $\hat{\boldsymbol{Y}} = F(\boldsymbol{X} \cdot \boldsymbol{w} + b)$. There are numerous activations that one could use, however, in this work we typically use the sigmoid ($\sigma(x)$) and leaky $\mathrm{ReLU}(x)$ activation functions defined as

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \tag{1.12}$$

$$\mathrm{ReLU}(x) = \begin{cases} lx & x < 0 \\ x & x > 0 \end{cases} \tag{1.13}$$

with a leaky parameter l, usually having $0.0 < l < 1.0$ where $l = 0$ recovers the standard $\mathrm{ReLU}(x)$ activation. Other common activation functions include standard ReLU, hyperbolic tangent, and softmax. Layers of a DNN in which each feature of $\boldsymbol{X}$ is multiplied by a unique entry of $\boldsymbol{w}$ are referred to as fully connected layers (FCLs). An example of a simple DNN with FCL is presented in Figure 1.1. This simple model has two inputs, x_i, two hidden layers, a, b, with one output, Y. This small DNN is the foundation of how the other models in this work function, so we present it as a minimal worked example of making predictions ("forward

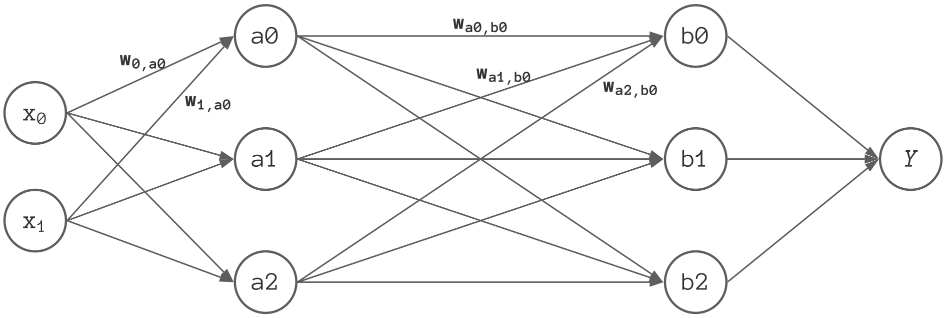


Figure 1.1. A simple DNN

pass") and adjusting weights to improve future predictions ("backward pass"). The forward pass consists of a series of matrix multiplications progressing through the model. As a concrete example, consider a forward pass of the model in Figure 1.1. The inputs, x_i are multiplied by the weights, $w_{i,a0}$, and then activated to give the value at the node a_0[3]. At layer a with ReLU activation,

$$a_i = \mathrm{ReLU}\big(\sum_j x_j w_{j,ai}\big). \tag{1.14}$$

There is a similar expression for the values in layer b:

$$b_i = \mathrm{ReLU}\big(\sum_j a_j w_{j,bi}\big). \tag{1.15}$$

And, the final prediction is the output at Y (assuming a $\sigma(x)$ activation function):

$$\hat{Y} = \sigma(\sum_j b_j w_j) \tag{1.16}$$

With a prediction from the model, we now need to judge its quality. This is done in the form of *loss functions* or *objective functions*. For classification problems, a common objective function (and used in this work) is cross-entropy:

$$L(\hat{y}, y, c) = \left\{ w(c) \left[-y_c \hat{y}_c + \log\left(\sum_j \exp(\hat{y}_j) \right) \right] \right\}, \tag{1.17}$$

$$L = \frac{\sum_c L(\hat{y}, y, c)}{\sum_c w(c)} \tag{1.18}$$

---

[3] Generally, there is a bias term added to the multiplication, which is set to zero in this section for clarity.

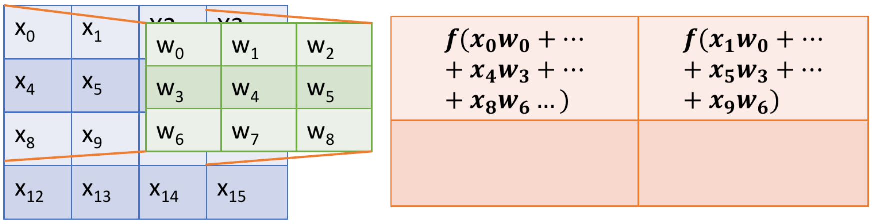


Figure 1.2. Convolutional kernel example.

where $w(c)$ represents weights given to the c^{th} class, the network output ($\hat{y}_j$) and true label (y_j). In regression problems, distance functions are commonly used, such as Manhattan distance, $L_1 = |y - \hat{y}|$ or L_2 distance, $L_2 = (y - \hat{y})^2$, where the final value is the mean over all predictions in the sample (e.g., when predicting multiple voxel values, the mean L for all voxels).

Finally, with the loss L in hand, we perform back propagation to improve the weights of the model. Ideally, we want to minimize L by solving an equation such as $dL/d\boldsymbol{w} = 0$, however with many weight/input configurations, this is not analytically tractable. Instead, we will adjust the weights of the model as directed by the gradient of the loss with respect to the weights. Specifically, we would adjust the weights by δw determined by:

$$\delta w_i = \varepsilon \frac{\partial L}{\partial w_i} \tag{1.19}$$

where i refers to the specific weight we would like to change and ε is the learning rate. To find the weight update in layer a or other interior layer weights, we would apply the chain rule using the expressions for outputs in Equations 1.14 and 1.15. In this work, we heavily rely on convolutional filters to process 3D data. Shown for the 2D case in Figure 1.2, the $\boldsymbol{w}$ are a N^3 convolutional kernel, K, with learnable weights. The kernel is passed over the input data, skipping stride, s, entries between multiplications. As with FCL, the final output from a convolutional layer is $\hat{Y} = f(K(\boldsymbol{X}) + b)$, where the bias is optional depending on the normalization chosen, with activation function $f(x)$.

There are a great number of modifications that can be made to the basic structure outlined in this section. When used, these more unique architectures will be discussed in their relevant chapters, e.g., the Inception and Dense modules of Chapter 2. In addition, there are very complicated designs that are barely recognizable from this base structure, such as the dueling generators and discriminators of generative adversarial networks (GANs) (Goodfellow et al., 2014) and time-evolution prediction using long-short term memory networks (LSTMs) (Hochreiter & Schmidhuber, 1997).

Chapter 2

Predicting Localized Primordial Star Formation With Deep Convolutional Neural Networks

2.1 Abstract

We investigate applying 3D deep convolutional neural networks as fast surrogate models of the formation and feedback effects of primordial stars in hydrodynamic cosmological simulations of the first galaxies. Here, we present the surrogate model to predict localized primordial star formation; the feedback model will be presented in a subsequent paper. The star formation prediction model consists of two sub-models: the first is a 3D volume classifier that predicts which (10 comoving kpc)^3 volumes will host star formation, followed by a 3D Inception-based U-net voxel segmentation model that predicts which voxels will form primordial stars. We find that the combined model predicts primordial star forming volumes with high skill, with $F_1 > 0.995$ and true skill score > 0.994. The star formation is localized within the volume to $\lesssim 5^3$ voxels (~ 1.6 comoving kpc^3) with $F_1 > 0.399$ and true skill score > 0.857. Applied to simulations with low spatial resolution, the model predicts star forming regions in the same locations and at similar redshifts as sites in resolved full-physics simulations that explicitly model primordial star formation and feedback. When applied to simulations with lower mass resolution, we find that the model predicts star forming regions at later redshift due to delayed structure formation resulting from lower mass resolution. Our model predicts primordial star formation without halo finding, so will be useful in spatially under-resolved simulations that cannot resolve primordial star forming halos. To our knowledge, this is the first model that can predict primordial star forming regions that match highly-resolved cosmological simulations.

2.2 Introduction

Despite the rise of petascale computing, astrophysical simulations continue to push the limits of the most advanced supercomputers. While dark matter (DM) or hydrodynamic gas-only simulations can now simulate massive volumes (e.g., Vogelsberger et al. 2014), the inclusion of more complete physical models, such as resolved star formation and feedback (SFF) processes, severely limits the volume of feasible simulations (Hopkins et al., 2018; Smith et al., 2015). When attempting precision modelling of high redshift galaxy formation with associated SFF, one must usually choose whether to simulate the primordial star (Population III) formation era, or to adopt a simplification scheme to compensate for the lack thereof. Here, we investigate a third option: use deep learning to train a surrogate model with data from simulations which resolve the primordial star formation and feedback processes directly. Then, use the surrogate model in inference to predict the coarse-grained effects of primordial stellar feedback in a large volume cosmological simulation. With SFF modeled rather than simulated, spatial resolution and timestepping requirements are greatly relaxed, thus accelerating the time to solution without sacrificing essential feedback effects such as chemical enrichment by primordial supernovae. In the following, we motivate this approach by reviewing primordial SFF as well as recent applications of deep learning to astrophysical applications.

Starting at $z \simeq 30$, Population III stars begin to form from pristine (H, He, H_2) gas in mini-halos with virial mass $M_{vir} \gtrsim 10^{5-6.5} M_\odot$ (Bromm, 2013). After formation, Population III stars may directly collapse to black holes (BHs), or if formed in the right mass range, live out a main sequence followed by a supernova (SN) of some type (Woosley & Heger, 2015). The SNe considered in this work fall into three categories determined by the stellar mass of their progenitor: Type-II supernovae (SNe), hypernovae (HNe), and pair-instability supernovae (PISNe). Depending on the mass of the star and mass of its host halo, the SN may completely disrupt the halo, ejecting the majority of gas and metal (Whalen et al., 2008b), effectively preventing any further star formation until the gas has recycled back into the halo and cooled (Tumlinson et al., 2017).

Even if a Population III star is outside the mass range for SN, it will still emit ionizing radiation for its main sequence lifetime before collapsing to a BH. Such radiative feedback has been shown to shut down continuing star formation in its vicinity (Whalen et al., 2008a; Wise et al., 2012; Hopkins et al., 2019), limiting the conversion of gas to stars in star forming regions. Unfortunately, having the extremely high resolution ($\lesssim$ 20 pccm[1], $M_{DM} \lesssim 10^4$ $M_\odot$) required to precisely model Population III SFF means that simulations of large (i.e., statistically relevant to the observable universe) volumes have not been able to approach modern redshifts (Wise et al., 2012; Xu et al., 2016; Smith et al., 2015; Hopkins et al., 2019). To avoid the computational expense of precisely modelling the Population III era, some practitioners adopt a metallicity floor (Hopkins et al., 2018), while others altogether disregard the effect of Pop III pre-enrichment on star formation (Vogelsberger et al., 2014). Neither of these simplifications account for the fact that enrichment by Population III stars is (a) non-uniformly distributed in space, (b) rare, and (c) necessary for enriched star formation. Multiple recent works have highlighted the need for a more intelligent model for the Population III era: Jeon et al. 2017 and Hicks et al. 2021 have shown that extremely low metallicity stars may form in halos that have been enriched by an external Pop III SN event, which represents a sequence of star formation that is impossible to model using the above simplifications. The lack of precision modelling of the Population III era is additionally cited as a potential cause for mismatches between simulations and observations of dwarf galaxies presented in Wheeler et al. 2019. Regan et al. 2017 attempted to determine how super-massive black holes could form from primordial halos. They determine that the environment of a protogalaxy must avoid Population III star formation to enable the collapse of the protogalaxy into a black hole–simplistic models that assume a metallicity floor or ignore the metal contribution to enriched star formation will never be able to capture such a phenomenon. These prior works and the limitations of currently used simplifications both highlight the need for a model that can bridge the gap between precision full-physics simulations and the current methods that ignore the “initial conditions” of metallicity in the universe.

While artificial neural networks (ANN), and particularly deep convolutional neural networks (DCNN), have been used for image recognition for nearly a decade since AlexNet (Krizhevsky et al., 2012) was used to classify images in the ImageNet dataset, they have also begun to foray into varied scientific applications, e.g., classifying observational images of radio galaxies (Aniyan & Thorat, 2017), predicting hydrodynamics quantities and modelling turbulence (Jin et al., 2018; Mohan et al., 2019), and determining the halo occupation distribution of galaxies from a dark matter distribution (Zhang et al., 2019). Another class of emerging scientific ANN are emulators–models that supplant all or part of the application, of which, one generalized approach (DENSE) has been shown to accelerate some computations by up to $10^9\times$ (Kasim et al., 2020).

---

[1] Comoving units have -cm appended to the base unit throughout this paper; the base unit is assumed to be proper.

Table 2.1. Source simulations for train/validation/test data.

Name	SFF	dx_{root}[kpccm]	dx_{min} [pccm]	L_{max}	M_{DM}[$\mathrm{M}_\odot$]	z_{final}	Cosmology
PHX256-1	w2012	10.183	19.89	9	2.384×10^4	13.57	$\Omega_\lambda = 0.6889$
PHX256-2	w2012	10.183	19.89	9	2.384×10^4	13.86	$\Omega_m = 0.3111$
P3N-128	None	20.366	156.8	7	1.910×10^5	10.0	$\sigma_8 = 0.811$
PHX256-HYD	None	10.183	156.8	6	2.384×10^4	18.05	$n = 0.968$

Note: The width of the root grid cell, dx_{root}, width of a cell at the most refined level, dx_{min}, dark matter particle mass (M_{DM}) and maximum AMR level L_{max} are shown for each simulation. All analysis in this work used data from these simulations up to the final redshift z_{final}. SFF denotes whether the simulation includes star formation (w2012 model from Wise et al. 2012) or not (None). The final column shows the cosmological parameters common to all simulations (The Planck Collaboration et al., 2014).

The DENSE model shows great promise in scientific application, but is limited to scalar inputs; a model that evaluates the hydrodynamic state of a running simulation would require a different design paradigm. Motivated by the monumental advancements in the field, this work presents a novel surrogate model for pre-enrichment using trained DCNN to identify primoridal star forming regions. This work will be coupled with a following work to predict Population III stellar feedback influence, where the entire model will function without resorting to halo finding and relax the severe resolution requirements of modelling the Population III era, accelerating the precision modelling of cosmological simulations.

Precisely predicting the location of star formation is a two-fold task: a small 3D binary classifier based on seminal image classification architectures (He et al., 2015; Huang et al., 2016; Szegedy et al., 2014) will quickly analyze potential 3D regions; if one is classified as star forming, an inception based (Szegedy et al., 2014) U-Net (Ronneberger et al., 2015; Zhang et al., 2020) adapted to 3D will predict star forming voxels. The rest of this paper outlines the development of the first phase of this feedback algorithm, to identify star forming regions, as follows: In Section 2.3 we outline the simulations and methods used to generate training data; Section 2.4 presents the network architectures used and discusses the parameter exploration that led to our design; we test the DCNN designs and present the results in Section 2.5, discuss the results in depth in Section 2.6, and summarize our findings in Section 2.7.

2.3 Simulations and Data

2.3.1 PHX256 Simulations

Training neural networks depends on copious amounts of data. To that end, we used the astrophysical adaptive mesh refinement (AMR) simulation code `Enzo` (Bryan et al., 2014; Brummel-Smith et al., 2019b) to produce two simulations (PHX256-1 and PHX256-2) from which to draw training, validation, and test data. Table 2.1 summarizes the simulations used within this work. The PHX256-1,2 simulations are modeled after the Renaissance Simulations (Xu et al., 2016) in terms of included physics, parameters, and resolution. Both simulations have identical Planck 2014 cosmological parameters: $\Omega_\lambda = 0.6889, \Omega_m = 0.3111, \Omega_b = 0.04898, \sigma_8 = 0.811, n = 0.965$ (The Planck Collaboration et al., 2014), but use different random seeds in the initial conditions, generated using `MUSIC` (Hahn & Abel, 2011). While the Renaissance Simulations used a zoom-in simulation setup with a hierarchy of static nested grids, we simulate a periodic box of size comparable to the finest nested grid in the former, at identical mass and spatial resolution. The simulation volume for both PHX256 simulations is 2.61^3 (Mpc)^3 in extent, with 256^3 root grid cells and dark matter (DM) particles and 9 levels AMR; the cell width at the deepest AMR level is 19 pc. With the given cosmology, the DM particles have mass 2.38×10^4 M$_\odot$ (M_{dm}), and the baryon mass (M_b) in a root grid cell of average density is 1.17×10^3 M$_\odot$.

Refinement of the AMR grid occurs based on M_b and M_{dm}, the baryonic and dark matter mass of a cell, respectively, with the minimum mass for refinement being 3 times the initial values above (M_{min}). At each iteration, any grid cell on level l with $M_{cell} \geq M_{min} \times 2^{-0.4l}$ will be refined to the next level; the grid refinement is super-lagrangian, so that the AMR levels will have $M_{cell} < M_{root\ grid}$. In addition to the mass criteria, any cells with Population III star particles are refined to the level such that the supernova radius parameter (10 pc) is resolved by ≥ 4 cell widths. With resolved halos defined as those with ≥ 100 DM particles, the minimum mass of a resolved halo in the PHX256 simulations is 2.3×10^6 M$_\odot$. We include radiation hydrodynamics using the Moray ray-tracing solver (Wise & Abel, 2011), with Population III and enriched star clusters as radiating sources. A uniform, z-dependent Lyman-Werner background is included to account for H$_2$ dissociating radiation sources originating from outside the simulation volume. Nonequilibrium primordial gas chemistry for the 9 species H, H$^+$, H$^-$, H$_2$, H$_2^+$, e$^-$, He, He$^+$, He^{++} is computed, and radiative heating and cooling of the gas includes both primordial and metal-line cooling contrbutions as in Smith et al. 2008. All simulations in this work were run on the TACC-Frontera supercomputer, with the PHX256-1,2 consuming $\sim$ 500 k cpu-hours total.

Although the PHX256-1,2 simulations are distinct simulations, PHX256-2 uses identical physical models and resolution with different initial conditions as compared with PHX256-1. The identical parameters include the total mass in the box, where the primary difference is in the spatial distribution of density peaks and troughs, whose magnitude is determined by the cosmological parameters common to both simulations. For this reason, we treat the two simulations as part of the same distribution of star formation examples.

2.3.2 Star Formation

Since this work aims to predict star formation, we will briefly review the relevant algorithms here. For a more detailed review of star formation in `Enzo`, please refer to the `Enzo` documentation[2].

The PHX256 simulations both include Population III single star and Population II star cluster formation. At each grid timestep, the finest grid cell at each location is evaluated for star formation. The Population III formation criteria and their parameter values which are checked are:

- Number density $n \geq 100$.
- H_2 density: $\rho_{H2}/\rho_b \geq 10^{-3}$.
- Metallicity[3]: $Z \leq Z_c$ with $Z_c = -5.5$ for Population III formation. Population II formation requires $Z \geq Z_c$.
- The freefall time should be less than the cooling time: $t_{ff} < t_{cool}$
- Converging gas flow: $\nabla \cdot \boldsymbol{v}_{gas} < 0$.

If these criteria are met, a Population II ($Z > Z_c$) or Population III ($Z < Z_c$) star particle is formed from a sphere containing twice the mass of the star centered on the star forming grid cell. In the Population III case, the particle represents a single star with mass taken from the a modified salpeter IMF of the form

$$f(\log M)dM = M^{-1.3}\exp\left[-\left(\frac{M_{\text{char}}}{M}\right)^{1.6}\right]dM, \tag{2.1}$$

with $M_{\text{char}} = 20\ \mathrm{M}_\odot$. Population II particles are formed if $Z > Z_c$ (the H_2 requirement is ignored). In this case, a single particle represents a radiating stellar cluster with $M_{min} = 1000\ \mathrm{M}_\odot$, assuming an unmodified Salpeter IMF. Although the grid cell that was identified for Population II star creation must have $Z > Z_{crit}$, the gas surrounding likely has lower or higher metallicity: when mass-averaged into the star particle, the resulting particle may have $Z \neq Z_{cell}$.

---

[2] https://enzo.readthedocs.io/en/latest

[3] Z denotes log metallicity relative to solar. With metal mass M_z, $Z = \log\frac{M_z}{M_b} - \log\frac{M_{z,\odot}}{M_\odot}$

2.3.3 Data Reduction and Preparation

Both simulations output the simulation state every 200 kyr, from $z = 30$ to the final redshift noted in Table 2.1. Between $z = 30$ to $z = 10$, there will be 1806 individual data outputs that are a snapshot of the entire simulation domain. This project serves as a proof of concept, so uses all simulation outputs up to their most progressed state (z_{final} in Table 2.1). Future work will progress simulations to $z \leq 10$ to incorporate lower-z data into the datasets. In this method, we assume that Population III formation is independent of z. In other words, we train our models on star formation events across a range of redshifts without treating data at different redshifts as an isolated distribution of star formation. The justification for this is the set of criteria that must be met to form a Population III star in a fully resolved simulation (bulleted list above) depends only on local conditions which are decoupled from global conditions. The one exception is the globally evolving UV Lyman-Werner background, which affects the H_2 fraction of the gas, which in turn enables the gas to cool and condense into stars. However, since our model includes the H_2 fraction, this global influence is taken into account. In support of our assertion that redshift is not an important parameter of the problem is the analysis of Xu et al. 2016, who showed that the statistical properties of high-z dwarf galaxies in the *Renaissance Simulations* were insensitive to redshift, but rather principally dependent on the halo virial mass. There, galaxies form from gas pre-enriched by Population III stars whose formation process is modeled directly.

The simulation outputs occupy > 75 TB in their unprocessed state at the current redshifts. Although copious, this raw data is not acceptable for input to a neural network, so the both simulations are post-processed to generate training, validation, and testing data, as well as to reduce the size of the data to a more manageable footprint. The final dataset counts, as derived from the following process, are tabulated in Table 2.2.

To generate model input data, we use pairs of snapshots from a single simulation, $\{D, D_{-1}\}$, where D is the current output and D_{-1} is the output immediately prior. Each D output is checked for new star particles since D_{-1}. If found, we generate a uniform grid with volume (9.98 kpccm)^3 centered on the new star particle in D using `YT` (Turk et al., 2011). Each hydrodynamic and color field in the region in D_{-1} is saved to this cube. To label the star location, we flag cells of a 3^3 cube centered on the star forming grid cell as star forming voxels. For sample augmentation, we additionally generate n_{shifted} volumes that are centered randomly, but still contain the target star particle. After all new stars are accounted for, we generate random samples of regions in D_{-1} with the restriction that any candidate volume must have a volume average density

Table 2.2. Summary of final dataset for training, testing, and validation

Partition	N_{stars}	N_{SFR}	N_{nSFR}	N_{total}
Train	1564	9380	72132	81512
Validation	67	412	6027	6439
Test	630	3782	23449	27231

Note: The raw number of stars in each partition is given by N_{stars}, the number of star forming regions after data augmentation by N_{SFR}, and the total number of non-star forming regions by N_{nSFR}

($\langle\rho\rangle$) greater than the mean($\bar{\rho}$): $\langle\rho\rangle/\bar{\rho} \geq 1.0$. There are $n_{\text{star}}(1+n_{\text{shifted}})+n_{\text{star}}\times 50$ samples generated for each snapshot of the simulation, where we set $n_{\text{shifted}} = 5$.

For PHX256-1, samples are separated into training and validation, which is randomly determined for each volume during initial data reduction while ensuring that volumes with the same star particle belong to the same split. This proof of concept uses 81,512 (9380 with stars) training samples with 6,439 (412 with stars) validation samples. There is a lower ratio of star containing volumes in validation because we take care to remove volumes that also contain a star in the training dataset. To ensure a pristine test dataset, we perform data reduction as above on PHX256-2, generating 27,231 (3,782 with stars) samples exclusively for the testing dataset.

In the final step of preparation, each sample is scaled by the standard deviation and mean of all training data, i.e., for density, $\rho_{scaled} = (\rho - \langle\rho\rangle)/\langle\sigma(\rho)\rangle$ for $\langle\rho\rangle$, the average voxel density in training and $\langle\sigma(\rho)\rangle$, the averaged standard deviation of voxel density in training data.

We choose which hydrodynamic fields to train on based on our knowledge of star formation criteria in `Enzo`. Each volume has 5 channels as input: baryon density (ρ_b), H_2 density (ρ_{H2}), gas velocity divergence ($\nabla \cdot v$), total metallicity (Z_{sum}), and total (kinetic + thermal) energy (E_T). Our selection of fields can also be based on physical intuition: Population III star formation occurs only in regions with high ρ_B, ρ_{H2}, and very low Z_{sum}. E_T serves as a dual probe: strong radiation fields will increase temperature, thereby increasing thermal energy, while fast-moving gas would increase the kinetic energy. Star formation is not expected in either fast moving or hot gas, so having a high value of E_T should disqualify the region for Population III star formation. The effect of Z_{sum} is essentially binary: $Z_{sum} \leq Z_{crit}$ should enable Population III star formation, but any other case should immediately disqualify the region. The $\nabla \cdot v$ field should immediately disqualify star formation if the gas is not converging.

Although a model with fewer fields may appear effective, losing any of these probes into the

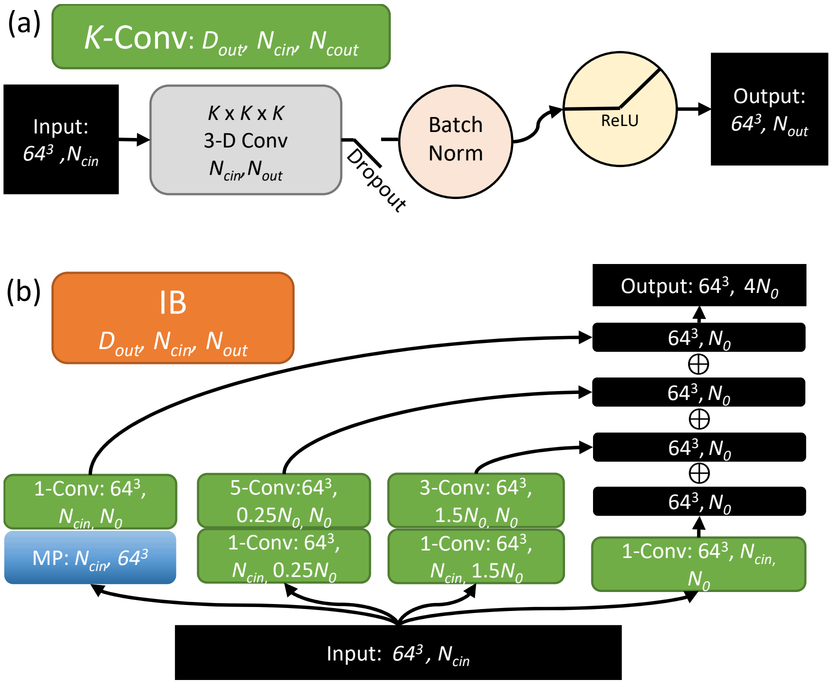


Figure 2.1. The basic convolution block and inception block used in this work. (a): 3D convolution with K^3 filter (K-Conv) used throughout this work. By default, every convolution is followed by dropout, batch normalization, and ReLU activation. The output dimension is unchanged, with N_{out} channels. (b): An example Inception block (IB) used throughout this work. A copy of the input is processed in four branches with the outputs from each branch concatenated ($\oplus$) together to form the final output. The convolutions follow the convention of K-Conv in (a), except that the 1-Conv blocks in the middle two branches have no activation function and dropout is applied only after the last convolution in each branch. Network width can be tuned by the parameter N_0. The channel numbers in the IB are used in SINet (section 2.4.1), IBs in IUNet (section 2.4.2) use simplified values with each branch outputting $N_0/4$ channels.

hydrodynamical state would be expected to not be as robust across all redshifts. For example, if ρ_{H2} is ignored, the model may start to fail after the LW background becomes strong enough to dissociate H_2 in relatively dense gas. If E_T were removed, the model would have to learn to infer the energy of the gas from $\nabla \cdot v$, while losing all probes into the temperature, except that ρ_{H2} and ρ would likely be lower at higher temperature.

2.4 Network Design

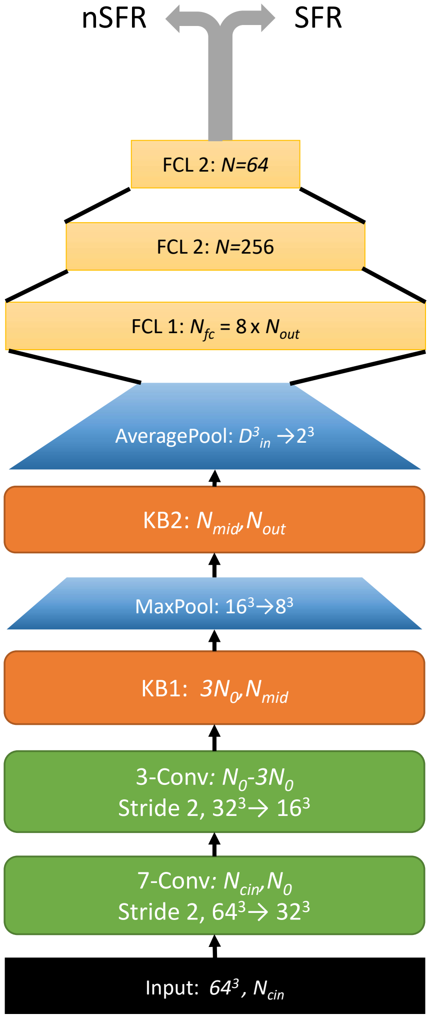


Figure 2.2. Small classifier design used in this work. Blocks indicate a layer of processing, with input and output channels notated. Data is initially processed in two standard convolutions with filter 7^3 and 3^3, with stride 2 to reduce dimension by half. These are followed by two key blocks(KB1, 2), with another max-pooling after KB1. Before the fully connected layers (FCL), the output of KB2 is average pooled to $\{N_{batch}, N_{out}, 2, 2, 2\}$. All layers use ReLU activation. The final output is two classes, {nSFR, SFR}. The parameter N_0 can be used to tune the width of the network while maintaining the same architecture. In principle, the KB can any specialized processing block, e.g., an inception, densely connected, or residual block, as seen in Figures 2.1 and 2.3.

In a grid-based simulation code like `Enzo`, the state of the simulation is stored on a fixed grid, where quantities like baryon density are stored as cell-centered quantities. Every quantity that is advanced by the simulation is tracked on the grid, or derived from quantities that are tracked there. Therefore, when inspecting a hydrodynamical simulation, it can be viewed as a 3-dimensional volumetric image, where each hydrodynamic quantity is analogous to the RGB color channels of a typical image. Of course, different hydrodynamic fields may carry nearly independent information, so the analogy to color channels is only surface deep; nonetheless, this logic leads us to classify volumes of a simulation as if they were volumetric images with hydrodynamic information as channels. Such an approach lets us take advantage of the numerous developments in the field of computer vision for this problem.

We use a two-stage approach to predict localized Population III star formation. The first stage ($S1$) is a classifier used to quickly decide if a potential region is capable of forming stars in any of its volume. If the classifier identifies a star forming region, it is followed by a more complicated voxel segmentation network ($S2$), used to identify which voxels within the volume are forming stars. The module composed of $S1$ and $S2$ (`StarFind`) must agree on the star forming state of the volume in order to classify the volume as star forming. All of the architectures presented here are implemented in Pytorch (Paszke et al., 2019) and were trained using 4 K80 GPUs on the SDSC-Comet supercomputer.

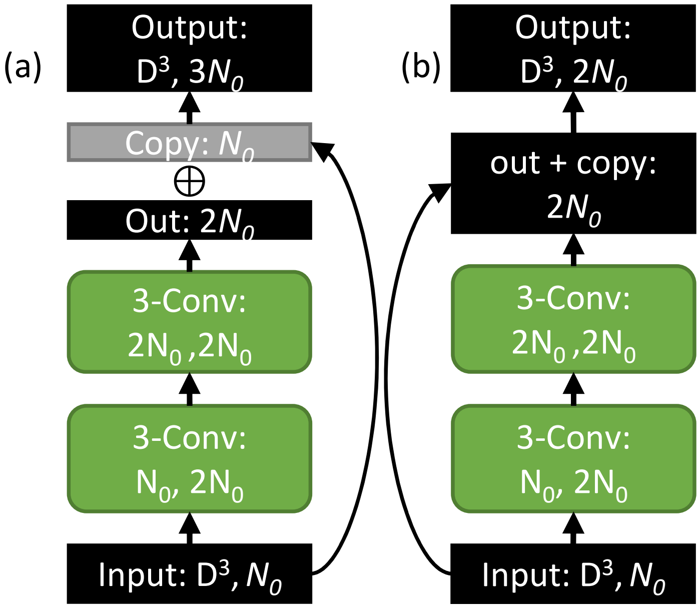


Figure 2.3. Key blocks for small classifier. Both versions accept a D^3 dimensional input with N_0 channels. (a): The DenseNet block that replaces KBs in SDNet. The input is copied, and concatenated ($\oplus$) to the output after the convolutions. (b): The ResNet block that replaces KBs in SRNet. The input is copied and added to the output after the convolutions.

2.4.1 Stage 1: Classification Network

We tested various network architectures to be used for classifying regions as star forming (SFR) or not star forming (nSFR). Initial tests included 3D adapted versions of 16-layer ResNet (He et al., 2015), 16-layer DenseNet (Huang et al., 2016), and GoogLeNet as described in Szegedy et al. 2014, however it quickly became apparent that these models were too complex for the task at hand. To reduce the model complexity, this work uses small classifiers based on ideas from those seminal papers, with vastly reduced network depth. In addition to changing the depth of the network, this work also implements several changes to the above architecture designs:

- Number of input channels represents input hydrodynamic fields and is a hyperparameter.
- The network width is another hyperparameter, tunable using the number of channels at the first layer, N_0.
- Each convolution, regardless of base architecture, is followed by dropout, batch normalization, and ReLU activation, with exceptions in Inception blocks (IBs) (Figure 2.1).

Figure 2.1 describes the basic convolution and IBs used throughout this work. By default, the convolution block has a 3D K^3 convolutional filter, dropout with probability 0.2, batch normalization (Ioffe & Szegedy, 2015), and ReLU activation. The default stride and padding are set such that the input dimension is unchanged. In the interest of preserving spatial relationships while reducing over-fitting, dropout zeros the output of an entire channel in the output of the layer (Tompson et al., 2014). The IBs presented in Figure 2.1 are used in the classifier architecture. In principle, all channel numbers in the IB are tunable hyperparameters, however this parameter space was not explored in this work.

To create a modular and minimal effective classifier architecture, we designed the small classifier networks as shown in Figure 2.2. These architectures are inspired by GoogLeNet, DenseNet, and ResNet, but are significantly shallower. After initial convolutions common to each architecture, each small classifier network is composed of only two key blocks, as seen in Figure 2.2. The key blocks may be inception modules (SINet), standard convolutions with residual skips (SRNet), or densely connected convolutions(SDNet)(see Figure 2.3). The rest of this work will focus exclusively on the use of small SINet, SDNet, and SRNet, as there was no appreciable gain in accuracy to using the full architectures from ResNet, DenseNet, or GoogLeNet, while inference and training were much slower with significantly higher memory consumption.

Before final classification, we remove dependence on input dimension by reducing to a $\{N_{channel}, 2, 2, 2\}$ volume via average pooling before the fully-connected layers. All network variations use Adam optimization (Kingma & Ba, 2014) with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Although this is a physics oriented problem, the prediction is a simple binary classification. With this in mind, we minimize the cross entropy loss given by

$$L(\hat{y}, y, c) = \left\{ w(c) \left[-y_c \hat{y}_c + \log \left(\sum_j \exp(\hat{y}_j) \right) \right] \right\}, \tag{2.2}$$

$$L = \frac{\sum_c L(\hat{y}, y, c)}{\sum_c w(c)} \tag{2.3}$$

where $w(c)$ represents weights given to the c^{th} class (here, classes = $\{0, 1\} = \{\text{nSFR, SFR}\}$), the network output ($\hat{y}_j$) and true label ($y_j$).

A uniformly sampled dataset of the simulations would have $\lll 0.001\%$ of the volume classified as SFR. We deal with this extreme bias via multiple methods: a) sampling is selective, as described in Section 2.3.3, and b) we use weights in the loss function with $w = [1, 4]$ for $\{\text{nSFR, SFR}\}$. Several combinations of weights were explored via random search in hyperparameter tuning, where we chose the weight value to minimize false positive results while maximizing accuracy (i.e., we prefer false negative predictions over false positive predictions). We additionally employ L_2 regularization to guard against overfitting with L_2 parameter $\lambda = 10^{-4}$. We use a plateau method to reduce the learning rate: the rate is reduced by half if the loss on the validation dataset has not reduced for 10 epochs until a minimum learning rate of 10^{-8}. We also employ a checkpointing method, saving model weights and parameters each time a new record low loss is achieved on validation data. The final model used in the following sections is the checkpoint using these best-loss weights.

Table 2.3. Tested classifier architectures for S1.

Model	Key Block	N_0	N_{tp}[M]	N_b	E_{BL}
SINet	IB	16	0.41	480	160
SDNet	Dense	32	14.67	460	156
SRNet	Residual	32	21.29	460	193

Note: All models were trained on input data with dimension dim^3=64^3. N_{tp} is the number of trainable parameters in the model, and N_b is the batch size. The Key Block describes the architecture used in place of KB1 and KB2 in Figure 2.2. The epoch of the checkpoint with the best validation loss, which is used for evaluation in this work, is recorded in E_{BL}.

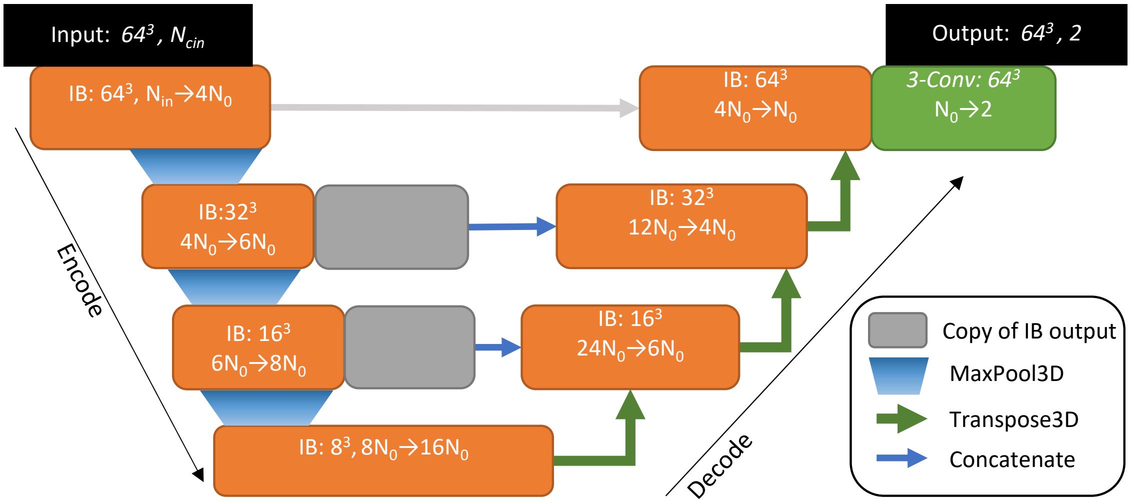


Figure 2.4. Customized Inception U-net used throughout this work. The top concatenation arrow (grey) is not used in this work to force processing at deeper layers of the network. The parameter N_0 is used to determine the width of the network as a whole, where we use $N_0 = 4$. The final convolution takes the network output of N_0 channels to 2 classes, $\{$nSFV, SFV$\}$. Each orange block is an IB (Figure 2.1) annotated with data dimension, input channel number and output channel number. The output of the encoder branch IBs is copied and concatenated to the appropriate level of the decoder branch before the IB is processed. The MaxPool3D operations use filter 3, stride 2, padding 1, so that the output has dimension reduced by half. The Transpose3D blocks on the right side have kernel size 2, stride 2, and padding 0 to mirror the reduction of the max pooling in the encoder branch. Each IB here follows the construction presented in Figure 2.1.

2.4.2 Stage 2: Segmentation Network

After classification, we use deep neural network architectures designed for pixel segmentation, which we have adapted to 3D voxel segmentation, to identify star forming voxels (SFVs) in stage 2 (S2). Two architechtures were tested here: U-Net (Ronneberger et al., 2015), and a variation of U-Net motivated by Zhang et al. 2020 that uses IBs instead of standard convolutions (IUNet). In early testing, U-net was plagued by false positives, and was abandoned in favor of the IUNet architecture presented in Figure 2.4. As outlined in section 2.3.3, each star in a SFR is labelled by 27 cells in a 64^3 box, so the data is extremely biased. To deal with this bias, we use class weights in the cross entropy loss function, with $w = [1, 9]$ for classes $\{$SFV, non-SFV(nSFV)$\}$. As with $S1$, this value is the result of minimizing false positive results, while maximizing volumetric accuracy during a random search centered at $w = [1, 8]$ as used in Zhang et al. (2019). As with the classifier networks, we use the Adam optimizer with L_2 regularization using $\lambda = 10^{-5}$. We use an initial learning rate of 5×10^{-3} with the same learning rate plateau adaptation method as S1.

IUNet functions similarly to an autoencoder, replacing fully connected layers with convolutional IB.

The encoding branch reduces dimensionality via max pooling while increasing the number of channels via concatenation and convolution operations until the lower bottleneck, where the process is reversed in the decoding branch. The encode and decode layers are connected by concatenating the encode output to the decode input at a given layer. Finally, the output is reduced to two channels, representing nSFV and SFV, for each voxel in the region.

The skips in IUNet allow more efficient back-propagation and allow information to flow directly from encoding branch to decoding branch, potentially reducing the size of the training set required to attain a robust and generalizable model (Ronneberger et al., 2015). The skip connections may also have a detrimental effect here though, as star formation is highly correlated to peaks in the density field, so the entire network may be skipped with predictions being made directly from the input fields. As seen in Zhang et al. 2019, to reduce the direct communication from the input fields to the final prediction, we removed the top most skip connection to force processing at deeper layers of the network. The IB architecture has several different sizes of convolutional filters ($K = 1^3, 3^3, 5^3$), which can give sensitivity to different scales of features in the data, e.g., 100 pc scale infalling gas toward a density peak 1-10 pc in radius. We hypothesize that this multi-scale sensitivity from IBs is likely why IUNet outperformed our standard U-net architechture early on.

Although $S2$ minimizes cross-entropy loss for training, we also calculate the intersection over union (IoU) metric to judge the quality of the $S2$ voxel-wise predictions. If P_t is the set of positive predicted voxels, and P_T is the set of ground truth (GT) positive voxels, the IoU is given by

$$IoU = 1 - \frac{P_t \cap P_T}{P_t \cup P_T}. \tag{2.4}$$

With this representation, $IoU = 0$ is a pefect prediction with no false negatives or false positives.

2.5 Results

2.5.1 S1 Volume Classifier

We use several primary metrics to quantify the success of an $S1$ model: (a) Raw accuracy as simply $A = N_{correct}/N_{total}$, (b) precision, $P = P_T/(P_T + P_f)$, with true positives P_T and false positives P_f, and (c) recall, $R = P_T/(P_T + N_f)$ with false negatives N_f. We will also extend the P and R metrics with the F_1 score, given by $F_1 = P_T/(P_T + 0.5(P_f + N_f) = 2 \times P \times R/(P + R)$ to provide another measure of classification

Table 2.4. Results of $S1$ on testing data from PHX256-2.

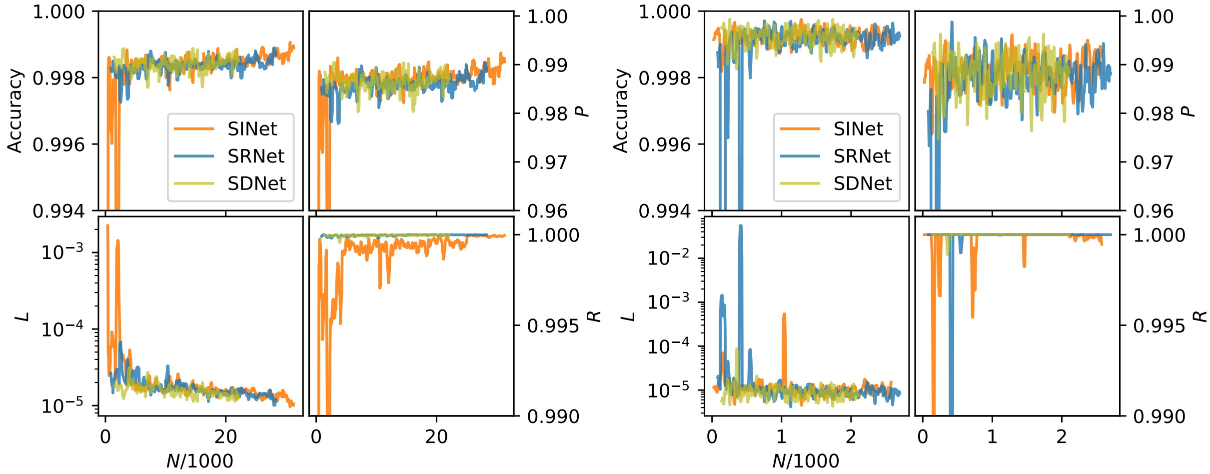
Model	Accuracy	P	R	F_1
SINet	27201/27231 (0.9989)	0.9952	0.9947	0.9950
SDNet	27201/27231 (0.9989)	0.9927	0.9976	0.9951
SRNet	27198/27231 (0.9988)	0.9924	0.9968	0.9946
IUNet	27112/27231 (0.9956)	0.9932	0.9960	0.9946

Note: We also present $S2$ using IUNet as if it were used to classify SFR. Every classifier tested is able to achieve high accuracy $\gtrsim 99.8\%$, reinforced by very high precision P, recall R, and F_1 score (as defined in Section 2.5.1).

ability. With these definitions, perfect result would have $P = R = F_1 = A = 1$. We finally present the completeness (Rosenberg & Hirschberg, 2007) considering SFR and nSFR as two clusters that must be separated by the $S1$ classifier. Although A is of limited use in these biased datasets, we present it as evidence that the models far outperform blanket SFR or nSFR predictions. Figure 2.5 presents the P and R measures along with the loss for training and validation of all tested $S1$ architectures. SDNet and SRNet, despite having similar memory requirements as SINet in training, have significantly more trainable parameters. Their increased complexity appears to affect how quickly they converge to a trained state, as they seem to have similar accuracy to SINet at much earlier epochs. This effect is particularly noticeable in the recall. On testing data results presented in Table 2.4, we find that all three S1 models perform extremely well, with all three having $A \geq 0.9988$. All three additionally have very similar precision and recall, as measured on their volume-wise classification. In F_1 score, SDNet performs best ($F_1 = 0.9951$), followed closely by SINet ($F_1 = 0.9950$) and SRNet ($F_1 = 0.9946$). We do not measure the inference rate of $S1$ as it does not produce a desirable prediction in this work; it needs to be coupled with $S2$ to localize star formation to a precise region.

2.5.2 S2 Voxel Segmentation

Loss, $\langle IoU \rangle$, and volumetric accuracy (Acc) for training and validation for $S2$ are presented in Figure 2.6. $S2$ is a capable volume classification model, with $A > 0.994$ in both training and validation. After training, $S2$ was also applied independently to the test dataset (Table 2.4), where $S2$ classified regions with $A = 0.9956$ with $P = 0.9932$, $R = 0.9960$, and $F_1 = 0.9946$, indicating that $S2$ is as capable as the $S1$ models at volumetric classification for SFRs. Despite this high accuracy in classifying regions, $S2$ struggles to match individual star forming voxels, with $\langle IoU \rangle > 0.819$. Indeed, if the voxel-wise accuracy is quantified by averaged precision and recall, we find $\langle P \rangle \geq 0.258$ and $\langle R \rangle \geq 0.875$ yeilding $\langle F_1 \rangle \geq 0.399$. These combined



(a) Training **(b)** Validation

Figure 2.5. Training (a) and validation (b) results for the training of small classifier architectures. We further quantify the accuracy with the precision (P) and recall (R) measures, as defined in section 2.5.1. Despite the bias in our samples, the errors in S1 are dominated by false positives indicated by the relatively lower P across all architectures. $N/1000$ represents one output every 10 iterations within each epoch–validation has fewer data points because there are fewer validation iterations in each epoch.

Table 2.5. Results of different configurations of `StarFind` modules on testing data from PHX256-2.

$S1$	$\langle IoU \rangle$	P	R	F_1	Completeness	TSS	HSS
SINet	0.8191	0.2590	0.8750	0.9950(0.3997)	0.9752 (0.2354)	0.9940 (0.8571)	0.9942(0.3626)
SDNet	0.8196	0.2584	0.8759	0.9954(0.3991)	0.9746 (0.2349)	0.9964(0.8571)	0.9943(0.3626)
SRNet	0.8195	0.2592	0.8754	0.9950(0.3999)	0.9724 (0.2356)	0.9956(0.8571)	0.9937(0.3626)

Note: Each row represents the `StarFind` module using a different architecture for $S1$. The averaged IoU ($\langle IoU \rangle$), P, and R pertain to voxel-wise predictions within the regions. F_1, Completeness, true skill score (TSS), and Heidke skill scores (HSS) are reported in volumetric classification and voxel-wise classification in parenthesis.

measures show a significant propensity to false positive SFV. This is reinforced by simple statistics; there are an average of 36.34 ($\sim 3.3^3$ voxels) true SFV in each SFR, with $\sigma = 14.78$[4] while $S2$ predictions average 89.86 ($\sim 4.4^3$) SFV per SFR with $\sigma = 45.9$.

The predicted region still covers the GT voxels, as indicated by the high value of the voxel-wise recall. Aside from recall, this behaviour can be quantified in a spatial sense: if we define a star forming center as the average location of SFV, the euclidean distance from the prediction center to GT center is another measure of prediction quality. Here, we measure the euclidean distance in voxel widths, and find that the average distance between SFV center and GT center is 1.90 voxel-widths. Further, the standard deviation of the distribution of distances is 2.368 voxel-widths, while the largest observed distance is 23.72 voxel-widths.

---

[4] σ here represents the standard deviation of the distribution

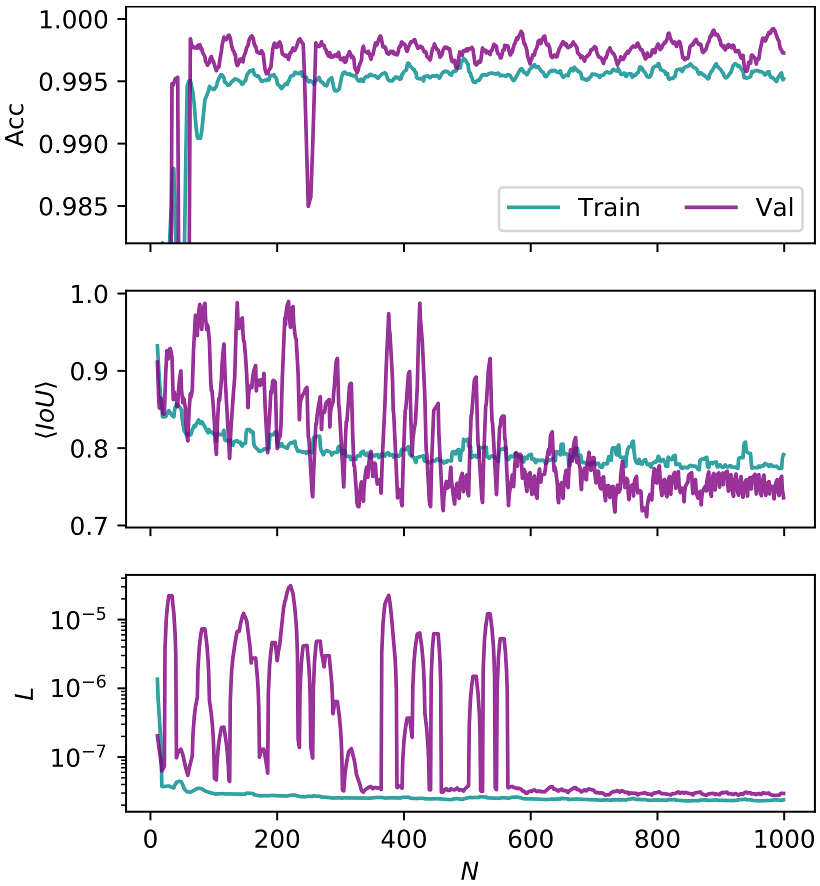
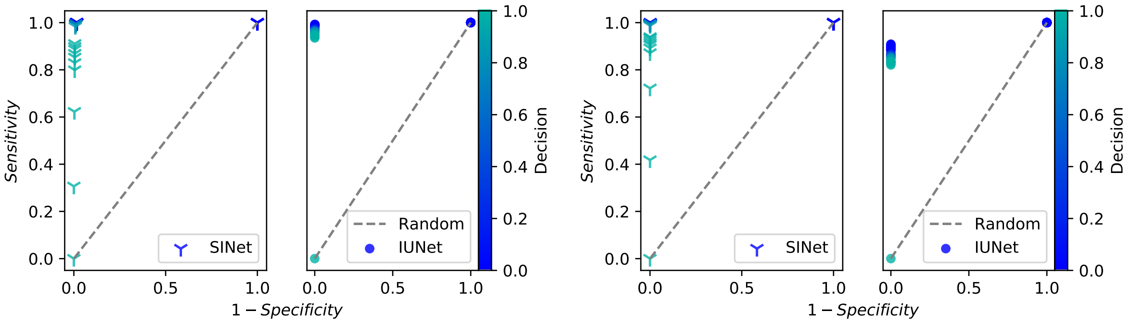


Figure 2.6. Accuracy (top, *Acc*), averaged *IoU* (middle, $\langle IoU \rangle$), and cross entropy loss (lower, *L*) from training and validation of IUNet. *N* represents intra-epoch recordings, not end of epoch data, and validation data has been expanded to align with training. These data reflect epochs of training up to the best recorded validation loss.

In examining the distribution of distances, we find that 98.1% of SFV centers are identified within 10 voxels of the GT central point, and that 0.796% of centers are more than 15 voxel-widths from the ground-truth center. This metric will likely be improved with further training and inclusion of more data.

The single largest shortcoming of using *S*2 for all processing is in its inherently slower computation time. Processing every volume through *S*2 can only be done at a rate R_{proc} of 54.67 volumes/second, when only timing the inference of the model. Given that $> 90\%$ of volumes contain no stars, a quick, simple model that can discard obvious nSFR would greatly expedite the inference of the final model. For this reason, we chain together *S*1 and *S*2 to form the `StarFind` module.

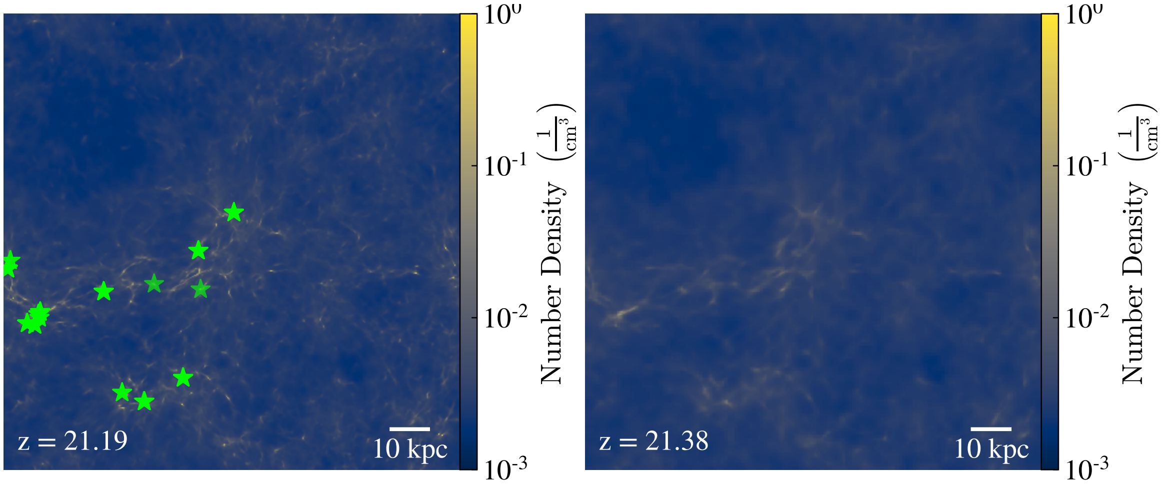
2.5.3 `StarFind` Module



(a) Train **(b)** Test

Figure 2.7. Receiver operator characteristic plots for volumetric (SINet) and voxel-wise (IUNet) classification. The decision boundary represented by various points is given by the color scale. Both models far exceed the performance of a random classifier, given by the hashed line. The AUC is 0.9977, 0.9961 in training for SINet and IUNet respectively, and 0.9996, 0.9544 in testing.

To process a sample through the `StarFind` module, we filter the sample volumes using $S1$, only passing those where $S1$ predicts SFR on to $S2$. The two models are designed to be independent to test the efficacy of different combinations of architectures. The results of this test are tabulated in Table 2.5. The completeness of volumetric classification is 0.9724, while voxel-wise completeness is lower, at 0.2349. This lower classification skill in voxel segmentation is obvious in all presented metrics including true skill score (TSS) and Heidke skill score (HSS). The low voxel-wise P and high R imply that all `StarFind` variations suffer from false positive voxel predictions, reinforced by relatively lower F_1 and voxel-wise skill metrics. $S2$ has high volumetric classification skill overshadowed by slow processing; implementing an $S1$ model to filter regions increases the rate of processing samples (R_{proc}), by up to $8\times$, and may increase some or all of the F_1, P, and R scores for IUNet seen in Table 2.4. Also of note is that the actual architecture of $S1$ is largely irrelevant given the models developed here; SINet, SDNet, and SRNet all perform very similarly in all evaluated metrics. Unfortunately, the inclusion of $S1$ was unable to significantly improve the voxel-wise predictions (e.g., $\langle IoU \rangle$ or completeness) of $S2$ by prefiltering nSFR–this implies that there are very few false positive regions being identified by $S2$ that are not also classified as SFR by the various $S1$ models.



(a) PHX256-2

(b) P3N-128

Figure 2.8. Comparing the density field of simulations with the same ICs. 2.8a shows the PHX256-2 simulation at $z = 21.19$ with Population III star forming regions annotated. P3N-128 is shown at a similar redshift in 2.8b–the density peaks in P3N-128 are both less extreme and more diffuse due to lower mass and force resolution and `StarFind` predicts no star forming regions.

The quality of the `StarFind` module is further quantified in receiver operator characteristic (ROC) curves in Figure 2.7. Given the similar behavior for $S1$ models, we only present these results using SINet as $S1$. The ROC curve plots the true positive rate (TPR, Sensitivity, R) and false positive rate (FPR, 1-Specificity, FPR = $P_f/(P_f + N_t$ for true negatives N_t) while varying the classification decision boundary, given by the color scale. A completely random classifier would follow the hashed line, where SINet and IUNet drastically outperform that behaviour. These plots reinforce that SINet and IUNet are capable and accurate classifiers on both training and testing data, which is further reinforced by the area under the curve (AUC, AUROC) $AUC = 0.9977, 0.9961$ on the training dataset for $S1$ and $S2$ respectively. The AUC is similarly high on the test dataset with $AUC = 0.9996, 0.9544$ respectively.

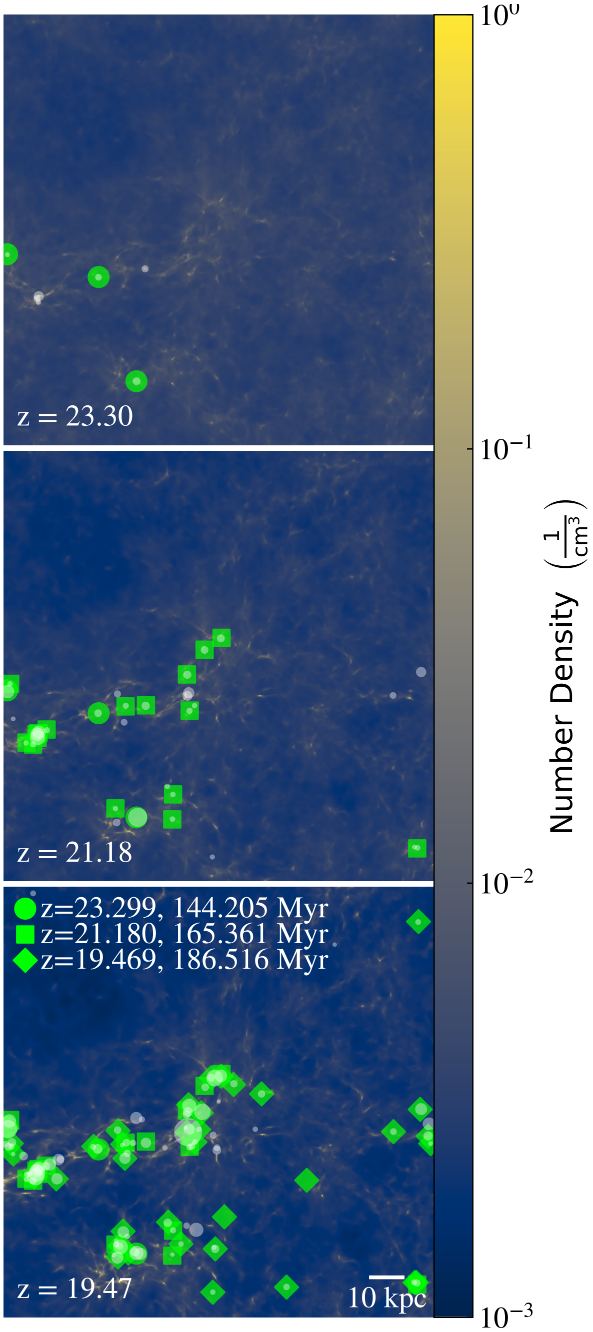


Figure 2.9. SFRs identified in PHX256-HYD at progressing redshifts. The middle panel is at the same redshift as PHX256-2 presented in Figure 2.8a. There are more SFR predicted, e.g., at $z = 21.18$, than in PHX256-2, however we have verified that the missing SFRs are present in PHX256-2 by $z = 20.86$.

2.5.4 Generalizability

In all training, testing and validation data, the data was prepared identically, from simulations that have the required resolution and physical models for Population III star formation. As a generalization test,

Table 2.6. Found volumes using the `StarFind` module as designed for simulations.

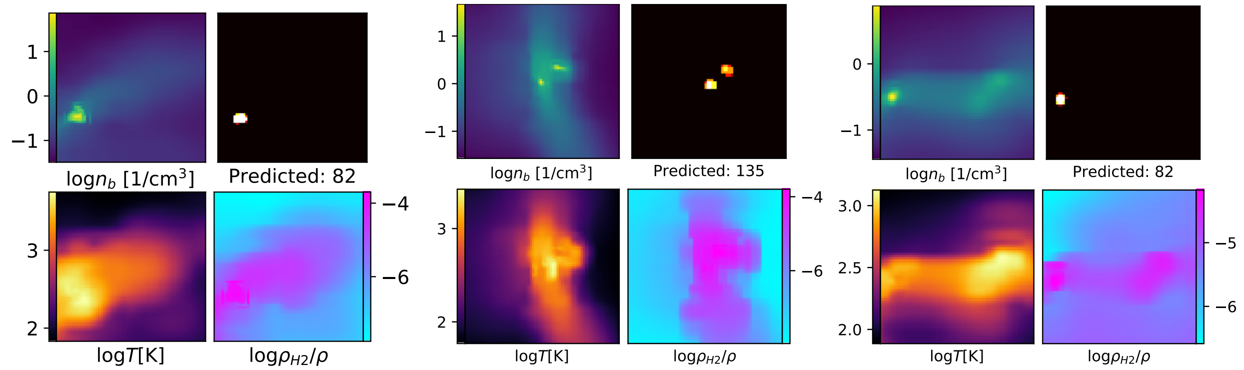
Module	Dataset	z	N_f
SINet+S2	PHX256-2	21.18	17
SINet+S2	PHX-HYD	21.18	21
SINet+S2	P3N-128	19.0	1

Note: We annotate the redshift (z) of the output, as well as the number of volumes (N_f) found at that redshift. Note that P3N-128 only has SFV found at a later z: this is due to resolution effects of the actual simulation, highlighting a limitation of applying `StarFind` to low-resolution data.

`StarFind` was applied to simulations that a) do not have the same physics models, and b) do not have the same resolution. PHX256-HYD has the same cosmological parameters as the numbered PHX256 series, however features fewer AMR refinement levels (see Table 2.1) without Population III star formation enabled. This simulation will test the models dependence on finest-grid resolution, and since it shares ICs with PHX256-2, we would expect `StarFind` to identify SFR in the same areas at early times, before Population III feedback has a chance to pollute the environment. P3N-128 has half the root-grid resolution of the PHX-series simulations (20^3 kpccm^3/root-grid cell), and fewer AMR refinement levels so that $dx_{min} = 156$ pccm. P3N-128 has no Population III star formation, and has mass resolution $\sim 1/8$ that of the PHX series. Ten DM particles in P3N-128 have more mass than a halo expected to form Population III stars: this simulation will test how `StarFind` performs in simulations that cannot resolve these first star forming halos.

For this test, we select a data output, and iterate through all AMR grids in the simulation hierarchy. Any grids less than level 3 are automatically skipped, as they cannot qualify for star formation. Grids at a deep enough AMR level are tiled in 10^3 kpccm^3 volumes and each volume with $\langle \rho_B \rangle / \bar{\rho} > 2$ is passed through the `StarFind` module. If $S1$ returns nSFR, the rest of the computation is skipped. If $S1$ returns SFR, then the volume is passed to $S2$, which will return a SFV classification for each of the 64^3 voxels in the volume.

In the simulations without SFF, those regions that do collapse to be identified as SFR will continue to collapse: there is no star formation to provide a sink for the gas, nor feedback to disperse or photoionize the it. If we identify all star forming regions at, e.g., $z = 21.18$, then this will represent all the star forming regions that have formed since the simulation start. Since the actual star formation rate (SF rate) will depend heavily on the stellar initial mass function (IMF), here we aim to match the number of star forming regions, with the assumption that a stellar IMF can then be applied to match the SF rate of higher-resolution simulations (i.e., PHX256).



(a) PHX256-HYD: $z = 21.18$ **(b)** PHX256-HYD: $z = 21.18$ **(c)** P3N-128: $z = 19.00$

Figure 2.10. Test result examples from generalizability test. Fields presented are the projection of $\log_{10}$ of baryon number density (n_b), temperature (T), and H_2 neutral fraction (ρ_{H2}/ρ); the "Predicted" panel shows voxels identified as SFV by $S2$ with the number of positive voxels annotated. (a) shows a nearly ideal result with tightly clustered cells predicting a star forming region at a peak of density and H_2 fraction. (b) shows the interesting possiblity of finding more than one star forming region per volume. (c) shows identification of the first identified star forming volume (8 kpccm from that in 2.10a) in an under-resolved simulation. Since the dynamics of structure formation are less resolved, the star forming region is not found until a much later redshift.

By $z = 21.18$, PHX256-2 has formed 17 clusters[5] of Population III stars, with 226 individual star particles (see Figure 2.8a): this early star formation is strongly clustered, averaging > 10 star particles per 10 kpccm cluster. With a baseline for the number of star forming regions in hand from PHX256-2, we apply the module to PHX256-HYD. The results of applying `StarFind` are presented in Figure 2.9. At $z = 21.18$, `StarFind` identified 21 regions. If compared to Figure 2.8a, we immediately identify star regions in PHX256-HYD that are not yet present in PHX256-2. These regions however, have been verified to begin forming stars in PHX256-2 by $z = 20.86$. `StarFind` is identifying SFR in PHX256-HYD at locations reflected in PHX256-2, and all of the regions identified by `StarFind` are mirrored by or precede star formation in PHX256-2.

Since there is no feedback implemented, a full analysis of errors incurred by using `StarFind` and its feedback method is deferred to future work. Of primary interest will be the errors in star forming locations and the metal distribution resulting from Population III stars, however we will also investigate, e.g., the impact of early or late star formation predictions such as those seen in the PHX256-HYD.

Example projections of the predictions from our generalization tests are shown in Figure 2.10. Shown are fields that correspond more easily to physical intuition than the fields which generated the predictions;

---

[5] Here, a cluster is defined as any group of stars all within 10 kpccm of the first identified star.

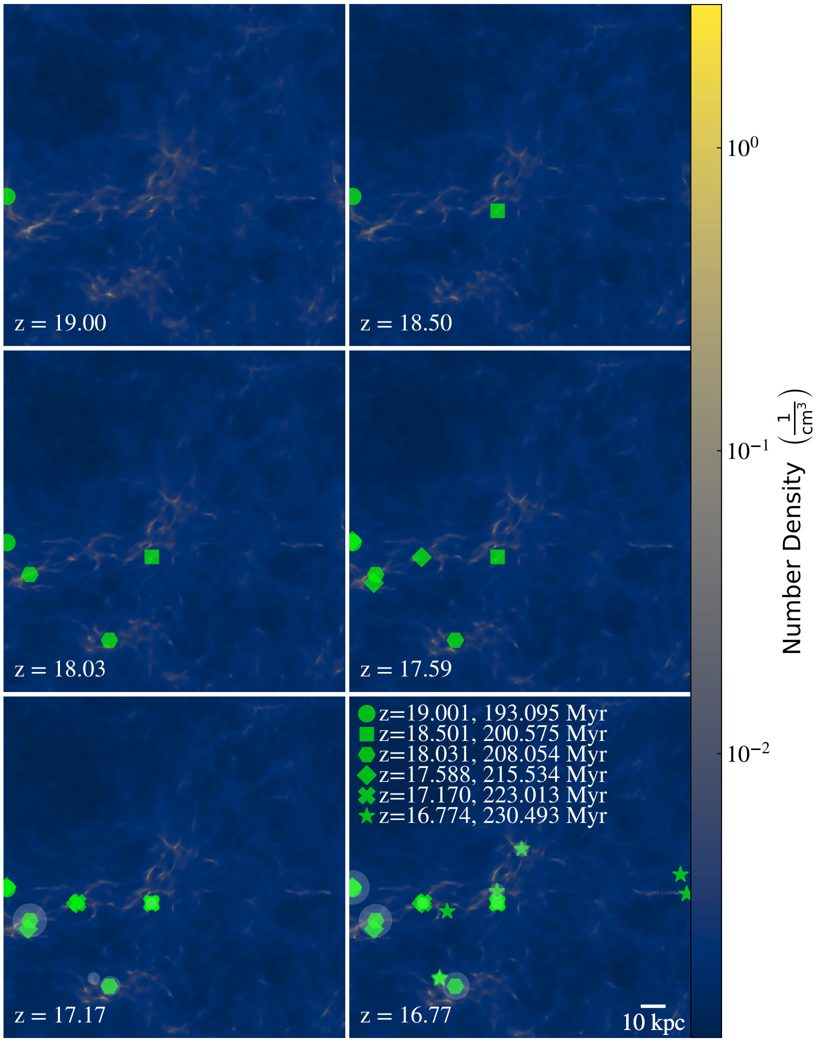


Figure 2.11. Redshift evolution of identified SFR in P3N-128. Halos found by the HOP algorithm are identified by light circles (scaled by virial radius). Most SFR are identified before their host halo is identified–by up to ∼37 Myr for the first SFR predicted at $z = 19.0$ whose host halo is identified at $z = 16.77$.

we show number density n_b in place of baryon density, temperature, T, in place of energy, and H_2 neutral fraction (ρ_{H2}/ρ_b) in place of neutral H_2 density. Figure 2.10a shows an ideal prediction; the predicted region is compact and well defined, and appears to agree with peaks in both n_b and ρ_{H2}/ρ_b, which are both to be expected from the star formation algorithms in `Enzo`. Figure 2.10b shows the prediction of two star forming regions in one volume. While examples like this are possible, they become more common as the collapse dynamics proceed without stellar feedback and enriched star formation to begin disqualifying regions for Population III star formation. The dual predictions still coincide with sensible locations in the projected fields. Figure 2.10c shows the first prediction of SFR in P3N-128, in a neighboring region near

2.10a at $z = 19.0$. The relative lateness of this prediction is an artifact of structure formation dynamics in under-resolved simulations; as seen in Figure 2.8b, at $z \simeq 21$, the density field in P3N-128 is amorphous, showing no sharp, distinct features as seen in the higher resolution simulations.

In applying the `StarFind` module to P3N-128, the module identifies no star forming regions at $z = 21.38$. To analyze how star formation as identified by `StarFind` may proceed, we iterate through all available outputs of P3N-128 to find both the first SFR, and subsequent SFR. The results of this test are presented in Figure 2.11. The first SFR is identified at $z = 19$, with more found at each subsequent output. Figure 2.11 does not include every SFR–those that are very close together are only plotted once to aid readability: at $z = 16.77$, there are 17 distinct SFR identified in P3N-128. This raw number of SFR and their locations in the projection agrees well with those in PHX256-2 at $z = 21.19$, as seen in Figure 2.8a.

2.6 Discussion

2.6.1 Further Discussions of Generalizability

The `StarFind` module has one major goal: to identify Population III star forming regions in simulations without the necessary resolution for star formation. The lack of resolution could present itself in a spatial sense, as in PHX256-HYD, or in a mass sense, as in P3N-128. We find that the module cannot completely compensate for the lack of mass resolution in P3N-128–the density peaks that would lead to star formation are evolving at a different rate and in a different manner, with peaks being more broad and potential star forming regions being less well-defined than a similar redshift in PHX256-2. Figure 2.8 shows a comparison between PHX256-2 and P3N-128. Even by eye, the dynamics of collapse in 2.8b are very different than 2.8a, with density peaks being both less extreme and more diffuse.

We can further illustrate the effect of structure formation from a halo mass function (HMF) perspective. Given that PHX256-2 and P3N-128 share identical initial conditions and box size, then without resolution effects, we should expect that they have the same HMF at a given redshift. To get the HMF (represented here by halo count as a function of halo mass), we use the HOP algorithm (Eisenstein & Hu, 1998) with overdensity threshold 100 to identify DM halos within both simulations. Comparisons of halo number are shown in Figure 2.12. At $z \simeq 21$, P3N-128 has no identifiable halos; the first ones start to appear with virial mass $M_{vir} \sim 10^{7.2-7.6}$ $\mathrm{M}_{\odot}$ at $z = 17.17$. P3N-128 never creates low mass halos with $M_{vir} \sim 10^{5-6.5}$ that would be expected to form Population III stars, as the particle mass is simply too high and those halos

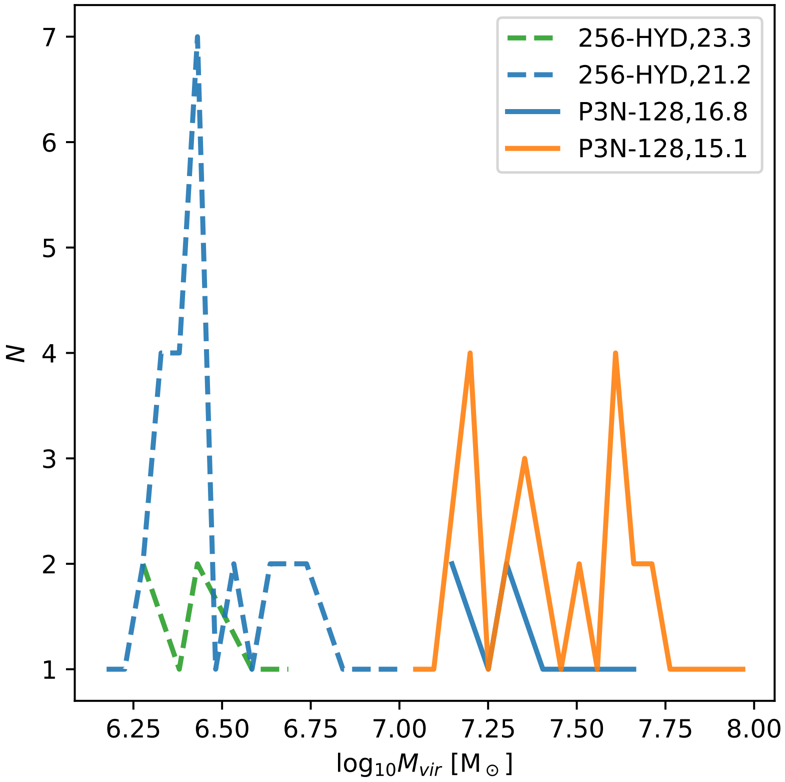


Figure 2.12. Halo number (N) as a function of log halo virial mass comparison between PHX256-2 and P3N-128 at redshift annotated in the legend. Note that halo formation in P3N-128 begins at $z \lesssim 17.17$, and low-mass, Population III forming halos ($M_{vir} \lesssim 10^{6.5}$ M$_{\odot}$) are never formed in P3N-128 due to its low mass resolution.

are unresolved. Despite this lack of resolution, as presented in Figure 2.11, `StarFind` predicts stars that will be in halos that form 10-30 Myr after the SFR identification.

The mass (particle mass or gas mass in a grid cell) is a determining factor in the effectiveness of stellar feedback (Hopkins et al., 2019). To determine the potential difference of feedback applied at $z = 19$ or $z = 16.77$, we analyzed the mass distribution of the SFR (M_{SFR}) as compared to the later halo (M_{halo}). We found that $M_{SFR} < 0.5M_{halo}$, when using the virial radius (R_{vir}) of the halo at $z = 16.77$ to define a volume at both redshifts. Importantly, the gas is less compact in the central regions e.g., at $R = R_{vir}/7$, $M_{SFR} < 0.2M_{halo}$. When we extend this analysis to all SFR identified by $z = 17.59$, comparing each SFR to the later halo it is nearest to, we find that the SFR has overdensity $[\langle\rho\rangle/\bar{\rho}]_{SFR} \lesssim 0.5[\langle\rho\rangle/\bar{\rho}]_{halo}$. If halos are too massive when they start forming Population III stars, the feedback will be confined and unable to pollute the local environment (Whalen et al., 2008b), so applying feedback at the less dense states identified by `StarFind` may reduce the effect of the massive halos found in P3N-128.

2.6.2 Error Analysis

One of the primary concerns when applying a deep learning model to a production application is to understand how the model fails. If the failure modes are predictable, then they can be accounted for in the deployment of the model, but if they are seemingly random, the model may not even be usable. First, we are concerned with failures to classify in a volumetric sense, i.e., identifying "nearby" star formation is acceptable, but misclassifying an entire region is a much more egregious failure. To analyse these critical failures of the `StarFind` module, we passed every training/test/validation sample through the model, and plotted a) the SFV labels if $S1$ was incorrect, or b) the SFV predictions and labels if $S2$ incorrectly classified the volume. 65% of the false positives in $S1$ were cases where the star voxels were within 3 voxel-widths of the border of the volume. All false positive cases of $S2$ were on the edge of the volume, and 75% of the false negatives had star forming labels on the border. The star forming border is a major avenue of failure for the module, but also an easy one to remedy. In a production pipeline, two obvious possibilities could alleviate these failures: a) ignore the SFR if the flagged voxels are on the edge, as the routine will be called again in < 1 Myr, at which point the SFV will have likely moved to a more central location, or b) recenter the volume on the suspicious SFV and re-run the module to receive a more reliable result. The results in this work used the second option to reduce false positives and negatives in analysis of the PHX256-HYD and

P3N-128 simulations.

We can additionally examine the predictions of `StarFind` by comparing voxel-based quantities for ground truth and false positive voxels. We collected the voxel field quantities for positive predictions within the testing data, to generate Figure 2.13, which shows 2-D histograms of $\rho{-}H_2$ fraction (top), $\rho - Z_{sum}$ (middle), and $\rho - E_{tot}$ (bottom), where the total energy given is quoted in code units and metallicity is given in $Z_{\odot}$. Qualitatively, the distributions look similar, however there are visible artifacts in the false positive (right) panels. Also visible in the ρ_{H_2} plot is a bimodal distribution, which is an artifact of labelling a 3^3 cloud as star forming, as opposed to including only the single voxel that hosts the star particle. Distributions such as those in Figure 2.13 can be used in the future to check that the predictions from `StarFind` are reasonable, or to reduce false positives simply by screening out voxels that fall far outside the distribution from training data. We can additionally use the diffuse points in the ground truth distribution to identify where the prototype training/test dataset is insufficient. For example, examination of the $\rho - Z_{sum}$ ground truth histogram (middle-left) shows very few voxels with finite but low metallicity ($Z > 10^{-10}$), where the false positive panel shows a significant number ($\gtrsim 10^3$) with $10^{-10} < Z_{sum} < Z_{crit}$. Focusing future dataset generation on these poorly sampled spaces will increase the robustness of the final model. Guided by Figure 2.13, we tested a simple screen that requires any SFV to have $n_b > 1.0$. Such a simple filter, on average, reduced false positive voxels by 11 and decreased the $\langle IoU \rangle$ by 0.02, while having no effect on false negative or true positive predictions. Future work will increase the robustness of `StarFind` by increasing training samples from poorly-sampled regions and quantifying the failures so that they can be compensated for in a real-time simulation.

2.7 Conclusions

We have designed a classification-segmentation model using deep ANN that is capable of predicting Population III star formation sites in cosmological simulations. This is the only method known to the authors that can accomplish such a feat. Our findings can be summarized as follows:

1. We have found that well-known image recognition architectures (adapted to 3D) are effective at this task, as are common pixel segmentation architectures. The choice of architecture had little impact on classification capability given the small classifiers developed here.

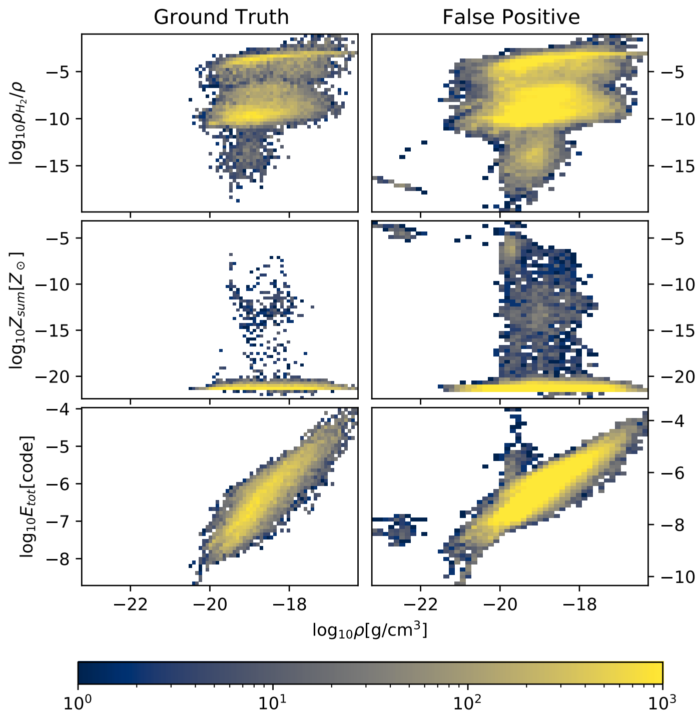


Figure 2.13. Voxel properties of ground truth voxels and false positive predictions. We present 2D histograms of $\log_{10}$ density (ρ) with other hydrodynamic quantities: $\log_{10}H_2$ fraction, $\log_{10}$ total metallicity (Z_{sum}), and $\log_{10}$ total energy (E_{tot}) from top to bottom. While false positives qualitatively fall into a similar distribution as the ground truth, there are obvious outliers.

2. The `StarFind` module predicts 9.98^3 kpccm^3 star forming volumes with $> 99.8\%$ accuracy, reinforced by volumetric classification completeness, TSS, HSS, precision, and recall.

3. The module has been applied to hydrodynamic simulations that have no star formation routines enabled and predicted star formation in the same regions at similar redshift as high resolution full-physics simulations.

4. In many cases, `StarFind` predicts star formation regions well before the formation of the host halo, particularly in under-resolved simulations such as P3N-128. More testing will be needed to quantify the success of `StarFind` in those cases.

5. By many classical measures of classification algorithms, i.e., ROC, AUC, precision, recall, F_1, and completeness, `StarFind` is a capable classifier in a volumetric sense. Correctly classifying volumes is much more important (in our application) than predicting the exact location of star formation.

6. Voxel-wise predictions IUNet suffers from a high number of false positive voxel predictions, as

indicated by low precision. It over predicts the size of localized regions in SFR, which is reinforced by the high $\langle IoU \rangle$ and simple statistical counting of SFV. Predicted star forming regions appear well defined (e.g., Figure 2.10), but cover a larger region than exists in the ground truth. The predicted region still generally covers the ground truth SFV, as indicated by the high value of the voxel-wise recall in Table 2.5.

7. Cosmology and star formation simulations are inherently multi-scale problems, and we hypothesize that the multi-scale sensitivity of IBs is contributing to the strong performance of IUNet, as well as SINet being quantitatively as capable as SDNet and SRNet, despite having $\sim 20\times$ fewer trainable parameters.

8. Utilizing voxel statistics from our early datasets is one way to increase the quality of the $S2$ classification without further training or expensive calculation, or to guide further dataset generation as the PHX-series simulations progress to lower redshift.

9. Despite these successes, fully quantifying the success of the model by comparing model-assisted simulations with ground-truth conventional sub-grid simulations will require an active stellar feedback method so that regions can move beyond primordial stars and into enriched stellar population production.

`StarFind` as a whole may be improved most by increasing the reliability and robustness of $S2$. This is the primary goal of on-going development; it may be improved through testing different loss functions, a combination of loss functions (Asgari Taghanaki et al., 2018; Hajiabadi et al., 2020), or improving the representation of poorly sampled cases, as seen in Figure 2.13. Future work will also focus on analyzing the failures of `StarFind` so that they can be more accurately estimated and guarded against.

This method will ultimately be used in a novel sub-grid feedback method, where SFR identified using this algorithm will then be evolved to a post-star state using another series of deep neural networks. The goal throughout this work was to design the proof of concept model that could be used in a production capacity (further training and development not withstanding): to that end, here we estimate the run-time implications of using this module to perform Population III SFF. Evolving PHX256-2 from z=21.19 to z=17.53 consumed 7K cpu-hours. Most of this time is spent evolving the hydrodynamic and radiation fields from hot supernova remnants or Population III main sequence stars in the deep AMR levels of the volume. PHX256-HYD evolves the same region in 14 cpu-hours by excluding Population III star formation and restricting the maximum AMR level. Applying the `StarFind` module to locate star forming regions once (i.e., one panel in Figure 2.9) consumes ∼4 CPU-hours, processing 5.2 volumes/sec on average. To localize Population III star forming regions every Myr would require performing this evaluation ~ 30 times, consuming a total of 120 CPU-hours. Most of the `StarFind` processing time is spent iterating through grids and disqualifying regions before they even enter $S1$: only $\lesssim 5\%$ of potential regions are classified by $S1$, with 3% continuing to $S2$, where only 2% finally qualify as SFR (0.003% of potential volumes), so we estimate that adding a feedback routine after $S2$ would not significantly increase the processing time of the algorithm as a whole. This method, when finalized with a feedback algorithm, may produce $\gtrsim 50\times$ speedup as compared to explicitly star forming simulations, with more speedup likely from optimizing the final method.

This rate makes this method feasible to use inline with an `Enzo` simulation using inline python with `YT`, particularly since the tests run here do not include optimization of any kind. We could expect further speedup by incorporating the method into `Enzo`'s source C/C++ code and/or optimizing `StarFind` using MPI-parallelism on CPU architecture (e.g., Mathuriya et al. 2018). Future work will focus on evolving Population III stellar remnants in regions identified using this method, and incorporating the framework into an `Enzo` simulation, along with further development and training of the `StarFind` module.

This research was supported by National Science Foundation grant CDS&E grant AST-1615848 to M.L.N. The simulations were performed using `Enzo` on the Frontera supercomputer operated by the Texas Advanced Computing Center (TACC) with LRAC allocation AST-20007. Data analysis and model development, training, and testing were performed on the Comet supercomputer operated for XSEDE by the San Diego Supercomputer Center under XRAC allocation TG-AST200019. Computations and analysis within this work were performed using the publicly available Enzo (Bryan et al., 2014),

Chapter 2, in full, is a formatted reprint of the material as it appears in Wells & Norman 2021, “Predicting Localized Primordial Star Formation with Deep Convolutional Neural Networks”. *The Astrophysical Journal Supplement Series*, 254:41. © American Astronomical society. Reproduced with permission. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Connecting Primordial Star Forming Regions and Second Generation Star Formation in the Phoenix Simulations

3.1 Abstract

We introduce the Phoenix Simulations, a suite of highly resolved cosmological simulations featuring hydrodynamics, primordial gas chemistry, primordial and enriched star formation and feedback, UV radiative transfer, and saved outputs with Δt=200 kyr. We observe 73,523 individual primordial stars within 3,313 distinct regions, forming 2,110 second-generation enriched star clusters by $z \geq 12$ within a combined 177.25 Mpc^3 volume across three simulations. The regions that lead to enriched star formation can contain $\gtrsim 150$ primordial stars, with 80% of regions having experienced combinations of primordial Type-II, hypernovae, and/or pair-instability supernovae. 0.8% of the volume was enriched by primordial supernovae, with 2% of enriched gas enriched by later generation stars. We determine the extent of a primordial stellar region by its metal-rich or ionized hydrogen surrounding cloud; the metal-rich and ionized regions have time-dependent average radii $r \lesssim 3$ kpc, with 7, 17% of regions having $r > 7$ kpc for metal-rich and ionized radii respectively. We find that the metallicity distribution function of second-generation stars overlaps that of subsequent Population II star formation, spanning metal-deficient ($\sim 7.94 \times 10^{-8}$ Z$_\odot$) to super-solar (~ 3.71 Z$_\odot$) and that 30.5% of second-generation stars have $Z > 10^{-2}$ Z$_\odot$. We find that the metallicity of second-generation stars depends on progenitor configuration, with metals from pair-instability supernovae contributing to the most metal-rich clusters; these clusters form promptly after the supernova event. Finally, we create an interpretable regression model to predict the radius of metal-rich influence of Population III star systems within the first 7-18 Myr after the first Population III star to form in the region.

3.2 Introduction

The first galaxies form from gas that has been enriched by an earlier generation of Population III supernovae. The initial conditions of metallicity in the universe after these first supernova events remains a difficult problem to model in astrophysical simulations. Researchers conducting astrophysical simulations have three practical options to determine the initial metallicity field prior to enriched (Population II) star formation: assume a metallicity floor (e.g., Hopkins et al., 2018); assume initial star formation rates that are independent of metallicity (e.g., Vogelsberger et al., 2013; Vogelsberger et al., 2014); or explicitly simulate the primordial (Population III) star formation and feedback (Smith et al., 2015; Xu et al., 2016; Wise et al., 2012; Wise et al., 2012). Ideally, all researchers would choose the last option, however the extreme small scale ($\sim$ pc^3) of primordial molecular cloud formation (Abel et al., 2000; Bromm et al., 2002) is at odds with the $\sim$ Gpc^3 scale necessary to gain useful statistics of the observable universe; any simulation using current computing facilities that can fully resolve Population III star formation is severely limited in volume, with the largest being only $\sim$ 300 Mpc^3 (Xu et al., 2016).

The Phoenix (PHX) suite of simulations is designed to facilitate the exploration of a fourth option to model Population III star formation and feedback: to develop surrogate models based on deep neural networks (DNN) as a new sub-grid method to create an heterogeneous metallicity initial condition that reflects the spatially irregular formation of Population III star formation and feedback. Data from the PHX suite have already been used to train `StarFind`, a predictive DNN based surrogate model that identifies Population III star formation sites without resorting to halo-finding or pc-scale resolution (Wells & Norman, 2021). Training is enabled by the PHX unique time between outputs, such that any star formation or feedback event is recorded to disk with 200 kyr time resolution.

Although their extreme time resolution was intended to provide high resolution of star formation and feedback events for training DNNs, the PHX suite also provides a unique opportunity to study the transition between Population III and second generation (Population II.1)[1] star formation. Studying low metallicity stars or damped Lyman-alpha systems in observations is currently our only window into the Population III initial mass function (IMF) (Cooke et al., 2017; Welsh et al., 2019, 2020), but the uncertainty in the IMF (Nakamura & Umemura, 2002; Ishigaki et al., 2018) and metallicity of Population II.1 stars that could form

---

[1] Population II.1 is our designation for the first generation of metal-enriched star formation that occurs in gas enriched exclusively by Population III supernovae. Population II stars formed from gas also enriched by earlier Population II stars is referred to as Population II.2.

Table 3.1. The Phoenix Simulations

Designation	D	$V\,(\mathrm{Mpc}^3)$	z_{final}
PHX512	512	141.8	14.04
PHX256-1	256	17.7	12.45
PHX256-2	256	17.7	13.27

Note: Each simulation includes the root-grid dimension (D), volume (V), and final redshift (z_{final}), which is completely determined by run time. All parameters aside from D and V between the three simulations are identical, see Section 3.3 for detail.

from enriching events make inferences about the Population III era difficult. Due to the fine time resolution in the PHX suite outputs, we use them to study the formation state of small Population II star clusters, as well as the evolution of Population III star forming regions. This will guide future studies that connect Population III and Population II stars by providing a reference of how many Population III stars can be related to a Population II cluster, as well as the range of metallicities that may be observable in a second generation star cluster.

The remainder of this paper is organized as follows: Section 3.3 presents a summary of the simulations and included physical models; Section 3.4 showcases summary statistics such as star formation rates and halo mass functions of the final simulation states; Section 3.5 presents a time-resolved study to determine the origins of the first generation of enriched star formation as well as quantification of primordial stellar systems; Section 3.6 presents an interpretable regression model that predicts the region of influence for a Population III system, given the stellar masses and birth times within that system; Finally, Section 3.7 consolidates our results and presents final notes on both this study and future directions.

3.3 The Phoenix Simulations

The PHX suite consists of three simulations performed with the ENZO (Bryan et al., 2014; Brummel-Smith et al., 2019a) adaptive mesh refinement hydrodynamic cosmology code designed to study the time and spatial evolution of early structure formation and galaxy assembly. Projections of the largest (512^3 root-grid) simulation are shown in Figure 3.1. In addition to the simulation shown in Figure 3.1, two other smaller simulations were also performed, as noted in Table 3.1. The 512^3 simulation (PHX512) has volume (5.21 Mpc)^3, the PHX suite includes two smaller 256^3 root-grid simulations with volume (2.61 Mpc)^3 each (PHX256-1,2). Prior efforts have produced simulations with similar redshifts as the target range of the PHX suite (Wise et al., 2012; Wise et al., 2012; Xu et al., 2016), however the saved outputs of these earlier works

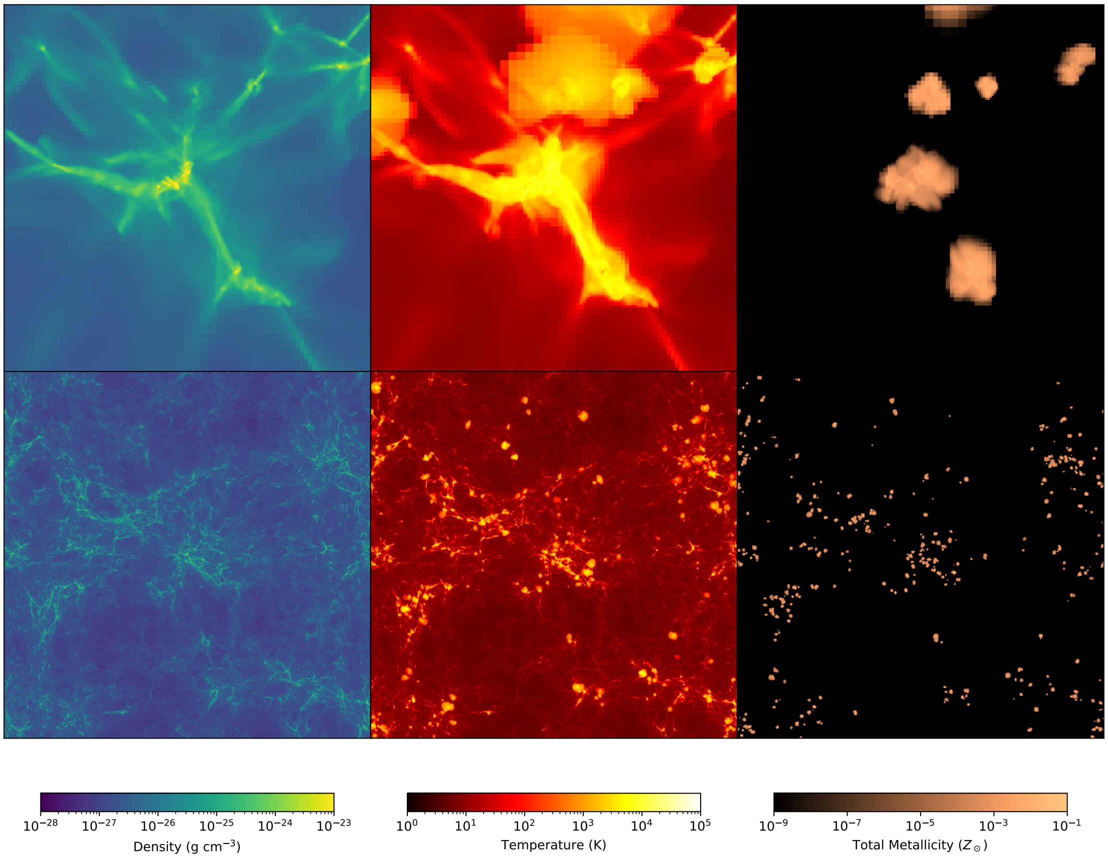


Figure 3.1. Projections of PHX512 simulation at $z = 14.04$. Top row: Projections through a 15 comoving kpc volume surrounding a complex of star forming halos. Bottom row: 1 Mpc thick slab projection of the entire simulation domain ($L_{box} = 5.21$ Mpc) at the same redshift centered on the star forming halo in the top row.

have time resolution $\Delta t \gtrsim 1-5$ Myr, whereas the PHX suite has 200 kyr between all saved outputs for redshifts $30 < z < z_{\rm final}$.

All simulations share identical cosmological parameters with $\{\Omega_m = 0.3111, \Omega_b = 0.048975, \Omega_k = 0, \Omega_\lambda = 0.6889, H_0 = 0.6766, \sigma_8 = 0.811, n = 0.965\}$ (The Planck Collaboration et al., 2014). The cosmological initial conditions are generated at $z = 99$ using `MUSIC` (Hahn & Abel, 2011), where each simulation uses a unique random seed to generate an initial state that is consistent with the given cosmological parameters. All simulations have identical mass and spatial resolution and the same refinement criteria as the *Renaissance Simulations* (Xu et al., 2013), with dark matter particle mass $M_{DM} = 2.34 \times 10^4$ $\mathrm{M}_\odot$, initial average baryon mass per cell $M_{b,i} = 1.17 \times 10^3$ $\mathrm{M}_\odot$. The root grid can be refined up to 9 levels of adaptive mesh refinement (AMR), where refinement occurs on dark matter density, baryon density, and regions surrounding Population III star particles such that the supernova radius (10 pc) is resolved by at least 4 cells, or is at the maximum AMR level. Refinement on densities is super-Lagrangian; cells are flagged for refinement at level l where the cell mass $M_c \geq M_i \times 2^{-0.4l}$, where M_i refers to M_{DM} or $M_{b,i}$ for refinement on dark matter or baryon density respectively. With these resolution parameters, and assuming ~ 100 dark matter particles for a resolved dark matter halo, the least massive resolved halos have virial masses 2.34×10^6 $\mathrm{M}_\odot$, while the most massive halos ($\approx 10^9 \mathrm{M}_\odot$) is limited by the total mass within the volume: 3.93×10^{11} $\mathrm{M}_\odot$ and 3.14×10^{12} $\mathrm{M}_\odot$ for 256^3 and 512^2 root grids respectively. The finest spatial resolution achieved on level 9 subgrids is 19.53 comoving pc, providing < 2 proper pc resolution for $z \gtrsim 9$.

For hydrodynamic and chemical evolution, each simulation includes 9 species non-equilibrium chemistry for primordial gas species H, H^+, H^-, H_2, H_2^+, H_2^{++}, He, He^+, He^{++}, and e^-, radiative heating, cooling, and metal-line cooling as in Smith et al. 2008, and hydrodynamics evolved using the piecewise parabolic method (Colella & Woodward, 1984). Radiation interactions are included via a uniform, redshift-dependent Lyman-Werner H_2 dissociating radiation background as documented in Xu et al. 2016 as well as photo-dissociating and ionizing radiation from point-sources using the rates and MORAY ray-tracing solver method in Wise & Abel 2011. The radiation is coupled to chemistry through heating and ionization rates, and to hydrodynamic evolution via momentum coupling from photons to the gas.

Each simulation includes two types of star formation events: individual Population III stars and Population II star clusters, each tracked with a star particle. Feedback (SNe, stellar winds) from each type of star particle contributes to a unique metal density field so that the contributions from Population III and

Population II stars can be tracked independently. The star formation and feedback algorithms used in the Phoenix suite are well documented[2] and described in several prior works (Wise & Cen, 2009; Wise et al., 2012; Wise et al., 2012; Xu et al., 2013; Hicks et al., 2021), so the following is restricted to a high-level overview that includes the specific parameters used in the PHX suite. At each time step, every grid cell is evaluated for Population III star formation according to the following criteria: 1) baryon number density: $n_b > 100$ cm^{-3}, 2) H_2 fraction: $n_{H2}/n_H > 10^{-4}$, 3) metallicity[3]: $Z < Z_c$ for $Z_c = -5.5$, 4) AMR grid level is most refined for that point in space, 5) the cooling time should be less than the freefall time, and 6) converging gas flow ($\nabla \cdot \vec{v} < 0$). If a cell qualifies for Population III star formation, a particle representing a single star is formed centered on the host cell with mass taken from a modified Salpeter initial mass function (IMF),

$$f(\log M)dM = M^{-1.3}\exp\left[-\left(\frac{M_{\rm char}}{M}\right)^{1.6}\right]dM, \tag{3.1}$$

with characteristic mass $M_{char} = 20$ M$_\odot$ with masses in the range $1 \leq M_*/\mathrm{M}_\odot \leq 300$. The final mass contributing to the star formation is taken from the grid in a sphere containing twice the mass of the star.

Population II star cluster formation involves similar criteria: 1) baryon overdensity, relative to the simulation volume, $\rho/\bar{\rho} > 100$, 2) $Z \geq Z_c$, 3) AMR grid level is most refined for that point in space, 4) the cooling time should be less than the freefall time, and 5) converging gas flow ($\nabla \cdot \vec{v} < 0$). At each timestep after the initial particle formation, the mass of the cluster is accreted from the surrounding cold gas mass, $M_{\rm cold}$, estimated as 7% of the gas mass in a sphere with mean gas number density $n_b > 10^3$ cm^{-3} until $M_* \geq 1000$ M$_\odot$ as described in Wise & Cen 2009. If after one dynamical time, the particle has not accreted 1000 M$_\odot$, a low-mass particle is formed with the current mass to prevent loss of the ionizing radiation due to lower-mass star clusters. The metallicity of the formed star is taken as the mass-averaged metallicity of the cells that contributed to its formation, therefore the final metallicity may be lower or higher than the cell that initially qualified for cluster formation.

Stellar feedback for Population III stars is included in two forms: supernovae of varying mass, and point-source radiative feedback. The supernova channel includes Type-II supernovae (SNe, $11 < M_*/\mathrm{M}_\odot < 20$), hypernovae (HNe, $20 \leq M_*/\mathrm{M}_\odot < 40$), and pair-instability supernovae (PISNe, $140 < M_*/\mathrm{M}_\odot < 260$). For SNe, the ejecta mass, energy, and metal yields are taken from Nomoto et al. 2006; HNe event energy and

---

[2] https://enzo.readthedocs.io/en/latest/

[3] with metal mass M_z and cell baryon mass M_c, metallicity is given by $Z = \log(M_z/M_c) - \log(M_{z,\odot}/\mathrm{M}_\odot)$.

metal yields are linearly interpolated from these values. PISNe have ejecta mass, metal yield, and energy taken from Heger & Woosley 2002. For each type of supernova, the resulting mass, metal yield, and energy are deposited to the grid in a sphere of 10 pc, or a cube of 3^3 cell-widths if 10 pc is unresolved.

Population II stars, modeled as coeval clusters of stars, use a continuous injection model of energy, mass, and metal deposition that represents both supernova and stellar winds. At each time step during the 20 Myr lifetime of the cluster, mass is returned to the computational grid as

$$m_{\rm ej} = \frac{0.25\Delta t \times M_*/{\rm M}_\odot}{16\ {\rm Myr}}, \tag{3.2}$$

where Δt is the grid timestep. The ejecta has a metallicity fraction matching solar metallicity ($Z_\odot = 0.01295$). The ejecta has energy 1.12×10^{49} erg/M$_\odot$, which is coupled to the grid as thermal energy along with mass and metal ejecta in a 10 pc sphere surrounding the source particle, again depositing to a 3^3 cube if if 10 pc is unresolved.

3.4 General Observations from the Phoenix Suite

Halo finding was performed using `ROCKSTAR` (Behroozi et al., 2013), requiring 50 DM particles per identified halo, however halo-based analyses are restricted to those halos with ≥ 100 particles. Figure 3.2 shows the total halo mass count (HMC), the HMC of halos with non-zero stellar mass ($M_* > 0$), the HMC of halos with active Population III stars ($N_{\rm PIII} > 0$), and the HMF of halos with Population III supernova remnants ($N_{\rm rem} > 0$). All simulations show that halos having $M_{vir} > 2 \times 10^7$ M$_\odot$ are universally forming Population II stars at their final redshifts. While some halos with $M_{vir} > 2 \times 10^7$ M$_\odot$ are also forming Population III stars, the number is quickly diminishing beyond. Since we qualify a halo with $M_{vir} > 2.34 \times 10^6$ M$_\odot$ as well-resolved, Figure 3.2 also shows that Population III star formation is occurring in under-resolved halos.

The fact that Population III stars still exist in halos forming Population II stars begs the question: is there a redshift dependence on the mass of halo that will contain Population III stars? Figure 3.3 displays the fraction of halos hosting active Population III stars at various redshifts including data from all simulations. The fraction of halos with mass $\sim 10^{6-7}$ M$_\odot$ hosting Population III stars is notably higher at higher redshift. Higher fractions of halos hosting Population III stars with $M_{\rm vir} > 5 \times 10^7$ M$_\odot$ are visible at lower redshift, e.g., $z \lesssim 17$. However, this observation is also likely due to the fact that higher mass halos occur with

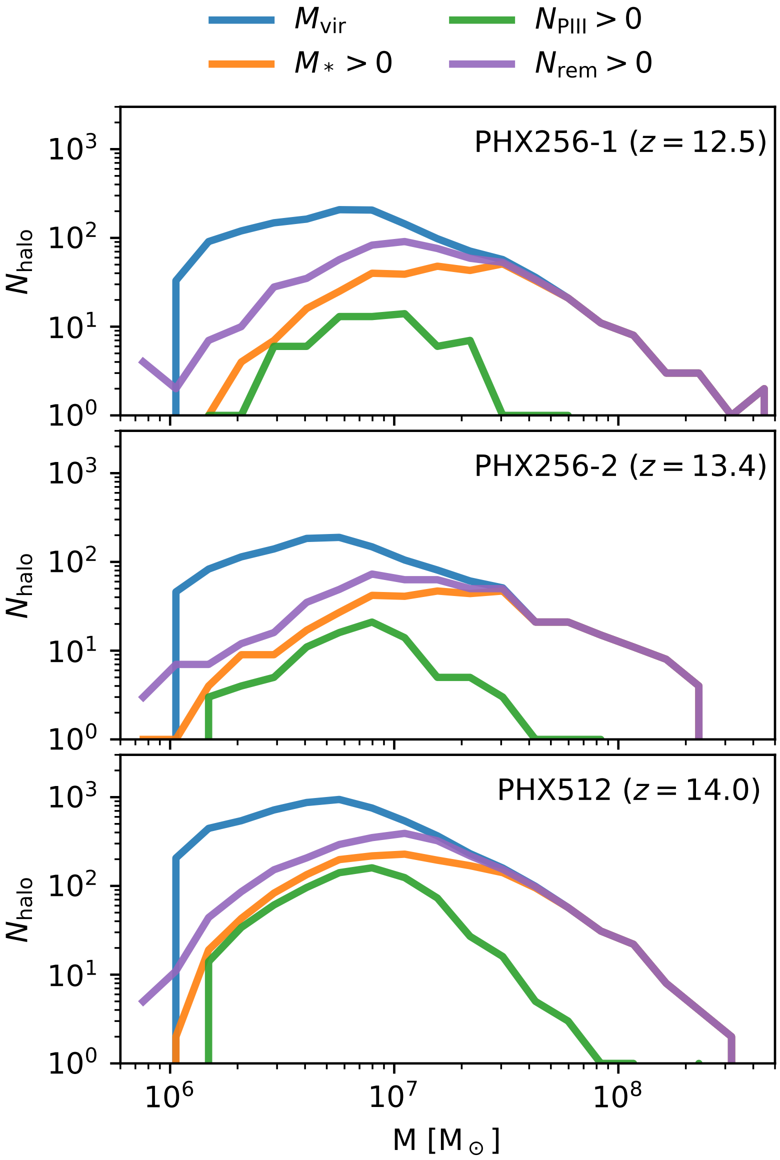


Figure 3.2. Halo counts at the final redshift of each simulation, with simulation and redshift annotated within each panel. $M_* > 0$ counts halos that have Population II stellar mass, $N_{PIII} > 0$ counts halos that contain active Population III stars, and $N_{rem} > 0$ counts those halos with Population III supernova remnants of any type. At the final redshift, halos with mass $M_{vir} \gtrsim 3 \times 10^7$ $\mathrm{M}_\odot$ all contain remnants, while some halos above $M_{vir} = 2 \times 10^7$ still host active Population III stars. Population III stars found in halos with $M_{vir} < 2 \times 10^6$ indicate that Population III formation does occur in under-resolved halos.

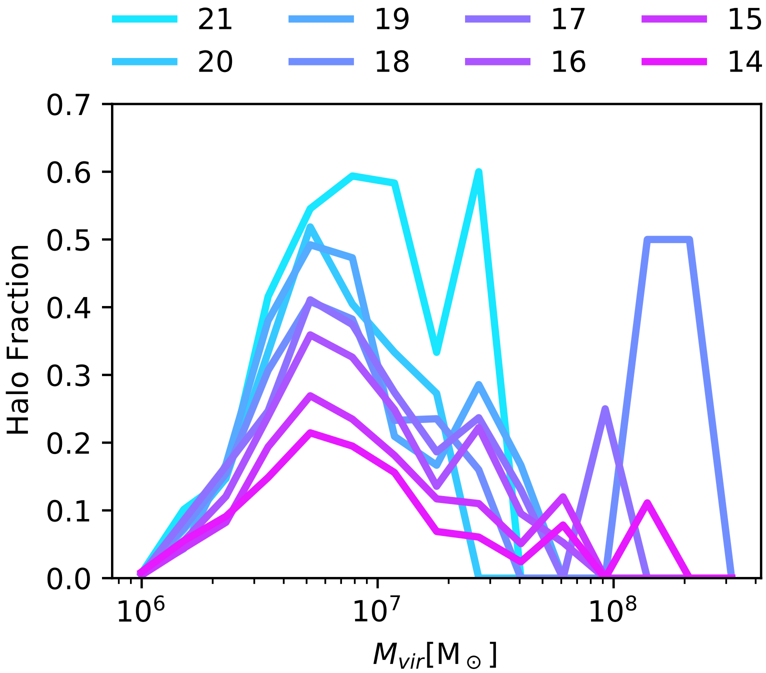


Figure 3.3. Fraction of halos occupied by active Population III stars for various redshifts, including all simulations. Halos with $M_{\rm vir} \sim 5 \times 10^6 \mathrm{M}_\odot$ are the primary hosts of Population III stars, and the highest mass halo with Population III stars is $M_{\rm vir} = 3 \times 10^8 \mathrm{M}_\odot$.

more frequency at lower redshift. The fraction of halos hosting Population III stars is highest for halos with $3 \times 10^6 \lesssim M_{\rm vir} \lesssim 2 \times 10^7$ $\mathrm{M}_\odot$: it is likely that halos with higher mass have generally been enriched by Population III stars earlier in their assembly history and no longer host high-density reservoirs of pristine gas to fuel primordial star formation. The visible outlier case is that of high-mass, high-redshift halos with $M_{\rm vir} \simeq 2 \times 10^8$ $\mathrm{M}_\odot$ at $z = 18$ with active Population III stars. These result from a lower ($M_{\rm vir} = 7.8 \times 10^7$ $\mathrm{M}_\odot$) mass halo with a low-mass Population III star merging into a larger halo–the Population III star did not form in the larger, $M_{\rm vir} > 10^8$ $\mathrm{M}_\odot$ halo. There are only two halos in this mass bracket at $z = 18$, and the above scenario happened in one of them to result in the high fraction observed in the plot.

Figure 3.4 shows the fraction of volume with gas enriched above varying cut-off Z as a function of z for $14.04 < z < 24$ in PHX512. The IGM is largely unenriched at the lowest redshift, as indicated by the low fraction enriched to $Z \gtrsim -5.5$. This suggests that early Population III stellar feedback is not responsible for large-scale enrichment of the IGM, having affected only $\sim 0.08\%$ of the simulation volume. We additionally analyzed the fraction of volume enriched by Population III versus Population II sources: if the Population III enriched fraction is f_3, and the fraction from all stars is f_Z, the quantity $\delta f = (f_Z - f_3)/f_Z$ shows the

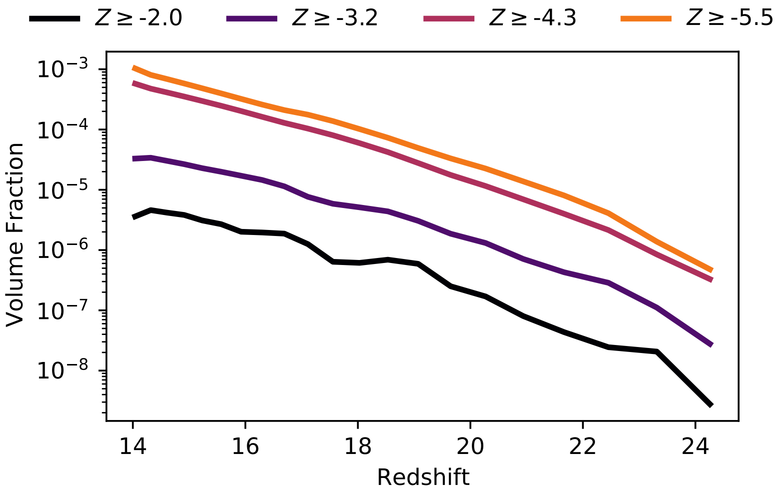


Figure 3.4. Fraction of volume enriched above varying Z values in PHX512 according to redshift. Less than 0.1% of the volume has been enriched to Z_c required for Population II star formation.

fraction of gas enriched by Population II sources. Enrichment from Population II stars has pervaded even less of the volume: we find that $\lesssim 2\%$ of enriched gas is enriched by Population II sources at the final redshift, emphasizing the importance of Population III chemical enrichment at high redshifts.

The fraction of volume that is ionized to varying degrees in PHX512 is shown in Figure 3.5. The ionized fraction is presented as $f_{\text{ion}} = n_{\text{H}^+}/(n_{\text{H}} + n_{\text{H}^+})$; less than 0.5% of the volume is ionized to $f_{\text{ion}} > 0.5$ at the final redshift, indicating that the point-source radiation feedback from ionizing sources has not yet escaped the dense clumps of halo or galactic gas. The highest levels of ionized gas, $f_{\text{ion}} \gtrsim 0.9$, likely result from ionizing radiation from Population III stars at these redshifts, with subsequent drops in f_{ion} resulting from recombination after the Population III main-sequence phase. Lower to middling values, $f_{\text{ion}} \leq 0.5$, reflect the increasing volume affected by hydrodynamic shock-heating from infalling gas and evolved Population III supernova remnant gas, as seen in the halo zoom-in temperature panel of Figure 3.1.

Star formation statistics are presented in Figure 3.6. The star formation rate densities (SFRDs) before $z \simeq 22$ are similar for Population II and Population III, however the SFRD of Population II eclipses that of Population III stars by $z = 22$. This is also the point at which the total mass in Population II stars surpasses that of the cumulative formed Population III mass. The Population II SFRD is fit by $SFR_{\text{est}} = 0.65 \exp(-z/2.45)$ $\text{M}_\odot$ yr^{-1} Mpc^{-3}, shown on the figure. The $z \lesssim 18$ Population III SFRD approaches a relatively constant 10^{-4} $\text{M}_\odot$ yr^{-1} Mpc^{-3}, representing the contribution from both newly collapsing halos in pristine gas or pristine regions of partially-enriched halos; since so little of the volume has been enriched, this will continue

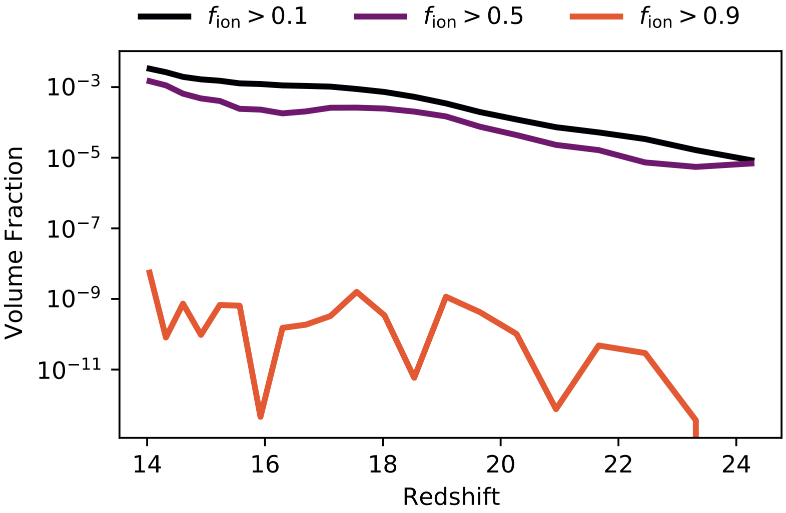


Figure 3.5. Fraction of volume that is ionized to varying degrees for PHX512, as measured by $f_{\text{ion}} = n_{\text{H}^+}/(n_{\text{H}} + n_{\text{H}^+})$. At the final redshift, a negligible fraction of the volume is fully ionized ($f_{\text{ion}} > 0.9$), indicating that IGM reionization has not begun to any substantial degree.

until more IGM gas has been enriched and the newly collapsed halos are enriched prior to their formation (Hicks et al., 2021).

The stellar mass weighted metallicity distribution function (MDF) for Population II stars across all PHX simulations is presented in Figure 3.7. The MDF of all Population II stars shows a large range of possible metallicities, with the most metal deficient cluster having $Z = -7.1$. Although the simulation parameter for Population II cluster creation requires $Z > -5.5$ at the cell hosting cluster formation, the mass-averaged metallicity of the gas that contributed to star formation allows these low-Z clusters to exist. The highest Z cluster has $Z = 0.57$; the mean metallicity of all Population II stars is $\langle Z \rangle = -1.89$. The unfilled blue curve shows only Population II.1 stars, with mean $\langle Z \rangle = -2.46$. For an observational reference point, we include the distribution retrieved from the JINA database (Abohalima & Frebel, 2018), shown in orange[4]. The overlap between the JINA observations and Population II stars is reassuring, however we also note that our Population II.1 stars can in general have metallicities above the JINA observations. This comparison implies that a great many Population II.1 stars may go unrecognized due to the *a priori* assumption that Population II.1 stars necessarily have low metallicity. Indeed, the MDF of Population II.1 stars and Population II.2 stars

---

[4] The JINA database does not quote mass-abundance metallicity, as it is tracked in Enzo. Assuming a solar metal abundance mass ratios with fixed H abundances, [Fe/H] = $\log(M_Z/M_H)_*$ - $\log(M_Z/M_H)_\odot$ = $\log(Z_*/Z_\odot)$. Despite that this approximation will not always be true, e.g., carbon-enhanced metal-poor stars common at low metallicity (Frebel & Norris, 2015), we quote the values as presented in the database to provide a point of comparison against the simulated metallicities.

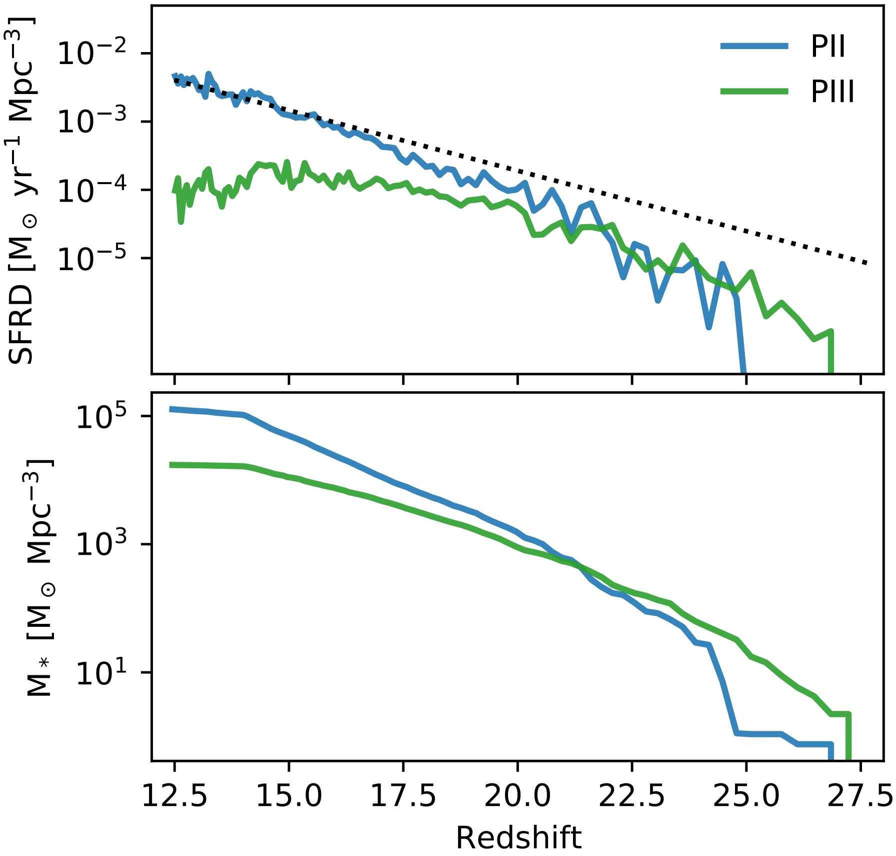


Figure 3.6. Star formation rate densities (top) and cumulative mass formed (bottom), volume normalized across all PHX simulations. Population II SFRD and total mass surpass that of Population III stars by $z \sim 22$. The black dotted SFRD line shows an approximate fit to the Population II SFRD as $SFR_{\rm est} = 0.65 \exp(-z/2.45)$ $M_\odot$ yr^{-1} Mpc^{-3}. Note that the change in slope for cumulative stellar mass past $z = 14$ is due to the limited redshift range of PHX512

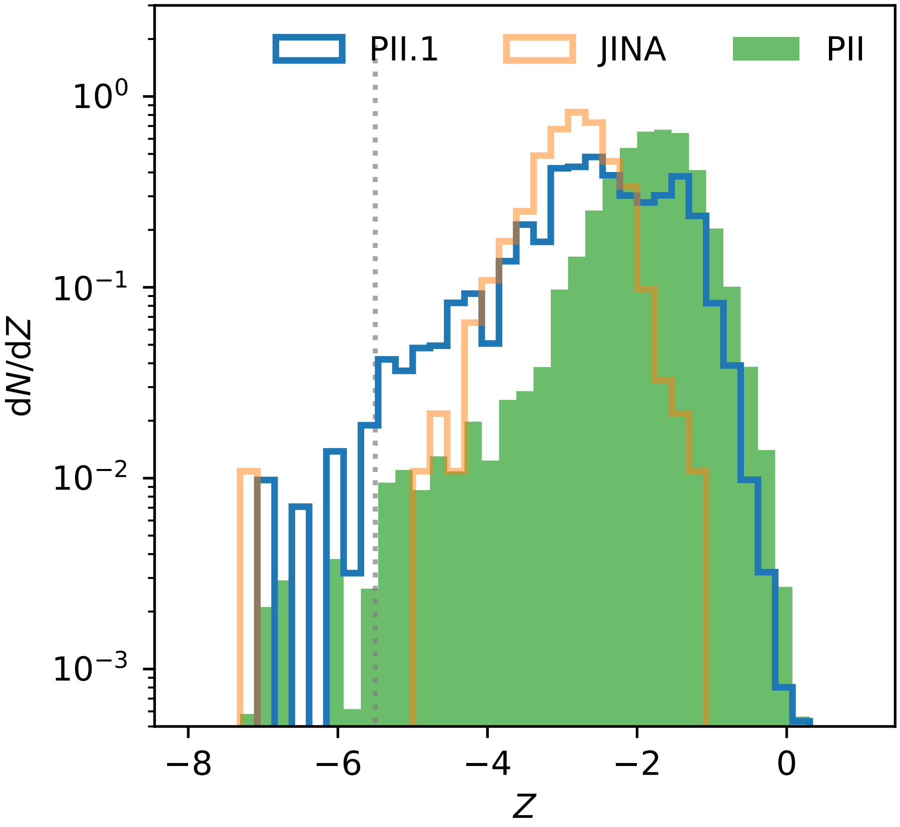


Figure 3.7. Population II metallicity distribution including all PHX simulations. Solid: The MDF of all Population II stars at the final redshift across all simulations. Blue, unfilled: The MDF of all Population II.1 stars from all simulations. The vertical line denotes the simulation $Z_c = -5.5$. Orange, unfilled: Sample MDF of observed low-metallicity stars from the JINA database (Abohalima & Frebel, 2018), including all stars from the Milky Way halo or dwarfs; the upper end ($Z > -2.5$) of this MDF is suppressed due to selection effects, and is not expected to resemble the simulation data. Notably, there is no obvious metallicity distinction between Population II.1 and Population II.2 clusters.

is completely overlapping, suggesting that while cases of stars with low metallicity may indicate very old Population II.1 stars, many more may have high, even super-solar, metallicity. Having these high-metallicity cases excluded from studies that attempt to reconstruct the Population III IMF will make it much more difficult to decipher the full range of Population III masses and their relative frequency.

3.5 Analysis: The First Stars and the Second Generation

To study the origin of Population II.1 star formation, we take two primary frames of reference: A) we analyze Population III star forming regions, studying the evolution of the region as it leads to Population II star formation, and B) we examine the region about the Population II.1 cluster immediately after formation, to study the events that immediately contributed to its formation. Each frame of reference uses separate analyses of the simulations.

3.5.1 Method: Population III Frame

Population III star formation within the PHX suite is clustered, which is not unexpected (e.g., Stacy et al., 2010). If we follow a single Population III star forming region as defined in this section, we find that there can be $\lesssim 200$ individual Population III stars per region. Although clustered, there are not generally enough individual stars to qualify as a star cluster in the canonical sense: we will therefore refer to them as PIII associations to avoid ambiguity with modern or Population II star clusters, but in analogy to modern O-B associations.

To examine the simulations from the perspective of Population III stars and define the extent of a group of coeval primordial stars (PIII assocation), we iterate each output of the simulation to find new Population III star particles that formed within 200 kyr by iterating dark matter halos, searching for new particles within $3R_{\mathrm{vir}}$ of the center of the halo. When found, and if that star is only accompanied by coeval Population III star formation (i.e., other new Population III stars only formed within 200 kyr, no supernova remnants, black holes, or Population II stars within the radius), we form a sphere to measure mean metallicity from new Population III stars ($\bar{Z}_{\mathrm{III}}$) and H^+ fraction($\bar{f}_{\mathrm{H+}}$), using the analysis software `yt` (Turk et al., 2011). Starting from a smaller radius than expected for either metal-rich bubbles or ionized regions, $R = 250$ pc, the average value for a field, $\bar{X}$, is measured. If $\bar{X}$ is not less than some critical value, R is increased to $R_{i+1} = 1.1R_i$, and a new average is taken. This procedure is repeated until the value of $\bar{X}$ is below our chosen critical values: $\bar{Z}_{\rm III} < -5.5$ and $\bar{f}_{\rm H+} < 0.05$. The "edge" of the region is then defined by this final radius. The final product of this analysis is an effective radius as function of time for the $Z_{\rm III}$ and $f_{\rm H+}$ variables.

3.5.2 Method: Population II.1 Frame

From the perspective of Population II clusters, we again iterate through simulation outputs to identify Population II star formation events. When a new Population II star particle is formed, and it occurs in gas enriched by only Population III stars, i.e., the metallicity from Population II stars, Z_2, meets the criteria $\bar{Z}_2 < -6$ in the cell hosting star cluster formation, we evaluate a sphere centered on that star with $r = 200$ comoving kpc. This initial radius assumption stems from the analysis of Section 3.5.1: greater than 95% of Population III star forming regions are contained within 8 proper kpc, corresponding to < 200 comoving kpc at $z = 20$. Within this sphere, we connect any Population III supernova remnant to the Population II cluster via a ray. We then verify that the ray has $Z > Z_c$ for all cells it intersects and require that distance between particles, d, satisfies $d \leq vt$ for $v = 100$ km/s SN remnant expansion speed, and t, the time between SN and Population II formation; $d \leq vt$ implies that metals from the Population III star could feasibly have reached the forming star cluster and acts as a filter for overlapping metal clouds from separate PIII associations. If these criteria are satisfied, then that Population III event is "connected" to the Population II formation and is considered to be part of the same "metal system". The extremely fine time resolution between outputs ensures that Population III events connected to the formed Population II cluster were connected at the time of formation (≤ 200 kyr prior).

3.5.3 The Population III Frame

The time evolution of metal and ionized radii of PIII associations including all PHX simulations is presented in Figure 3.8 using B-spline quantile regression fits[5] to the $q = \{0.2, 0.5, 0.95\}$ quantiles. After the initial formation, both radii take on small starting values, reflecting that $Z_{\rm III}$ is sourced from supernovae that have not occurred yet, and that H ionization requires time to first ionize the dense cloud that the particle formed within. The average radius for $Z_{\rm III}$ ($R_{\rm III}$) is maximum at 9-12 Myr after the first formation: This is the time-frame expected for most Population III stars to have reached their main sequence endpoint if we are considering a coeval system of stars formed at $t = 0$. The reduction in $R_{\rm III}$ likely reflects gravitational collapse, with the radius supported at later times by increased temperatures and the feedback from Population

---

[5] Using the python `patsy` package, https://patsy.readthedocs.io

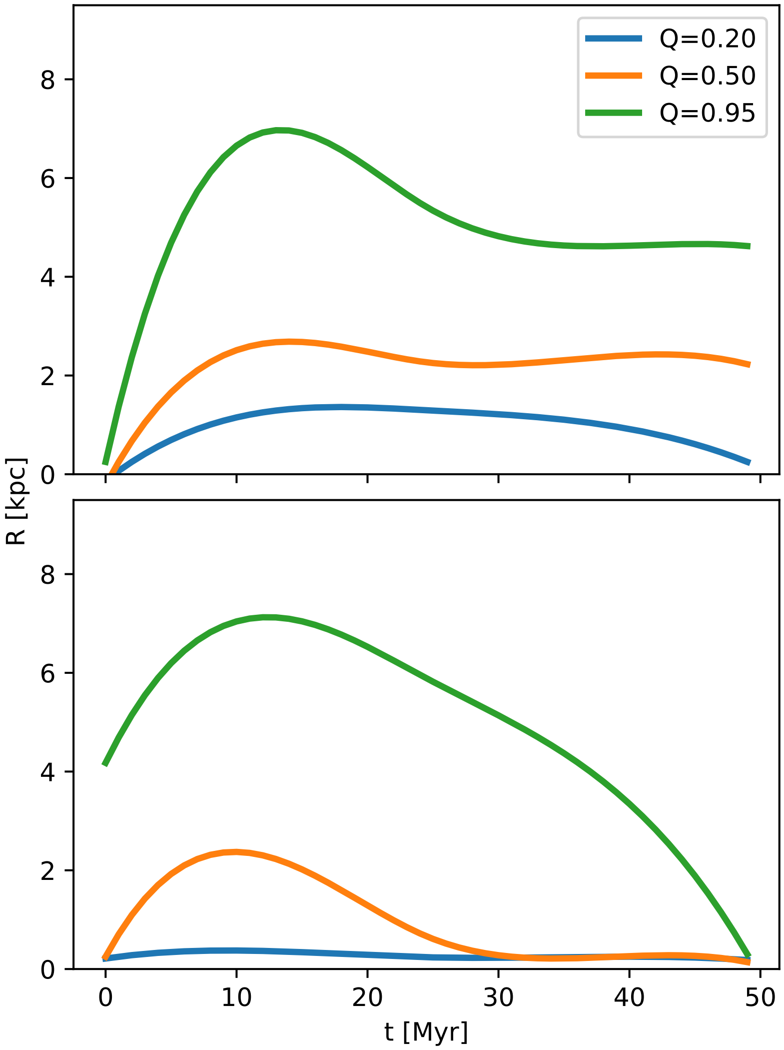
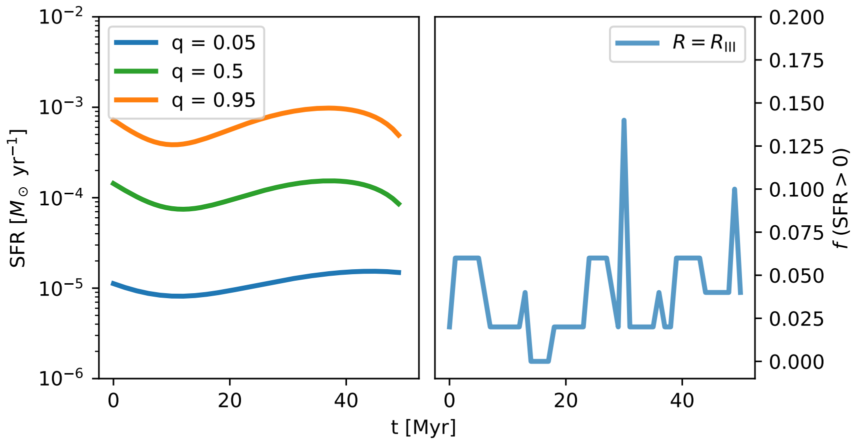
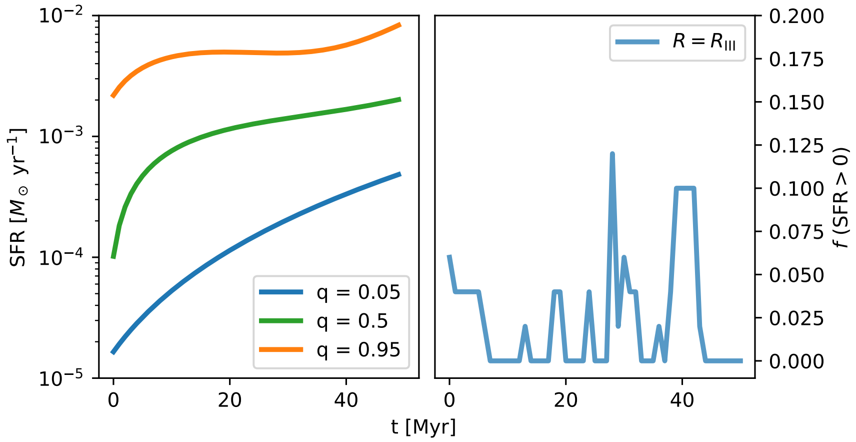


Figure 3.8. Population III star forming region characteristics spanning the first 50 Myr past the first particles' formation including data from all PHX simulations. Top: radius of the metal-rich bubble (top) Bottom: H^+ ionized region radius. B-spline quantile regression fits are shown, with $q = \{0.2, 0.50, 0.95\}$. We present $q = 0.2$ instead of the symmetric value of $q = 0.05$ because the $q = 0.05$ quantile fit is nearly zero for all times. Considering the entire dataset for all times, 7% of points fall above the 0.95 quantile fit for the radius of metals, while ionization radius shows more scatter, with 17% of points above the 0.95 quantile fit.



(a) Population III SFR



(b) Population II SFR

Figure 3.9. Star formation rates from all PHX simulations as quantile fits to regions with non-zero SFRs. Most regions have zero SFR at a given time; we show the fraction with non-zero SFR in the right panel of each plot ($f(\mathrm{SFR} > 0)$.

II stars. The radius of H ionization, $R_{\mathrm{H+}}$, increases quickly to a maximum average within 10 Myr. The reduction beyond this point reflects recombination after the most massive and ionizing Population III sources of radiation have extinguished. Since the quantiles represent fits, they do not exactly bound the, e.g., 0.95 quantile of observed data: 7% of R_{III} fall above the 0.95 quantile fit, while 17% of $R_{\mathrm{H+}}$ fall above.

In Figure 3.9 (left panel), we present quantile fits to the SFR of 3,313 distinct primordial star forming regions considering regions from all PHX simulations. The SFR at these early times is roughly constant for Population III stars, i.e., regions that are creating stars are doing so at similar rates, regardless of time. In contrast, the Population II SFR is increasing in time, with $q = 0.5$ approaching 10^{-3} $\mathrm{M}_\odot$ yr^{-1} by 50 Myr. In the $q = 0.95$ fit, 5.05% and 5.07% of regions fall above the 0.95 quantile fit for Population III, Population

Table 3.2. Progenitors of Population II.1 cluster formation

Configuration	$N_{\rm II}$	$f_{\rm II}$	$\langle Z \rangle$	$\langle R_f \rangle$ [pc]	$N_{\rm III}$	$M_{\rm III}[{\rm M}_\odot]$
SN	17	0.011	-3.27 ± 1.15	95.38 ± 101.56	2.3 ± 1.1	35.0 ± 16.0
HN	188	0.045	-2.79 ± 0.64	67.19 ± 102.57	2.3 ± 1.9	73.6 ± 54.4
PISN	618	0.144	-1.54 ± 0.57	27.55 ± 17.88	1.1 ± 0.3	201.3 ± 71.6
SN-HN	356	0.266	-2.77 ± 0.85	225.51 ± 404.98	12.0 ± 8.1	286.4 ± 204.4
SN-PISN	11	0.008	-2.75 ± 1.71	162.64 ± 146.85	3.2 ± 1.5	223.8 ± 54.8
HN-PISN	133	0.032	-2.06 ± 0.83	92.92 ± 153.19	5.1 ± 3.7	352.6 ± 171.9
SN-HN-PISN	787	0.494	-2.27 ± 0.94	558.50 ± 815.70	21.9 ± 15.6	829.8 ± 504.8

Note: Characteristics of 2,110 Population II.1 star clusters from all PHX simulations given the connected Population III events within the region. We categorize each region containing a new Population II star by the type of Population III supernovae it contained: SN, HN, PISN, or combinations thereof. For each configuration, we present the number of Population II.1 stars formed ($N_{\rm II}$, the fraction of Population II.1 stellar mass generated ($f_{\rm II}$), the mean metallicity ($\langle Z \rangle$), the mean radius from progenitor to forming cluster (R_f), the mean number of progenitors in the region ($N_{\rm III}$), and the mean total progenitor mass ($M_{\rm III}$). Note that 61.0% (by mass) of Population II.1 stars have progenitors of multiple types.

II SFRs respectively. In the right panel, we show the fraction of regions that are forming stars at each time. Within the first 50 Myr, the Population II SFR is sporadic, with $> 87.5\%$ of regions having zero SFR at any time, including several intervals where no region is forming Population II stars. The Population III SFR is simiarly sporadic, with $< 85\%$ of regions forming stars at any given time. The Population III star formation rate is surprisingly constant throughout the 50 Myr period for regions that are forming Population III stars, however, we do observe that the average Population II SFR ($q = 0.5$) increases with the region age. The early non-zero Population II SFR results from prompt star formation around the earliest and most massive Population III stars.

3.5.4 The Population II.1 Frame

Generated using Section 3.5.2, we present the statistics of Population II.1 star clusters including all PHX simulations immediately after the cluster's formation in Table 3.2. The region surrounding the Population II.1 star particle is categorized based on the type of Population III progenitors it contains: SNe, HNe, PISNe, or any combination of the three. There are several immediate observations worth noting: for single progenitor-type regions, there is an inverse correlation between the distance between progenitor and Population II.1 star, R_f, and the SN energy of the progenitor, while the metallicity of the resulting Population II.1 cluster is correlated to the energy. The mass of the Population III stars that contributed to formation is highly variable; and finally, 14.4% of Population II.1 stars were enriched by an average of $N_{\rm III} = 1.1$ PISN, while other Population II.1 stars are enriched by $N_{\rm III} > 2$ Population III progenitors. The average distance from progenitor to Population II.1 cluster is maximized if all types of Population III progenitors are present and connected to the Population II.1 star, in fact, we find generally higher R_f for the more complicated regions with high $N_{\rm III}$ and many progenitor types. Outside the single-type of progenitor cases of the SNe and HNe, the average Population II.1 star is connected to $M_{\rm III} > 200$ $\rm M_\odot$ of Population III progenitors, with some Population II.1 stars connecting to $M_{\rm III} > 1000$ $\rm M_\odot$ of Population III supernova generating stars. Interestingly, the mean metallicity of Population II.1 clusters enriched by PISNe is above the observed metallicity in, e.g., the JINA database samples presented in Figure 3.7, suggesting that observational campaigns seeking Population II.1 stars by their low metallicity will miss a very large fraction of these Population II.1 stars. In fact, the metallicity of regions with PISNe, HN-PISN, and SN-HN-PISN progenitors all have mean metallicity falling above the selection effect cutoff of the JINA database ($Z \gtrsim -2.5$).

We do note that $M_{\rm III}$ and $N_{\rm III}$ are both highly dependent on the chosen IMF, and this influence will definitely have an impact on R_f. We also note that the lack of correlation between cluster metallicity and progenitor configuration may be an artifact of the star formation algorithm: since the Population II clusters take on the mass-averaged metallicity of the cold gas that formed them, we lose the ability to track extreme cases that may occur with higher resolution and modelling Population II stars as single stars. That said, within the framework of the PHX simulations, the highest metallicity star clusters $(Z = -1.57)$ result from regions with only PISNe, while the lowest metallicity $(Z = -3.27)$ have SNe progenitors. 80% of Population II.1 stars have multiple types of Population III supernovae as progenitors, suggesting that theoretical constraints on the Population III IMF, as measured by metal spectra of Population II.1 stars, will require determining the metallicity contributions from each progenitor type.

The combination of high metallicity and small radii from PISNe is interesting. In particular, there are two feasible avenues that could result in Population II.1 stars forming in such close proximity to a PISN remnant: the Population II.1 stars could be the result of triggered star formation, where the expanding SNR causes nearby regions to qualify for star formation. Alternatively, the Population II.1 stars could be forming long after the SNR, after the gas has re-collapsed into the original region that hosted the PISN. We find that the average time between a PISN and the Population II.1 cluster formation is $\Delta t = 1.26$ Myr, with a significant number of Population II.1 formation events having $\Delta t \leq 0.25$ Myr, suggesting that these Population II.1 stars can result from prompt star formation very close to the SNR. Our measured Δt is similar to high-resolution simulations of prompt star formation (Chiaki et al., 2013), however, we observe scatter in Δt due to the varying density and halo masses that occur organically in the PHX. This phenomena is not observed in the SNe case, and is less prevalent for HNe, which implies that the extreme energy of the PISN is instrumental in early prompt Population II.1 formation at the resolution of the PHX simulations.

3.6 An Interpretable Regression Model of Primordial Stars' Influence

In this section, we develop and describe an interpretable model designed to predict the range of influence of PIII associations. Using data generated via the method of Section 3.5.1, we create a model to learn the extent of primordial metals from a PIII assocation based on the composition of SN events and the duration we wish to model, $t_{\rm final}$. To translate the continuum of available Population III masses and creation times, we transform the stellar information into a simple series of features as binned mass and creation times. The mass features ($\boldsymbol{X}_M$) are defined by edges $\{1, 11, 20, 40, 100, 140, 200, 260, 300\}$ M$_\odot$, and creation time bins ($\boldsymbol{X}_C$) have edges $\{0, \delta t, 2\delta t, ..., n\delta t\}$ Myr, with the time between bin edges (δt) and the final edge defined by $n\delta t = {\rm floor}(t_{\rm final}/\delta t)\delta t$. As a concrete example of mass and creation time features, consider a region with five stars. They have masses $\{14, 12, 86, 94, 210\}$ M$_\odot$ with creation times $\{4, 4, 3, 0, 7\}$ Myr, as measured from the first star created. With this hypothetical sample, $\boldsymbol{X}_M = \{0, 2, 0, 2, 0, 0, 1, 0\}$. If we take $\delta t = 6$ Myr with $t_{\rm final} = 16$ Myr, the time features bins are $\{0, 6, 12\}$ Myr, so that $\boldsymbol{X}_t = \{4, 1, 0\}$ Myr. This idea is motivated by a method commonly known as tokenization; we use it here to create an input that can accommodate regions with any number of stars without altering the model. We split samples in a spatial sense to prevent information leaking between training and testing splits: if the center of the region ($\boldsymbol{r}$) at the first star's formation is $\boldsymbol{r} > \{0.4, 0.4, 0.5\} \times L_{box}$ relative to the simulation volume, then the sample is assigned to the test split, $\boldsymbol{r} < \{0.6, 0.6, 0.5\} \times L_{box}$ is assigned to validation, while all others are assigned to training. This splitting of data yields 2,273 training, 722 validation, and 318 testing samples.

3.6.1 Data

Figure 3.10 shows $\log(R_{\rm III})$ (with $R_{\rm III}$ measured in kpc) according to the number of Population III stars (N_*) within the region for the training examples at $t_{\rm final} = 16$ Myr. The side histogram shows a PDF of $\log(R_{\rm III})$ as measured in kpc; the distribution of $\log(R_{\rm III})$ is distinctly more evenly distributed about the peak ($\log(R_{\rm III}) \simeq 0.3$) than without the logarithm: we therefore make our predictions on $\log(R_{\rm III})$, as regression

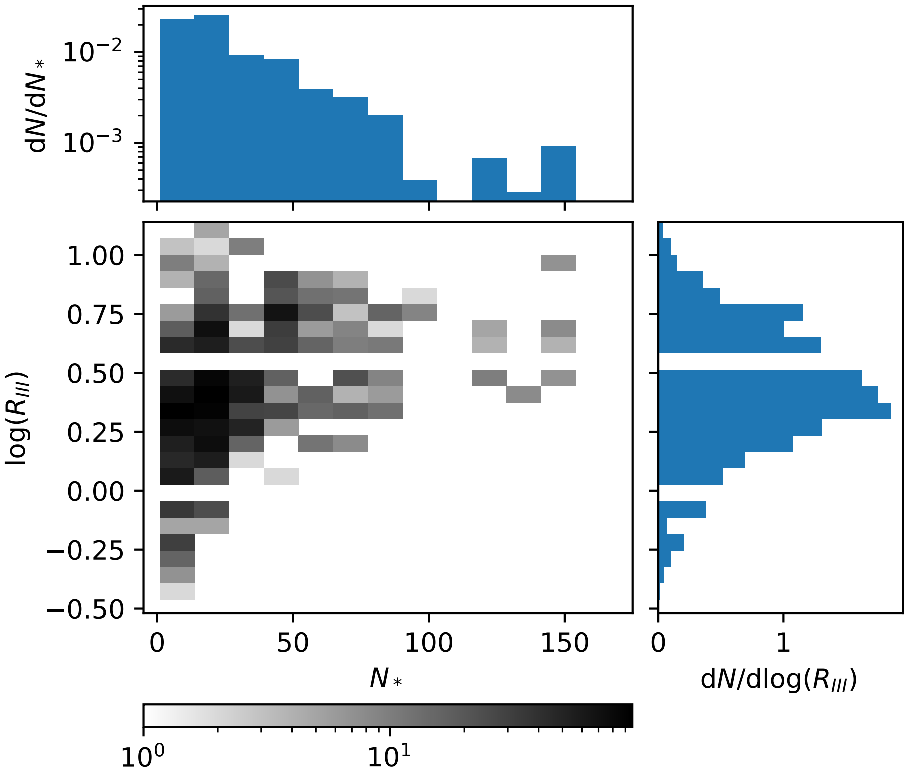


Figure 3.10. Training data for $t_{\text{final}} = 16$ Myr after first Population III formation. Histograms show PDFs of the number of Population III stars, N_* (top), and radius of metal influence in kpc, $\log(R_{III})$ (side). Counts of regions with N_*, $\log(R_{\text{III}})$ shown in center plot are given by the color bar. The gaps in the histogram of $\log(R_{\text{III}})$ are an artifact of the sampling method described in Section 3.5.1.

is more successful at modelling scatter from a mean than long tails of asymmetric distributions. We also observe significant variation in the number of Population III stars contained within R_{III}; regions contain $1 < N_* \lesssim 155$ within that radii. For each t_{final}, we define $\bar{N}_* = \mu(\log N_*)$, and $\sigma_N = \sigma(\log N_*)$[6], then for 16 Myr, $\bar{N}_* = 1.337$ and $\sigma_N = 0.374$. In addition, while we observe a general trend that increasing N_* increases R_{III}, however this relationship is by no means linear in N_*, and is not single-valued, i.e., $N_* = 10$ could lead to many values in the range $-0.5 < \log(R_{\mathrm{III}}) < 0.8$. $\bar{N}_*$ and σ_N have a very small dependence on t_{final}: when fit with a simple polynomial, we find that they are fit with $\lesssim 1\%$ error by

$$\bar{N}_*(t_{\mathrm{final}}) \simeq 1.197 + 0.0195 t_{\mathrm{final}} - 6.57 \times 10^{-4} t_{\mathrm{final}}^2 \tag{3.3}$$

$$\sigma_N(t_{\mathrm{final}}) \simeq 0.314 + 7.25 \times 10^{-3} t_{\mathrm{final}} - 2.87 \times 10^{-4} t_{\mathrm{final}}^2. \tag{3.4}$$

3.6.2 Model and hyperparameters

To generate an interpretable model, we use a simple linear regression model, where we wish to minimize the error

$$L = \frac{1}{2}(\boldsymbol{Y} - \hat{\boldsymbol{Y}})^2, \tag{3.5}$$

with ground-truth radii $\boldsymbol{Y}$ and predicted radii $\hat{\boldsymbol{Y}}$ represented as vectors. The prediction is simply solved by $\hat{\boldsymbol{Y}} = \boldsymbol{X} \cdot \boldsymbol{w}$. The learnable weights are denoted by $\boldsymbol{w}$, and $\boldsymbol{X}$ is composed of mass and creation time features with a bias feature: $\boldsymbol{X} \equiv \{1, X_M^1, X_M^2, ..., X_M^8, X_C^1, X_C^2, ..X_C^{\mathrm{final}}\}$, with, e.g., X_M^n referring to the n^{th} feature in the $\boldsymbol{X}_M$ vector. Since equation 3.5 is convex, the solution we seek is $dL/d\boldsymbol{w} = 0$, which is solved by

$$\boldsymbol{w} = (\boldsymbol{X}\boldsymbol{X}^T)^{-1}\boldsymbol{X}^T\boldsymbol{Y}. \tag{3.6}$$

This regression only admits linear behavior, which as outlined in Section 3.6.1, is not a reasonable expectation in this case. The non-linearity is not only influenced by N_*, but also their varying lifetimes, explosion energies, metal yields, birth times, and supernovae times. Instead of abandoning the linear model for a less interpretable deep learning architecture, we split the training samples into subsets based on N_* and apply the

---

[6] μ and σ represent the mean and standard deviation of the distribution, respectively

linear model to each subset: we use the samples $\mathbf{X}_n$ to train only the model M_n according to

$$M_n = \begin{cases} M_0, & \mathbf{X} | \log N_* < \bar{N}_* - \sigma_N \\ M_1, & \mathbf{X} | \bar{N}_* - \sigma_N < \log N_* < \bar{N}_* \\ M_2, & \mathbf{X} | \bar{N}_* < \log N_* < \bar{N}_* + \sigma_N \\ M_3, & \mathbf{X} | \log N_* > \bar{N}_* + \sigma_N . \end{cases} \tag{3.7}$$

In this way, we have piece-wise defined models specifically for small systems with few stars, average systems, and highly populated systems.

What follows is a concrete example of using this model for observed data, as used in the rest of this paper, or for predictions. For this example, we choose $\delta t = 6$ Myr and $t_{\text{final}} = 16$ Myr. To generate a sample, we statistically estimate the number of stars by sampling a random number described by a log-normal distribution with $\bar{N}_*(16\text{ Myr})$ and $\sigma_N(16\text{ Myr})$. Given the number of stars in this sample, we then determine the mass of each star by sampling the IMF described in Section 3.3 and determine the creation time of each star by sampling the SFR of Figure 3.9a. Using this information, we can build $\boldsymbol{X}$ as described above and determine which M_n to use. Using the weights of the M_n model, $\boldsymbol{w}_n$, we can finally predict the radius of the enriched region as $\log R_{\text{III}} = \boldsymbol{X} \cdot \boldsymbol{w}_n$.

Finally, we can use the weights to gain insight into the relative importance of different features of $\boldsymbol{X}$. For example, in Table 3.4, we present the final trained weights of the M_n models for $t_{\text{final}} = 16$ Myr: by examining the relative values of w_1 (which multiplies the 1-11M$_\odot$ feature) with w_6 and w_7 (which multiply the 140-200 and 200-260 M$_\odot$ features respectively), we observe that $w_6 \gg w_1$ and $w_7 \gg w_1$, implying that the 140-260 M$_\odot$ stars have a much stronger effect on the final radius of the metal cloud. This is, of course, intuitively true as well: 140-260 M$_\odot$ stars generate PISNe, which distribute more metals with more supernovae energy than any other mechanism in these simulations.

To evaluate the model predicted radii, we employ two metrics: we use the R_2 score for N samples, given by

$$R_2(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) = 1 - \frac{1}{N} \frac{\sum_{i=1}^{N} y_i - \hat{y}_i}{\sum_{i=1}^{N} y_i - \bar{y}}, \tag{3.8}$$

for $\bar{y} = 1/N \sum_{i=1}^{N} y_i$ and y_i are the individual components of $\boldsymbol{Y}$. The R_2 score is informative in comparing the

Table 3.3. Testing dataset performance varying modelled time.

Time	R_2	J	$R_2/(1+J)$	$\bar{L}_1$
7 Myr	0.493	0.205	0.409	0.192
8 Myr	0.600	0.150	0.521	0.147
9 Myr	0.503	0.187	0.424	0.158
10 Myr	0.438	0.133	0.386	0.161
11 Myr	0.448	0.196	0.374	0.152
12 Myr	0.331	0.189	0.278	0.158
13 Myr	0.362	0.216	0.298	0.159
14 Myr	0.408	0.202	0.339	0.167
15 Myr	0.356	0.230	0.290	0.174
16 Myr	0.499	0.190	0.419	0.153
17 Myr	0.383	0.212	0.316	0.162
18 Myr	0.393	0.167	0.337	0.162
19 Myr	0.385	0.213	0.317	0.161

Note: These models use $\delta t = 6$ Myr. The optimal model will maximize R_2, while simultaneously minimizing J: the shown quantity $R_2/(1+J)$ will be maximized for the best performing models.

quality of the model as compared to simply predicting the mean value for of $\boldsymbol{Y}$; 1 is a perfect R_2 score, $R_2 = 0$ indicates predicting $\hat{y} = \bar{y}$ for all inputs, and $R_2 < 0$ indicates arbitrarily worse performance. In addition to the individual predictions, we would also like the PDF of predicted radii to match that of the ground truth radii. We define the PDF of $\boldsymbol{Y}$ as P_y, $\hat{\boldsymbol{Y}}$ as $P_{\hat{y}}$, and the Kullback-Leibler divergence $D(P|Q) = P\log(P|Q)$ for probability distributions P and Q, to compare the PDFs using the Jensen-Shannon distance given by

$$J(P_y, P_{\hat{y}}) = \sqrt{\frac{D(P_y|\bar{P}_y) + D(P_{\hat{y}}|\bar{P}_y)}{2}}, \tag{3.9}$$

where $\bar{P}_x$ represents the average of the distribution P_x. $J = 0$ represents two identical distributions, with higher values indicating mismatches in the PDFs. Although not used in the initial evaluation of models, we include reports on the average L_1 distance as $\bar{L}_1 = 1/N \sum_{i=0}^{N} |y_i - \hat{y}_i|$.

3.6.3 Model results

The only true hyperparameter of this model is δt, the time width of the $\boldsymbol{X}_C$ bins. We tested several widths, and found no significant improvement beyond the inclusion of two time features: early coeval formation, and all other formation, so all results here use $\delta t = 6$ Myr. J and R_2 as functions of predicted time in the all datasets are shown in Figure 3.11, while tabulated quantification of errors are presented in Table 3.3. All times have $0.12 < J < 0.23$, while R_2 has more variation with $0.33 < R_2 < 0.60$. The overall performance

Table 3.4. Final parameters for linear regression models with 6 Myr time bin width and $t_{\text{final}} = 8, 16$ Myr.

t_{final}	$\boldsymbol{w}(M_0, M_1, M_2, M_3)$
8 Myr	-0.439249, -0.044625, -0.029246, 0.073477, 0.027912, -0.045580, 0.263906, 0.429911, 0.133443, 0.030199 -0.055685, -0.027469, -0.014668, 0.030979, -0.016111, 0.006372, 0.191418, 0.329448, 0.073871, 0.009292 0.321754, -0.017447, -0.002649, 0.012913, -0.003512, -0.028743, 0.036156, 0.177121, 0.057620, 0.000865 0.450866, 0.000590, -0.003458, 0.004811, -0.000099, -0.007944, -0.013434, 0.062136, -0.012985, 0.001748
16 Myr	-0.191580, -0.093447, -0.027977, 0.001763, 0.016985, -0.065883, 0.148267, 0.337160, -0.042382, 0.047274, 0.098253 0.025503, -0.003428, 0.008574, 0.014688, -0.020077, 0.012643, 0.141135, 0.240971, 0.032025, 0.010795, 0.011036 0.226557, -0.012466, -0.000646, 0.021068, -0.000813, -0.026686, 0.038700, 0.146538, -0.009889, 0.000068, -0.002796 0.553367, -0.021249, 0.007781, 0.004529, -0.010159, -0.011814, -0.002303, 0.023666, 0.024624, 0.000682, -0.000766

Note: These exemplary weights are only for 8 Myr and 16 Myr, however the data table with full machine precision for all models presented in Table 3.3 will be available at www.rensimlab.github.io.

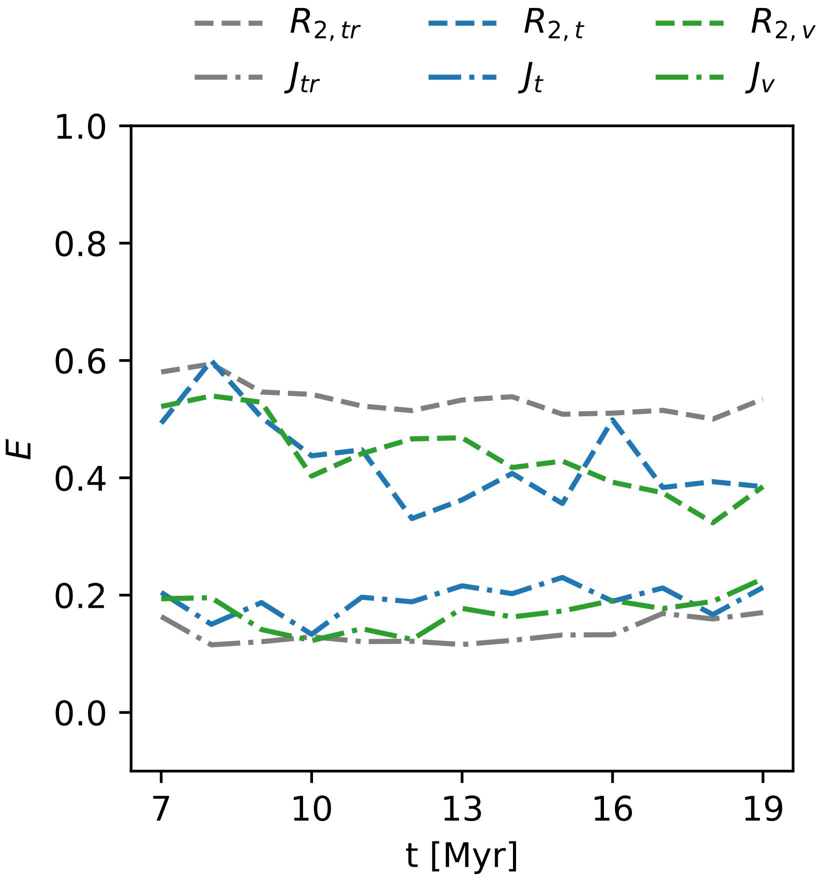
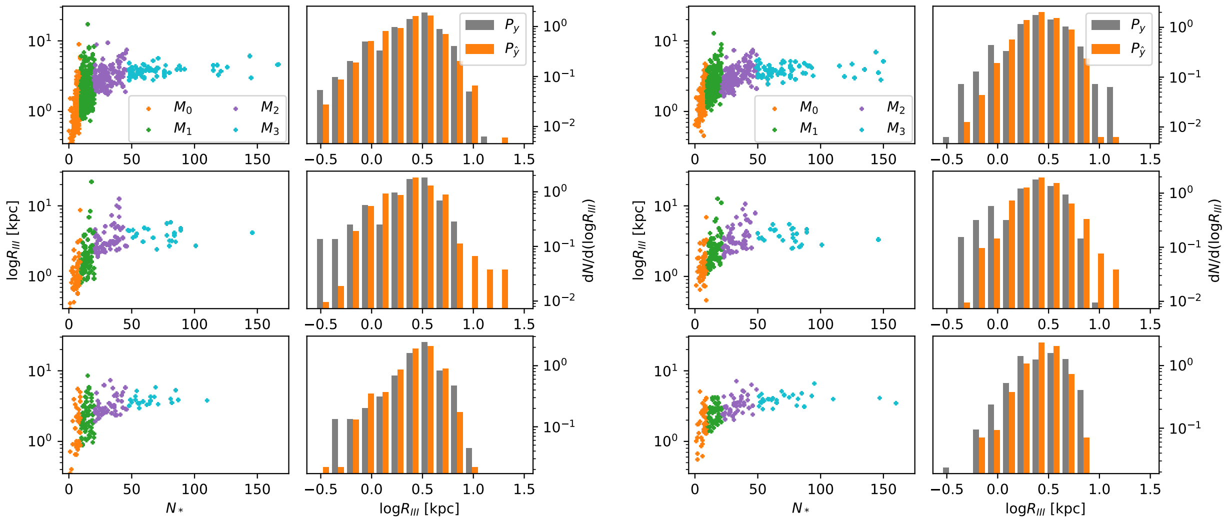


Figure 3.11. Error (E) as measured by J (dash-dot lines) and R_2 (dashed lines) on the training (grey), validation (blue) and testing (green) data splits. This method is generally capable if the modelled time is ≥ 7 Myr, but struggles by these metrics if $t < 5$ Myr, as the early dynamic evolution and dependence on specific hydrodynamic state of the system is making the modelling task more difficult at those early times.



(a) 8 Myr **(b)** 16 Myr

Figure 3.12. For predicting the metal radii at two timeframes, 8 and 16 Myr, we present the training (top), validation (middle), and testing (bottom) performance as plots of true radii (grey) with the predictions for each model (M_0, M_1, M_2, M_3 in orange, green, purple and cyan, respectively) according to the number of stars in the system (N_*). The Left histograms compare the PDF of true radii (grey) with model predictions (orange).

is quantified by combining R_2 and J as $R_2/(1+J)$, which will be maximized if R_2 is maximized and J is minimized. These results show that while the model may struggle, e.g., with the lowest R_2 score at $T = 18$ Myr, to reproduce exact predictions, it does well at reproducing the distribution of possible radii reflected in the relatively low value of $J = 0.167$. Based on Table 3.3, if we take the "best" performing model as the one that maximizes $R_2/(1+J)$, we find that models with $t_{\text{final}} = \{7, 8, 9, 16\}$ Myr are the best performing, while $t_{\text{final}} = \{12, 13, 15\}$ have the worst performance. The remainder of discussion in this section considers the 8 Myr and 16 Myr models; while 8 Myr is the most performant model, the 16 Myr model is a desirable time frame to evaluate: given the stellar lifetimes and IMF used in the PHX, we expect any coeval stars formed near $t = 0$ to have reached their respective endpoints by $t \sim 16$ Myr. We can also observe that the metal radius is maximized around 16 Myr as seen in Figure 3.8.

The final weights of the exemplary $t_{\text{final}} = \{8, 16\}$ Myr models are presented in Table 3.4. These weights represent the entirety of the trained model. Due to their simplicity, erroneous predictions could be traced through the model to find the offending weight and determine why the model made such an error very easily.

Exemplary results from the 8 Myr and 16 Myr model are shown in Figure 3.12. On the left side we plot the model predictions, color coded by which model made the prediction. While these models score well with $J = 0.189, 0.162$ for 8 and 16 Myr respectively, we can still observe mismatches on the tails of the PDFs for low and high $\log R_{III}$. However, it is reassuring that there are no erroneous massive predictions given by the validation or testing dataset, e.g., predicting $R_{III} > 50$ kpc. This is largely due to our splitting of the dataset among different models–using a single model for all N_* leads to linear fits that had erroneously high estimates of R_{III} at high N_*, however the M_3 sub-model, which models high N_*, has a nearly flat slope, reducing the errors from very high N_* systems. As well, a single model struggled to reproduce the low-end of radii from very low N_* systems, making no predictions of $R_{III} < 1$ kpc. Restricting M_0 to predict low-N_* systems allows the fit to have a high slope that can predict both $R_{III} < 1$ kpc and $R_{III} \gtrsim 10$ kpc, while still avoiding falsely large regions that may happen if the same model was responsible for predictions in the higher N_* regime.

In the scatter plots, we can identify that the high predictions stem from the M_1 and M_2 models, however even these results are plausible for the radius of the metal bubble. As a final stress-test of the model, we generated 3000 samples that match the N_* and mass distribution of the training dataset at 16 Myr, using creation times derived from the SFRs presented in Figure 3.9a. We compare the PDF of the synthetic sample and full dataset in Figure 3.13, where the full dataset is in grey and synthetic dataset in orange. There is no ground truth for the synthetic dataset (so there are no R_2, L_1, L_2 scores), but it produced a PDF that is in agreement with the ground truth dataset, with $J = 0.150$.

3.7 Conclusions

All halos in the Phoenix suite with $M_{vir} \geq 2 \times 10^7$ $M_\odot$ have Population III star remnants within their radii and have some finite mass of stars (either Population III or Population II). However, all three simulations also contain several halos without stellar mass above $M_{vir} \geq 10^7$ $M_\odot$; these may be candidates for further study to analyze whether they may be super-massive black hole candidates as in the analysis of Regan et al. 2017. At $z \sim 12 - 14$, all simulations are dominated by Population II star formation, however there continue to be pockets of pristine gas that can form more Population III stars. As well, we do not observe large-scale enrichment or ionization of the IGM at these redshifts, with only $\sim 0.8\%$ of the gas being enriched to Z_c, and only 2% of that volume being primarily enriched by Population II sources. Both of these findings are in agreement with prior works (Wise et al., 2012; Xu et al., 2016)

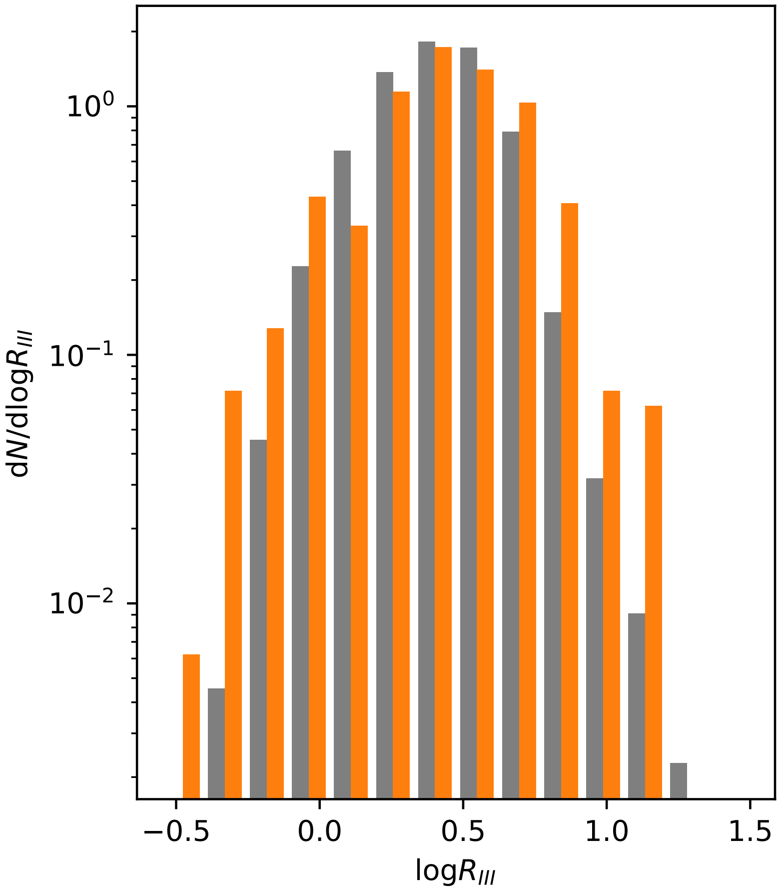


Figure 3.13. PDF of radii predicted from synthetic dataset (orange) with those from the full ground truth dataset (grey) at $t_{\text{final}} = 16$ Myr. The two PDFs agree well, with $J = 0.150$.

Population III stars in the Phoenix suite form in PIII associations. These are very diverse, with up to ~ 155 individual stars and system Population III stellar masses up to $M_* \gtrsim 10^3$ $\mathrm{M}_\odot$; 49.4% of Population II.1 stars are connected to at least one progenitor of each supernovae type (SN-HN-PISN in Table 3.2). A PIII association's influence is large, but limited, with 5%, 17% of PIII associations radii having $R > 8$ kpc for Z_{III} and H^+ respectively.

Several important observations have been noted regarding Population II.1 stars within the PHX suite: their MDF is not distinguishable from Population II.2 cluster formation, i.e., we cannot state with any certainty that a star with metallicity Z is of the Population II.1 or Population II.2 group; 80% of Population II.1 clusters were created from gas that had been enriched by combinations of Population III supernova endpoints, 14.4% were enriched by only PISNe, and, despite being the endpoint for the IMF peak, only 4.5% were enriched by only HNe; the highest Z clusters were created from gas enriched by PISNe, while the lowest Z gained metals from only SNe; the number of progenitors is highly variable, averaging only 1 progenitor in PISN case, up to > 20 progenitors if all endpoints were found connected to the forming star cluster. We also find that the average radius and time between SN event and Population II.1 star formation is smaller and shorter for PISNe, suggesting that prompt star formation due to the expanding remnant is a vector of early Population II.1 star formation, consistent with earlier simulation works (Machida et al., 2005; Ritter et al., 2012) of smaller systems at high resolution.

Both Population II.1 and Population II.2 stars exhibit low metallicity that includes the ranges of observations of stars found in the JINA database (Abohalima & Frebel, 2018), where most stars have [Fe/H] $\gtrsim -5$. Although the JINA database MDF peaks at [Fe/H] ~ -2.5, the high-metallicity distribution is inhibited by selection effects, so is not comparable to the PHX Population II distribution at any metallicity above this level. The MDF of Population II.1 stars has significant representation above the selection effect cutoff of the JINA MDF. The range of observed metallicity for Population II.1 stars here suggests that observational campaigns seeking to determine the Population III IMF via stellar archaeology (Frebel & Norris, 2015) will have incomplete data due to missing high-metallicity stars e.g., 30.5% of Population II.1 star clusters in the PHX simulations have $Z \gtrsim -2$. The peak of the MDF in the PHX also is similar to the observed metallicity of, e.g., Ursa Minor, with $\langle[\mathrm{Fe/H}]\rangle = -2.13 \pm 0.01$ (Kirby et al., 2013). Keller et al. 2014 studies the carbon-enhanced metal-poor star SMSS J0313-6708, with [Fe/H] < -7.1; while uncommon, this metallicity represents the extreme low metallicity end of the MDF in Figure 3.7. Interestingly, there are no Population II.1

stars with this low metallicity, where the minimum is $Z = -5.92$, however, several Population II.2 clusters are observed with comparable metallicity.

To evaluate the reasonableness of our Population III SFR and examine the Population III star formation efficiency (SFE), neither of which are constrained observationally, we examine high density cells in comparison to earlier Population III star formation work. Bromm et al. 2002 presented that the formation of molecular gas clouds proceed at a rate of $\dot{M}_g = 3.70 \times 10^{-7}$ $\mathrm{M}_\odot$ yr^{-1} pc^{-3}. $\dot{M}_g$ is the upper limit on the star formation rate, since a Population III star cannot form without a molecular cloud source. If we estimate the volume which could form unresolved molecular clouds in PHX256-1 as that which meets the star formation density criteria ($n_b \geq 100$ cm^{-3}), then we find that the PHX256-1 simulation has Population III SFRD $= 7.43 \times 10^{-9}$ $\mathrm{M}_\odot$ yr^{-1} pc^{-3} within those regions, implying a Population III SFE of $\sim 2\%$.

Data in Table 3.2 can be compared to prior works and observational campaigns. Welsh et al. 2020 estimates 5^{+13}_{-3} progenitor stars per metal-poor enriched star formation. Our presentation of N_{III} is largely consistent with their estimate, and highlights the dependence on the IMF to generate reasonable estimates of the number of enriching events. Average star forming halos in Hicks et al. 2021 contained $\lesssim 20$ enriching events, with the most massive halo having > 100. Table 3.2 would indicate that the average halo contained a few discrete PIII associations that led to the final state of the halo, while the most massive may have had several. Welsh et al. 2019 places an upper limit of 70 Population III enriching events in a metal-poor DLA system; our results would indicate that that system would likely have several PIII associations contained within, but is also small enough to be a single system.

Progenitor configurations that lead to Population II.1 star cluster formation are diverse so it would be difficult to analytically build an accurate metal abundance model for a Population II.1 star. Future work may be able to track separate metallicity fields from each progenitor type, allowing an estimation of the Population II.1 metal abundances given the composition of progenitors and mixing of the ISM that has occurred.Within the PHX, we find that while PIII associations have an average radius $R_{\mathrm{III}} \sim 3$ kpc, the low-mass, high-redshift halos we observe have $R_{\mathrm{vir}} \sim 0.4$ kpc–because the PIII associations tend to be much larger than their host halos, and that halos tend to be clustered together, we find that only 17.5% of PIII associations in the PHX512 simulation encompass a single halo, 64.1% contain > 5 halos, and 81.2% contain ≤ 10 halos; the remaining PIII associations cover up to 58 halos. Hence, these PIII associations can serve as an metal enriching source for many potential proto-galaxies, including those hosted by neighboring halos.

PIII assocation influence can be modelled reasonably well by piecewise linear regression fits, with our best-performing models having $R_2 \gtrsim 0.5$, and reproducing the PDF of observed radii with $J \lesssim 0.2$. We also find that these models are dependent on the time since the first stars' formation. This model has been used to generate a distribution of enrichment radii that agrees well with our observations in the PHX suite, while affording both simplicity and interpretability. Given a halo distribution, or sites of Population III star formation, this model can be used to estimate the extent of enriched regions from primordial stars, providing a new alternative to generating the metallicity initial conditions of cosmological simulations. A more complicated model with parameters for inferring hidden variables or learning from hydrodynamic inputs may be more successful in the short evolution time frame, or be able to generate non-spherical metallicity fields that reflect the complexity of the density distribution of a star forming region.

This research was supported by National Science Foundation CDS&E grants AST-1615848 and AST-2108076 to M.L.N. The simulations were performed on the Frontera supercomputer operated by the Texas Advanced Computing Center (TACC) with LRAC allocation AST20007. Data analysis was performed on the Comet and Expanse supercomputers operated for XSEDE by the San Diego Supercomputer Center. Simulations were performed with Enzo (Bryan et al., 2014; Brummel-Smith et al., 2019a) and analysis with YT (Turk et al., 2011), both of which are collaborative open source codes representing efforts from many independent scientists around the world. The authors thank John Wise for his insight and guidance regarding the Population III and Population II methods in Enzo.

Chapter 3, in full, is a formatted reprint of the material as it appears in Wells & Norman 2022, "Connecting Primordial Star-forming Regions and Second-generation Star Formation in the Phoenix Simulations". *The Astrophysical Journal*, 932:71. © American Astronomical society. Reproduced with permission. The dissertation author was the primary investigator and author of this paper.

Chapter 4

The First Galaxies and the Effect of Heterogeneous Enrichment from Primordial Stars

4.1 Abstract

We introduce the **S**cale-intelligent **T**erminal-momentum **A**lgorithm for **R**ealistic **S**tellar **S**ources (`STARSS`) in the Eulerian adaptive mesh astrophysical simulation code, `Enzo`. `STARSS` emphasizes the use of prior simulation and analytic inputs to model star formation and feedback in astrophysical simulations. After validation, `STARSS` is employed in simulations that use deep-learning and regression-based surrogate models (`StarNet`) to produce the heterogeneous metal enrichment that would result from primordial star formation. We analyze the impact of heterogeneous metal enrichment by studying the earliest generations of stars, and the protogalaxies ($10^6 \lesssim M_v/M_\odot \lesssim 10^8$) containing them. We find that ignoring metallicity requirements for enriched star formation results in a up to 30% excess in stellar mass created. Alternatively, employing a metallicity floor causes an early underproduction of stars before $z = 21$ that reverses to overproduction by $z = 18$, creating $\sim 20\%$ excess stellar mass and 8.5% exess in protogalaxy number by $z = 14.95$. Heterogeneous metallicity conditions greatly increase the range of halo observables, e.g., stellar metallicity, stellar mass, and absolute magnitude. The increased range leads to better agreement with observations of ultra-faint dwarf galaxies when compared to metallicity-floor simulations. `StarNet` generates protogalaxies with low stellar mass, $M_* \lesssim 10^3 M_\odot$, so also models low-luminosity protogalaxies more effectively than a metallicity floor criterion at similar mass and spatial resolution.

4.2 Introduction

Historically, astrophysical simulations have been restricted to single, purpose-motivated scales, e.g, using gravity only or gravity+hydrodynamics simulations to study cosmology (Vogelsberger et al., 2013; Vogelsberger et al., 2014; Emberson et al., 2019), increasing resolution and adding sub-grid star formation and feedback recipes to study single galaxies (Hopkins et al., 2018; Wheeler et al., 2019; Emerick et al., 2020), or monumentally increasing the resolution and including recipes for the formation and feedback of the first stars and star forming regions (Wise & Abel, 2011; Wise et al., 2012; Smith et al., 2015).

Because of the extreme difference in scale between star formation and large-scale structure formation, even next generation simulation codes running on exascale systems will struggle to comprehensively model single-star formation and feedback in a large-scale cosmological volume. For example, in the *Phoenix simulations* (Wells & Norman, 2022)(W22), $> 50\%$ of simulation time is spent evolving high-resolution regions required for Population III star formation and feedback. This problem is further exacerbated by evolving supernova remnants (SNR) at an early and hot stage at high resolution. In addition, both the Population III and Population II stellar feedback routines used in the PHX are *not* scale insensitive–they deposit supernova (SN) energy in thermal form, which places severe limits on resolution (Martizzi et al., 2015; Rosdahl et al., 2016). Taking full advantage of the next generation of high-performance computing at exascale will require not only new paradigms in parallel code development, but also novel accelerations within those new codes to address these difficulties.

The problems outlined above suggest two major avenues of progress for next-generation simulations in `Enzo`[1] and its massively parallel, CHARM++-based successor, `Enzo-E`[2](Bordner & Norman, 2018). First, a scale-intelligent star formation and feedback algorithm. Currently, star formation is bound by hydrodynamic evolution and force resolution, so we first devote our focus to the future of stellar feedback in `Enzo`. Motivated by other recent work using algorithms that adapt to resolution in some sense (Kimm & Cen, 2014; Rosdahl et al., 2017; Hopkins et al., 2018) , we have developed the **S**cale-intelligent **T**erminal-momentum **A**lgorithm for **R**ealistic **S**tellar **S**ources (`STARSS`). Using physically motivated parameterizations and prior simulation results, `STARSS` uses a minimum of user-defined parameters to implement star formation and feedback that does not require tuning by the user prior to thier next simulation. Section 4.3, 4.3.2, and 4.3.2 describe the

---

[1] https://enzo.readthedocs.io

[2] https://enzo-e.readthedocs.io/en/main/

`STARSS` algorithm, while Section 4.4 describes how the model performs in testing at various resolutions.

Second,we address the issue of Population III star formation and feedback. Despite the fact that Population III star formation is not uniform or isotropic, many current works assume a uniform prior distribution of metals (e.g., Hopkins et al., 2018), thereby negating the need for the Population III era. Other works simply ignore the first generation completely, allowing enriched star formation to occur in pristine gas (e.g., Vogelsberger et al., 2014). Both methods would negate the effect of Population III stars, e.g., prompt star formation from powerful SN (Machida et al., 2005; Ritter et al., 2012; Chiaki et al., 2013; Wells & Norman, 2022), destruction of early halos from Population III SN events (Whalen et al., 2008a), and the possibility of halos that go unenriched by Population III star formation (Regan et al., 2017). This work combines prior efforts in Wells & Norman (2021)(W21) and W22 to create a deep learning accelerated surrogate model that will address the non-uniformity of halo enrichment from Population III SNe. The framework developed for this purpose is referred to as the StarNetRuntime or simply `StarNet`. Although we developed it for the purpose of modelling PIII associations[3], `StarNet` is a framework that would be readily adaptable to modelling single stars, enriched star clusters, or even galaxy level star formation and feedback. `StarNet` is described in Section 4.6, while Section 4.6.2 describes the results of `StarNet` simulations and compares them to simulations with a metallicity floor using `STARSS` in order to quantify the impact of intelligently heterogeneous enrichment on the first protogalaxies and their minihalos.

4.3 Star particle creation and feedback

4.3.1 Star Particle Formation

Star particle formation follows a very similar prescription as described in Hopkins et al. (2018)(FIRE-2), and can optionally follow the updated methods of Hopkins et al. (2022)(FIRE-3). Those criteria that can be ignored or modified to follow FIRE-3 are denoted with a "*" below. Several criteria are checked at each grid cell to see if that cell qualifies for star formation:

1. *$\rho/\bar{\rho} > \rho_c$, where ρ_c is the minimum overdensity relative to the simulation volume that is allowed star formation and $\bar{\rho}$ is the simulation volume average density. Alternatively, the density parameter can consider number density: $n_b > n_{b,c}$.

---

[3] As noted in W22, this term refers to a group of coeval Population III stars, too small to form a canonical cluster, but too large to be a simple binary/trinary system.

2. Converging gas flow: $\nabla \cdot \mathbf{v}_{cell} < 0$ for $\mathbf{v}$, the cell-centered velocity.

3. The virial parameter (α), or ratio of kinetic + internal energy to gravitational potential energy is checked using Enzo's total energy field (available since Enzo uses a dual-energy formalism to track the energy of the gas). We require $\alpha < 1$ with

$$\alpha = \frac{E_{\text{total}}}{E_{\text{gravity}}},$$

where all quantities refer to the cell-centered value. Canonically, $\alpha = 2E_k/E_g$ for the kinetic energy, E_k (Bertoldi & McKee, 1992). Here, we use E_{total} to explicitly include contributions from all energy sources.

4. The cooling time must be less than the freefall time: $t_c < t_{\text{ff}}$ with $t_{\text{ff}} = \sqrt{(3\pi)/(32G\rho)}$, or temperature $T < 10^4$ K.

5. The gas mass in the cell must be greater than the critical Jeans mass: $m_b > \max(m_j, 10^3 \mathrm{M}_\odot)$.

6. * The gas must have self shielded hydrogen fraction, $f_s > 0$. This is checked via analytic approximation (Krumholz & Gnedin, 2011). If using 9-species chemistry, this requirement can also be explicitly checked by comparing neutral to ionized molecular hydrogen as evolved in the `Enzo` chemistry routines.

7. The metallicity must be above some user-defined critical value, Z_c [4]. Alternatively, the metallicity can be ignored, although it is still included in the analytic approximation of f_s.

If the above relevant criteria are satisfied, the creation routine assigns a integrated star formation rate

$$\dot{M}_* = \frac{f_s \eta_{\text{sf}} M_b}{t_{\text{ff}}}, \tag{4.1}$$

and maximum allowed star formation efficiency[5] η_{sf}. The probability of star formation is then given for the grid timestep δt as

$$p_{\text{sf}} = 1 - \exp\left[-\frac{\dot{M}_* \delta t}{\eta_{\text{sf}} M_b}\right]. \tag{4.2}$$

---

[4] Z is defined as the log of metal abundance relative to solar metallicity; for metal mass M_z and baryon mass M_b, $Z = \log(M_z/M_b) - \log(M_{z,\odot}/M_\odot)$

[5] This efficiency is simply the maximum fraction of gas within a cell that would be allowed to convert to stars within a timestep; this avoids converting all gas to stars and generating simulation failures from the resulting discontinuity in the density field. It does not imply a global efficiency, e.g., of gas to star conversion within a galaxy.

A star particle is formed only if sampling a binomial distribution with p_{sf} yields success. With probability satisfied, a new star is formed with $M_* = \min(f_s \eta_{\text{sf}} M_b, M_{*,\text{max}})$, where $M_{*,\text{max}}$ is an optional user-defined maximum star mass. The user may specify no maximum mass, in which case $M_* = f_s \eta_{\text{sf}} M_b$. However, if M_* is large, there may be times where one would expect > 1 SN per particle in a timestep–such a situation is undesirable, since `STARSS` feedback is designed to couple single-events to the grid and its analytic models do not in general hold true for combined events that would result in a "super-bubble" remnant (SBR). To avoid many SNe per timestep, we split large particles into several smaller sub-cluster particles. Each sub-cluster maintains the metallicity, position and velocity of the parent, however we offset the creation time (t_c) by factors of the dynamical time: for the i^{th} of n sub-clusters flagged for creation at time $t_{c,0}$, the modified creation time is set as $t_{c,i} = t_{c,0} + 3i \times t_{\text{ff}}/n$ Myr so that the creation of all sub-clusters is distributed across three t_{ff} (Murray, 2011).

4.3.2 Star Particle Feedback

Determination of Feedback Quantities

This section details the feedback of `STARSS` particles. For each formed particle, at each timestep, we calculate the age-based rates for supernovae (type II (R_{ii} and type Ia (R_{ia})) and winds. The probability (P_x) of a SN of type x is then $P_x = M_* R_x dt$, for dt being the timestep measured in Myr. If sampling a binomial distribution with P_x returns success, then a SN event will be modeled for this timestep. The rates used in this implementation are taken from FIRE-2, which were generated using Starburst 99 (Leitherer et al., 1999) simulations.

There are three modes of feedback coupled from star particles to the computational grid: Supernovae, winds, and radiation. Supernovae rates are approximated by piecewise functions that depends on the age of the particle measured in Myr (a_{Myr}):

$$R_{\text{ii}} = \begin{cases} 0 & a_{\text{Myr}} < 3.401 \\ 5.408 \times 10^{-4} & 3.401 \leq a_{\text{Myr}} < 10.37 \\ 2.516 \times 10^{-4} & 10.37 \leq a_{\text{Myr}} < 37.53 \\ 0 & a_{\text{Myr}} \geq 37.53 \end{cases} \tag{4.3}$$

$$R_{\rm ia} = \begin{cases} 0 & a_{\rm Myr} < 37.53 \\ 5.2 \times 10^{-8} + 1.6 \times 10^{-5} & \\ \times \exp\left\{ -\left(\frac{[(a_{\rm Myr}-50)/10]^2}{2} \right)^2 \right\} & a_{\rm Myr} \geq 37.53 \end{cases} \tag{4.4}$$

If either $P_{\rm ii}$ or $P_{\rm ia}$ result in a SNe event, we assign an ejecta mass of 10.5 or 1.5 $\rm M_\odot$ to Type-II and Type-Ia SN respectively. The metal ejecta for Type-II is $M_{z,\rm ej} = 1.91 + 0.0479 \times \max(Z_*, 1.65)$ or 1.4 $\rm M_\odot$ for Type-Ia.

At each timestep for the particle we additionally derive mass, energy and metal from stellar winds that must be coupled to the grid. The wind mass is given by $M_{\rm w} = M_* f_{\rm w} dt$ $\rm Gyr^{-1}$, with the wind loading factor

$$f_{\rm w} = \begin{cases} 4.763 \times \min(0.01 + Z, 1) & a_{\rm Myr} < 1 \\ 4.763 a_{\rm Myr}^{\kappa} \times \min(0.01 + Z, 1.0) & 1 \leq a_{\rm Myr} < 3.5 \\ 29.4 \left(\frac{a_{\rm Myr}}{3.5} \right)^{-13/4} + 0.0042 & 3.5 \leq a_{\rm Myr} < 100 \\ 0.43 \left(\frac{(a_{\rm Myr}/100)^{-1.1}}{19.81/\log(a_{\rm Myr})} \right) & 100 \leq a_{\rm Myr} \end{cases} \tag{4.5}$$

$$\kappa = 1.45 + 0.08 \times \min(Z, 1). \tag{4.6}$$

The mass of metals in the wind, $M_{z,\rm w}$ is then given by

$$M_{z,\rm w} = \max(0.02, 0.016 + f_z) \times M_{\rm w}, \tag{4.7}$$

with

$$f_z = 0.0041 \times (\max(Z, 1.65) + 0.0118). \tag{4.8}$$

Finally, the energy is determined as $E_{\rm w} = 10^{12} M_{\rm w} \varepsilon$, with ε is defined as:

$$\varepsilon = 4.83 + \frac{5.94 \times 10^4}{1 + (a_{\rm Myr}/2.5)^{1.4}} + \left(\frac{a_{\rm Myr}}{10} \right)^5$$

for $a_{\rm Myr} < 100$, or $\varepsilon = 4.83$ for all other times.

The last form of feedback associated with these star particles is via radiation. Each particle produces ionizing radiation according to

$$\Psi_{\rm ion} = \begin{cases} 500 & a_{\rm Myr} < 3.5 \\ 60\left(\frac{a_{\rm Myr}}{3.5}\right)^{-3.6} + 460\left(\frac{a_{\rm Myr}}{3.5}\right)^{\gamma}) & 3.5 \leq a_{\rm Myr} < 25 \end{cases}$$

$$\gamma = 0.045 - 1.82\log(a_{\rm Myr})$$

with Ψ in units $L_{\odot}/M_{\odot}$. With this parameterization, each particle will emit $\sim 8.94 \times 10^{60}$ ionizing photons/$M_{\odot}$ throughout its 25 Myr radiative lifetime. While SN and winds are coupled to the grid as prescribed in as follows in Section 4.3.2, radiation is coupled directly to the Moray ray-tracing radiation solver (Wise & Abel, 2011) in `Enzo`.

Coupling Feedback

Given an event from supernovae or stellar winds with ejecta energy $E_{\rm ej}$, mass $M_{\rm ej}$ and metal mass $M_{z,\rm ej}$, we couple the feedback to the computational domain as described in this section. We wish to separate $E_{\rm ej}$ into thermal and kinetic components based on physically motivated analytic expressions: further, we wish to distribute the kinetic energy via explicitly coupling momentum to the gas neighboring the star particle in an isotropic manner. Coupling the momenta in physically meaningful ways implies altering the amount of momenta coupled given the stage of the SNR, which is determined by the grid resolution. To determine the phase of the SNR, we compare the grid resolution, dx, to various quantities: the free expansion radius ($R_{\rm free}$), the cooling radius ($R_{\rm c}$), and the fading radius, ($R_{\rm fade}$). $R_{\rm free}$ is determined as in Kim & Ostriker (2015) as

$$R_{\rm free} = 2.75\left(\frac{M_{\rm ej}}{3M_{\odot}}\right)^{1/3} n_b^{-1/3}\ {\rm pc}. \tag{4.9}$$

$R_{\rm c}$ is given in FIRE-2,

$$R_{\rm c} = 28.4 n_b^{-3/7} E_{51}^{2/7} f(Z)\ {\rm pc} \tag{4.10}$$

with $E_{51} = E_{\rm ej}/(10^{51}\ {\rm erg})$ and $f(Z) = Z^{-0.14}$ if $Z \geq 0.01$; else $f(Z) = 2$, because there is little to no metallicity dependence below $Z \sim 0.01$ (Thornton et al., 1998). Finally, $R_{\rm fade}$ represents the radius at which we would expect the SNR to have merged with the ISM, i.e., the expected velocity of the shell is comparable to the local speed of sound. The expression is taken from Draine (2011) as

$$R_{\rm fade} = 66.0\, E_{51}^{0.32} n_b^{-0.37} \left(\frac{c_s}{10 \text{ km s}^{-1}}\right)^{-0.4} \text{pc}. \tag{4.11}$$

Since this simplified expression has no consideration for local metallicity, we finally take $R_{\rm fade} = \min(R_{\rm fade}, 1.5 \times R_{\rm c})$. In all the prior expressions, for SNRs, the energy is taken as 10^{51} erg, the mass is 10.5 and 1.5 $M_\odot$ for Type-II and Type-Ia SNe respectively.

After determining the phase of SNR, we calculate the expected momentum to couple (p_{cpl}) using the following expressions:

$$p_{\rm cpl} = \begin{cases} \sqrt{2M_{ej}E_{ej}} \times (dx/R_{\rm free})^3 & dx < R_{\rm free} \\ \min(p_{ST}, p_T \times dx/R_{\rm c}) & R_{\rm free} < dx < R_{\rm c} \\ p_T \times \eta & R_{\rm c} < dx \end{cases} \tag{4.12}$$

where $\eta = \left(1.0 - \tanh\left[(1.25 \times dx/R_{\rm fade})^2\right]\right)$ to smoothly connect the terminal phase to the fading phase where coupled momentum is zero. The expected momentum of the Sedov-Taylor(ST) phase ($p_{\rm ST}$) at the radius given by dx is

$$p_{\rm ST} = 2.21 \times 10^4\, E_{51}^{0.8} n_b^{0.2} t_3^{0.6} {\rm M}_\odot \text{ km s}^{-1} \tag{4.13}$$

where we solve for t_3 taking the radius as dx:

$$t_3 = \left(\frac{dx}{5\text{ pc}(E_{51}/n_b)^{1/5}}\right)^{5/2} \tag{4.14}$$

as derived from Kim & Ostriker (2015). The momentum in the momentum-driven snowplough phase (terminal phase, p_T) is taken from Thornton et al. (1998) as

$$p_{\rm T} = \begin{cases} 1.67 \times 10^5 E_{51}^{13/14} n_b^{-1/4} Z^{-0.36} & Z > 0.01 \\ 8.36 \times 10^5 E_{51}^{13/14} n_b^{-1/4} & \text{else} \end{cases} \tag{4.15}$$

Finally, Cioffi et al. (1988) notes that there is a low-density regime where there is no expected shell formation and the remnant merges with the ISM before the shell mass approaches the ejecta mass. This case is extremely important in our model, as the SNe originate from a single point in space; after the first few events, the interior region is very hot ($T \geq 10^7$ K) and low-density ($n_b \lesssim 10^{-2}$ cm^{-3}). To smoothly connect the phase of the SNR to this regime where we expect no momentum coupling, in cases where $n_b < 10n_\mathrm{C}$ we modify the coupled momentum as

$$p_\mathrm{cpl} = p_\mathrm{cpl} \times \left(1 - \tanh\left[\left(1.45\frac{n_\mathrm{C}}{n_b}\right)^{6.5}\right]\right) \tag{4.16}$$

with critical density

$$n_\mathrm{C} = 0.0038\frac{(P_4)^{7/9}(v_s/c_s)^{14/9}}{E_{51}^{1/9}}Z^{1/3} \tag{4.17}$$

for $P_4 = n_b T/10^4$ K and the expected shell velocity v_s. In order to calculate v_s, we utilize the following expressions for the mass of the SNR shell (M_s):

$$\frac{M_s}{\mathrm{M}_\odot} = \begin{cases} M_\mathrm{ej} & dx < R_\mathrm{free} \\ M_\mathrm{ej} + \frac{4}{3}\pi dx^3 \bar{\rho} & R_\mathrm{free} < dx < R_\mathrm{c} \\ M_\mathrm{ej} + 1.41 \times 10^4 \frac{E_{51}^{6/7} Z^{0.27}}{n_b^{0.24}} & R_\mathrm{c} < dx, Z > 0.01 \\ M_\mathrm{ej} + 4.89 \times 10^4 \frac{E_{51}^{6/7}}{n_b^{0.24}} & R_\mathrm{c} < dx, Z < 0.01 \end{cases} \tag{4.18}$$

he first case assumes no shell mass in the free expansion stage. The only mass coupled in this stage is the ejecta mass. The second assumes that the shell mass behaves as if sweeping up mass in a sphere with averaged density $\bar{\rho}$. The final two expressions describe the shell mass in the terminal phase, at different metallicity regimes (Thornton et al., 1998). Of course, these analytic expressions have to consider the mass that exists on the grid; we determine the mass that exists within the central cells (M_central) where the shell mass would be removed and the final shell mass is limited to $M_s = \min(0.75 M_\mathrm{central}, M_s)$. With M_s in hand, the velocity of the shell is given by $v_s = p_\mathrm{cpl}/M_s$ and the metal contribution from the shell is given as the mass averaged metallicity of the central cells from which the shell is being evacuated.

In principle the analytic forms of this section are also applied to stellar winds. In practice however, the wind energy is so low that the momentum is negligible. Despite the lack of momenta coupling at lower resolutions, the mass, gas energy and metal from winds are still deposited.

We present the metal contribution across the lifetime of a `STARSS` particle in Figure 4.1. In the right, we show the short-time evolution that includes O-B SNe and their associated winds, with SN events annotated

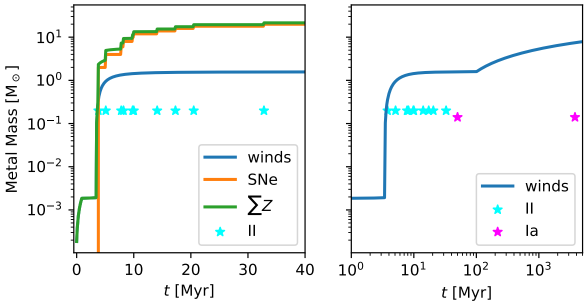


Figure 4.1. Metal contribution from the evolution of a 1000 M$_\odot$ `starss` particle. The left plot shows the 40 Myr evolution of early Type-II SNe era with SN events annotated (cyan stars) and the metal contributions from winds, SNe, and total contributions ($\sum Z$). Right: the contribution from stellar winds across 5 Byr, with Type-II and Type-Ia events annotated in cyan and magenta, respectively.

for the first 40 Myr. The left plot shows the long-time evolution up to 5 Gyr. It includes late winds from AGB branch stars starting at 100 Myr, and finite possibility for Type-Ia SNe, with occurrences annotated in magenta stars. At late times, the metal contribution from winds becomes comparable to that from Type-II SNe.

4.3.3 Coupling Method

Given $M_{\rm ej}$, $M_{z,\rm ej}$, $E_{\rm ej}$, and $p_{\rm cpl}$ from a source particle, S_a, we now describe coupling those quantities to the grid in an isotropic and conservative method that maintains very small linear error. Shown in the 2D example in Figure 4.2, we create a virtual cloud of 3^3-1 coupling target particles (S_b, purple circles) spaced at dx from the feedback source (S_a, red star). This method generates a fixed geometry, and maintains the unique position of S_a within its host cell. Each S_b receives an equal fraction of energy, mass, metal, and momenta to couple to the grid, so the final quantity coupled at each S_b is

$$M_b = \frac{M_{\rm ej}+M_{\rm s}}{26} \tag{4.19}$$

$$M_{z,b} = \frac{M_{z,\rm ej}+M_{z,\rm s}}{26} \tag{4.20}$$

$$\boldsymbol{p}_b = \frac{p_{\rm cpl}}{26}\hat{\boldsymbol{r}}_{b\to a} \tag{4.21}$$

where $\hat{\boldsymbol{r}}_{b\to a}$ is the unit vector from S_a to the S_b particle and the factor of 26 represents the 26 particles in the 3-dimensional virtual coupling cloud. The above expressions do not include energy: we couple kinetic

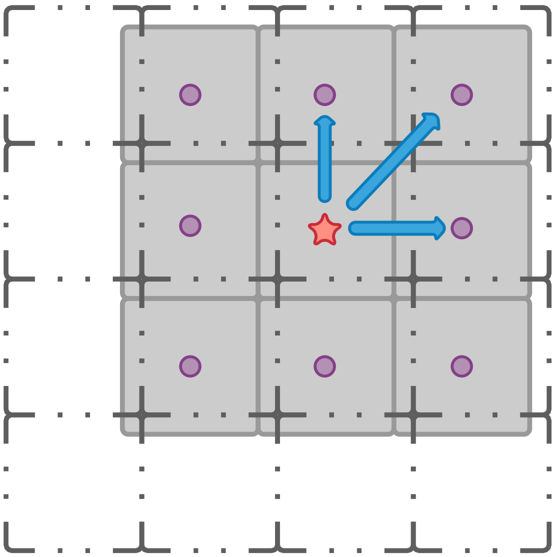


Figure 4.2. 2D example of deposition method. The feedback source (star) is coupled to neighboring cloud particles (circles) at spacing dx from the source; all feedback quantities are calculated at this radius, with momentum having a vector quantity indicated by blue arrows. The cloud particles are then cloud-in-cell deposited to the computational grid (hashed black grid). There is no quantity coupled at the source particle, however, the calculated shell mass is removed via CIC from the central cells centered on the source particle.

energy determined by the momenta coupled to the cell as $E_{\mathrm{k},b} = |\boldsymbol{p}_b|^2/2M_b$. The thermal energy ($E_t$) coupled is the remainder of the energy budget, i.e., $E_t = E_{\mathrm{ej}} - \sum_b E_{\mathrm{k},b}$. If $dx > R_\mathrm{c}$, then the thermal energy is reduced to account for unresolved PdV work using $E_t = E_t(dx/R_\mathrm{c})^{-6.5}$ (Thornton et al., 1998; Hopkins et al., 2018). Finally, the thermal energy is coupled in the same manner as kinetic with $E_{t,b} = E_t/26$. With known quantities for deposition, each S_b virtual particle is coupled to the computational domain via cloud-in-cell deposition.

4.4 Idealized Tests

To test `STARSS` supernova feedback algorithm, we used the `TestStarParticle` problem in `Enzo`. This problem sets the star particle at the center of the box (plus 1/2 cell width) with uniform density. To test resolution dependence, we used $\rho = 9.79 \times 10^{-23} \simeq 45\ \mathrm{cm}^{-3}$ in baryon density while varying the cell width, dx between $0.5\ \mathrm{pc} \leq dx \leq 50\ \mathrm{pc}$. The simulation domain was initialized with temperature $T = 1000$ K and $Z = -1$ (absolute metallicity 0.001295). The test utilizes all the standard `Enzo` physics capabilities: 6 species chemical evolution (H I, H II, e^-, He I, He II, He III), radiative cooling, and metal-line cooling using Cloudy lookup tables (Smith et al., 2009; Ferland et al., 2017).

At time $t = 0.00025$ Myr, a single SN as modeled by `STARSS` is coupled to the grid. The simulation

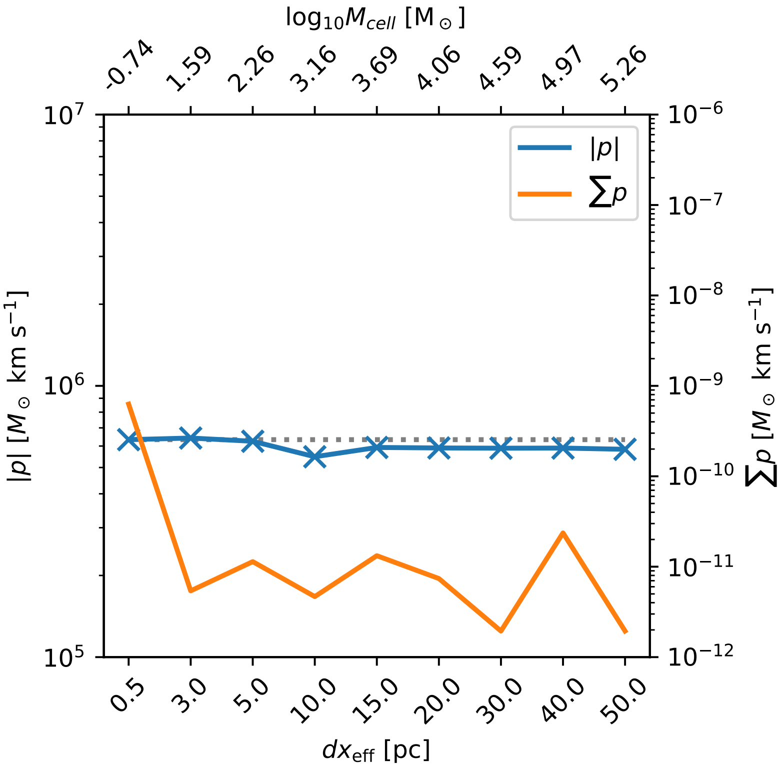


Figure 4.3. Terminal momentum of ideal tests while varying cell width with fixed n_b. Cell width (dx) annotated on lower axes, corresponding cell gas mass on upper. Linear momentum error is shown in orange corresponding to the right vertical axis. Since the density is fixed, we expect the same momentum across all samples, which is the behavior shown in blue. Expected momentum from the fully-resolved (0.5 pc) test is shown in grey. It is well matched by `STARSS` across all spatial resolutions tested.

continues for 1 Myr, and we record the terminal momentum as the maximum measured during that time. Taking the maximum observation is motivated by the behaviour in momentum while varying resolution: in resolved cases where the deposition represents the free-expansion or ST phases (e.g., $dx < 3$ pc) the momentum increases until the terminal value, and decreases afterward, however, in unresolved cases, $\boldsymbol{p}_t$ is directly coupled to the grid. The initial value of $\boldsymbol{p}$ represents the terminal solution exactly, which decreases after coupling to the mesh due to cooling.

Results from resolution tests are shown in Figure 4.3. The true solution (grey), obtained by a fully resolved simulation with $dx = 0.5$ pc, is closely matched by `STARSS` (blue) in all test cases. In addition, the linear momentum error ($\sum \boldsymbol{p}$) is $\leq 10^{-9}$ for all test cases. This test does not use the reduction in coupling beyond R_{fade}, which results in a sharp drop in coupled momenta for $dx > 30$, finally coupling negligible momenta by $dx = 50$.

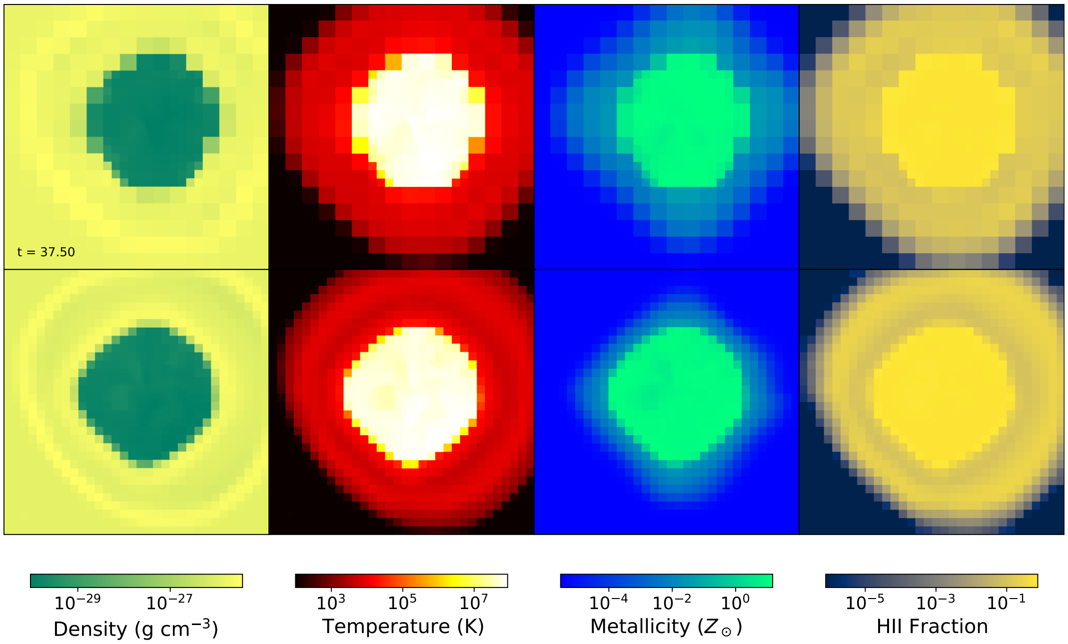


Figure 4.4. Ideal halo test comparing two spatial resolutions. Given identical initial density perturbations, we expect very similar results regardless of spatial resolution. Top shows a single particle 20 Myr after creation in a 10^7 M$_\odot$ halo with a NFW profile (Navarro et al., 1997) with 16^3 root-grid and 3 levels of AMR yielding a maximum resolution of 16.38 pc. Bottom shows the identical initial conditions using a 64^3 root grid and 3 levels AMR to achieve 4.1 pc maximum resolution. The final state across these tests is very consistent. The slight anisotropy is due to the particle being offset from the halo center in the initial conditions.

To test the full feedback framework of `STARSS` within a controlled environment, we use an ideal spherical halo. Specifically, the success of `STARSS` hinges on generating similar results irrespective of spatial resolution, so we performed several tests varying the spatial resolution of a test halo, but holding the baryon and dark matter density constant. This test uses identical cooling physics as the ideal single supernova test, but now includes radiation from the star particle in the form of ionizing radiation, heating, and photon momentum coupling to the gas. The star particle of 1000 M$_\odot$ is positioned at the center of the halo, slightly offset from the center of the box with width 800 pc. In the first test (LR), the domain has 16^3 root-grid cells and 3 levels of AMR on dark matter density so that the center of the halo has $dx = 16.38$ pc. The second test (HR) allows for 6 levels AMR, so that the maximum resolution is 4.1 pc. The refinement criteria of each simulation are tuned such that the center of the halo is at the maximum AMR level. The background density of the box is set to 7.0×10^{-26} g cm^{-3} giving the center of the halo number density $n_b = 2.2$ cm^{-3} at $t = 0$ in both cases. The simulation allows for the full probabilistic feedback of `STARSS`, and proceeds for the duration where Type-II SNe have finite probability, 37.53 Myr. Although the initial central number density of the halo has $n_b = 2.2$ cm^{-3}, early radiation pressure reduces the density surrounding the star particle to $n_b \sim 0.8$ cm^{-3} in both LR and HR. Since the number of SN events is not deterministic, the simulations were repeated until we observed a similar number of events during the lifetime of the particle. Figure 4.4 shows slices through the simulation domain at the final output, 37.50 Myr. We observed 16 events in the LR (top) and 15 in the HR. Both show qualitatively similar behavior, including in the magnitude of density, temperature, and metal density within the SBR. At the final output, we measure the momentum in the “shell” of the SBR and find that $|\boldsymbol{p}| = 7.12 \times 10^6$ and 7.42×10^6 M$_\odot$ km s^{-1} for the HR and LR respectively, representing a $\sim 4.2\%$ difference. The magnitude of difference in momenta, $\sim 3 \times 10^5$ M$_\odot$ km s^{-1}, is very similar to the expected momenta from a single terminal-phase SNR.

4.5 Cosmological Simulation Tests

We choose to explore the performance of `STARSS` within the framework of cosmological simulations. We prefer the multiple galaxy-halos of cosmological simulations so that we can compare `STARSS` across multiple star forming halos and multiple halo mass ranges within a single simulation. Each simulation uses identical cosmological parameters with with $\Omega_m = 0.3111$, $\Omega_b = 0.048975$, $\Omega_k = 0$, $\Omega_\lambda = 0.6889$, $H_0 = 0.6766$, $\sigma_8 = 0.811$, $n = 0.965$ (The Planck Collaboration et al., 2014). For easy comparison, each uses identical initial conditions generated using `MUSIC` (Hahn & Abel, 2011) on a 256^3 root-grid with (2.61 Mpc)^3 volume. With these parameters defined, the dark matter particle mass is 2.34×10^4 M$_\odot$ and the average initial baryon mass per cell is 1.17×10^3 M$_\odot$. For refinement criteria, we consider baryon and dark matter overdensity, where the cell is refined if $\rho_x/\bar{\rho}_x > 3$, for the baryon or dark matter density ρ_x, and $\bar{\rho}_x$ refers to the simulation averaged quantity. In addition, we use an exponential factor to enforce super-lagrangian refinement: the the mass within a cell to cause refinement is given by $M_x \geq 3M_i \times 2^{-0.6l}$ for level l and M_i as the initial average baryon or DM mass per cell in the simulation.

We use identical physical and chemical models across all `STARSS` simulations: 9-species chemistry including H, H$^+$, He, He$^+$, He^{++}, e$^-$, H$_2$, H$_2^+$, H$_2^-$; 4-dimensional cooling considering density, metallicity, electron fraction, and temperature as determined by Cloudy lookup tables (Smith et al., 2009); radiative feedback including photon momentum coupling to the gas using the Moray ray tracing solver (Wise & Abel, 2011) with `STARSS` particles as sources; finally, we include a redshift dependent Lyman-Werner H$_2$ dissociating radiation field to model sources from outside the simulation region.

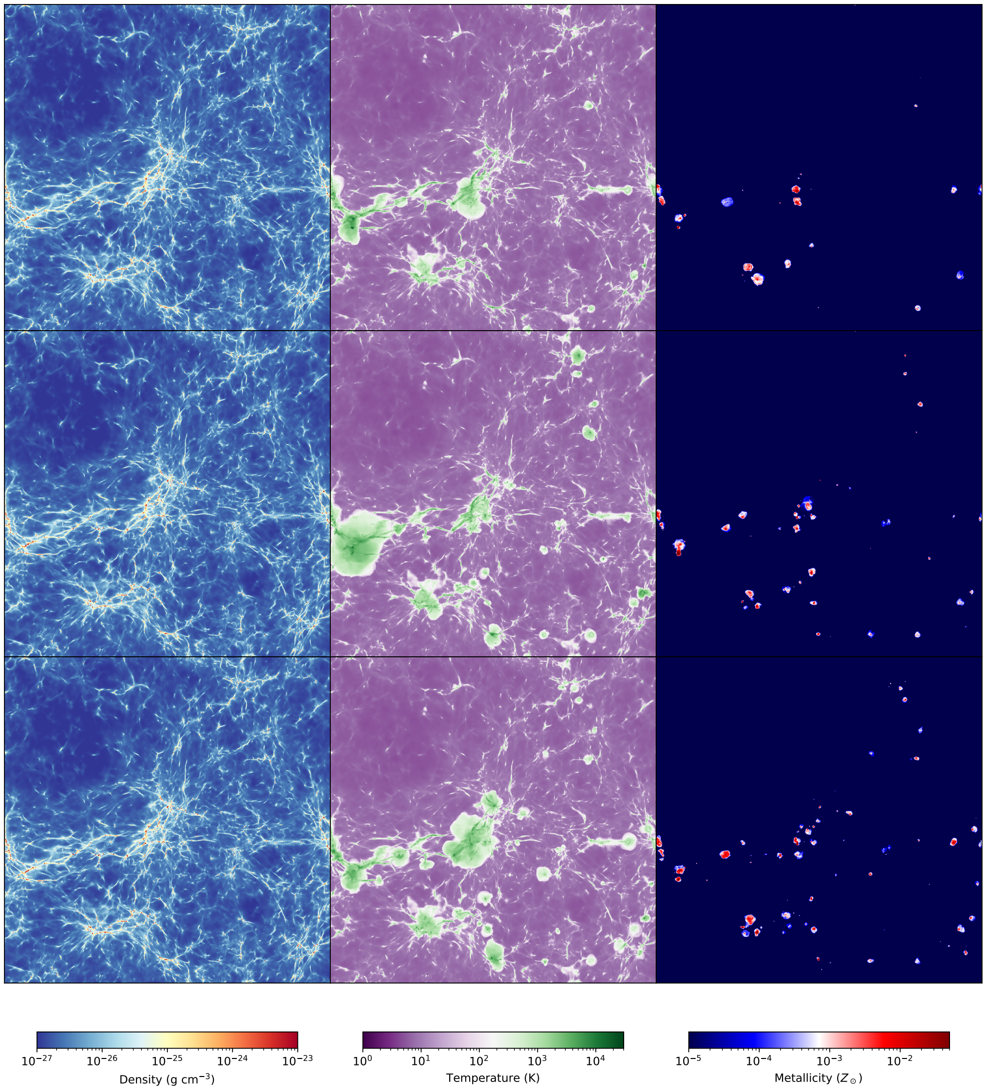


Figure 4.5. Comparing simulations with identical mass resolution varying maximum spatial resolutions at $z = 13.91$. The top, middle, and bottom show the 4L, 5L and 6L simulations respectively. Note that since there is a probabilistic element in both the star formation and supernova routines, we do not expect perfect agreement among the rows.

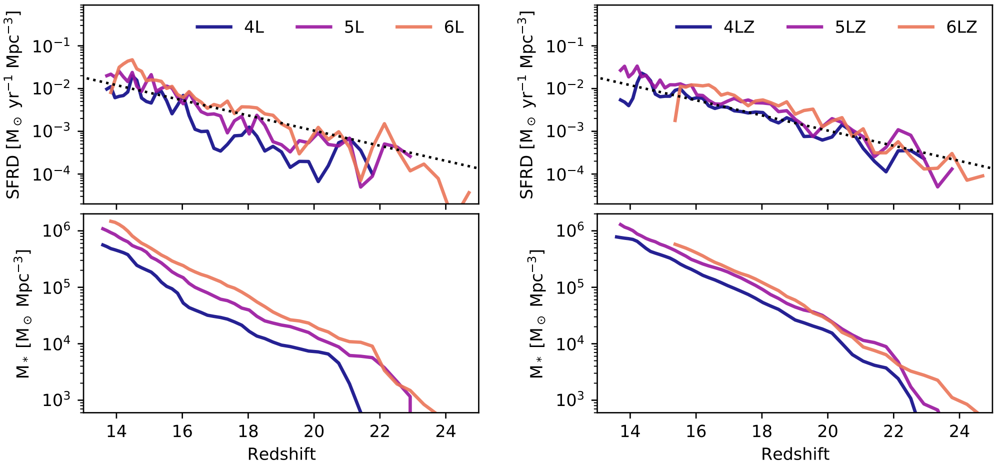
Table 4.1. Summary of `STARSS` simulations.

ID	L_{Max}	Z_{f}	dx_0	dx_{15}
4L	4	-5.5	624 pc	39 pc
5L	5	-5.5	312 pc	19.5 pc
6L	6	-5.5	156 pc	9.75 pc
4LZ	4	-3	624 pc	39 pc
5LZ	5	-3	312 pc	19.5 pc
6LZ	6	-3	156 pc	9.75 pc

Note: The differences between simulations is shown here; the metallicity floor (Z_{f}) and maximum AMR level varies, while all other parameters are identical throughout. The finest cell-width is shown for $z = 0$ (dx_0) and $z = 15$ (dx_{15}).

The cosmological simulations used to validate `STARSS` are shown as projections in Figure 4.5, and with parameters summarized in Table 4.1. Qualitatively, the simulations produce very similar results, however minor differences, e.g., in the size of bubbles in metallicity, are noticeable. We can see that the volume is largely unenriched by $z = 13.91$, and by using temperature as a proxy for ionization state, most of the volume remains neutral (in reference to H, assuming ionized H gas would have $T \gtrsim 10^4$ K). While the simulations are comparable in the metallicity field, signifying consistent SN feedback across resolutions, the temperature field shows systematic differences, particularly between 4L and 5L simulations, where 4L has noticeably smaller high-temperature regions that are fewer in number. These systemic differences are less pronounced when comparing the 5L and 6L simulations, indicating that the star formation algorithm is more consistent at those resolutions. The remainder of this section is dedicated to quantitative comparisons across these three resolutions.

We can quantify the convergence of the `STARSS` algorithm by examining SFR density (SFRD) and stellar mass formed from a global perspective that includes both initial conditions for a given resolution. In Figure 4.6, we present the SFRD and stellar mass density for all `STARSS` simulations. Figure 4.6a uses the same metallicity floor (Z_{f}) as the critical metallicity of the *Phoenix Simulations* with $Z_{\mathrm{f}} = -5.5$. While 5L and 6L have very similar cumulative stellar mass, 4L lags by ~ -0.3 dex. However, 4L maintains similar slope in cumulative mass, suggesting that the effect of resolution is to delay star formation, but not to change the rate of star formation after it has begun. Figure 4.6b shows the same set of simulations, but using a metallicity floor more common in modern work of $Z_{\mathrm{f}} = -3$. The higher metallicity floor enables more efficient collapse of star forming regions due to enhanced cooling, which has acted to negate most resolution effects seen in Figure 4.6a, but has limited effect on the resolved 6L case. With $Z_{\mathrm{f}} = -3$, 5LZ and 6LZ have converged to



(a) $Z_\mathrm{f} = -5.5$

(b) $Z_\mathrm{f} = -3$

Figure 4.6. SFR density and formed stellar mass density for varying resolution using `STARSS`. All resolutions use the same star formation and feedback parameters. Panel 4.6a uses a metallicity floor of $Z_\mathrm{f} = -5.5$, while 4.6b uses $Z_\mathrm{f} = -3$. 5LZ and 6LZ simulations converge to nearly identical behavior with the higher metallicity floor, but have slight differences in stellar mass with the low metallicity floor. The black dotted line is an approximate fit to the SFRDs with SFRD = $3.65 \exp(-z/2.45)$.

nearly identical behavior, and the stellar mass deficit in 4LZ is reduced to only ~ 0.2 dex below 5LZ and 6LZ. This behavior in Figure 4.6b is our primary motivation for using 5LZ in our comparisons of Section 4.6.2.

To describe how star forming regions interact with the environment from a global perspective, we show the volume fractions ionized and enriched to varying levels in Figure 4.7. There is much better agreement in the enriched volume fraction than ionized volume fractions. Lower resolution (4L, 5L) tend to enrich more of the volume at early times, but the difference between resolutions is largely negligible by $z \sim 14$. The ionized volume fraction still shows deviations at the lowest redshift, and shows large steps corresponding to star formation beginning in new regions. Exploration to lower redshift would be beneficial to determine whether the differences in ionized volume fraction are reduced or exacerbated by further evolution.

In Figure 4.8, we show a histogram of feedback region radius by redshift for all `STARSS` simulations. We additionally show the mean radius plotted over the histogram. To obtain this data, we iterate each data output and search for halos with finite stellar mass. If found, we center a 0.1 kpc sphere on the center of stellar mass and record the volume-weighted mean metallicity and HII fraction. If $\langle Z \rangle > -5$ or $f_\mathrm{HII} > 0.05$, the sphere is expanded by 0.1 kpc and the averaging is repeated until $\langle Z \rangle < -5$ and $f_\mathrm{HII} < 0.05$. Figure 4.8 clearly shows that there are fewer and fewer small feedback regions as resolution decreases, implying

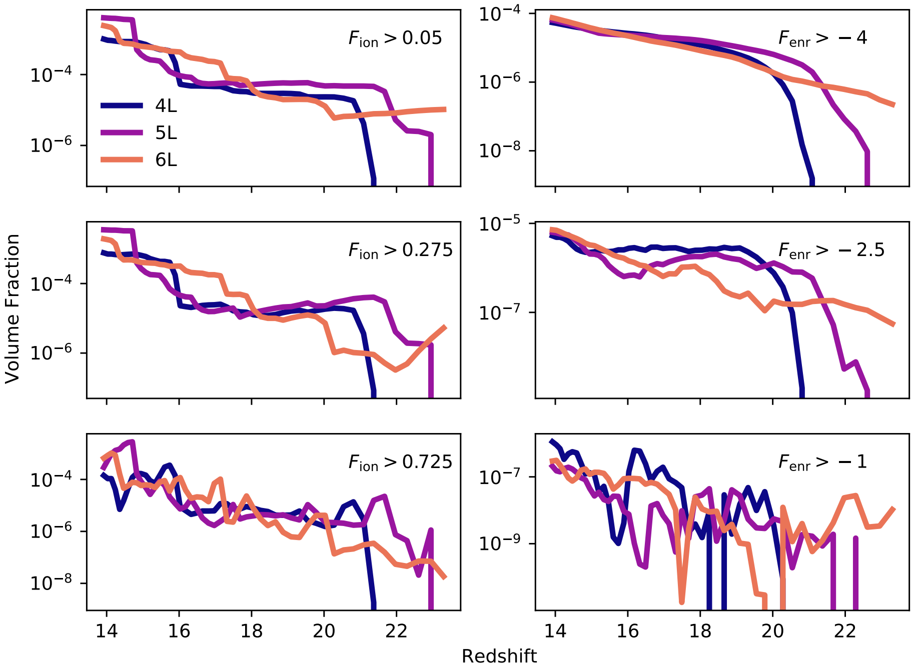


Figure 4.7. Volume fraction ionized to various levels (Left column), or enriched to varying levels (Right column) by redshift using the second initial conditions set for all resolutions.

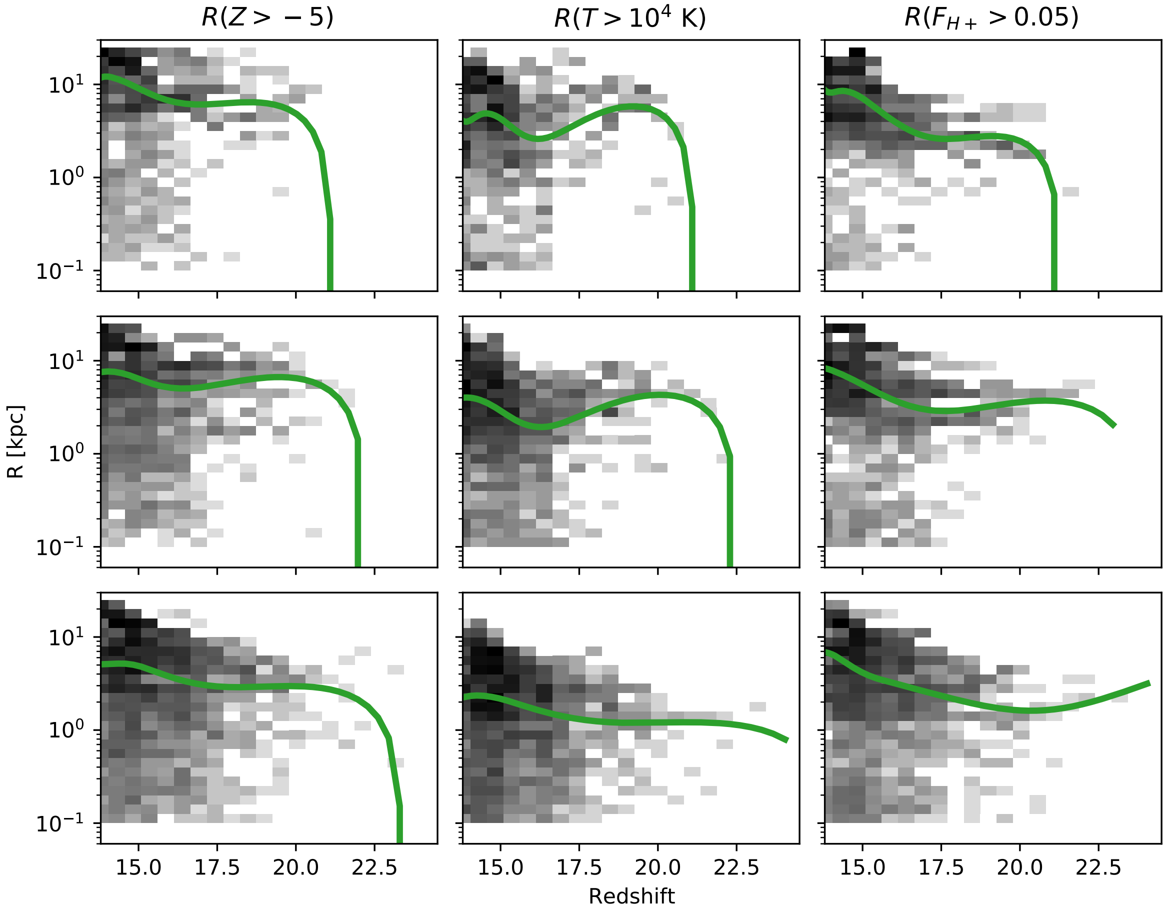


Figure 4.8. Comparing the size of feedback remnants across 4L (top), 5L (middle), and 6L (bottom) simulations. The background histogram logs each star forming region, while the plotted line shows a bicubic spline fit to the 50th quantile of remnants in each redshift bin.

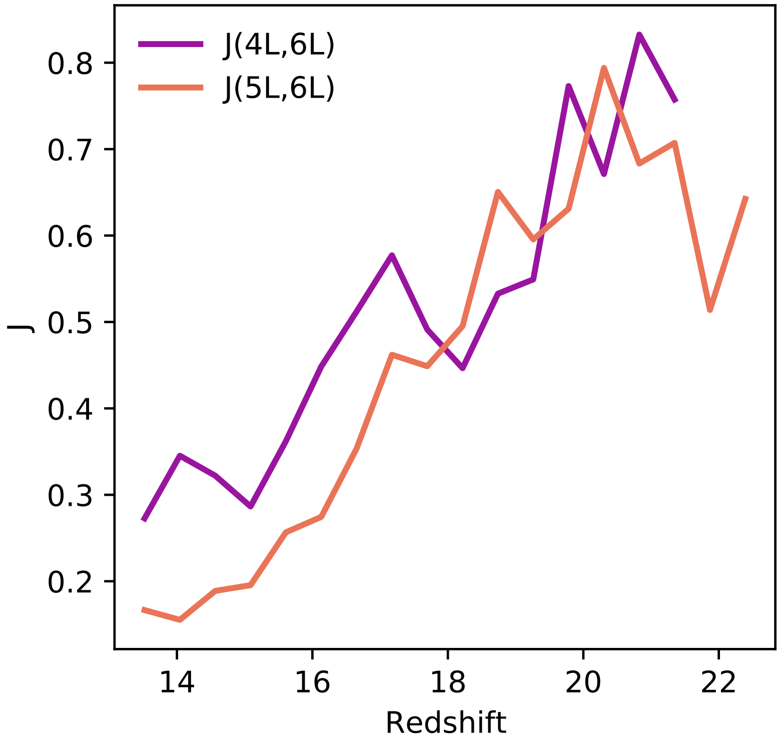
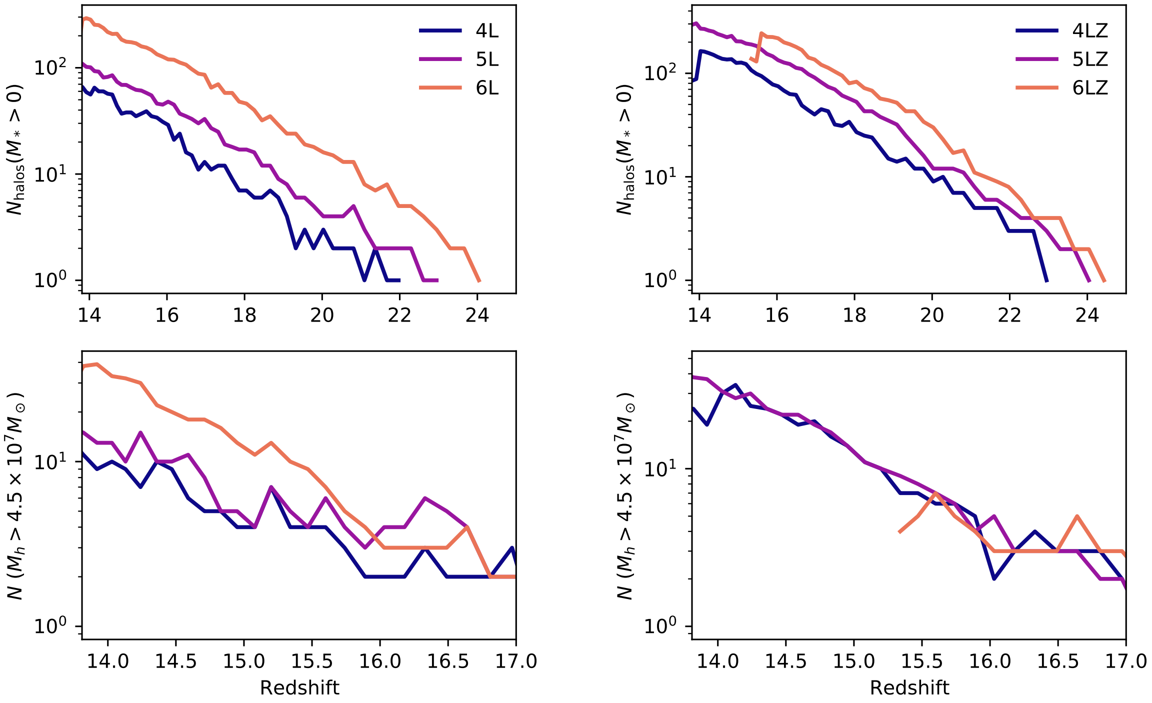


Figure 4.9. Comparing redshift bins of metallicity radius in Figure 4.8 by examining Jensen-Shannon distance. Lower distance implies more closely related PDFs of radii.

that lower-resolution runs struggle to model low stellar mass or young systems. The ionization fraction is consistent across the presented resolutions, however we note larger metal-rich remnants as resolution decreases. While noticable between 5L and 6L, the effect is much more pronounced in comparing 4L and 6L. This is likely due to 4L creating fewer but more clustered stellar particles instead of many small, scattered particles. Those few clustered particles deposit more SNe into the same region. In a positive feedback loop, the higher number of SNe are then deposited into less dense and hotter gas, resulting in higher-velocity expansion and larger super-bubble remnants. Despite this complication, the quantile fit to the size of remnants in 4L is still $\lesssim 2\times$ that of 6L, despite having $4\times$ worse spatial resolution. Examining the background histogram suggests that all resolutions have similar distributions at high radii, and the significant change is in the modeling of younger or less massive stellar systems with smaller radii.

The distribution of remnant regions in Figure 4.8 can be quantitatively compared by considering each redshift bin as an independent distribution of remnant sizes. We can then compare each bin using the Jensen-Shannon distance as a metric to compare two probability distributions. With the probability



(a) $Z_{\mathrm{f}} = -5.5$

(b) $Z_{\mathrm{f}} = -3$

Figure 4.10. Star forming halo counts including all IC variations. Top: Number of star forming halos by redshift. Bottom: Number of star forming halos with $M_h > 4.5 \times 10^7$ $\mathrm{M}_\odot$ by redshift. Note that the plots do not share range in redshift. 4.10a shows the low Z_{f} simulations, while 4.10b shows $Z_{\mathrm{f}} = -3$. Using $Z_{\mathrm{f}} = -3$ has negated much of the discrepancy noted in the $Z_{\mathrm{f}} = -5.5$ case. Factor of ~ 2 drops in counts at $z \sim 14$ (4LZ) and $z \sim 15.5$ (6LZ) are due to different end redshifts of simulations within that resolution suite.

distribution of radii from simulation A as P_a, and from simulation B as P_b, we can define the Kullback-Leibler divergence $D(P_a|P_b) = P_a \log(P_a/P_b)$ to compare the PDFs using the Jensen-Shannon distance given by

$$J(P_a, P_b) = \sqrt{\frac{D(P_a|\bar{P}_{ab}) + D(P_b|\bar{P}_{ab})}{2}}, \tag{4.22}$$

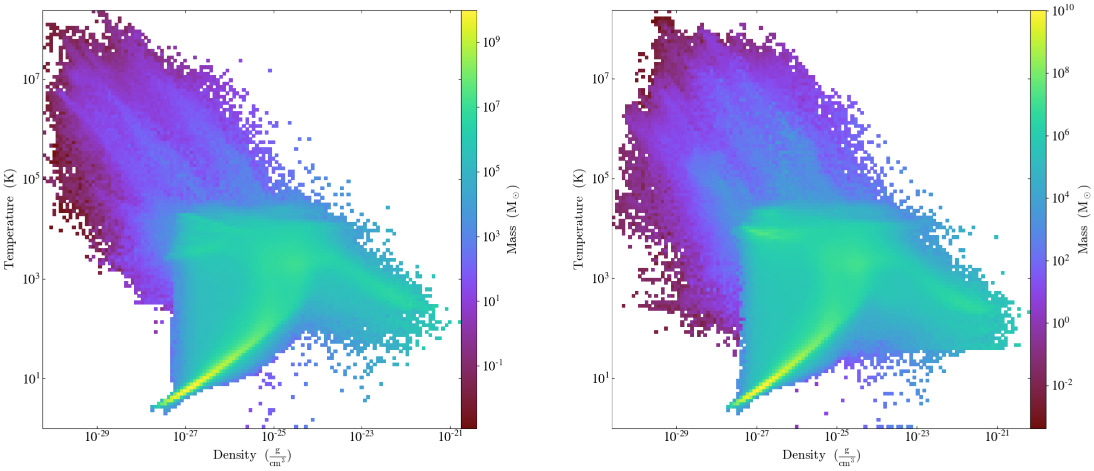
where P_{ab} is the pointwise mean of P_a and P_b. For this description of J, $J = 0$ implies identical PDFs, while $J = 1$ would indicate completely unrelated PDFs. We compare the 4L and 5L simulations to the 6L simulation using this method in Figure 4.9. In general, J is lower when comparing 5L and 6L, and the agreement between the three resolution increases as redshift decreases. Although 5L appears to be a qualitatively better match in Figure 4.8, this metric solidifies that observation.

The figures presented thus far indicate that while the feedback from `STARSS` is largely consistent across resolution scale, the star formation algorithm is not. This is obvious in Figure 4.10, where we present the number of star-forming halos (top) and number of massive $M_h > 4.5 \times 10^7$ $M_\odot$ star-forming halos (bottom) in each tested resolution for both $Z_f = -5.5$ and $Z_f = -3$. The effect of resolution is to host star formation in fewer halos, particularly for low Z_f (Panel 4.10a). The effect is reduced as the halo mass of interest is increased, however, even at $M_h \sim 4.5 \times 10^7$, 6L shows enhanced star forming halo counts. The effect is also greatly reduced when using $Z_f = -3$ (Panel 4.10b). While the difference between 5LZ and 6LZ is still ~ 0.1 dex for all halos, their behavior is converged for larger halos with $M_{vir} \geq 4.5 \times 10^7$ $M_\odot$. In `STARSS` current paradigm of predicting star formation, this relationship between resolution and star-forming halo counts is likely unavoidable. Since force resolution in `Enzo` decreases with spatial resolution, the simulation cannot resolve peaks and troughs of the hydrodynamic fields as accurately. This leads to lower magnitudes of density peaks reducing the ability of the gas to self-shield, as well as reduced accuracy in modeling gas flows, limiting star formation by our $\nabla \cdot \mathbf{v} < 0$ star formation criterion.

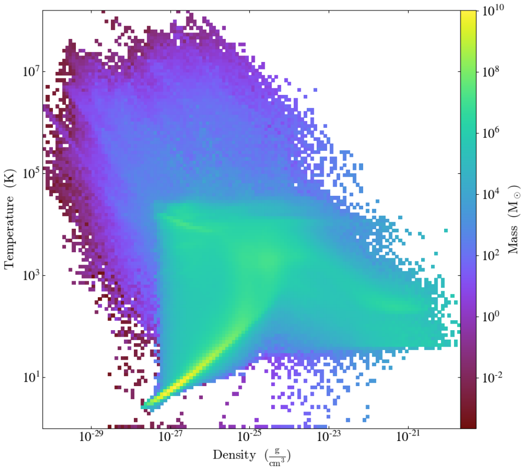
Finally, we present density-temperature phase diagrams in Figure 4.11. Again, 5L and 6L have converged to very similar behaviour, even in the high-temperature, low-density region that represents stellar feedback. All three simulations have similar high-density behaviour, suggesting that star formation and feedback is having a similar effect once gas reaches those high densities ($\rho \gtrsim 10^{-21}$ g cm^{-3}) and that the primary difference between the simulations is the gas dynamics due to resolution effects outside of stellar feedback.

4.6 `StarNet`: Surrogate Models of Primordial Star Formation and Feedback

In the following sections we present a new method of generating metallicity initial conditions in cosmological simulations. Where prior methods would set a “metallicity floor,” where the metallicity everywhere in the volume takes on a set value, or ignore metallicity effects on the first and second generations of stars, here we aim to generate an heterogeneous metallicity field by considering where primordial stars would form and modelling their subsequent feedback. Using the inline Python analysis capability in `Enzo`, we incorporate deep learning models to predict star formation (`StarFind` (W21)) and regression models of primordial stellar region effects (W22) into a single framework (`StarNet`) that can evaluate simulations *in situ* for primordial star formation sites and deposit a rudimentary approximation of their effects. Although every effort was made to streamline and optimize the inline Python for `StarNet`, the data structure of `Enzo`



(a) 4L **(b)** 5L



(c) 6L

Figure 4.11. Density-temperature phase profiles for the entire simulation volume in 4L, 5L and 6L `STARSS` simulations at z=13.71. The primary difference between resolutions is in the low-density high-temperature regime.

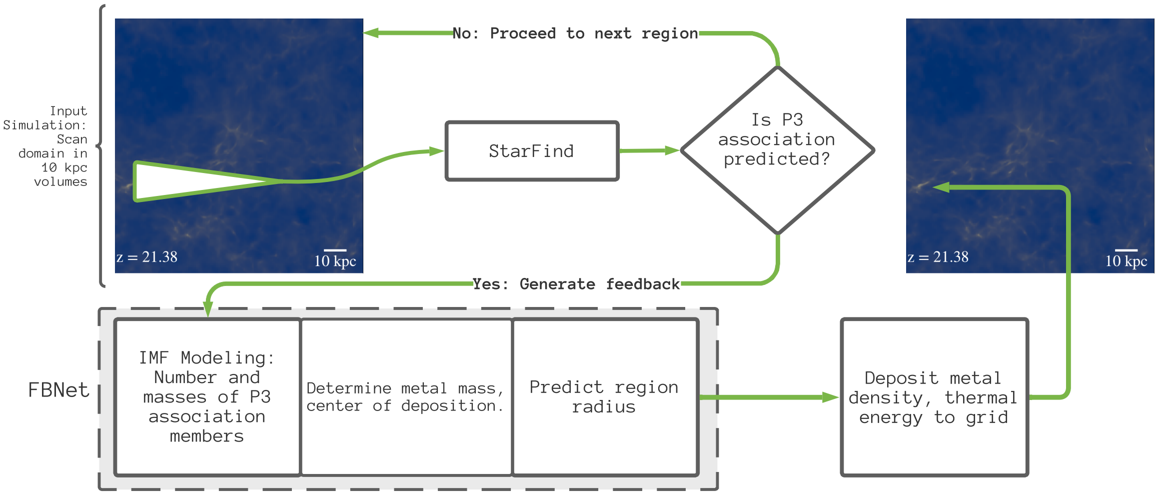


Figure 4.12. Simplified work-flow diagram of `StarNet`.

combined with implementation limitations on the inline Python means that these runs were limited in the following ways: each deposition of a Population III remnant region requires access to all levels of the computational grid, therefore we cannot utilize load balancing algorithms that would separate child grids from their parents. Since there is no communication across tasks, we must reduce the possibility of remnant regions that cross root-grid boundaries (each root-grid is a work-unit managed by one task in `Enzo`). We accomplish this by using the minimum number of tasks that can support a 256^3 root-grid simulation. Because of these limitations, this work only explores the very high-redshift regime of $z \gtrsim 15$. Future work to integrate the StarNetRuntime into `Enzo-E` will be able to fully explore lower-redshift simulations. This limitation also means that this work is a "proof-of-concept" and first attempt at exploring the application of DL methods to active simulations, and as such, will form the basis of, and inform the direction of, future similar simulations and surrogate modelling efforts.

4.6.1 `StarNet`

`StarNet` predicts the sites of PIII association formation, estimates the primordial stellar population from statistical models, and applies simplified spherical feedback expected to the region. A simplified work-flow of `StarNet` is presented in Figure 4.12. Every n megayears, the simulation is paused and the active computational domain is sent to `StarNet` for evaluation. First, each grid with AMR level $l = 2$ is tiled into (10 comoving kpc)^3 volumes. Each volume needs to to represent AMR level 6, despite that each volume may include data from various levels and is assuredly not uniformly the same resolution. To bring data from

$l = 2$ to $l = 6$, we copy simulation data at $l = 2$ to an inference grid, G, use trilinear interpolation to achieve $l = 3$, and then copy the simulation data at level 3 to G. This process is repeated at increasing level until the desired resolution, where G has 64^3 dimension. G is then passed into trained models for `StarFind` to perform inference.

If `StarFind` predicts a PIII association, we begin calculating a simplified feedback solution that represents the final state of the PIII association (from a metallicity perspective). `StarFind` predicts a group of voxels in G that are participating in star formation, where we assume that the center of the region is at the center-of-mass of those positive prediction voxels. We use the statistical relations described in W22 to determine the number and masses of Population III stars within the association and use the mass-SN yield parameterization of `Enzo` to derive the metal mass originating from the PIII association, $M_{\rm III}$. We then use the linear regression models of W22 to predict the radius of the feedback region, $R_{\rm P3}$. From the center of the predicted PIII association, the computational grid within $R_{\rm P3}$ is modified using the following:

- Primordial metal density, $\rho_{\rm ZIII} = M_{\rm III}/(4\pi R_{\rm P3}^3/3)$[6].
- Temperature, $T = 10^4$ K.
- All chemical species retain the same fraction of density that existed before to maintain mass conservation.

While these modifications are very crude, they satisfy several requirements: the metallicity field is not homogeneous on large scales (outside the feedback region) and the high temperature ionizes H gas and will reduce the H_2 fraction within the region. In addition, since $R_{\rm P3}$ and $\rho_{\rm ZIII}$ are calculated from prior simulations statistics, each "bubble" has a broad range of possible metallicities, which prevents each star forming region from attaining the exact same metallicity floor before Population II star formation takes hold.

`StarNet` is composed of the results of previous projects which have never been combined into a single module until this work. A visual projection of the primordial metallicity field is shown in Figure 4.13 comparing the metallicity field in the *Phoenix Simulations* PHX256-1 and PHX256-2 simulations (top) to the predicted regions found by using `StarNet` in a simulation matching the initial conditions of the PHX, but with no Population II star formation enabled. There is obvious agreement between many regions in the projections, however there are some notable differences. Some regions have been predicted early by `StarNet`,

---

[6] All spherical-volume based calculations are corrected at deposition to account for depositing into cubical grid cells

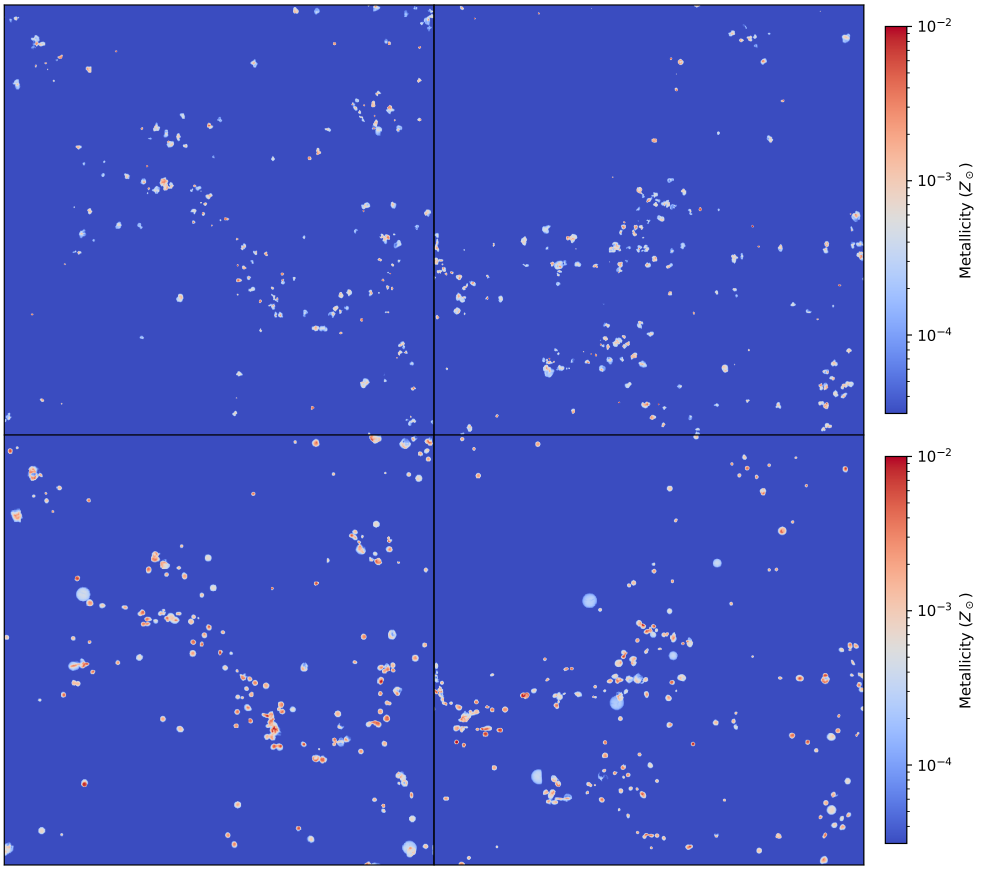


Figure 4.13. Top: Metallicity field sourced from Population III stars within the PHX256-1 and PHX256-2 simulations (W22). Bottom: Metallicity field as predicted by `StarNet`. Both rows show the state at $z = 14.95$.

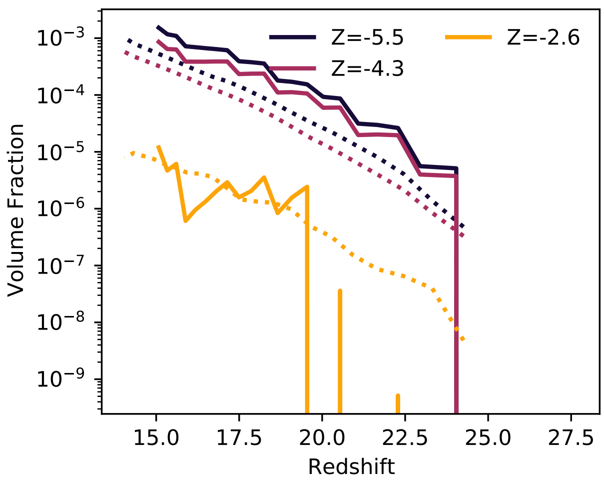


Figure 4.14. Volume of simulation enriched above the given metallicity, as generated by `StarNet` without including `STARSS` star formation and feedback. The hashed lines show the corresponding fraction as measured in the *Phoenix Simulations'* PHX512 (W22). Since there is no on-going star formation in the `StarNet` simulation, comparisons between ionized volume fractions are uninformative.

within the lower plot. As well, it appears that the regions are, generally, larger in `StarNet`. This is an artifact of `StarNet`'s design: it predicts a "final state" for the feedback region, estimating the impact of primordial stars 16 Myr after the first star forms. Since the regions in the PHX suite are at varying stages of evolution, there are many regions that have just formed in PHX, but have the final state predicted by `StarNet`. It is notable however, that the magnitude of metallicity seems to agree well between the two methods as well.

In order to quantify the effect of pre-enrichment by Population III stars, we perform five simulations:

1. sn5L-1: `StarNet` simulation that requires $Z \geq -5.5$ prior to `STARSS` particle formation.
2. sn5L-2: same as sn5L with different initial conditions.
3. sn5l-noZ: `StarNet` simulation with no metallicity prerequisite–`STARSS` particles can form regardless of gas metallicity in the host cell, however `StarNet` still generates a metallicity field due to Population III star formation.
4. sts5l-1: `STARSS` simulation from Section 4.5; this has no Population III treatment, instead initializing the metallicity field at $Z = -3$ and requiring $Z \geq -3$ for `STARSS` particles to form. This modified metallicity floor is the 5LZ simlulation set noted in Section 4.5, and is more representative of typical metallicity floors used in, i.e., FIRE-2.

5. sts5l-2: same as sts5l with changed initial conditions matching that of sn5l-2.

Broadly, these simulations fit two categories: dependent on or independent from initial enrichment. Below, the terms *SNET* or *STS* shall refer to the combined statistics of sn5l-1,2 or sts5l-1,2 respectively. Our primary comparison is between the SNET and STS suites to identify the difference between simulations using metallicity floors and those requiring pre-enrichment from Population III SFF. SNET uses a very low critical metallicity compared to the metallicity floor of STS. Our motivation for doing so is because the training data to create `StarFind` and `StarNet` were generated from the *Phoenix Simulations* suite where Population II star formation can only proceed in gas enriched to $Z_c \geq -5.5$. Modifying the critical metallicity floor of SNET would require further testing to ensure that `StarNet` can enrich gas appropriately to higher levels without further training. Hence, in this work, we compare the two as methodological approaches, but not as an attempt to compare similar critical metallicities or floors. The sn5L-noZ simulation provides a very unique dataset. Since each `STARSS` particle inherits its metallicity from the host cell where it formed, this simulation will identify stars that formed in metal-free or low-Z gas by their metallicity fraction, while still allowing for star formation to proceed in the pre-enriched regions created by `StarNet`. This simulation will, in particular, provide the fraction of stars in a final galaxy that would not have formed if we enforced a metallicity requirement for star formation.

To further validate that `StarNet` is producing expected size of feedback region and in a reasonable number of events, we compare the fraction of volume enriched to varying degrees in Figure 4.14. The volume fraction enriched to low levels, i.e., $Z = -4.3, -5.5$ is matched by `StarNet` to 0.1 dex. Notably, `StarNet` does a fair job, with more deviation, of modeling the higher-Z remnants, i.e., with $Z \gtrsim -2.6$. The existence of these relatively high-Z regions will enable a broad range of metallicity for second-generation star formation, further increasing the diversity of protogalaxies enabled by `StarNet`. There are many times where the discrepancy between `StarNet` and the PHX solution is larger. Many of these can be attributed to `StarNet` depositing the "final state": there is no point where `StarNet` models the early phases of expansion–the SBRs existence is binary–so the early phase of SBR is "fast-forwarded" to the final state that would exist 16 Myr later. This effect causes the enriched volume to be incremented in steps instead of a smooth transition between times, which leads to temporary larger errors.

To further discuss the distribution of enrichment by `StarNet`, we compare the two-point correlation functions ($\xi(r)$) for halos enriched by primordial gas both in PHX256-1 and cosmological simulations using

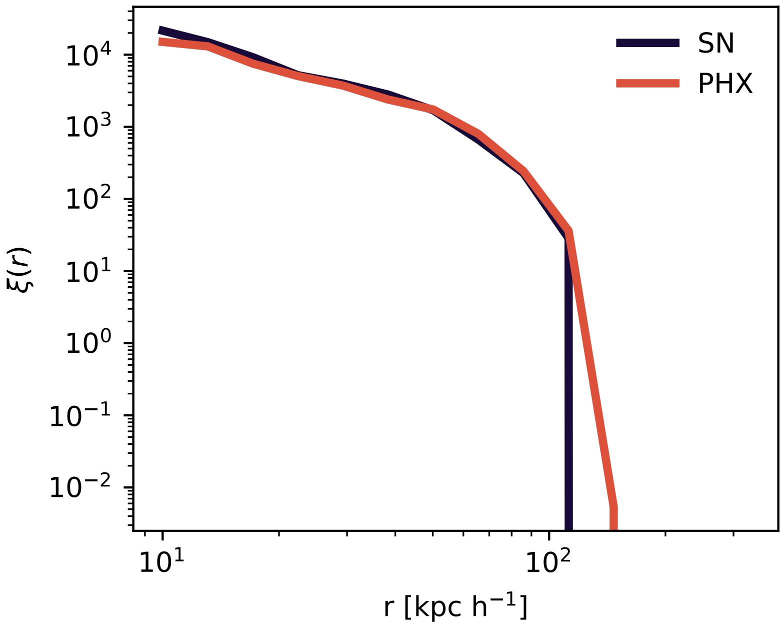


Figure 4.15. Two point correlations of protogalaxies enriched with primordial gas in the SNET and PHX256-1 simulations. With minor deviation at low-r, $\xi(r)$ is consistent across simulations.

`StarNet` for primordial enrichment, where both versions share identical initial conditions. We calculate $\xi(r)$ as

$$\xi(r) = DD(r)/RR(r) - 1 \tag{4.23}$$

where $DD(r)$ is the number of halo pairs with separation equal to r, $RR(r)$ is the number of randomly distributed pairs that would have separation r, with radii are split into bins of width δr[7]. Ideally, `StarNet` would duplicate $\xi(r)$ from PHX256-1, indicating that `StarNet` had enriched a very similar distribution of halos predicted by `StarNet`. Indeed, Figure 4.15 shows that sn5l-1 and PHX256-1 have nearly identical $\xi(r)$, with only minor deviation at high and low r.

4.6.2 Measuring the Impact of Heterogeneous Metallicity Initial Conditions from Population III Stars

In this section, we make comparisons between SNET and STS suites to quantify the impact of the heterogeneous metallicity field predicted by `StarNet` on the first generations of enriched star formation. We elected to use the 5LZ simulations from Section 4.5 because of their converged behavior when compared to 6LZ, despite requiring less than half the runtime. As well, despite the fact that there is no metallicity floor

---

[7] `halotools`: https://halotools.readthedocs.io

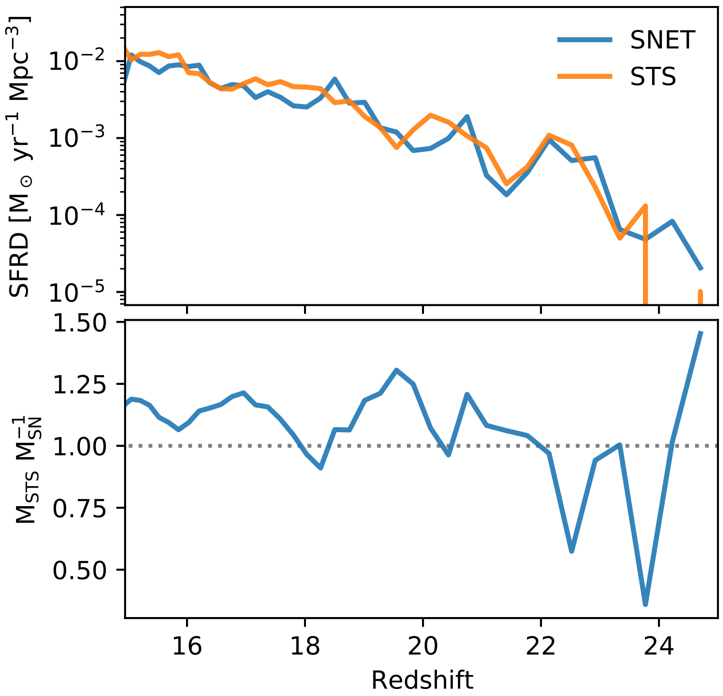


Figure 4.16. Comparing SFRD of SNET and STS. The lower panel shows the error in stellar mass obtained by using a uniform metal density IC as in the STS.

in SNET, the pre-enriched bubbles deposited by `StarNet` tend to enrich higher than $Z \sim -5.5$–the cooling imparted to the collapsing halos from the new metallicity field resembles the 5LZ simulations much more closely than the 5L simulations with $Z_f = -5.5$.

Global Simulation Statistics

In Figure 4.16, we show both SFRD and stellar mass difference across SNET and STS with stellar mass diference given by $M_{*,\text{STS}}/M_{*,\text{SN}}$. SNET and STS have comparable SFRD at all times (top), however the small differences in SFRD accumulate to rather large errors in cumulative stellar mass formed (bottom). At very early times ($z > 21$), STS drastically under-produces stars, having as much as 50% less stellar mass. This early trend is reversed at later times, where STS overproduces stars and eventually has $\sim 20\%$ excess stellar mass compared to SNET.

To further describe the volume SFRD, we present the counts of star-forming halos in SNET compared to STS in Figure 4.17. There are slightly fewer star forming halos in SNET suite at any given redshift. This, combined with the SFRD that matches STS, suggests that the individual halos in SNET undergo similar star formation at similar times, on average. However, since there are fewer halos in SNET, there exist some

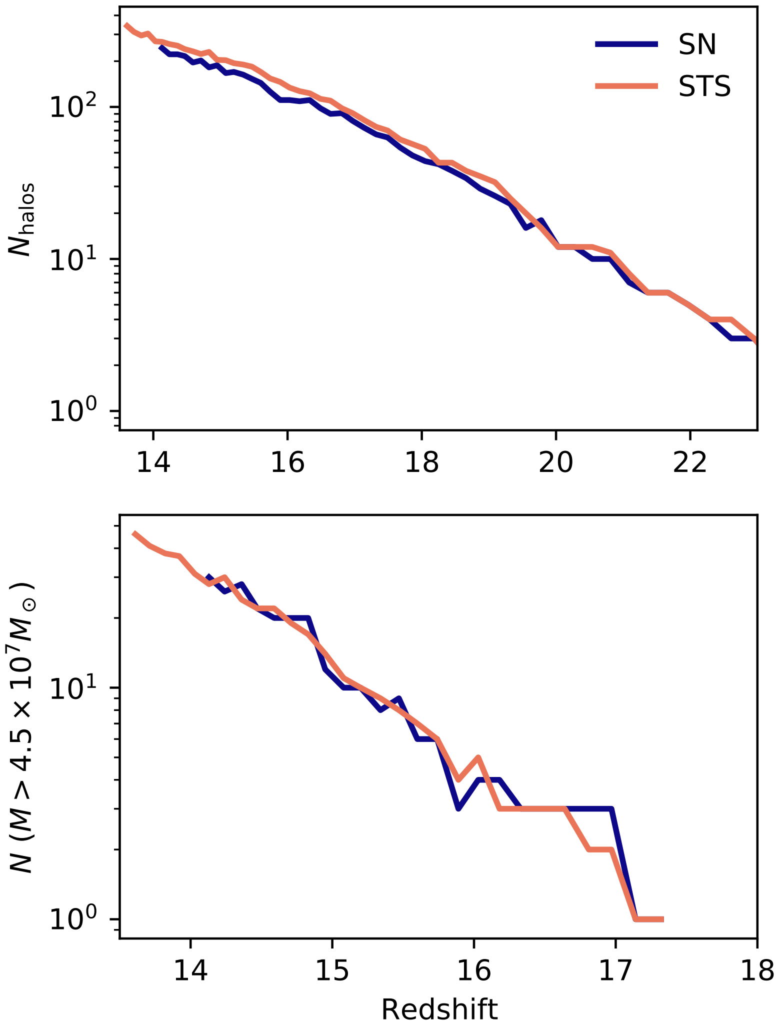


Figure 4.17. Counts of star-forming halos comparing SNET and STS suites. SNET halo counts are $\lesssim 0.1$ dex below STS for many redshifts, indicating a systemic but minor reduction in star-forming halo counts due to using `StarNet`. The lower panel only counts more massive protogalaxies with $M_h > 4.5 \times 10^7$ $\mathrm{M}_\odot$. The difference between SNET and STS is absent.

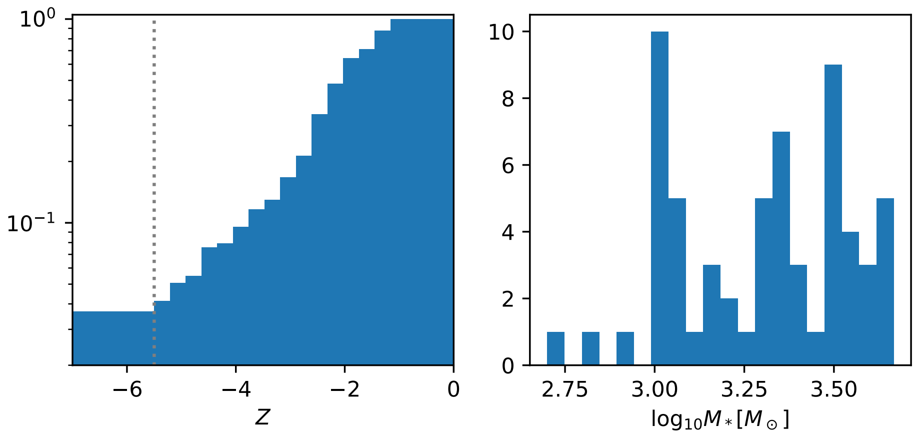


Figure 4.18. Cumulative MDF of `STARSS` particles formed in sn5l-noZ. A long tail of low-metallicity clusters have formed with $Z \leq -5.5$, accounting for $\sim 4\%$ of stars within the simulation. We also show the mass of these low-metallicity clusters: they are not low-mass errors, but full $\sim 10^3$ $\mathrm{M}_\odot$ clusters that have formed in primordial gas.

pristine halos in SNET that are enriched star-forming protogalaxies in STS. In total, STS contains 8.6% more star forming halos than SNET. However, we can see by the lower panel that any difference between SNET and STS is absent when only considering more massive protogalaxies with $M_h > 4.5 \times 10^7$ $\mathrm{M}_\odot$.

In Figure 4.18, we show the cumulative MDF of the sn5l-noz simulation. Of note is the fact that $\sim 4\%$ of stars in sn5L-noZ formed in gas below $Z_c = -5.5$. This bulk of stars represents the error that would result from ignoring the requirement of metallicity prior to enriched star formation. However, there is nothing stopping `StarNet` from enriching gas well above $Z_c = -5.5$. We can therefore ask what the effect is when we vary Z_c and ask which stars have formed in error as a function of changing Z_c.

Figure 4.19 shows the fraction of stars that formed with $Z < Z_c$ while varying Z_c, at various redshifts. As redshift increases, the central values of Z_c have increasing error, i.e., stars forming with $Z < Z_c$. In particular, at $z > 18.1$, more than 15% of stars were formed with $Z < -6$, well below the threshold of $Z = -5.5$ used in the simulation. This is a compounding error upon further reflection: since some fraction of stars form with $Z < Z_c$, they also contribute to the enrichment of the surrounding gas, leading to further star formation within the region. However, we are not able to differentiate these second-generation errors in this analysis because the star particles do not retain the source of their metallicity (whether from Population II or Population III sources). The logical conclusion is that some fraction of stars with $Z > Z_c$ are second generation errors that also should not have formed.

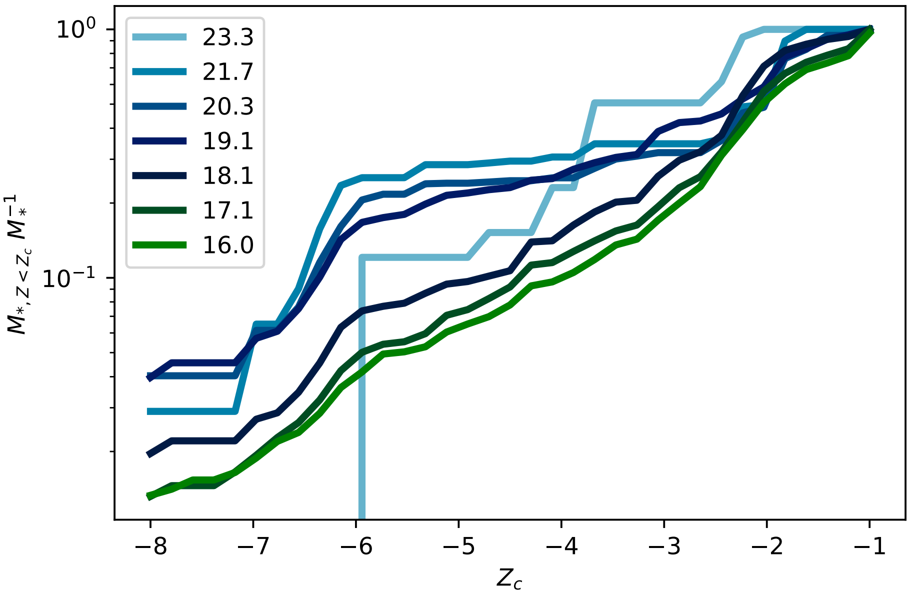


Figure 4.19. Effect of metallicity floor on stellar mass errors. The fraction of stars forming with $Z < Z_c$ for varying Z_c are shown at various redshifts.

Halo and Galaxy Statistics

Thus far, the presented statistics focus on a global perspective, disregarding individual halo or star forming regions statistics. In this final section, we will zoom in and discuss single halos and their proto-galaxies. To generate our galaxy catalog, we use `Rockstar` (Behroozi et al., 2013), requiring at least 100 DM particles within the halo. Each halo is then iterated, and we log various quantities beyond virial radius and mass: stellar mass, averaged historical star formation rate, bolometric luminosity, ionized and neutral hydrogen gas masses, and the mean metallicity of the gas within R_{1000}[8]. The bolometric luminosity and optical luminosity is as inherited to the `STARSS` algorithm from FIRE-2. Figure 4.20 shows plots of all quantities discussed above. At the final redshift, SNET and STS contain $> 10^3$ halos each. So many individual points would not be legible on any plot, so we represent the full distribution of halos by gaussian kernel density estimation. To aid in identifying trends in the plotted halos, we also include a bicubic spline fit to the 50th quantile of halos. We can measure the error in the spline fit by comparing the spline prediction $\hat{\mathbf{Y}}$

---

[8] Defined as the radius, R, at which the mean density of the within R, $\bar{\rho}_h(R)$ satisfies $\bar{\rho}_h(R) \leq 1000\bar{\rho}_b(z)$, for the mean baryon density (ρ_b) in the universe at redshift z.

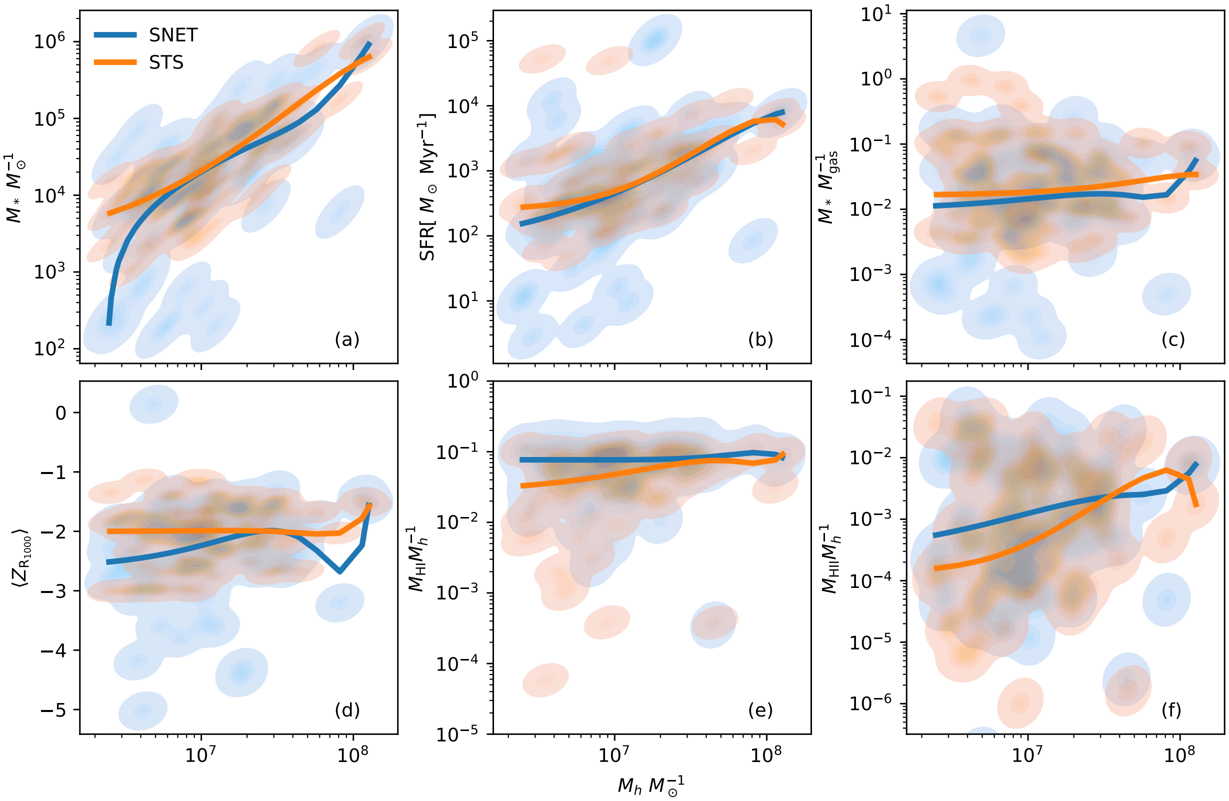


Figure 4.20. Summary statistics for single halos comparing STS (orange) and SNET (blue) as functions of halo mass. The full catalog of halos is represented by gaussian kernel density estimation, shown on the background of each plot. Lines represent a bicubic fit to the 50th quantile of halos, using the degree of fit that minimizes R_2 score.

to the known value $\mathbf{Y}$ by R_2 score given by:

$$R_2(\boldsymbol{Y},\hat{\boldsymbol{Y}}) = 1 - \frac{1}{N}\frac{\sum_{i=1}^{N} y_i - \hat{y}_i}{\sum_{i=1}^{N} y_i - \bar{y}}. \tag{4.24}$$

We then choose the degree of spline fit that maximizes the R_2 score; this prevents overfitting (by simply using high-degree fits) while preserving non-linear behavior that can be achieved with higher-degree fitting.

In general Figure 4.20 shows that SNET and STS converge to similar behavior for high-mass halos with $M_h \gtrsim 5\times 10^7$ M$_\odot$. SNET displays different behavior at lower mass halos, however, showing more low-mass halos with, e.g., low SFR, $M_*/M_{\rm gas}$, and optical luminosity (L_v). Also notable is the extremely lower metallicity gas, $Z_{\rm R_{1000}}$, found in the SNET suite. The ability of SNET to model low-mass halos is particularly apparent in M_* (Subplot a), where there are several halos with $M_* \lesssim 3\times 10^3$ M$_\odot$, below the minimum stellar mass observed in STS. This suggests that a future application for SNET may be to assist in filling in the faint-end of mass-luminosity relationships in simulations. Similar behavior is seen in f_* (Subplot c); the low-mass halos with $M_h < 6.8\times 10^6$ M$_\odot$ have significantly reduced star formation efficiency (given as M_*/M_g). SNET also increases the diversity in metallicity of dense regions in star forming halos as shown in Subplot d: while most halos in SNET have similar $\langle R_{1000}\rangle$ as STS, SNET allows for low-metallicity cores to exist, providing an avenue for metal-poor star formation even in previously enriched star forming halos.

One motivating factor to develop SNET was to model pristine halos that could collapse to supermassive black holes, as seen in Regan et al. (2017), without the high resolution requirements of the *Renaissance Simulations* (Xu et al., 2013) where such halos were found. To determine if SNET can in fact model these regions, we search STS for star forming halos and determine whether the corresponding halo in SNET is forming stars. These “error halos” in STS represent sites where the pristine gas could have ideally formed supermassive black holes, but also could identify large reserves of pristine gas that can fuel star formation in neighboring protogalaxies at later times.

As shown in Figure 4.21, the number of error halos, i.e., those that are forming stars in STS but not in SNET, is increasing with decreasing redshift. While interesting, the `StarNet` is known to have a “phase” error, predicting some star forming regions $\lesssim 35$ Myr early, or others $\lesssim 10$ Myr late (W21). Identifying halos that are truly erroneous would require identifying those halos that are non-star forming in SNET over periods of time > 30 Myr, in order to reduce the possibility that the error halo is simply due to phase error. To that

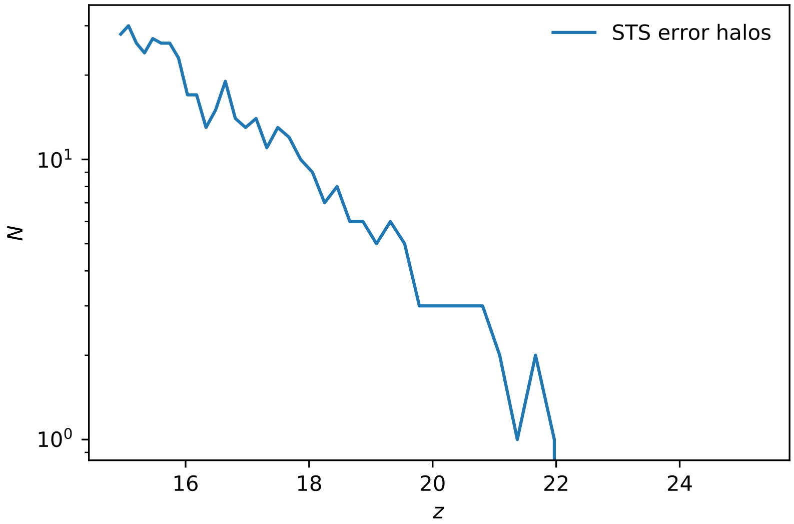


Figure 4.21. Counts of halos forming stars in STS where the corresponding halo in SNET has no stellar mass as a function of redshift.

end, we analyzed which error halos have existed for the longest times in SN; there are three halos at the final data output, $z = 14.95$, that have been forming stars in STS for $\gtrsim 30$ Myr and are not enriched by metals from either Population III or Population II sources. A prototypical example is presented in Figure 4.22. In this example, the halo of focus is centered in the frame with $M_h = 3.9 \times 10^6$ $\mathrm{M}_\odot$. Notably, this halo is in a busy region with many star forming halos nearby (as indicated by the metallicity field sourced from stellar feedback). If a halo such as this one continues without forming stars, it could be a candidate for supermassive black hole collapse or a substantial reservoir of pristine gas for star formation at at later times. Disregarding the phase error, there are 19 pristine halos in SNET corresponding to error halos at the final output, or 26 halos if we include those that have been enriched by `StarNet` but have not begun star formation. These 26 halos represent 8.6% of star-forming halos in STS that have not begun forming stars in SNET. Only one has $M_h > 10^7$ $\mathrm{M}_\odot$, suggesting that the error halos are less common at higher mass.

Observational quantities within `StarNet`

The following discusses the effect of including Population III via `StarNet` from the perspective of observational quantities. Although the simulations we present are far outside the observational range of even the James-Webb Space Telescope, these samples may be relatable to ancient protogalactic remnants known as ultra-faint dwarf (UFD) galaxies (Simon, 2019). In particular, we will examine the velocity dispersion ($\sigma(\nu)$), stellar metallicity (Z_*), absolute magnitude (M_V) and neutral and ionized hydrogen masses (M_HI

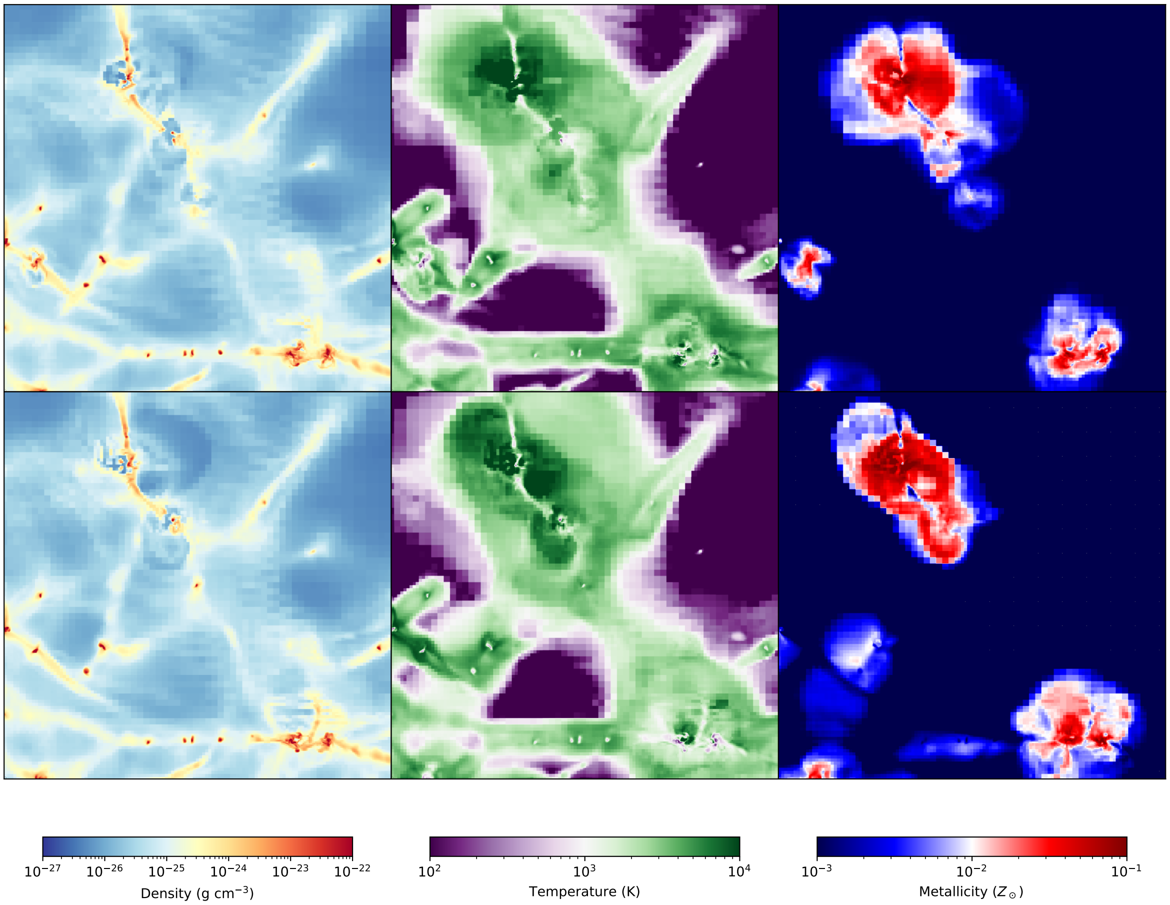


Figure 4.22. An exemplary halo that is forming stars in STS (top), but has no stellar mass in SNET (bottom). The halo of interest is centered in all plots. Metallicity includes metals from all `STARSS` particles and StarNetRuntime depositions.

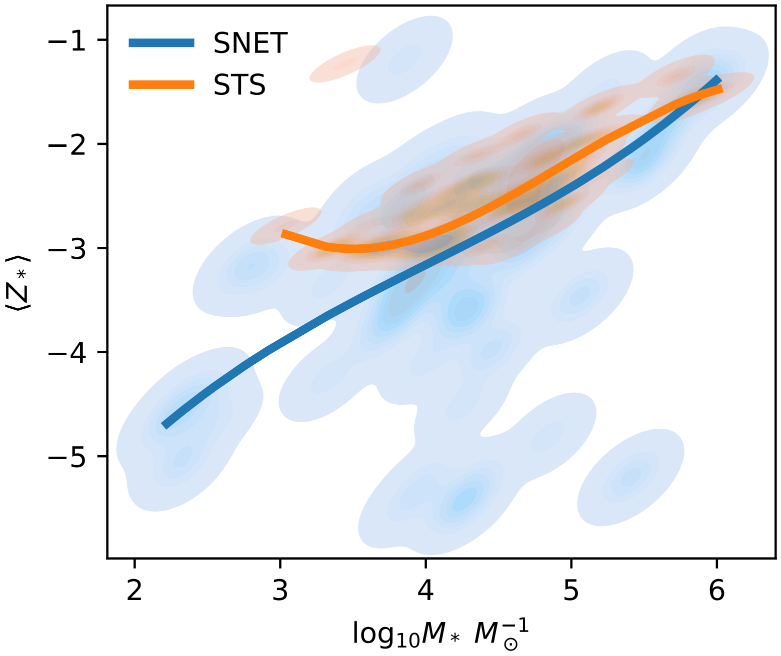


Figure 4.23. Metallicity of stars as function of halo stellar mass.

and M_{HII}). Each quantity is shown as a function of stellar mass, another easily studied quantity from an observational perspective. We study these quantities to determine whether the inclusion of `StarNet` can generate protogalaxies that may evolve into UFDs.

Figure 4.23 shows average stellar metallicity as a function of halo stellar mass. This dataset continues the trend that SNET generates more diverse behavior than STS, allowing for exceptionally low-metallicity systems to form at lower stellar mass than is observed in STS. The 50th quantile fit shows an increasing trend correlated to stellar mass, however, there is significant scatter allowing for both high and low metallicity examples. Particularly in SNET, there are middling mass protogalaxies with exceptionally low metallicity. There is also a much more exaggerated low-metallicity, low-stellar-mass tail present in the SNET suite that is not present in STS.

Early results from James Webb Space Telescope (JWST) have found several galaxies at $z \sim 8$ with low metallicity, albeit at higher stellar mass than the protogalaxies observed in this work. (Curti et al., 2022). Figure 4.24 shows the metallicity quoted in O/H, as in Curti et al. (2022), with the approximation that oxygen comprises 35% of the mass-weighted metallicity within R_{1000} of the halo (Torrey et al., 2019). Since our protogalaxies are of much lower stellar mass, we show the level of the observed metallicity by

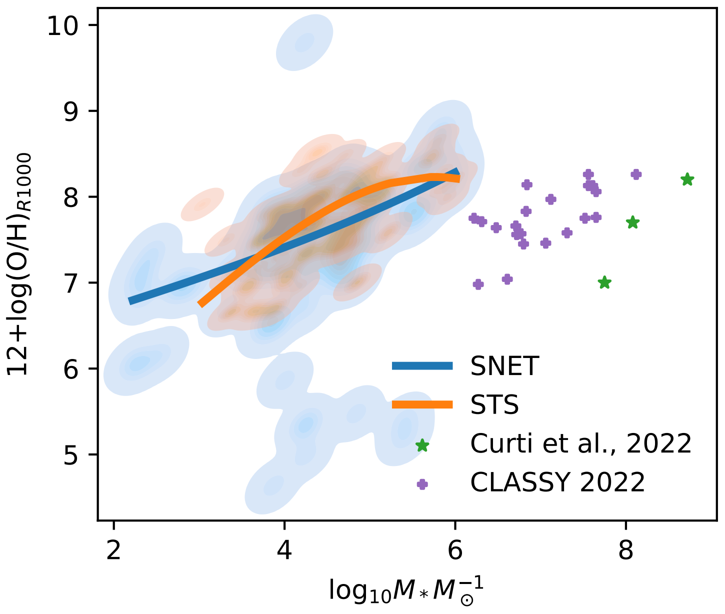


Figure 4.24. O/H as function of stellar mass. Green stars represent the level of metallicity observed in Curti et al. (2022), although their observed galaxies had much higher stellar mass ($10^{7.5} < M_*/M_\odot < 10^8$). We also annotate a low-stellar mass subset of samples obtained from the CLASSY database in purple (Berg et al., 2022)

green stars, however, the stellar mass of their observations are $> 10^{7.5}$ $M_\odot$. Nonetheless, the inclusion of SNET again allows more diverse behaviour in our simulated protogalaxies, so that their low metallicity regime overlaps with that of the observations from JWST. The inclusion of `StarNet` also enables the range of metallicity seen in the CLASSY database (Berg et al., 2022), however, most samples there are also much higher in stellar mass than the protogalaxies observed in `StarNet`'s small volumes. In order to expand on this relationship, `StarNet` will need a more efficient implementation that can explore larger volumes and higher-mass protogalaxies.

Figure 4.25 shows the radial velocity dispersion as a function of stellar mass. All samples above the 50th quantile have $\sigma(\nu) > 2$, where most UFDs are found (Simon, 2019). This sample does have a significant population with lower dispersion, typically at low stellar mass ($M_* < 10^5$ $M_\odot$). Figure 4.26 builds on the study of $\sigma(\nu)$ by showing $\sigma(\nu)$ as a function of absolute magnitude (M_v). UFDs typically occupy the region with $M_v > -8$: this regime is probed by both SNET and STS, however there is a higher magnitude ($-2 < M_v < -5$) sample of halos with low velocity dispersion that is only present in SNET.

It may be possible that the inclusion of `StarNet` allows the simulation to capture the observed behavior of low-mass dwarf galaxies more effectively than simulations employing metallicity floors. To

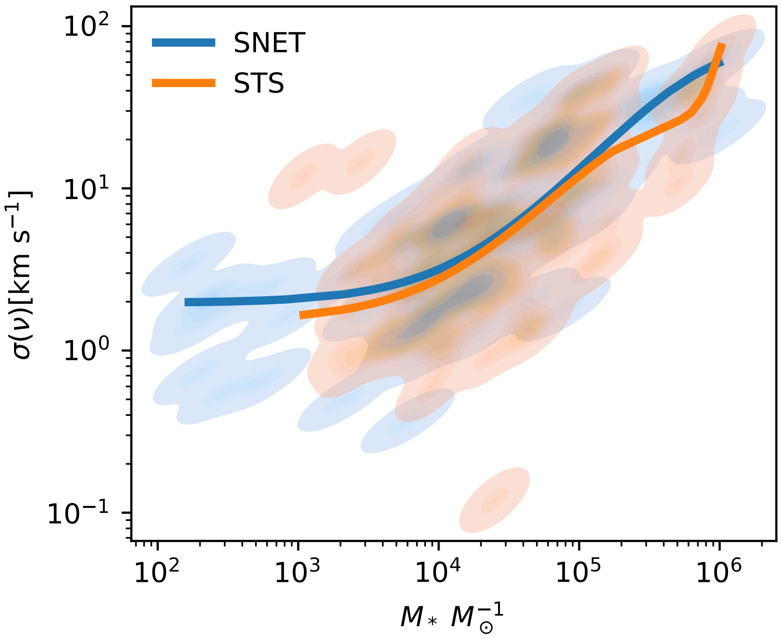


Figure 4.25. Velocity dispersion as function of stellar mass.

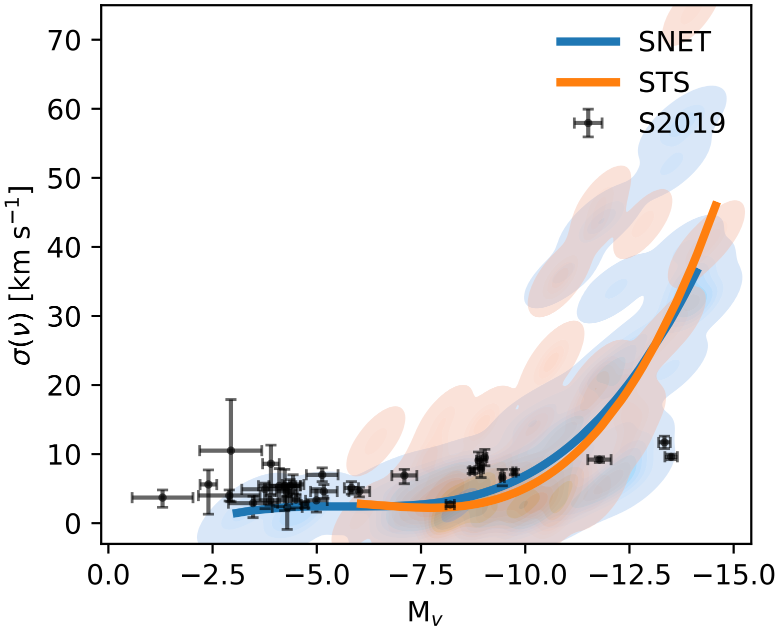


Figure 4.26. Velocity dispersion as function of absolute magnitude. Also shown are observational data of ultra-faint dwarf galaxies (see text for references).

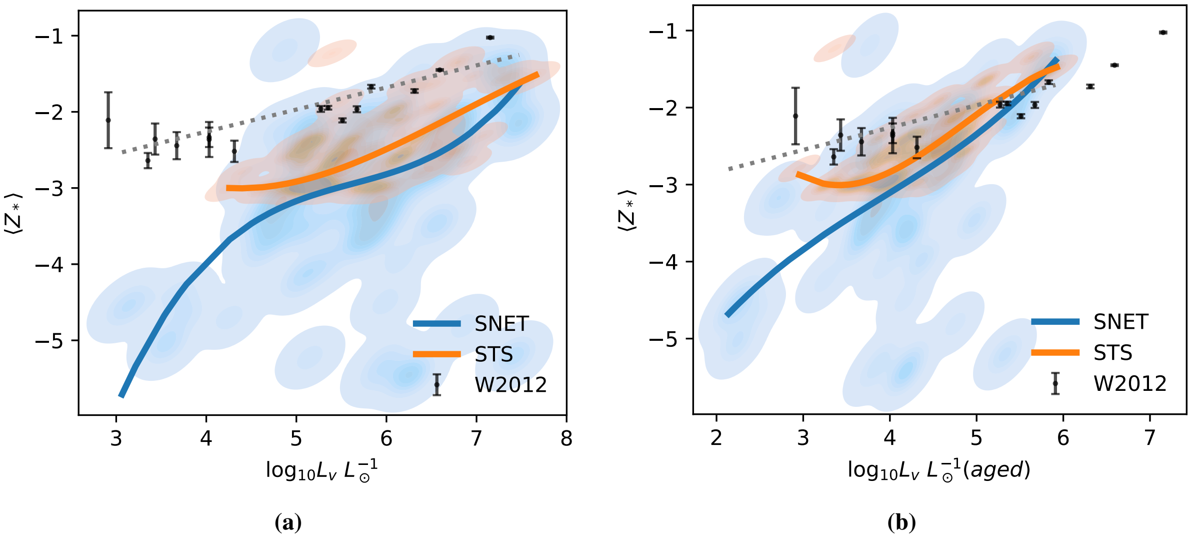


Figure 4.27. Mean stellar metallicity as function of luminosity. Panel 4.27a shows the total halo luminosity for the stellar ages and masses present within the virial radius. Panel 4.27b shows the same quantity, if we suppose the stellar population has aged by 4 Gyr. The expected luminosity-metallicity relationship (Equation 4.25) is also shown (gray hashed line). Local group dwarfs (black points) are also shown (Willman & Strader, 2012) with their corresponding error estimates.

explore SNET and STS compared to observations, we include observational data along with our simulated data. The observations were compiled by Simon (2019), and include the works of many authors (Simon & Geha, 2007; Walker et al., 2009b,a; Bechtol et al., 2015; Koposov et al., 2011, 2015; Crnojević et al., 2016; Simon et al., 2015; Li et al., 2017, 2018; Simon et al., 2011; Torrealba et al., 2018; Mateo et al., 2008; Willman et al., 2011; Spencer et al., 2017; Collins et al., 2017; Torrealba et al., 2016a; Caldwell et al., 2017; Kirby et al., 2015; Koch et al., 2009; Majewski et al., 2003; Bellazzini et al., 2008; Kim & Jerjen, 2015; Kim et al., 2015; Torrealba et al., 2016b; Walker et al., 2016; Muñoz et al., 2018). SNET extends the high-magnitude tail past $M_v > -6$, allowing the potential coverage of many observations with $\sigma(v) < 10$ and $M_v > -6$, where the bulk of UFDs reside. Further evolution of the SNET suite will be interesting to see if the high-magnitude, low $\sigma(\nu)$ region becomes more well-represented. We finish our study of observational quantities with Figure 4.27a, showing stellar metallicity as a function of luminosity. Also shown is the metallicity-luminosity relation given by

$$[\mathrm{Fe/H}] = -1.68 + 0.29 \log\left(\frac{L_v}{10^6 L_\odot}\right), \tag{4.25}$$

although we use total mass-weighted metallicity Z as a proxy for [Fe/H][9] (Simon, 2019). Both SNET and STS have a broad, 3-5 dex spread of metallicity that is not represented by Equation 4.25, however, both suites have samples that overlap the analytic model. As has been the trend, the inclusion of SNET permits a much broader range of halo observables, increasing the metallicity range from ~ 3 dex to ~ 5.5 dex.

To compare SNET to observations, we include dwarf and ultra-faint dwarf (UFD) observations of the Local Group (Willman & Strader, 2012) (W2012). The dwarf galaxies are well represented by both SNET and STS, however, both simulation suites tend toward lower metallicity and do not duplicate the low-luminosity end of observational data. However, SNET does have a low-luminosity tail that extends almost as low as observations, albeit with much lower mean metallicity. The time evolution of this low-luminosity tail will be of great interest in studies where simulations approach reionization, $z \lesssim 6-7$.

Since our sample of halos is taken from $z = 14.95$, one may ask what the relationship would look like for a halo that stopped forming stars at that point and aged to more modern times. Since UFDs are thought to have ceased star formation very long ago, $z \sim 6-7$, then this may serve as a rough approximation of how some halos could age within the STS and SNET framework. In Figure 4.27b, we have plotted the luminosity-metallicity relationship, however, the luminosity of the stellar population has been aged by 4 Gyr. The distribution of luminosity has much more overlap with the UFD observational points, although a lack of metallicity evolution means that neither STS or SNET match the expected slope of Equation 4.25.

4.7 Discussion

In this work we have developed both a resolution-intelligent, physically-motivated feedback method for stellar sources in enzo (`STARSS`), and a new method to initialize the metallicity field of astrophysical simulations (`StarNet`). `STARSS` has shown success as a resolution-intelligent feedback algorithm, but still struggles with resolution effects regarding star formation. If cooling dynamics are substituted for force resolution, i.e., using $Z_\mathrm{f} = -3$ or using `StarNet` to seed the metallicity field, much of the resolution dependence is resolved. That said, creating a truely resolution-intelligent star formation algorithm will require a new paradigm of star formation recipes; the current criteria are *always* resolution sensitive to some degree, because the force resolution of the simulation is intrinsically related to the spatial resolution. One possibility is to use super-resolution techniques (e.g., Kappeler et al., 2016) as a surrogate model of high-resolution

---

[9] This assumption holds true for solar metallicity stars, but will break down for, e.g., carbon-enhanced metal poor stars whose abundances do not resemble solar.

gas dynamics in low-resolution regions, but this is left to future work. With `STARSS` current implementation, lowered resolution acts to delay star formation to later times, but maintains a similar SFR once star formation begins. This suggests that at lower redshift, the effect may be disregarded, depending on the focus of the simulation.

`StarNet` in its current state is a proof-of-concept work. It uses the inline Python capability of `Enzo`, which severely limits its scalability. Despite this restriction, we have simulated a significant portion of the canonical Population III era ($10 < z < 30$) in two 17.56 Mpc^3 simulation volumes. Using `StarNet` as a surrogate for Population III star formation and feedback enables substantial speedup, even in this unoptimized state. Using the *Phoenix Simulations* as a reference, PHX256-2 required ~ 5600 node-hours on the TACC-Frontera supercomputer using 6 nodes. There is significant acceleration using `StarNet`: sn5l-2-v3 required only 1025 node-hours running on 2 nodes, for a total speedup of 5.46×. It is useful to note that this speedup is the minimum to expect from future implementations of `StarNet`: incorporating `StarNet` into `Enzo-E` will enable load balancing and will benefit from the framework being translated to C++.

The inclusion of `StarNet` significantly modifies simulation behavior, not from a global perspective, but from the perspective of outliers and rare events. SNET and STS show similar SFRD within their respective volumes, but the slight differences in SFRD generate large differences in cumulative stellar mass; STS overproduces stars by $\sim 20\%$ by $z = 15$, despite under-producing stars at early times ($z \gtrsim 21$). The global similarity between SNET and STS is reinforced by examining star-forming halo counts, where we observe very similar numbers of star-forming halos in both SNET and STS. Any difference in regards to halo counts is largely resolved if we only consider higher-mass halos ($M_h \gtrsim 4.5 \times 10^7$ M$_\odot$). The difference in halo counts, resulting in those forming stars in STS but not in SNET, is increasing with decreasing redshift, suggesting that this is an error that would become more significant at lower redshifts.

One simulation in this work was designed solely to identify which stars would form in error if we removed the Z_c criterion for star formation. We find that this simulation overproduces stars by $\sim 4\%$ at $z \sim 16$. Although the difference is minor, this effect is more pronounced at higher redshift, with errors as high as 30% at $z \sim 23$. The error also increases if we increase Z_c, with as much as 15% of stars having $Z < -3$ at $z \sim 16$. The errors we note here are likely compounding: a single cluster formed in error can enrich neighboring gas, leading to more star formation that is also in error.

The net effect of including `StarNet` to model the initial metallicity field is the increased range of behaviors that can be observed in Figures 4.20-4.27. Particularly when we include observational data of dwarf galaxies in Figures 4.26 and 4.27, the inclusion of `StarNet` leads to halos whose characteristics overlap more substantially with observational data. While the shown observations are near to $z = 0$, there is suspicion that the UFDs have been essentially static since high-redshift (Simon, 2019) from a star formation and baryonic gas perspective, suggesting that they may be relics of a high-redshift universe. If so, then there should be a redshift at which our halo distributions begin to overlap with the observational data. This may not happen until reionization, $z \sim 7$, since many of these halos do not have enough gas to self-shield from external ionizing radiation. Such a simulation will be attainable once `StarNet` has been incorporated into `Enzo-E`. For now, if we synthetically age the population of dwarf galaxies in STS, the resulting $L_{\rm v}$-Z_* relationship overlaps much more with observations of UFDs. This quick and dirty comparison suggests that the age and star formation cutoff will have a significant impact on modeling UFDs, and modeling them with `StarNet` enables a much lower $L_{\rm v}$ cutoff at $L_{\rm v} \sim 10^2$ $\mathrm{L}_\odot$. Further reinforcing the importance of the dynamic range in metallicity achieved by incorporating `StarNet`, protogalaxies in SNET show the full range of stellar metallicity as observed in early results from the JWST, where $z \sim 8$ galaxies have been identified with $-5 < Z < -3.75$ (Curti et al., 2022). Although the observed galaxies have higher stellar mass than our simulated protogalaxies, a larger simulation domain and incorporation of `StarNet` into `Enzo-E` will enable us to explore galaxies in the stellar mass range of the JWST observations.

`StarNet` has also shown the ability to model rare events, such as a halo that, while forming stars in STS, has not been pre-enriched by a PIII association and so remains pristine in SNET. These halos are uncommon, with only nine examples at the final redshift of SNET and only three examples that have existed for $> 35 Myr$. Protogalaxies in STS that have yet to begin star formation in SNET are more common, with 8.6% (26 examples) of protogalaxies in STS having no twin in SNET. However, these samples may be extremely important to later evolution, serving as pristine reservoirs of gas to fuel star formation or as potential sites of supermassive black hole formation. Optimizing and integrating `StarNet` in future SNET simulations with larger volume to lower redshift will better elucidate the evolution, lifetime, and impact of these pristine halos.

This research was supported by National Science Foundation grant CDS&E grant AST-1615848 to M.L.N. The simulations were performed using ENZO on the Blue Waters supercomputer operated by the National Center for Supercomputing Applications (NCSA) with PRAC allocation support by the NSF (award number

ACI-0832662). Data analysis was performed on the Comet supercomputer operated for XSEDE by the San Diego Supercomputer Center.

Chapter 4, in full, is being prepared for submission for publication of the material as it may appear in Wells, A.I. & Norman, M.L. The First Galaxies and the Effect of Heterogeneous Enrichment from Primordial Stars. *The Astrophysical Journal* (2022). The dissertation author was the primary investigator and author of this paper.

Chapter 5

Future work

`StarFind` and `StarNet` have been resoundingly successful in their original intent: to model the heterogeneous metallicity field that would result from PIII associations. That said, this work has identified several avenues for improving upon the proof-of-concept methods.

5.1 Improving `StarFind`

The most improvement in `StarFind` can be achieved by further advancement of the PHX suite simulations, providing more training data in more varied environments. Further studies into interpretability may also be beneficial, such as measuring uncertainty with dropout (Lakshminarayanan et al., 2017), may also be useful in ensuring `StarFind` predictions are robust. Currently, `StarFind` operates on a fixed spatial resolution, with 156 comoving parsecs per grid cell. An extraordinary improvement would be to generalize the method so that it can operate on a range of resolutions, or irrespective of resolution. This would likely involve more manipulation of simulation data, e.g., using super-resolution methods (Kappeler et al., 2016) to increase resolution to the currently trained value.

5.2 Improving `STARSS`

`STARSS` has been very successful at resolution-intelligent stellar feedback. Star cluster formation, on the other hand, still poses a challenge. This is not a problem that can be solved easily, as the star formation depends on hydrodynamic field variables that are strongly affected by force resolution. Since the hydrodynamics of low-resolution simulations behave quantitatively differently than high-resolution, solutions to the star cluster formation problem will require either new criteria that are resolution insensitive, or new methods to super-resolve grid quantities, similarly to super-resolution methods discussed above.

Although problematic, we have also noted that the star formation problem is less pronounced if using either a higher-metallicity floor or the heterogeneous enrichment of `StarNet`.

A major shortcoming of `STARSS` also exists in the feedback algorithm. Since SNe are stochastic, and winds/radiation are determined by particle age for a stellar cluster, small clusters in highly resolved ($dx \lesssim 1$ pc) do not make sense. It is also a gross approximation that all stellar mass forms at the same instant in time. While very hard to address in `Enzo`, these problems are quite tractable in `Enzo-E` by assigning cluster population information–stellar masses, creation times, lifetimes, luminosities, and metal yields–as particle attributes. In this way, the composition of the particle can be stochastic, but the evolution will not be, and can be scaled to even model single stars without algorithmic modification.

5.3 Improving `StarNet`

`StarNet` has coalesced as a very good first step in surrogate modeling of the effects of the first stars. The current version of `StarNet` does have significant shortcomings that will need to be addressed to create a more robust and realistic model. There are two groups of major issues with the current paradigm `StarNet` uses (namely modeling PIII association regions as simple spheres of metallicity and warm-hot gas):

1. `StarNet` only models a simple sphere of metallicity, where we also apply a heightened gas temperature. While this captures the heterogeneity of the metallicity field on large scales, we lose small-scale heterogeneity within a single feedback region. High-resolution works have shown that the mixing of SN-sourced metals is inhomogeneous, and even within neighboring halos, will have a wide range of metallicity within the gas (Smith et al., 2015). Further, the PIII association will also ionize the gas to varying degrees–combined with SNe feedback, there will be significant disruption to the host halo (Whalen et al., 2008a). This is also behavior that the current `StarNet` cannot model.

2. One consequence of modeling single primordial stars is prompt second-generation star formation (Wells & Norman, 2022). `StarNet` currently loses all information pertaining to whether a second-generation star can form within the remnant region.

Discussed briefly in Appendix C, many of these issues could be overcome by implementing a deep-learning based model of feedback instead of the simple regression models of Chapter 3 and 4. In particular, there are two beneficial perspecives to view our problem from: one, of a translation, changing a pristine halo into a

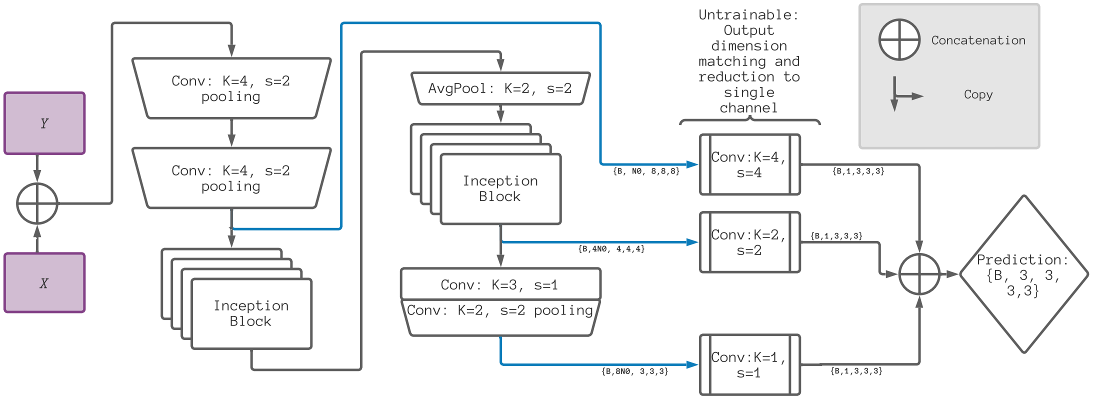


Figure 5.1. Multi-scale adversarial discriminator network. Sensitivity to varying length scales is achieved by using Inception architecture in addition to having paths to output at varying levels of processing.

PIII association remnant. Two, as a time series of individual steps that can slowly transform the region into a PIII association remnant. The first perspective lends itself to the ideas of image-to-image translation with generative adversarial networks (GANs) as used in other areas of deep learning (Goodfellow et al., 2014; Isola et al., 2016; Zhu et al., 2017), while the second perspective is adaptable to time-series modelling with long-short term memory networks (LSTMs) (Hochreiter & Schmidhuber, 1997; Mohan et al., 2019).

A GAN is composed of a pair of deep learning networks; a generator, tasked with generating realistic samples, and a discriminator, tasked with distinguishing real and generated data. One newer application of GANs is to translate images from one domain to another, e.g., a photo at noon to the same photo at midnight, or similarly, summer to winter. It is not hard to imagine a region before/after PIII association fitting into this conceptual framework. Several important pieces of this idea have already been developed: The generator can use a modified version of `pix2pix` (Isola et al., 2016) that can operate on 3D data, and is further modified to predict additional quantities, e.g., Population II stellar mass. Such a network has been implemented in the development of `StarNet` and is exactly the U-net and its variations from Chapter 2. The generator acts as an encoder-decoder but has additional connections between encoding and decoding levels to help background information flow from input to output, which assists with identity transformations (i.e., the part of the image that does not change). The discriminator developed for this purpose is unique to this work. Intuitively, PIII association feedback regions should have multiple scales of importance: large scale temperature and metallicity, but also small scale perturbations, such as extremely hot outflow channels and regions of varying metallicity within the remnant. The Multi-Scale Adversarial Discriminator (MSAD) network in Figure 5.1

achieves this fairly well in testing. MSAD has several paths to the output prediction that have sensitivity to varying length scales, while inception blocks (Chapter 2) give additional multi-scale sensitivity.

An example output from early development of GAN methods is shown in Figure 5.2. A great advantage of this method over the linear regression models of Chapter 4 is its ability to model several fields. Although, of course, the model must be trained on which fields to predict, it could predict any field of interest carried on the computational grid (or derived from it). This example goes several steps beyond the linear regression models to predict the final baryon density field and velocity fields, resulting in a momentum solution for the remnant.

Using GANs or LSTMs for this purpose is immature, lacking in both development and validation for future use, but the development thus far serves as a solid launching pad from which to vastly improve the `StarNet` method.

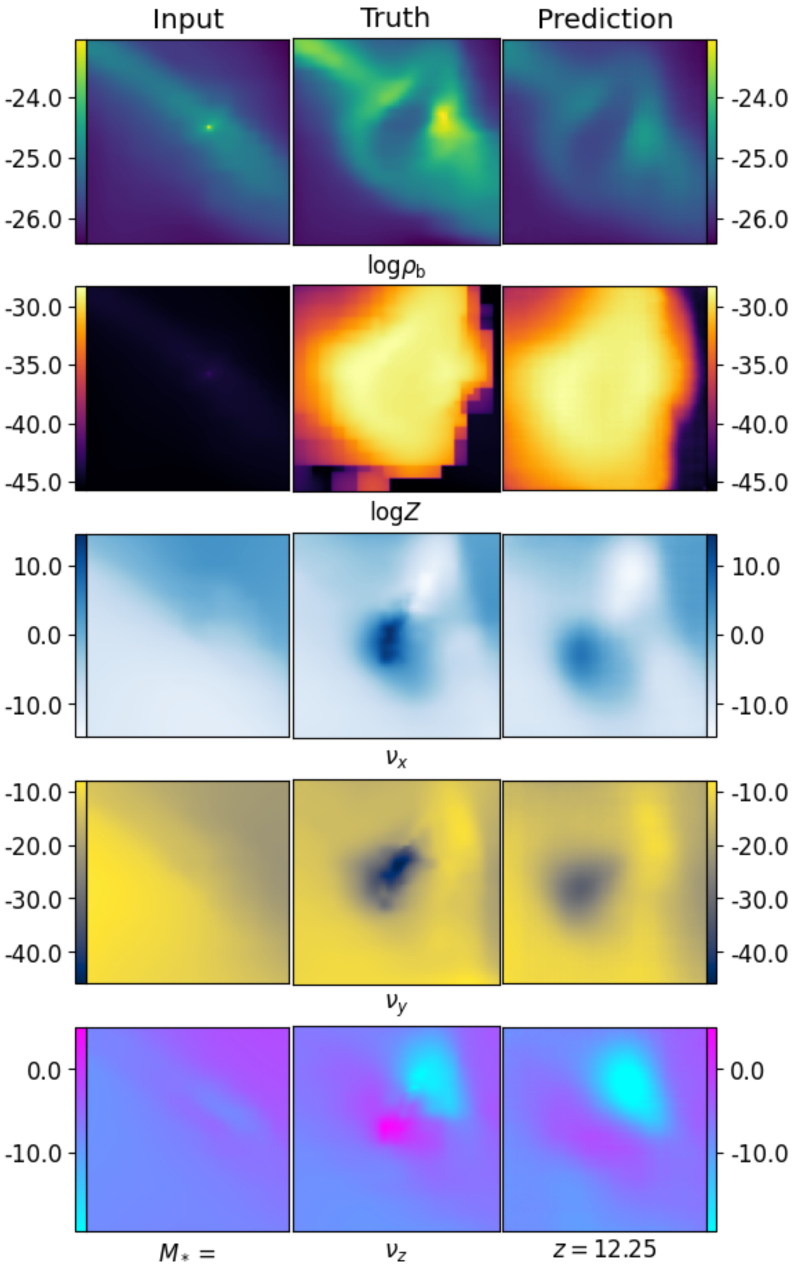


Figure 5.2. Example using GANs to model PIII association regions. The input hydrodynamic fields are shown in the Input column, the actual state of the region after 16 Myr in the Truth column, and the output from the generator is shown in the Prediction column. The predicted fields include baryon density (top), metal density from Population III sources (row 2), and velocity (rows 3-5). This particular example was achieved with 35 epochs of training.

Appendix A

Initial Conditions and Parameters

A.1 MUSIC Initial Conditions

The initial conditions files to run simulations in this work were generated with `MUSIC`[1] (Hahn & Abel, 2011). After compiling `MUSIC`, the initial conditions for the PHX256-1,2, and PHX256-HYD simulations can be remade with the following configuration file:

```
[setup]
boxlength = 1.765
zstart = 99
levelmin = 8
levelmin_TF = 8
levelmax = 8
padding = 16
overlap = 4
align_top = yes
baryons = yes
use_2LPT = yes
use_LLA = yes
periodic_TF = yes
avg_fine = yes

[cosmology]
Omega_m = 0.3111
Omega_L = 0.6889
Omega_b = 0.048975
H0     = 67.66
sigma_8 = 0.811
nspec = 0.965
transfer = eisenstein

[random]
```

---

[1] https://www-n.oca.eu/ohahn/MUSIC/

```
# add
#    seed[8]             = 86753009
# to seed PHX256-2 initial conditions
seed[9] = 201104190
seed[10] = 301104190
seed[11] = 401104190
seed[12] = 501104190

[output]
# ENZO - also outputs the settings for the parameter file
format = enzo
filename = ic.phoenix256

[poisson]
fft_fine = yes
accuracy = 1e-5
pre_smooth = 3
post_smooth = 3
smoother = gs
laplace_order = 6
grad_order = 6
kspace = no
```

or for the PHX512 simulation:

```
[setup]
boxlength = 3.53
zstart = 99
levelmin = 9
levelmin_TF = 9
levelmax = 9
padding = 16
overlap = 4
align_top = yes
baryons = yes
use_2LPT = yes
use_LLA = yes
periodic_TF = yes
avg_fine = yes

[cosmology]
Omega_m = 0.3111
Omega_L = 0.6889
Omega_b = 0.048975
H0     = 67.66
sigma_8 = 0.811
nspec = 0.965
transfer = eisenstein
```

```
[random]
seed[9] = 201104190
seed[10] = 301104190
seed[11] = 401104190
seed[12] = 501104190

[output]

##ENZO - also outputs the settings for the parameter file
format = enzo
filename = ic.phoenixL0

[poisson]
fft_fine = yes
accuracy = 1e-5
pre_smooth = 3
post_smooth = 3
smoother = gs
laplace_order = 6
grad_order = 6
kspace = no
```

The final required files will be generated and deposited into the ic.phoenix* directory specified in the configuration file.

Appendix B

Parameter Files

B.1 Enzo Parameter Files

`Enzo` simulations start with a set parameter file. It contains, e.g., radiation and chemistry parameters, grid information, and refinement criteria to be followed in the simulation. Aside from the simulation box size and resolution, the PHX256-1,2 and PHX512 all share identical parameters:

```
#
StopCycle                              = 100000
StopCPUTime                            = 420000

ParallelRootGridIO                     = 1
ParallelParticleIO                     = 1

UnigridTranspose                       = 2
HierarchyFileOutputFormat              = 1
CorrectParentBoundaryFlux              = 1


# AMR PROBLEM DEFINITION FILE: Cosmology Simulation (amr version)

#
#  define problem
#

ProblemType                            = 30
TopGridRank                            = 3
// change to match grid dimension in initial conditions:
        // 512 512 512 for PHX512 simulation
TopGridDimensions                      = 256 256 256
PotentialIterations                    = 10
SelfGravity                            = 1
```

```
TopGridGravityBoundary                  = 0
LeftFaceBoundaryCondition               = 3 3 3
RightFaceBoundaryCondition              = 3 3 3

#
#  problem parameters
#

CosmologySimulationOmegaBaryonNow       = 0.048975
CosmologySimulationOmegaCDMNow          = 0.262125
CosmologySimulationNumberOfInitialGrids = 1
CosmologySimulationDensityName          = GridDensity
CosmologySimulationVelocity1Name        = GridVelocities_x
CosmologySimulationVelocity2Name        = GridVelocities_y
CosmologySimulationVelocity3Name        = GridVelocities_z
CosmologySimulationCalculatePositions   = 1
CosmologySimulationParticleVelocity1Name = ParticleVelocities_x
CosmologySimulationParticleVelocity2Name = ParticleVelocities_y
CosmologySimulationParticleVelocity3Name = ParticleVelocities_z
CosmologySimulationParticleDisplacement1Name = ParticleDisplacements_x
CosmologySimulationParticleDisplacement2Name = ParticleDisplacements_y
CosmologySimulationParticleDisplacement3Name = ParticleDisplacements_z

#
#  define cosmology parameters
#

ComovingCoordinates                 = 1
CosmologyOmegaMatterNow                 = 0.3111
CosmologyOmegaDarkMatterNow          = 0.262125
CosmologyOmegaLambdaNow                 = 0.6889
CosmologyHubbleConstantNow              = 0.6766
// change this to match the box size in initial conditions:
//      3.53 for PHX512
CosmologyComovingBoxSize                = 1.765
CosmologyMaxExpansionRate               = 0.015
CosmologyInitialRedshift              = 99.000
# alter here for desired final redshift
CosmologyFinalRedshift                = 25.000
GravitationalConstant                   = 1
#
#  set I/O and stop/start parameters
#

DataDumpName                            = SO
RedshiftDumpName                        = RD

#
```

```
#  set hydro parameters
#
Gamma                                  = 1.6667
PPMDiffusionParameter                  = 0
DualEnergyFormalism                    = 1
InterpolationMethod                    = 1
FluxCorrection                         = 2
ConservativeInterpolation              = 0
CourantSafetyNumber                    = 0.3
ParticleCourantSafetyNumber            = 0.8
RadiativeCooling                       = 1
MultiSpecies                           = 2
MetalCooling                           = 3
CloudyCoolingGridFile                  = solar_2009_4D_metals.h5
UseMinimumPressureSupport              = 0
RefineByJeansLengthSafetyFactor        = 4.0

#
#  set grid refinement parameters
#

StaticHierarchy            = 0
MaximumRefinementLevel           = 9
MaximumGravityRefinementLevel        = 9
MaximumParticleRefinementLevel        = 9
RefineBy                          = 2
CellFlaggingMethod                = 2 4 20
PopIIIMustRefineResolution          = 4
PopIIIMustRefineRegionLifetime        = 3
MinimumEfficiency                 = 0.3
MinimumOverDensityForRefinement       = 3.0 3.0

MinimumMassForRefinementLevelExponent  = -0.2 -0.2
MinimumEnergyRatioForRefinement       = 0.4

#
#  set some global parameters
#

GreensFunctionMaxNumber                = 30

#
# Star formation and feedback
#

StarParticleCreation                   = 40
StarParticleFeedback                   = 40
```

```
RadiativeTransfer                       = 1
RadiativeTransferRaysPerCell            = 3.1
RadiativeTransferInitialHEALPixLevel    = 1
RadiativeTransferHydrogenOnly           = 1
RadiativeTransferOpticallyThinH2        = 1
RadiativeTransferPeriodicBoundary       = 1
RadiativeTransferAdaptiveTimestep       = 1
RadiativeTransferRadiationPressure      = 1
RadiativeTransferHubbleTimeFraction     = 0.01
RadiativeTransferPhotonMergeRadius      = 3.0
RadiativeTransferSourceClustering     = 1

RadiationFieldType = 14  // z-dependent LW background
RadiationShield = 2  // H2 self-shielding approximation

#
# PopIII/II star parameters
#

PopIIIOverDensityThreshold            = -100
PopIIIMetalCriticalFraction           = 4.1e-8
PopIIIH2CriticalFraction              = 1e-3
PopIIISupernovaRadius                 = 10
PopIIISupernovaUseColour              = 1
PopIIIHeliumIonization                = 1
PopIIIOutputOnFeedback       = 1


PopIIIStarMass                        = 20
PopIIIInitialMassFunction             = 1
PopIIIInitialMassFunctionSeed         = 1823
PopIIIMassRange                       = 1.00 300.00

StarClusterUseMetalField              = 1
StarClusterMinDynamicalTime           = 3e+06
StarClusterIonizingLuminosity         = 1.12e+46
StarClusterSNEnergy                   = 1e49
StarClusterSNRadius                   = 10
StarClusterFormEfficiency             = 0.07
StarClusterMinimumMass                = 1000

CosmologyOutputRedshift[0]            = 30.00
# ... other outputs requested go here ...
CosmologyOutputRedshift[1866]         = 10.000874
```

The interested reader is referred to https://enzo.readthedocs.io/en/latest/ for helpful descriptions of all involved parameters listed above.

B.2 STARSS Parameters and Descriptions

`STARSS` introduces a number of parameters for the user to set at runtime. The following includes all parameters and a short description of their purpose.

1. `StarMakerOverDensityThreshold`: If greater than zero, sets the criteria that the overdensity $\rho/\bar{\rho} \geq$`StarMakerOverDe` is needed to qualify for star formation. If less than zero, sets the critira that the baryon number density, $n_b \geq |\texttt{StarMakerOverDensityThreshold}|$. Finally, if set to zero, this parameter is ignored.

2. `StarMakerMassEfficiency`: The absolute upper limit on how much gas can convert to stars within one cell within one formation event. The probability of star formation,

$$p_{\rm sf} = 1 - \exp\left[-\frac{\dot{M}_{\rm SFR}\delta t}{(\eta_{\rm sf}M_{\rm cell})}\right],$$

   where $\eta_{\rm sf} =$`StarMakerMassEfficiency` and δt is the grid time step. The bulk integrated SFR is $\dot{M}_{\rm SFR} = f_s\eta_{\rm sf}M_{\rm cell}/t_{\rm ff}$, with the shielded fraction, f_s, and freefall time, $t_{\rm ff}$.

3. `StarMakerMinimumMass`: The lower limit of stellar cluster mass. If

$$\min(f_sM_{\rm cell}, \eta_{\rm sf}M_{\rm cell}) < \texttt{StarMakerMinimumMass},$$

   then the star formation will not be allowed to proceed. Since `STARSS` star particles represent star clusters, having "clusters" with very low mass makes little sense.

4. `StellarWinds`: Turns on stellar winds using the parameterized equations of Chapter 4. Mostly used for testing, there is not a good reason to turn this off in general.

5. `SingleSN`: Turns on probablistic supernova rates described in Chapter 4. Mostly used for testing, there is not a good reason to turn this off in general.

6. `StarMakerMaximumFormationMass`: Sets the maximum limit on a `STARSS` particle. This limit can be used to prevent very massive stars from forming (e.g., if the user prefers to form several lower-mass

clusters instead of a single massive cluster), which can in turn limit the number of SN per timestep. Set to -1 to allow unlimited stellar mass, with forming M_* set by $f_s \eta_{\rm sf} M_{\rm cell}$. If the predicted star mass has $M_* > 5 \times 10^3$ M$_\odot$, it is split into N-10^3 M$_\odot$ particles which have creation time evenly distributed across $3t_{\rm ff}$.

7. `NEvents`: For testing, with `NEvents` = N, one event per `STARSS` particle will be set off every timestep until $N = 0$. To enable full probabilistic SN rates, set `NEvents` = -1. Testing and debugging parameter.

8. `UnrestrictedSN`: Used to enable multiple SN per timestep per particle, but beware that the momentum formulations of Chapter 4 are only considering single events and do not properly model, e.g., superbubbles with multiple merged remnants. Combine with `StarMakerMaximumFormationMass` to ensure that the mass of clusters only leads to < 1 SN per timestep per particle. 1) Allow multiple SNe. 0) Fail if rates predict more than one SNe, i.e., if $P(SN) > 1$ at a timestep.

9. `AnalyticSNRShellMass`: Determines SNR shell mass analytically depending on phase of SNR. 1) Shell mass analytically determined. 0) The shell mass will only be that of the ejecta. It doesnt make sense to turn this off unless the simulation does not resolve any phase where shell mass $\gtrsim$ ejecta mass, e.g., all SNR will be coupled in free expansion or early ST phases.

10. `MechStarsRadiationSpectrum`: Whether to couple stars as point sources of ionizing radiation according to luminosity in Chapter 4. This minimally requires

    `RadiativeTransfer = 1`

    and `StarParticleRadiativeFeedback = 1.`

    1) `STARSS` particles as radiative sources. 0) Do not couple.

11. `MechStarsCriticalMetallicity`: If positive, will not form stars if the gas has absolute metal fraction $Z_{\rm abs} = M_{metal}/M_{gas}$ below this value. If zero or negative, this criterion is ignored.

12. `MechStarsMetallicityFloor`: If using

    `MechStarsCriticalMetallicity`

    and `MechStarsRadiationSpectrum`

to set a metallicity floor with radiative feedback, this must be set to ensure that the floor is enforced when adding radiation fields. The normal metallicity floor parameter,

`CosmologySimulationInitialFractionMetal`

is overwritten by the routine that initializes radiation fields. This parameter should be set in units of absolute metal fraction, i.e., $Z_{\rm abs} = M_{metal}/M_{gas}$.

13. `MechStarsUseAnalyticFS`: Analytically checking f_s is resolution sensitive; lower-resolution ($dx \gtrsim$ 320 comoving pc) runs will have significantly delayed or suppressed star formation if using this check. 0) Off, 1) On.

14. `MechStarsUseMeasuredFS`: Check f_s by explicitly calculating $\rm H_2$ shielded fraction as $n_{\rm H_2}/(n_{\rm H_2} + n_{\rm H_2^+})$.

15. `MechStarsFadeSNR`: If the expected merging radius, $R_{\rm merge}$, is unresolved, i.e., $R_{\rm merge} < dx$, then this parameter will reduce the coupled momenta of the SNR. At the extreme case, $R_{\rm merge} \lll dx$, the coupled momenta goes to zero. 0) Always couple full momenta. 1) Reduce momenta in unresolved case.

Appendix C

A Graduate Student's Guide to StarNet

In this appendix, I will describe the components, operation, and training using the framework I have built in Python. The goal, of course, is that a graduate student or other novice deep learning practitioner could use this resource to get up and running with reproducing the work of Chapters 2 and 4. There are three major components to the deep learning task of generating the feedback influence of Population III stars. First, we must identify regions that host star formation. This is accomplished via region-wide classification using the small classifier architecture of Chapter 2. Given that a region is predicted to host star formation, we aim to predict a precise location that can participate in the star formation, i.e., using the segmentation model, IUNet, of Chapter 2. With a positive region and star forming locus identified, we need only to generate a reasonable approximation of the feedback influence. This can be accomplished in two ways: the first-generation spherical remnant approximation presented in Chapter 4, or the deep learning using GANs (the feedback approximation framework is referred to as `FBNet` below) methods discussed in this appendix, Section C.4. The remainder of this appendix will offer a bit of technical instruction for all aspects of training the models in this thesis. Section C.1 details setting up the software environment that was used in this work; Section C.2.1 details the procedure for training the `StarFind` module, including generating training data; Section C.3 details the StarNetRuntime, where we interface with a running simulation using `Enzo`'s inline Python analysis capability; Finally, Section C.4 details generating training data and training procedures for `FBNet`.

C.1 Software Environment

Training and model development was performed on the SDSC-Comet and SDSC-Expanse supercomputers using NVIDIA K80 (Comet) and Nvidia V100 (Expanse) GPUs. A consistent development environment was obtained using `conda`[1]. It is a well-known issue that changes in deep-learning framework can cause inconsistent results in the final trained models, so below is the `conda` environment definition. This can be put into a file, `environment.yml`, and the desired environment can be built using `conda create --file environment.yml` (be sure to change the final line to point to your `conda` installation). The environment can then be activated by entering `conda activate DeepAI`, and all `pytorch, yt`, etc., libraries should be available.

```
environment.yml:
        name: DeepAI
        channels:
          - pytorch
          - conda-forge
          - defaults
        dependencies:
          - _libgcc_mutex=0.1=conda_forge
          - _openmp_mutex=4.5=1_llvm
          - backcall=0.2.0=pyhd3eb1b0_0
          - backports=1.0=pyhd3eb1b0_2
          - backports.functools_lru_cache=1.6.4=pyhd8ed1ab_0
          - blas=1.0=mkl
          - bzip2=1.0.8=h7b6447c_0
          - ca-certificates=2021.10.26=h06a4308_2
          - certifi=2021.10.8=py38h06a4308_0
          - cudatoolkit=10.2.89=hfd86e86_1
          - cycler=0.11.0=pyhd3eb1b0_0
          - dbus=1.13.18=hb2f20db_0
          - decorator=5.1.0=pyhd3eb1b0_0
          - expat=2.4.1=h2531618_2
          - ffmpeg=4.3=hf484d3e_0
          - fontconfig=2.13.1=hba837de_1005
          - freetype=2.11.0=h70c0345_0
          - giflib=5.2.1=h7b6447c_0
          - glib=2.69.1=h5202010_0
          - gmp=6.2.1=h2531618_2
          - gmpy2=2.1.0b5=py38h8384b0a_0
          - gnutls=3.6.15=he1e5248_0
          - gst-plugins-base=1.14.0=h8213a91_2
          - gstreamer=1.14.0=h28cd5cc_2
          - icu=58.2=hf484d3e_1000
          - intel-openmp=2021.4.0=h06a4308_3561
          - ipython=7.29.0=py38hb070fc8_0
          - jedi=0.18.0=py38h578d9bd_2
          - jpeg=9d=h7f8727e_0
```

---

[1] https://docs.anaconda.com/anaconda/install/linux/

```
- kiwisolver=1.3.1=py38h1fd1430_1
- lame=3.100=h7b6447c_0
- lcms2=2.12=h3be6417_0
- ld_impl_linux-64=2.35.1=h7274673_9
- libffi=3.3=he6710b0_2
- libgcc-ng=11.1.0=hc902ee8_8
- libiconv=1.16=h516909a_0
- libidn2=2.3.2=h7f8727e_0
- libpng=1.6.37=hbc83047_0
- libstdcxx-ng=11.1.0=h56837e0_8
- libtasn1=4.16.0=h27cfd23_0
- libtiff=4.2.0=h85742a9_0
- libunistring=0.9.10=h27cfd23_0
- libuuid=2.32.1=h7f98852_1000
- libuv=1.40.0=h7b6447c_0
- libwebp=1.2.0=h89dd481_0
- libwebp-base=1.2.0=h27cfd23_0
- libxcb=1.14=h7b6447c_0
- libxml2=2.9.12=h03d6c58_0
- llvm-openmp=12.0.1=h4bd325d_1
- lz4-c=1.9.3=h295c915_1
- matplotlib=3.4.3=py38h578d9bd_0
- matplotlib-base=3.4.3=py38hf4fb855_0
- matplotlib-inline=0.1.2=pyhd3eb1b0_2
- mkl=2021.4.0=h06a4308_640
- mkl-service=2.4.0=py38h7f8727e_0
- mkl_fft=1.3.1=py38hd3c417c_0
- mkl_random=1.2.2=py38h51133e4_0
- more-itertools=8.12.0=pyhd3eb1b0_0
- mpc=1.1.0=h04dde30_1009
- mpfr=4.0.2=hb69a4c5_1
- mpmath=1.2.1=py38h06a4308_0
- ncurses=6.3=h7f8727e_2
- nettle=3.7.3=hbbd107a_1
- numpy=1.21.2=py38h20f2e39_0
- numpy-base=1.21.2=py38h79a1101_0
- olefile=0.46=pyhd3eb1b0_0
- openh264=2.1.1=h4ff587b_0
- openssl=1.1.1l=h7f8727e_0
- packaging=21.3=pyhd3eb1b0_0
- parso=0.8.2=pyhd3eb1b0_0
- pcre=8.45=h295c915_0
- pexpect=4.8.0=pyhd3eb1b0_3
- pickleshare=0.7.5=pyhd3eb1b0_1003
- pillow=8.4.0=py38h5aabda8_0
- pip=21.2.4=py38h06a4308_0
- prompt-toolkit=3.0.20=pyhd3eb1b0_0
- ptyprocess=0.7.0=pyhd3eb1b0_2
```

```
- pygments=2.10.0=pyhd3eb1b0_0
- pyparsing=3.0.4=pyhd3eb1b0_0
- pyqt=5.9.2=py38h05f1152_4
- python=3.8.8=hdb3f193_4
- python-dateutil=2.8.2=pyhd3eb1b0_0
- python_abi=3.8=2_cp38
- pytorch=1.10.1=py3.8_cuda10.2_cudnn7.6.5_0
- pytorch-mutex=1.0=cuda
- qt=5.9.7=h5867ecd_1
- readline=8.1=h27cfd23_0
- setuptools=58.0.4=py38h06a4308_0
- sip=4.19.13=py38he6710b0_0
- six=1.16.0=pyhd3eb1b0_0
- sqlite=3.37.0=hc218d9a_0
- sympy=1.9=py38h06a4308_0
- tk=8.6.11=h1ccaba5_0
- toml=0.10.2=pyhd3eb1b0_0
- torchaudio=0.10.1=py38_cu102
- torchvision=0.11.2=py38_cu102
- tornado=6.1=py38h497a2fe_1
- traitlets=5.1.1=pyhd3eb1b0_0
- typing_extensions=3.10.0.2=pyh06a4308_0
- unyt=2.8.0=py_0
- wcwidth=0.2.5=pyh9f0ad1d_2
- wheel=0.37.0=pyhd3eb1b0_1
- xz=5.2.5=h7b6447c_0
- yaml=0.2.5=h7b6447c_0
- yt=4.0.1=py38h1abd341_0
- zlib=1.2.11=h7f8727e_4
- zstd=1.4.9=haebb681_0
- pip:
  - absl-py==0.10.0
  - aiohttp==3.7.4.post0
  - argparse==1.4.0
  - astroid==2.5
  - astropy==4.2.1
  - async-timeout==3.0.1
  - attrs==20.3.0
  - cachetools==4.1.1
  - chardet==3.0.4
  - cython==0.29.24
  - fsspec==0.8.4
  - future==0.18.2
  - google-auth==1.22.1
  - google-auth-oauthlib==0.4.1
  - grpcio==1.32.0
  - h5py==3.3.0
  - idna==2.10
```

```
    - imageio==2.9.0
    - isort==5.7.0
    - lazy-object-proxy==1.5.2
    - markdown==3.3.1
    - mccabe==0.6.1
    - mpi4py==3.0.3
    - multidict==5.1.0
    - oauthlib==3.1.0
    - pandas==1.1.3
    - progressbar2==3.53.1
    - protobuf==3.13.0
    - pyasn1==0.4.8
    - pyasn1-modules==0.2.8
    - pyerfa==2.0.0
    - pylint==2.7.1
    - python-utils==2.4.0
    - pytorch-lightning==1.2.6
    - pytorch-msssim==0.2.1
    - pytz==2020.1
    - pyyaml==5.3.1
    - requests==2.24.0
    - requests-oauthlib==1.3.0
    - rsa==4.6
    - tensorboard==2.3.0
    - tensorboard-plugin-wit==1.7.0
    - test-tube==0.7.5
    - torchmetrics==0.2.0
    - tqdm==4.50.2
    - trident==1.2.1
    - urllib3==1.25.10
    - werkzeug==1.0.1
    - wrapt==1.12.1
    - yarl==1.6.3
    - yt-astro-analysis==1.0.0
prefix: path/to/conda/envs/DeepAI
```

All files that were developed with the aim of creating Population III feedback using DNN are available on my github; `https://github.com/azton/StarNet`. The file structure within the repository is divided as follows: `data_generation` contains scripts to generate training data and scaling information (discussed below); network definitions are found in `network_modules` including all small classifier, Inception U-net, and GAN architectures; many helper functions, including custom data loaders, are defined in `network_utilities`; training and validation scripts can be found in their respective sub-directories of `run_networks`: `stage1` and `stage2` for `StarFind` and `stage3` for `FBNet`.

C.2 `StarFind`

C.2.1 Data Generation and Scaling

The first step in training a deep learning model is to generate training data, and, if necessary, take steps to understand the training data fully. Here, I will discuss the first. `StarFind_dataset_generator.py`, was run on the TACC-Frontera supercomputer and can accept a variety of parameters to control the resolution and size of resulting samples. There should be no additional requirements beyond those listed in Section C.1. An example submission script that lists relevant command line arguments and submission procedure is as follows:

```
#!/bin/bash
#SBATCH --job-name="Form256-1"
#SBATCH --partition=normal
#SBATCH --nodes=5
#SBATCH --ntasks-per-node=28
#SBATCH -t 2:00:00
#SBATCH --export=ALL
#SBATCH -A <account_number>

module load hdf5 intel/18.0.5 python3/3.7.0
# dataset output to begin generation at
d0=200
# final output to consider
d1=1269
# simulation name
sim=phoenix_256_IC1
# path to simulation directory
stg="/path/to/phoenix simulation outputs/phoenix_256_IC1"
# dimension of output data, eg, the volume dimensions
# that will be input to starfind module
dim=128
# desired resolution per grid cell, in comoving pc
res=160

ibrun python3 -u StarFind_dataset_generator.py \
    $d0 $d1 $sim $stg $dim $res stars \
    > ${sim}--${d0}-${d1}-${dim}_formationset.out 2>&1
```

The script can also be run to generate training data for other simulations with appropriate modification to the simulation path and name. The training, validation, and test files will be saved at `<scratch directory>/network_data_inputs/<split name>/<simulation name>/` with each particle identifier having its own directory and set of input and final states.

After generating training data, it must be scaled accordingly (the final step in Chapter 2). The script `gen_scaling.py` performs this task, and saves the relevant file to `StarNetworks/<output name>` as a `.torch` file that can be loaded using the `torch.load()` method, where `<output name>` is a command-line argument.

C.2.2 Training Stage 1

With training data and scaling file in hand, training can take place on the GPU acclerated system of your choice. `StarFind` was originally trained on the SDSC-Comet supercomputer, however later training and refinement were performed on SDSC-Expanse. The training logic is performed using `pytorch-lightning`[2] to abstract the minutia of parallelization and boilerplate code logic. The training takes both command line arguments and a configuration file–while command line parameters are more fluid and will be iterated upon in a hyperparameter exploration, the configuration file contains set or finalized parameters. A sample batch file to run a small hyperparameter exploration is given here:

```
#!/bin/bash
#SBATCH --job-name="S1Htune"
#SBATCH --output=lightning_4GPU.out
#SBATCH --partition=gpu
#SBATCH --nodes=1
## one task per gpu, with threads working for dataloading
#SBATCH --ntasks-per-node=40
#SBATCH --gpus=4
#SBATCH -A <account>
#SBATCH -t 48:00:00
#SBATCH --export=ALL


gpus=4
module load gpu

config=Lightning.conf
# learning rate is divided by 1e6 in run_lightning_model.py
LR=1
# modified weight given to positive samples in loss
w=2
# number of channels at first layer of input
n0=2
# which sub-model to use in S1
model=smalldense
```

---

[2] https://pytorch-lightning.readthedocs.io/en/latest/

```
# where to write checkpoints
writer_out=${model}_hypertune
# there are other ways, but heres a simple way to
# iterate a small set of hyperparameters.
for lr in {0..4}; do
    for w in {2..5}; do
        for N0 in {1..6..2}; do
            LR=$((1*10**lr))
            n0=$((2*2**N0))
            python3 -u run_lightning_model.py
                            --conf $config
                            --model $model
                            --gpus $gpus
                            --nnodes $SLURM_JOB_NUM_NODES
                            --nprocs $gpus
                            -fl $n0
                            -lr $LR
                            -w $w > lightning_classifier.out 2>&1
        done
    done
done
```

Within the `stage1` directory, a new directory will be created which hosts checkpoints. Checkpoints are the individual files containing network definitions and weights that can be reloaded to either continue training or to perform inference later. More efficiency can be attained by exporting models to, e.g., ONNX[3] or `torchscript`[4], however such optimization is beyond the scope of this guide. After the training job starts execution, the progress of training can be observed using either `tail -f <log file>` or the `tensorboard`[5] command. A further list of the parameters used at runtime can be found within `StarNetworks/run_networks/stage1/Lightning.conf`; note that the command line arguments take precedence over the configuration file parameters.

C.2.3 Training Stage 2

Training stage 2 is very similar to stage 1. This module was originally written without the `pytorch-lightning` layer, so `run_segmentation.py` is the version used in Chapter 2. The `lightning` version, `run_seg_lightning.py` has been written, but not fully tested. Since `run_segmentation.py` runs without using command-line arguments, the batch file is simpler:

---

[3] https://onnx.ai/

[4] https://pytorch.org/docs/stable/jit.html

[5] https://pytorch.org/docs/stable/tensorboard.html

```
#!/bin/bash
#SBATCH --job-name="S1Htune"
#SBATCH --output=lightning_4GPU.out
#SBATCH --partition=gpu
#SBATCH --nodes=1
## one task per gpu, with threads working for dataloading
#SBATCH --ntasks-per-node=40
#SBATCH --gpus=4
#SBATCH -A <account>
#SBATCH -t 48:00:00
#SBATCH --export=ALL


srun python run_segmentation.py configuration_file.conf > outputlog.out
```

The limitation of this `lightning`-free approach is that it uses a less efficient parallelization method (*data distributed parallel*). If aiming to improve on the training of stage 2, I recommend using the `lightning` implementation, whose command-line parameters are available in `run_seg_lightning.py`; it is readily adaptable to any number of GPUs using the fully *data parallel* training approach[6].

C.3 `StarNet`

`StarNet` is the framework created for the work in Chapter 4. It combines `StarFind` with the linear regression models of Chapter 3 to predict both the location and effect of Population III star formation. It operates on a running simulation to generate the metallicity initial conditions *in situ*. The framework can be found at https://github.com/azton/StarNetRuntime, however the required model checkpoints exceed the file size limitation on github. The checkpoints will be available on the *Renaissance Simulations Laboratory*[7] or via direct contact[8], and should be placed in `StarNetRuntime/model_checkpoints`. In order to use `StarNet`, Enzo must be compiled with a special make-flag, enabled by `make python-yes`, which also currently requires Python 2.xx (as of Enzo v2.6, 6/2/2022). An example make file for Enzo that will build correctly on SDSC-Expanse is given here:

```
#=======================================================================
#
# FILE:        Make.mach.expanse
#
```

---

[6] https://pytorch.org/tutorials/beginner/blitz/data_parallel_tutorial.html

[7] https://rensimlab.github.io/

[8] aiwells@ucsd.edu or mlnorman@ucsd.edu

```
# DESCRIPTION: Makefile settings for the Expanse Resource at SDSC/UCSD with Python2
#
# AUTHOR:      Azton Wells
#
# DATE:        2021
#
#
#
#=======================================================================

MACH_TEXT  = Expanse
MACH_VALID = 1
MACH_FILE  = Make.mach.sdsc-expanse

MACHINE_NOTES = "MACHINE_NOTES for Expanse at SDSC/UCSD: \
Load these modules, \
’module load cpu intel intel-mpi hdf5’\
"

#-----------------------------------------------------------------------
# Compiler settings
#-----------------------------------------------------------------------

LOCAL_PYTHON_INSTALL = /path/to/anaconda3/envs/python2/bin
LOCAL_COMPILER = ${INTELHOME}/compilers_and_libraries/linux
LOCAL_MPI_INSTALL = ${MPIHOME}
LOCAL_HDF5_INSTALL = ${HDF5HOME}
LOCAL_HYPRE_INSTALL =

# With MPI

MACH_CPP       = cpp
MACH_CC_MPI    = mpicc # C compiler when using MPI
MACH_CXX_MPI   = mpicxx # C++ compiler when using MPI
MACH_FC_MPI    = ifort # Fortran 77 compiler when using MPI
MACH_F90_MPI   = ifort # Fortran 90 compiler when using MPI
MACH_LD_MPI    = mpicxx # Linker when using MPI

# Without MPI

MACH_CC_NOMPI  = icc # C compiler when not using MPI
MACH_CXX_NOMPI = icpc # C++ compiler when not using MPI
MACH_FC_NOMPI  = ifort # Fortran 77 compiler when not using MPI
MACH_F90_NOMPI = ifort # Fortran 90 compiler when not using MPI
MACH_LD_NOMPI  = icpc # Linker when not using MPI

#-----------------------------------------------------------------------
# Machine-dependent defines
```

```
#-----------------------------------------------------------------------
# Defines for the architecture; e.g. -DSUN, -DLINUX, etc.
MACH_DEFINES   = -DLINUX -DH5_USE_16_API

#-----------------------------------------------------------------------
# Compiler flag settings
#-----------------------------------------------------------------------


MACH_CPPFLAGS = -P -traditional
MACH_CFLAGS   =
MACH_CXXFLAGS = -DMPICH_SKIP_MPICXX
MACH_FFLAGS   = -132
MACH_F90FLAGS =
MACH_LDFLAGS  =
#MACH_OPENMP   = -mp
#MACH_OPENMP   = -openmp

#-----------------------------------------------------------------------
# Precision-related flags
#-----------------------------------------------------------------------

MACH_FFLAGS_INTEGER_32 = -i4
MACH_FFLAGS_INTEGER_64 = -i8
MACH_FFLAGS_REAL_32    = -r4
MACH_FFLAGS_REAL_64    = -r8

#-----------------------------------------------------------------------
# Optimization flags
#-----------------------------------------------------------------------

MACH_OPT_WARN        = -Wall -g # Flags for verbose compiler warnings
MACH_OPT_DEBUG       = -O0 -g # Flags for debugging
# Flags for high conservative optimization
MACH_OPT_HIGH        = -O2
# Note that this breaks determinism, which is why it’s commented out!
#
MACH_OPT_AGGRESSIVE  = -O3 -march=native # Flags for aggressive optimization

#-----------------------------------------------------------------------
# Includes
#-----------------------------------------------------------------------

LOCAL_INCLUDES_MPI    = -I$(LOCAL_MPI_INSTALL)/include
LOCAL_INCLUDES_HDF5   = -I$(LOCAL_HDF5_INSTALL)/include # HDF5 includes
LOCAL_INCLUDES_HYPRE  =
LOCAL_INCLUDES_PAPI   = # PAPI includes
LOCAL_INCLUDES_PYTHON = -I/path/to/include/python2.7 \
```

```
-I/path/to/.local/lib/python2.7/site-packages/numpy/core/include/

MACH_INCLUDES         = $(LOCAL_INCLUDES_HDF5)
MACH_INCLUDES_PYTHON  = $(LOCAL_INCLUDES_PYTHON)
MACH_INCLUDES_MPI     = $(LOCAL_INCLUDES_MPI)
MACH_INCLUDES_HYPRE   = $(LOCAL_INCLUDES_HYPRE)

#-----------------------------------------------------------------------
# Libraries
#-----------------------------------------------------------------------

LOCAL_LIBS_MPI    =
LOCAL_LIBS_HDF5   = -L$(LOCAL_HDF5_INSTALL)/lib -lhdf5 -lz # HDF5 libraries
LOCAL_LIBS_HYPRE  =
LOCAL_LIBS_PYTHON  = /path/to/libpython2.7.so
LOCAL_LIBS_MACH  = -L$(LOCAL_COMPILER)/lib/intel64 -lcilkrts -lifcore -lifport -limf -lsvml -lint


MACH_LIBS         = $(LOCAL_LIBS_HDF5) $(LOCAL_LIBS_MACH)
MACH_LIBS_MPI     = $(LOCAL_LIBS_MPI)
MACH_LIBS_HYPRE   = $(LOCAL_LIBS_HYPRE)
MACH_LIBS_PYTHON  = $(LOCAL_LIBS_PYTHON)
```

However, since Enzo will require Python 2.xx, so establish a new conda environment using Python 2.7 and use the following `environment.yaml` to get the dependencies in order:

```
name: python2
channels:
  - pytorch
  - conda-forge
  - defaults
dependencies:
  - _libgcc_mutex=0.1=main
  - _openmp_mutex=4.5=1_gnu
  - backports=1.1=pyhd3eb1b0_0
  - backports.functools_lru_cache=1.6.4=pyhd3eb1b0_0
  - backports.shutil_get_terminal_size=1.0.0=pyhd3eb1b0_3
  - backports_abc=0.5=py_1
  - basemap=1.2.1=py27hd1be537_2
  - blas=1.0=mkl
  - ca-certificates=2021.10.8=ha878542_0
  - certifi=2019.11.28=py27h8c360ce_1
  - cffi=1.13.2=py27h2e261b9_0
  - configparser=4.0.2=py27_0
  - cudatoolkit=10.1.243=h6bb024c_0
  - cycler=0.10.0=py27_0
```

```
- dbus=1.13.18=hb2f20db_0
- decorator=5.1.0=pyhd3eb1b0_0
- enum34=1.1.6=py27_1
- expat=2.4.1=h2531618_2
- fastcache=1.1.0=py27h7b6447c_0
- fontconfig=2.13.1=h6c09931_0
- freetype=2.11.0=h70c0345_0
- functools32=3.2.3.2=py27_1
- future=0.18.2=py27_0
- futures=3.3.0=py27_0
- geos=3.8.0=he1b5a44_1
- glib=2.63.1=h5a9c865_0
- gmp=6.2.1=h2531618_2
- gmpy2=2.0.8=py27h10f8cd9_2
- gst-plugins-base=1.14.0=hbbd80ab_1
- gstreamer=1.14.0=hb453b48_1
- h5py=2.9.0=py27h7918eee_0
- hdf5=1.10.4=hb1b8bf9_0
- icu=58.2=he6710b0_3
- intel-openmp=2022.0.1=h06a4308_3633
- ipython=5.8.0=py27_0
- ipython_genutils=0.2.0=pyhd3eb1b0_1
- jpeg=9d=h7f8727e_0
- kiwisolver=1.1.0=py27he6710b0_0
- libedit=3.1.20210910=h7f8727e_0
- libffi=3.2.1=hf484d3e_1007
- libgcc-ng=9.3.0=h5101ec6_17
- libgfortran-ng=7.5.0=ha8ba4b0_17
- libgfortran4=7.5.0=ha8ba4b0_17
- libgomp=9.3.0=h5101ec6_17
- libpng=1.6.37=hbc83047_0
- libstdcxx-ng=9.3.0=hd4cf53a_17
- libtiff=4.2.0=h85742a9_0
- libuuid=1.0.3=h7f8727e_2
- libwebp-base=1.2.0=h27cfd23_0
- libxcb=1.14=h7b6447c_0
- libxml2=2.9.12=h03d6c58_0
- linecache2=1.0.0=py_1
- lz4-c=1.9.3=h295c915_1
- matplotlib=2.2.3=py27hb69df0a_0
- matplotlib-base=2.2.4=py27hfd891ef_0
- mkl=2020.2=256
- mkl-service=2.3.0=py27he904b0f_0
- mkl_fft=1.0.15=py27ha843d7b_0
- mkl_random=1.1.0=py27hd6b4f25_0
- mpc=1.1.0=h10f8cd9_1
- mpfr=4.0.2=hb69a4c5_1
- mpi=1.0=mpich
```

```
- mpi4py=3.0.3=py27h028fd6f_0
- mpich=3.3.2=hc856adb_0
- mpmath=1.1.0=py27_0
- ncurses=6.3=h7f8727e_2
- ninja=1.10.2=h5e70eb0_2
- numpy=1.16.6=py27hbc911f0_0
- numpy-base=1.16.6=py27hde5b4d6_0
- olefile=0.46=py27_0
- openssl=1.1.1k=h7f98852_0
- pathlib2=2.3.5=py27_0
- pcre=8.45=h295c915_0
- pexpect=4.8.0=pyhd3eb1b0_3
- pickleshare=0.7.5=py27_0
- pillow=6.2.1=py27h34e0f95_0
- pip=19.3.1=py27_0
- proj4=5.2.0=he1b5a44_1006
- prompt_toolkit=1.0.15=py27_0
- ptyprocess=0.7.0=pyhd3eb1b0_2
- pycparser=2.20=py_2
- pygments=2.5.2=py_0
- pyparsing=2.4.7=pyhd3eb1b0_0
- pyproj=1.9.6=py27h516909a_1002
- pyqt=5.9.2=py27h05f1152_2
- pyshp=2.1.3=pyh44b312d_0
- python=2.7.15=h9bab390_6
- python-dateutil=2.8.2=pyhd3eb1b0_0
- python_abi=2.7=1_cp27mu
- pytorch=1.4.0=py2.7_cuda10.1.243_cudnn7.6.3_0
- pytz=2021.3=pyhd3eb1b0_0
- qt=5.9.7=h5867ecd_1
- readline=7.0=h7b6447c_5
- scandir=1.10.0=pyh5d7bf9c_3
- setuptools=44.0.0=py27_0
- simplegeneric=0.8.1=py27_2
- singledispatch=3.7.0=pyhd3eb1b0_1001
- sip=4.19.8=py27hf484d3e_0
- six=1.16.0=pyhd3eb1b0_0
- sqlite=3.33.0=h62c20be_0
- subprocess32=3.5.4=py27h7b6447c_0
- sympy=1.5.1=py27_0
- tk=8.6.11=h1ccaba5_0
- torchvision=0.5.0=py27_cu101
- tornado=5.1.1=py27h7b6447c_0
- traceback2=1.4.0=py27_0
- traitlets=4.3.3=py27_0
- typing=3.7.4.1=py27_0
- unittest2=1.1.0=py27_0
- wcwidth=0.2.5=pyhd3eb1b0_0
```

```
  - wheel=0.37.1=pyhd3eb1b0_0
  - xz=5.2.5=h7b6447c_0
  - yt=3.4.1=py27h14c3975_1
  - zlib=1.2.11=h7f8727e_4
  - zstd=1.4.9=haebb681_0
prefix: /path/to/anaconda3/envs/python2
```

With the correct environment, running a simulation requires a `user_script.py` to be present in the simulation directory. All of the StarNetRuntime logic is contained in `StarNetRuntime/ex_user_script.py`, so the only portion required in the simulation directory is:

```
user_script.py:
    import sys
    sys.path.append('./')
    from StarNetRuntime.ex_user_script import run_main

    def main():
        run_main()
```

As well, an example configuration file is available in `StarNetRuntime/ex_Pop3Net.conf`. This should be copied directly to the simulation directory, and modified to contain the correct file paths for the current file structure. The simulation such as those presented in Chapter 4 can be run via batch file:

```
#!/bin/bash
#SBATCH -N 1
#SBATCH --ntasks-per-node=64
#SBATCH -p compute
#SBATCH --mem=249320M
#SBATCH -t 48:00:00
#SBATCH --job-name="5L-inline"
#SBATCH --export=ALL
#SBATCH --mail-type=ALL
#SBATCH -A <account>
#SBATCH --mail-user=user@email.com




module load cpu intel intel-mpi hdf5

conda activate python2
export PYTHONPATH=${HOME}anaconda3/envs/python2/lib/python2.7
export ENZO=${HOME}/enzo-python/src/enzo/enzo.exe
```

```
    # uncomment this to allow core files
    # ulimit -c unlimited

    ibrun $ENZO -d run.enzo > estd.out 2>&1
```

While running the simulation, we need several parameters to ensure that the inline python routine is called and used. A sample parameter file to run Enzo using inline Python with StarNetRuntime is:

```
StopCycle           = 100000
StopCPUTime         = 420000
PythonTopGridSkip = 15 // run user script every Nth top-grid timestep
                        // regardless, StarNetRuntime will only execute
                        // if there has been enough time since the last run
                        // given by minimum_time_python_skip in Pop3Net.conf
PythonSubcycleSkip = 0 // dont run on lower levels at all
ParallelRootGridIO = 1
ParallelParticleIO = 1
UnigridTranspose = 2
HierarchyFileOutputFormat = 1
CorrectParentBoundaryFlux = 1


ProblemType                = 30
TopGridRank                = 3
TopGridDimensions          = 256 256 256
PotentialIterations        = 10
SelfGravity                = 1
TopGridGravityBoundary     = 0
LeftFaceBoundaryCondition  = 3 3 3
RightFaceBoundaryCondition = 3 3 3



CosmologySimulationOmegaBaryonNow       = 0.048975
CosmologySimulationOmegaCDMNow          = 0.262125
CosmologySimulationNumberOfInitialGrids = 1
CosmologySimulationDensityName           = GridDensity
CosmologySimulationVelocity1Name         = GridVelocities_x
CosmologySimulationVelocity2Name         = GridVelocities_y
CosmologySimulationVelocity3Name         = GridVelocities_z
CosmologySimulationCalculatePositions    = 1
CosmologySimulationParticleVelocity1Name = ParticleVelocities_x
CosmologySimulationParticleVelocity2Name = ParticleVelocities_y
CosmologySimulationParticleVelocity3Name = ParticleVelocities_z
CosmologySimulationParticleDisplacement1Name = ParticleDisplacements_x
CosmologySimulationParticleDisplacement2Name = ParticleDisplacements_y
```

```
CosmologySimulationParticleDisplacement3Name = ParticleDisplacements_z

#
#  define cosmology parameters
#

ComovingCoordinates     = 1
CosmologyOmegaMatterNow     = 0.3111
CosmologyOmegaDarkMatterNow = 0.262125
CosmologyOmegaLambdaNow     = 0.6889
CosmologyHubbleConstantNow  = 0.6766
CosmologyComovingBoxSize    = 1.765
CosmologyMaxExpansionRate   = 0.015
CosmologyInitialRedshift  = 99.000000
CosmologyFinalRedshift    = 13.000000
GravitationalConstant       = 1

#
#  set I/O and stop/start parameters
#

DataDumpName            = RD
dtDataDump              = 0.15
RedshiftDumpName        = RD

#
#  set hydro parameters
#

Gamma                       = 1.6667
PPMDiffusionParameter       = 0
DualEnergyFormalism         = 1
InterpolationMethod         = 1
FluxCorrection              = 2
ConservativeInterpolation   = 0
CourantSafetyNumber         = 0.3
ParticleCourantSafetyNumber = 0.8
RadiativeCooling            = 1
MultiSpecies                = 2
MetalCooling                = 3
CloudyCoolingGridFile       = solar_2009_4D_metals.h5
UseMinimumPressureSupport   = 0
RefineByJeansLengthSafetyFactor  = 4.0

#
#  set grid refinement parameters
#
```

```
StaticHierarchy                    = 0   \\ Use AMR
MaximumRefinementLevel             = 5
MaximumGravityRefinementLevel      = 5
MaximumParticleRefinementLevel     = 5
RefineBy                           = 2
CellFlaggingMethod                 = 2 4 \\ DM and baryon mass
MinimumEfficiency                  = 0.3
MinimumOverDensityForRefinement    = 3.0 3.0
MinimumMassForRefinementLevelExponent = -0.3 -0.3
MinimumEnergyRatioForRefinement    = 0.4

GreensFunctionMaxNumber            = 30

#
# Stars and radiation parameters
#

StarParticleCreation                   = 32768
StarParticleFeedback                   = 32768

RadiativeTransfer                      = 1
RadiativeTransferRaysPerCell           = 3.1
RadiativeTransferInitialHEALPixLevel   = 1
RadiativeTransferHydrogenOnly          = 1
RadiativeTransferOpticallyThinH2       = 1
RadiativeTransferPeriodicBoundary      = 1
RadiativeTransferAdaptiveTimestep      = 1
RadiativeTransferRadiationPressure     = 1
RadiativeTransferHubbleTimeFraction    = 0.01
RadiativeTransferPhotonMergeRadius     = 3.0
RadiativeTransferSourceClustering      = 1
StarParticleRadiativeFeedback          = 1

RadiationFieldType                     = 14
RadiationShield                        = 2


StarMakerOverDensityThreshold          = 75  // n/cc if using 32768
StarMakerMassEfficiency                = 0.25
StarMakerMinimumMass                   = 150
StellarWinds                           = 1
SingleSN                               = 1
StarMakerMaximumFormationMass          = 1000
NEvents                                = -1
UnrestrictedSN                         = 1
AnalyticSNRShellMass                   = 1
MechStarsRadiationSpectrum             = 1
MechStarsCriticalMetallicity           = 4.1e-8
```

```
MechStarsUseAnalyticFS              = 0
MechStarsFadeSNR                    = 1
MechStarsMetallicityFloor           = 1e-20
```

C.4 **FBNet**

`FBNet` is the least developed module, as it has not been used in any publication as of yet. It is a prospective method to improve upon the linear regression method of Chapter 4, however requires more work to produce usable models; hence, we will only briefly cover it here. Since the data for feedback is unique from star formation, we have a new script to generate training data, `ClusterSimulator_dataset_generator.py`. Each feedback region (i.e., region that has a unique initial star formation event) will have two files: the input state (`<particle_id>_00.h5` files), and final state (`<particle_id>_01.h5` files). Each state has the full set of `Enzo` fields computed in the *Phoenix Simulations* including ionized species, so that `FBNet` can be trained to predict any combination of fields. Scaling information is still required, so there is a similar script to generate standard scaling information, `gen_clusternet_scaling.py`. Data generation can be run on Frontera using the batch submission file:

```
#!/bin/bash
#SBATCH --job-name="CDS_512"
#SBATCH --partition=small
#SBATCH --nodes=2
#SBATCH --ntasks-per-node=50
#SBATCH -t 48:00:00
#SBATCH --export=ALL

module load hdf5 intel/18.0.5 python3/3.7.0

sim=phoenix_512
# grid dimension of training input
dim=96
# width in kpc of model samples
width=30.72
# level of grid(ish) at that resolution
level=5
outdir=/path/to/phoenix/simulation/directory
finaldir=/path/to/training/data/storage/network_data_inputs
# duration between input and output states
modeltime=16.60
```

```
# 3 dimension or 2 dimensional volumes
final_dim=3d
# list of first and last outputs to evaluate
outbounds="200 920"


ibrun python3 -u ClusterSimulator_dataset_generator.py \
    --output_dir $outdir \
    --final_dir $finaldir \
    --level $level \
    --width $width \
    -dim $dim \
    --output_bounds $outbounds \
    --sim_name $sim \
    --model_time $modeltime > ClusterDS_${sim}_${dim}x${width}.out 2>&1
```

Very similarly to `StarFind`, training can be performed on SDSC-Expanse using the following batch submission file:

```
#!/bin/bash
#SBATCH --job-name="cGAN-DSNC"
#SBATCH --partition=gpu-shared
#SBATCH --nodes=1
#SBATCH --ntasks=10
#SBATCH --gpus=1
#SBATCH --mem=89G
#SBATCH -A <account>
#SBATCH --mail-type=BEGIN,END
#SBATCH --mail-user=user@email.com
#SBATCH -t 48:00:00

module load gpu slurm

gpus=1
nnodes=1
nprocs=1

reload=/path/to/checkpoint/if/needed
conf=run_gan.conf
data_dim=96
match_weight=100

discriminator=pix2pix_ada
disc_out=4
disc_first_layer=16
disc_learn_rate=1e-4
```

```
ada_pfloor=0.1

generator=pix2pix_unet
batch_size=8
gen_first_layer=128
gen_learn_rate=1e-3
unique_name=buffered_30kpc_96cube
transforms=translate
field_list="Density SN_Colour x-velocity y-velocity z-velocity"

srun -n $nprocs --gpus-per-task=1 \
    python run_lightning_gan.py \
    --unique_name $unique_name \
    --conf $conf \
    --gpus $gpus \
    --nnodes $nnodes \
    --nprocs $nprocs \
    --data_dimension $data_dim \
    --gen_first_layer $gen_first_layer \
    --disc_first_layer $disc_first_layer \
    --disc_learn_rate $disc_learn_rate \
    --gen_learn_rate $gen_learn_rate \
    --match_weight $match_weight \
    --discriminator $discriminator \
    --disc_out $disc_out \
    --generator $generator \
    --batch_size $batch_size \
    --ada_pfloor $ada_pfloor \
    --field_list $field_list > GANTraining.out 2>&1
```

However, there are many more options, given the under-developed state, variety of defined architectures, and plethora of hyperparameters (intrinsic to GANs designs).

Bibliography

Abadi MG, JF Navarro, M Steinmetz, & VR Eke 2003 'Simulations of galaxy formation in a lambda cold dark matter universe. II. the fine structure of simulated galactic disks' *The Astrophysical Journal* 597(1):21–34 doi: 10.1086/378316 URL https://doi.org/10.1086/378316.

Abel T, GL Bryan, & ML Norman 2000 'The Formation and Fragmentation of Primordial Molecular Clouds' *The Astrophysical Journal* 540(1):39–44 doi: 10.1086/309295.

Abohalima A & A Frebel 2018 'JINAbase—a database for chemical abundances of metal-poor stars' *The Astrophysical Journal* 238(2):36 doi: 10.3847/1538-4365/aadfe9 URL https://doi.org/10.3847/1538-4365/aadfe9.

Aniyan AK & K Thorat 2017 'Classifying radio galaxies with the convolutional neural network' *The Astrophysical Journal Supplement Series* 230(2):20 doi: 10.3847/1538-4365/aa7333 URL https://doi.org/10.3847%2F1538-4365%2Faa7333.

Asgari Taghanaki S, Y Zheng, SK Zhou, B Georgescu, P Sharma, D Xu, D Comaniciu, & G Hamarneh 2018 'Combo Loss: Handling Input and Output Imbalance in Multi-Organ Segmentation' *arXiv e-prints* :arXiv:1805.02798.

Bechtol K, A Drlica-Wagner, E Balbinot, A Pieres, JD Simon, B Yanny, B Santiago, RH Wechsler, J Frieman, AR Walker, P Williams, E Rozo, ES Rykoff, A Queiroz, E Luque, A Benoit-Lévy, D Tucker, I Sevilla, RA Gruendl, LN da Costa, A Fausti Neto, MAG Maia, T Abbott, S Allam, R Armstrong, AH Bauer, GM Bernstein, RA Bernstein, E Bertin, D Brooks, E Buckley-Geer, DL Burke, A Carnero Rosell, FJ Castander, R Covarrubias, CB D'Andrea, DL DePoy, S Desai, HT Diehl, TF Eifler, J Estrada, AE Evrard, E Fernandez, DA Finley, B Flaugher, E Gaztanaga, D Gerdes, L Girardi, M Gladders, D Gruen, G Gutierrez, J Hao, K Honscheid, B Jain, D James, S Kent, R Kron, K Kuehn, N Kuropatkin, O Lahav, TS Li, H Lin, M Makler, M March, J Marshall, P Martini, KW Merritt, C Miller, R Miquel, J Mohr, E Neilsen, R Nichol, B Nord, R Ogando, J Peoples, D Petravick, AA Plazas, AK Romer, A Roodman, M Sako, E Sanchez, V Scarpine, M Schubnell, RC Smith, M Soares-Santos, F Sobreira, E Suchyta, MEC Swanson, G Tarle, J Thaler, D Thomas, W Wester, J Zuntz, & DES Collaboration 2015 'Eight New Milky Way Companions Discovered in First-year Dark Energy Survey Data' *The Astrophysical Journal* 807(1):50 doi: 10.1088/0004-637X/807/1/50.

Behroozi PS, RH Wechsler, & HY Wu 2013 'The ROCKSTAR phase-space temporal halo finder and the velocity offsets of cluster cores' *The Astrophysical Journal* 762(2):109 URL http://stacks.iop.org/0004-637X/762/i=2/a=109.

Bellazzini M, RA Ibata, SC Chapman, AD Mackey, L Monaco, MJ Irwin, NF Martin, GF Lewis, &

E Dalessandro 2008 ‘The Nucleus of the Sagittarius Dsph Galaxy and M54: a Window on the Process of Galaxy Nucleation’ *The Astronomical Journal* 136(3):1147–1170 doi: 10.1088/0004-6256/136/3/1147.

Berg DA, BL James, T King, M McDonald, Z Chen, J Chisholm, T Heckman, CL Martin, DP Stark, A Aloisi, RO Amorín, KZ Arellano-Córdova, M Bayliss, R Bordoloi, J Brinchmann, S Charlot, J Chevallard, I Clark, DK Erb, A Feltre, M Gronke, M Hayes, A Henry, S Hernandez, A Jaskot, T Jones, LJ Kewley, N Kumari, C Leitherer, M Llerena, M Maseda, M Mingozzi, T Nanayakkara, M Ouchi, A Plat, RW Pogge, S Ravindranath, JR Rigby, R Sanders, C Scarlata, P Senchyna, ED Skillman, CC Steidel, AL Strom, Y Sugahara, SM Wilkins, A Wofford, X Xu, & Classy Team 2022 ‘The COS Legacy Archive Spectroscopy Survey (CLASSY) Treasury Atlas’ *The Astrophysical Journal Supplement Series* 261(2):31 doi: 10.3847/1538-4365/ac6c03.

Bertoldi F & CF McKee 1992 ‘Pressure-confined Clumps in Magnetized Molecular Clouds’ *The Astrophysical Journal* 395:140 doi: 10.1086/171638.

Bordner J & ML Norman 2018 ‘Computational Cosmology and Astrophysics on Adaptive Meshes using Charm++’ *arXiv e-prints* :arXiv:1810.01319.

Bromm V 2013 ‘Formation of the first stars’ *Reports on Progress in Physics* 76(11):112901 doi: 10.1088/0034-4885/76/11/112901 URL https://doi.org/10.1088%2F0034-4885%2F76%2F11%2F112901.

Bromm V & RB Larson 2004 ‘The first stars’ *Annual Review of Astronomy and Astrophysics* 42(1):79–118 doi: 10.1146/annurev.astro.42.053102.134034 URL https://doi.org/10.1146/annurev.astro.42.053102.134034.

Bromm V, PS Coppi, & RB Larson 2002 ‘The formation of the first stars. i. the primordial star-forming cloud’ *The Astrophysical Journal* 564(1):23–51 doi: 10.1086/323947 URL https://doi.org/10.1086/323947.

Brummel-Smith C, G Bryan, I Butsky, L Corlies, A Emerick, J Forbes, Y Fujimoto, NJ Goldbaum, P Grete, CB Hummels, J hoon Kim, D Koh, M Li, Y Li, X Li, B O’Shea, MS Peeples, JA Regan, M Salem, W Schmidt, CM Simpson, BD Smith, J Tumlinson, MJ Turk, JH Wise, T Abel, J Bordner, R Cen, DC Collins, B Crosby, P Edelmann, O Hahn, R Harkness, E Harper-Clark, S Kong, AG Kritsuk, M Kuhlen, J Larrue, E Lee, G Meece, ML Norman, JS Oishi, P Paschos, C Peruta, A Razoumov, DR Reynolds, D Silvia, SW Skillman, S Skory, GC So, E Tasker, R Wagner, P Wang, H Xu, & F Zhao 2019a ‘Enzo: An adaptive mesh refinement code for astrophysics (version 2.6)’ *Journal of Open Source Software* 4(42):1636 doi: 10.21105/joss.01636 URL https://doi.org/10.21105/joss.01636.

Brummel-Smith C, G Bryan, I Butsky, L Corlies, A Emerick, J Forbes, Y Fujimoto, NJ Goldbaum, P Grete, CB Hummels, J hoon Kim, D Koh, M Li, Y Li, X Li, B O’Shea, MS Peeples, JA Regan, M Salem, W Schmidt, CM Simpson, BD Smith, J Tumlinson, MJ Turk, JH Wise, T Abel, J Bordner, R Cen, DC Collins, B Crosby, P Edelmann, O Hahn, R Harkness, E Harper-Clark, S Kong, AG Kritsuk, M Kuhlen, J Larrue, E Lee, G Meece, ML Norman, JS Oishi, P Paschos, C Peruta, A Razoumov, DR Reynolds, D Silvia, SW Skillman, S Skory, GC So, E Tasker, R Wagner, P Wang, H Xu, & F Zhao 2019b ‘ENZO: An adaptive mesh refinement code for astrophysics (version 2.6)’ *Journal of Open Source Software* 4(42):1636 doi: 10.21105/joss.01636 URL https://doi.org/10.21105/joss.01636.

Bryan GL, ML Norman, BW O’Shea, T Abel, JH Wise, MJ Turk, DR Reynolds, DC Collins, P Wang, SW Skillman, B Smith, RP Harkness, J Bordner, J hoon Kim, M Kuhlen, H Xu, N Goldbaum, C Hummels, AG Kritsuk, E Tasker, S Skory, CM Simpson, O Hahn, JS Oishi, GC So, F Zhao, R Cen, Y Li, & TE

Collaboration 2014 ‘Enzo: An adaptive mesh refinement code for astrophysics’ *The Astrophysical Journal* 211(2):19 URL http://stacks.iop.org/0067-0049/211/i=2/a=19.

Caldwell N, MG Walker, M Mateo, EW Olszewski, S Koposov, V Belokurov, G Torrealba, A Geringer-Sameth, & CI Johnson 2017 ‘Crater 2: An Extremely Cold Dark Matter Halo’ *The Astrophysical Journal* 839(1):20 doi: 10.3847/1538-4357/aa688e.

Chiaki G, N Yoshida, & T Kitayama 2013 ‘Low-mass Star Formation Triggered by Early Supernova Explosions’ *The Astrophysical Journal* 762(1):50 doi: 10.1088/0004-637X/762/1/50.

Cioffi DF, CF McKee, & E Bertschinger 1988 ‘Dynamics of Radiative Supernova Remnants’ *The Astrophysical Journal* 334:252 doi: 10.1086/166834.

Colella P & PR Woodward 1984 ‘The piecewise parabolic method (ppm) for gas-dynamical simulations’ *Journal of Computational Physics* 54(1):174–201 doi: https://doi.org/10.1016/0021-9991(84)90143-8 URL https://www.sciencedirect.com/science/article/pii/0021999184901438.

Collins MLM, EJ Tollerud, DJ Sand, A Bonaca, B Willman, & J Strader 2017 ‘Dynamical evidence for a strong tidal interaction between the Milky Way and its satellite, Leo V’ *Monthly Notices of the Royal Astronomical Society* 467(1):573–585 doi: 10.1093/mnras/stx067.

Cooke RJ, M Pettini, & CC Steidel 2017 ‘Discovery of the most metal-poor damped Lyman α system’ *Monthly Notices of the Royal Astronomical Society* 467(1):802–811 doi: 10.1093/mnras/stx037 URL https://doi.org/10.1093/mnras/stx037.

Crnojević D, DJ Sand, K Spekkens, N Caldwell, P Guhathakurta, B McLeod, A Seth, JD Simon, J Strader, & E Toloba 2016 ‘The Extended Halo of Centaurus A: Uncovering Satellites, Streams, and Substructures’ *The Astrophysical Journal* 823(1):19 doi: 10.3847/0004-637X/823/1/19.

Curti M, F D’Eugenio, S Carniani, R Maiolino, L Sandles, J Witstok, WM Baker, JS Bennett, JM Piotrowska, S Tacchella, S Charlot, K Nakajima, G Maheson, F Mannucci, S Arribas, F Belfiore, NR Bonaventura, AJ Bunker, J Chevallard, G Cresci, E Curtis-Lake, C Hayden-Pawson, N Kumari, I Laseter, TJ Looser, A Marconi, MV Maseda, GC Jones, J Scholtz, R Smit, H Ubler, & IEB Wallace 2022 ‘The chemical enrichment in the early Universe as probed by JWST via direct metallicity measurements at z~8’ *arXiv e-prints* :arXiv:2207.12375.

Draine BT 2011 *Physics of the Interstellar and Intergalactic Medium* Princeton University Press URL http://www.jstor.org/stable/j.ctvcm4hzr.

Eisenstein DJ & P Hu 1998 ‘HOP: A new group-finding algorithm forN-body simulations’ *The Astrophysical Journal* 498(1):137–142 doi: 10.1086/305535 URL https://doi.org/10.1086%2F305535.

Emberson JD, N Frontiere, S Habib, K Heitmann, P Larsen, H Finkel, & A Pope 2019 ‘The borg cube simulation: Cosmological hydrodynamics with CRK-SPH’ *The Astrophysical Journal* 877(2):85 doi: 10.3847/1538-4357/ab1b31 URL https://doi.org/10.3847/1538-4357/ab1b31.

Emerick A, GL Bryan, & MM Mac Low 2019 ‘Simulating an isolated dwarf galaxy with multichannel feedback and chemical yields from individual stars’ *Monthly Notices of the Royal Astronomical Society*

482(1):1304–1329 doi: 10.1093/mnras/sty2689.

Emerick A, GL Bryan, & MM Mac Low 2020 ‘Simulating Metal Mixing of Both Common and Rare Enrichment Sources in a Low-mass Dwarf Galaxy’ *The Astrophysical Journal* 890(2):155 doi: 10.3847/1538-4357/ab6efc.

Ferland GJ, M Chatzikos, F Guzmán, ML Lykins, PAM van Hoof, RJR Williams, NP Abel, NR Badnell, FP Keenan, RL Porter, & PC Stancil 2017 ‘The 2017 Release Cloudy’ *Revista Mexicana de Astronomía y Astrof’isica* 53:385–438.

Frebel A & JE Norris 2015 ‘Near-field cosmology with extremely metal-poor stars’ *Annual Review of Astronomy and Astrophysics* 53(1):631–688 doi: 10.1146/annurev-astro-082214-122423 URL https://doi.org/10.1146/annurev-astro-082214-122423.

Gnedin NY & JP Ostriker 1997 ‘Reionization of the universe and the early production of metals’ *The Astrophysical Journal* 486(2):581–598 doi: 10.1086/304548 URL https://doi.org/10.1086/304548.

Goodfellow I, J Pouget-Abadie, M Mirza, B Xu, D Warde-Farley, S Ozair, A Courville, & Y Bengio 2014 ‘Generative adversarial nets’ in Z Ghahramani, M Welling, C Cortes, N Lawrence, & K Weinberger (eds) *Advances in Neural Information Processing Systems* Curran Associates, Inc. volume 27 pp. 1–9 URL https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.

Hahn O & T Abel 2011 ‘Multi-scale initial conditions for cosmological simulations’ *Monthly Notices of the Royal Astronomical Society* 415(3):2101–2121 doi: 10.1111/j.1365-2966.2011.18820.x URL https://doi.org/10.1111/j.1365-2966.2011.18820.x.

Hajiabadi H, V Babaiyan, D Zabihzadeh, & M Hajiabadi 2020 ‘Combination of loss functions for robust breast cancer prediction’ *Computers & Electrical Engineering* 84:106624 doi: 10.1016/j.compeleceng.2020.106624.

He K, X Zhang, S Ren, & J Sun 2015 ‘Deep Residual Learning for Image Recognition’ *arXiv e-prints* :arXiv:1512.03385.

Heger A & SE Woosley 2002 ‘The Nucleosynthetic Signature of Population III’ *The Astrophysical Journal* 567:532–543 doi: 10.1086/338487.

Heger A, CL Fryer, SE Woosley, N Langer, & DH Hartmann 2003 ‘How massive single stars end their life’ *The Astrophysical Journal* 591(1):288–300 doi: 10.1086/375341 URL https://doi.org/10.1086%2F375341.

Hicks WM, AI Wells, ML Norman, JH Wise, BD Smith, & BW O’Shea 2021 ‘External enrichment of mini halos by the first supernovae’ *The Astrophysical Journal* 909(1):70 doi: 10.3847/1538-4357/abda3a URL https://doi.org/10.3847/1538-4357/abda3a.

Hochreiter S & Ju Schmidhuber 1997 ‘Long Short-Term Memory’ *Neural Computation* 9(8):1735–1780 doi: 10.1162/neco.1997.9.8.1735 URL https://doi.org/10.1162/neco.1997.9.8.1735.

Hockney RW & JW Eastwood 1988 *Computer Simulation Using Particles* USA: Taylor & Francis, Inc.

Hopkins PF, A Wetzel, D Kereš, CA Faucher-Giguére, E Quataert, M Boylan-Kolchin, N Murray, CC Hayward, & K El-Badry 2018 'How to model supernovae in simulations of star and galaxy formation' *Monthly Notices of the Royal Astronomical Society* 477(2):1578–1603 doi: 10.1093/mnras/sty674 URL https://doi.org/10.1093/mnras/sty674.

Hopkins PF, MY Grudić, A Wetzel, D Kereš, CA Faucher-Giguére, X Ma, N Murray, & N Butcher 2019 'Radiative stellar feedback in galaxy formation: Methods and physics' *Monthly Notices of the Royal Astronomical Society* 491(3):3702–3729 doi: 10.1093/mnras/stz3129 URL https://doi.org/10.1093/mnras/stz3129.

Hopkins PF, A Wetzel, C Wheeler, R Sanderson, MY Grudic, O Sameie, M Boylan-Kolchin, M Orr, X Ma, CA Faucher-Giguere, D Keres, E Quataert, KY Su, J Moreno, R Feldmann, JS Bullock, SR Loebman, D Angles-Alcazar, J Stern, L Necib, & CC Hayward 2022 'FIRE-3: Updated Stellar Evolution Models, Yields, & Microphysics and Fitting Functions for Applications in Galaxy Simulations' *arXiv e-prints* :arXiv:2203.00040.

Huang G, Z Liu, L van der Maaten, & KQ Weinberger 2016 'Densely Connected Convolutional Networks' *arXiv e-prints* :arXiv:1608.06993.

Ioffe S & C Szegedy 2015 'Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift' *arXiv e-prints* :arXiv:1502.03167.

Ishigaki MN, N Tominaga, C Kobayashi, & K Nomoto 2018 'The Initial Mass Function of the First Stars Inferred from Extremely Metal-poor Stars' *The Astrophysical Journal* 857(1):46 doi: 10.3847/1538-4357/aab3de.

Isola P, JY Zhu, T Zhou, & AA Efros 2016 'Image-to-Image Translation with Conditional Adversarial Networks' *arXiv e-prints* :arXiv:1611.07004.

Jeon M, G Besla, & V Bromm 2017 'Connecting the first galaxies with ultrafaint dwarfs in the local group: Chemical signatures of population iii stars' *The Astrophysical Journal* 848(2):85 URL http://stacks.iop.org/0004-637X/848/i=2/a=85.

Jin X, P Cheng, WL Chen, & H Li 2018 'Prediction model of velocity field around circular cylinder over various reynolds numbers by fusion convolutional neural networks based on pressure on the cylinder' *Physics of Fluids* 30(4):047105 doi: 10.1063/1.5024595 URL https://doi.org/10.1063/1.5024595.

Kappeler A, S Yoo, Q Dai, & AK Katsaggelos 2016 'Video super-resolution with convolutional neural networks' *IEEE Transactions on Computational Imaging* 2(2):109–122 doi: 10.1109/TCI.2016.2532323.

Kasim MF, D Watson-Parris, L Deaconu, S Oliver, P Hatfield, DH Froula, G Gregori, M Jarvis, S Khatiwala, J Korenaga, J Topp-Mugglestone, E Viezzer, & SM Vinko 2020 'Building high accuracy emulators for scientific simulations with deep neural architecture search' *arXiv e-prints* :arXiv:2001.08055.

Katz N 1992 'Dissipational Galaxy Formation. II. Effects of Star Formation' *The Astrophysical Journal* 391:502 doi: 10.1086/171366.

Keller SC, MS Bessell, A Frebel, AR Casey, M Asplund, HR Jacobson, K Lind, JE Norris, D Yong, A Heger, Z Magic, GS da Costa, BP Schmidt, & P Tisserand 2014 ‘A single low-energy, iron-poor supernova as the source of metals in the star SMSS J031300.36-670839.3’ *Nature* 506:463–466 doi: 10.1038/nature12990.

Kim CG & EC Ostriker 2015 ‘MOMENTUM INJECTION BY SUPERNOVAE IN THE INTERSTELLAR MEDIUM’ *The Astrophysical Journal* 802(2):99 doi: 10.1088/0004-637x/802/2/99 URL https://doi.org/10.1088/0004-637x/802/2/99.

Kim D & H Jerjen 2015 ‘Horologium II: A Second Ultra-faint Milky Way Satellite in the Horologium Constellation’ *The Astrophysical Journal Letters* 808(2):L39 doi: 10.1088/2041-8205/808/2/L39.

Kim D, H Jerjen, AP Milone, D Mackey, & GS Da Costa 2015 ‘Discovery of a Faint Outer Halo Milky Way Star Cluster in the Southern Sky’ *The Astrophysical Journal* 803(2):63 doi: 10.1088/0004-637X/803/2/63.

Kimm T & R Cen 2014 ‘Escape Fraction of Ionizing Photons during Reionization: Effects due to Supernova Feedback and Runaway OB Stars’ *The Astrophysical Journal* 788(2):121 doi: 10.1088/0004-637X/788/2/121.

Kingma DP & J Ba 2014 ‘Adam: A Method for Stochastic Optimization’ *arXiv e-prints* :arXiv:1412.6980.

Kirby EN, JG Cohen, P Guhathakurta, L Cheng, JS Bullock, & A Gallazzi 2013 ‘The Universal Stellar Mass-Stellar Metallicity Relation for Dwarf Galaxies’ *The Astrophysical Journal* 779(2):102 doi: 10.1088/0004-637X/779/2/102.

Kirby EN, JD Simon, & JG Cohen 2015 ‘Spectroscopic Confirmation of the Dwarf Galaxies Hydra II and Pisces II and the Globular Cluster Laevens 1’ *The Astrophysical Journal* 810(1):56 doi: 10.1088/0004-637X/810/1/56.

Koch A, MI Wilkinson, JT Kleyna, M Irwin, DB Zucker, V Belokurov, GF Gilmore, M Fellhauer, & NW Evans 2009 ‘A Spectroscopic Confirmation of the Bootes II Dwarf Spheroidal’ *The Astrophysical Journal* 690(1):453–462 doi: 10.1088/0004-637X/690/1/453.

Koposov SE, G Gilmore, MG Walker, V Belokurov, NW Evans, M Fellhauer, W Gieren, D Geisler, L Monaco, JE Norris, S Okamoto, J Peñarrubia, M Wilkinson, RFG Wyse, & DB Zucker 2011 ‘Accurate Stellar Kinematics at Faint Magnitudes: Application to the Boötes I Dwarf Spheroidal Galaxy’ *The Astrophysical Journal* 736(2):146 doi: 10.1088/0004-637X/736/2/146.

Koposov SE, AR Casey, V Belokurov, JR Lewis, G Gilmore, C Worley, A Hourihane, S Randich, T Bensby, A Bragaglia, M Bergemann, G Carraro, MT Costado, E Flaccomio, P Francois, U Heiter, V Hill, P Jofre, C Lando, AC Lanzafame, P de Laverny, L Monaco, L Morbidelli, L Sbordone, Š Mikolaitis, & N Ryde 2015 ‘Kinematics and Chemistry of Recently Discovered Reticulum 2 and Horologium 1 Dwarf Galaxies’ *The Astrophysical Journal* 811(1):62 doi: 10.1088/0004-637X/811/1/62.

Krizhevsky A, I Sutskever, & GE Hinton 2012 ‘Imagenet classification with deep convolutional neural networks’ in F Pereira, CJC Burges, L Bottou, & KQ Weinberger (eds) *Advances in Neural Information Processing Systems 25* Curran Associates, Inc. pp. 1097–1105.

Krumholz MR & NY Gnedin 2011 ‘A COMPARISON OF METHODS FOR DETERMINING THE MOLECULAR CONTENT OF MODEL GALAXIES’ *The Astrophysical Journal* 729(1):36 doi:

10.1088/0004-637x/729/1/36 URL https://doi.org/10.1088%2F0004-637x%2F729%2F1%2F36.

Lakshminarayanan B, A Pritzel, & C Blundell 2017 ‘Simple and scalable predictive uncertainty estimation using deep ensembles’ in I Guyon, UV Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan, & R Garnett (eds) *Advances in Neural Information Processing Systems* Curran Associates, Inc. volume 30 pp. 1–12 URL https://proceedings.neurips.cc/paper/2017/file/9ef2ed4b7fd2c810847ffa5fa85bce38-Paper.pdf.

Larson RB 2003 ‘The physics of star formation’ *Reports on Progress in Physics* 66(10):1651–1697 doi: 10.1088/0034-4885/66/10/R03.

Latif MA, D Whalen, & S Khochfar 2022 ‘The Birth Mass Function of Population III Stars’ *The Astrophysical Journal* 925(1):28 doi: 10.3847/1538-4357/ac3916.

Leitherer C, D Schaerer, JD Goldader, RMG Delgado, C Robert, DF Kune, DF de Mello, D Devost, & TM Heckman 1999 ‘Starburst99: Synthesis Models for Galaxies with Active Star Formation’ *The Astrophysical Journal Supplement Series* 123(1):3–40 doi: 10.1086/313233.

Li TS, JD Simon, A Drlica-Wagner, K Bechtol, MY Wang, J García-Bellido, J Frieman, JL Marshall, DJ James, L Strigari, AB Pace, E Balbinot, Y Zhang, TMC Abbott, S Allam, A Benoit-Lévy, GM Bernstein, E Bertin, D Brooks, DL Burke, A Carnero Rosell, M Carrasco Kind, J Carretero, CE Cunha, CB D’Andrea, LN da Costa, DL DePoy, S Desai, HT Diehl, TF Eifler, B Flaugher, DA Goldstein, D Gruen, RA Gruendl, J Gschwend, G Gutierrez, E Krause, K Kuehn, H Lin, MAG Maia, M March, F Menanteau, R Miquel, AA Plazas, AK Romer, E Sanchez, B Santiago, M Schubnell, I Sevilla-Noarbe, RC Smith, F Sobreira, E Suchyta, G Tarle, D Thomas, DL Tucker, AR Walker, RH Wechsler, W Wester, B Yanny, & DES Collaboration 2017 ‘Farthest Neighbor: The Distant Milky Way Satellite Eridanus II’ *The Astrophysical Journal* 838(1):8 doi: 10.3847/1538-4357/aa6113.

Li TS, JD Simon, AB Pace, G Torrealba, K Kuehn, A Drlica-Wagner, K Bechtol, AK Vivas, RP van der Marel, M Wood, B Yanny, V Belokurov, P Jethwa, DB Zucker, G Lewis, R Kron, DL Nidever, MA Sánchez-Conde, AP Ji, BC Conn, DJ James, NF Martin, D Martinez-Delgado, NED Noël, & MagLiteS Collaboration 2018 ‘Ships Passing in the Night: Spectroscopic Analysis of Two Ultra-faint Satellites in the Constellation Carina’ *The Astrophysical Journal* 857(2):145 doi: 10.3847/1538-4357/aab666.

Machida MN, K Tomisaka, F Nakamura, & MY Fujimoto 2005 ‘Low-Mass Star Formation Triggered by Supernovae in Primordial Clouds’ *The Astrophysical Journal* 622(1):39–57 doi: 10.1086/428030.

Majewski SR, MF Skrutskie, MD Weinberg, & JC Ostheimer 2003 ‘A Two Micron All Sky Survey View of the Sagittarius Dwarf Galaxy. I. Morphology of the Sagittarius Core and Tidal Arms’ *The Astrophysical Journal* 599(2):1082–1115 doi: 10.1086/379504.

Martizzi D, CA Faucher-Giguère, & E Quataert 2015 ‘Supernova feedback in an inhomogeneous interstellar medium’ *Monthly Notices of the Royal Astronomical Society* 450(1):504–522 doi: 10.1093/mnras/stv562.

Mateo M, EW Olszewski, & MG Walker 2008 ‘The Velocity Dispersion Profile of the Remote Dwarf Spheroidal Galaxy Leo I: A Tidal Hit and Run?’ *The Astrophysical Journal* 675(1):201–233 doi: 10.1086/522326.

Mathuriya A, D Bard, P Mendygral, L Meadows, J Arnemann, L Shao, S He, T Karna, D Moise, SJ

Pennycook, K Maschoff, J Sewall, N Kumar, S Ho, M Ringenburg, Prabhat, & V Lee 2018 ‘CosmoFlow: Using Deep Learning to Learn the Universe at Scale’ *arXiv e-prints* :arXiv:1808.04728.

Mohan A, D Daniel, M Chertkov, & D Livescu 2019 ‘Compressed Convolutional LSTM: An Efficient Deep Learning framework to Model High Fidelity 3D Turbulence’ *arXiv e-prints* :arXiv:1903.00033.

Muñoz RR, P Côté, FA Santana, M Geha, JD Simon, GA Oyarzún, PB Stetson, & SG Djorgovski 2018 ‘A MegaCam Survey of Outer Halo Satellites. III. Photometric and Structural Parameters’ *The Astrophysical Journal* 860(1):66 doi: 10.3847/1538-4357/aac16b.

Murray N 2011 ‘STAR FORMATION EFFICIENCIES AND LIFETIMES OF GIANT MOLECULAR CLOUDS IN THE MILKY WAY’ *The Astrophysical Journal* 729(2):133 doi: 10.1088/0004-637x/729/2/133 URL https://doi.org/10.1088/0004-637x/729/2/133.

Nakamura F & M Umemura 2002 ‘The stellar initial mass function in primordial galaxies’ *The Astrophysical Journal* 569(2):549–557 doi: 10.1086/339392 URL https://doi.org/10.1086/339392.

Navarro JF, CS Frenk, & SDM White 1997 ‘A Universal Density Profile from Hierarchical Clustering’ *The Astrophysical Journal* 490(2):493–508 doi: 10.1086/304888.

Nomoto K, N Tominaga, H Umeda, C Kobayashi, & K Maeda 2006 ‘Nucleosynthesis yields of core-collapse supernovae and hypernovae, and galactic chemical evolution’ *Nuclear Physics A* 777:424 – 458 doi: https://doi.org/10.1016/j.nuclphysa.2006.05.008 URL http://www.sciencedirect.com/science/article/pii/S0375947406001953 special Issue on Nuclear Astrophysics.

Oppenheimer BD & R Davé 2006 ‘Cosmological simulations of intergalactic medium enrichment from galactic outflows’ *Monthly Notices of the Royal Astronomical Society* 373(4):1265–1292 doi: 10.1111/j.1365-2966.2006.10989.x.

Park J, M Ricotti, & K Sugimura 2021 ‘Population III star formation in an X-ray background - II. Protostellar discs, multiplicity, and mass function of the stars’ *Monthly Notices of the Royal Astronomical Society* 508(4):6193–6208 doi: 10.1093/mnras/stab3000.

Paszke A, S Gross, F Massa, A Lerer, J Bradbury, G Chanan, T Killeen, Z Lin, N Gimelshein, L Antiga, A Desmaison, A Kopf, E Yang, Z DeVito, M Raison, A Tejani, S Chilamkurthy, B Steiner, L Fang, J Bai, & S Chintala 2019 ‘Pytorch: An imperative style, high-performance deep learning library’ in H Wallach, H Larochelle, A Beygelzimer, F Alché-Buc, E Fox, & R Garnett (eds) *Advances in Neural Information Processing Systems 32* Curran Associates, Inc. pp. 8024–8035.

Regan JA, E Visbal, JH Wise, Z Haiman, PH Johansson, & GL Bryan 2017 ‘Rapid formation of massive black holes in close proximity to embryonic protogalaxies’ *Nature Astronomy* 1:0075 doi: 10.1038/s41550-017-0075.

Ritter JS, C Safranek-Shrader, O Gnat, M Milosavljević, & V Bromm 2012 ‘Confined Population III Enrichment and the Prospects for Prompt Second-generation Star Formation’ *The Astrophysical Journal* 761(1):56 doi: 10.1088/0004-637X/761/1/56.

Ronneberger O, P Fischer, & T Brox 2015 ‘U-Net: Convolutional Networks for Biomedical Image Segmentation' *arXiv e-prints* :arXiv:1505.04597.

Rosdahl J, J Schaye, Y Dubois, T Kimm, & R Teyssier 2016 'Snap, crackle, pop: sub-grid supernova feedback in AMR simulations of disc galaxies' *Monthly Notices of the Royal Astronomical Society* 466(1):11–33 doi: 10.1093/mnras/stw3034 URL https://doi.org/10.1093/mnras/stw3034.

Rosdahl J, J Schaye, Y Dubois, T Kimm, & R Teyssier 2017 'Snap, crackle, pop: sub-grid supernova feedback in AMR simulations of disc galaxies' *Monthly Notices of the Royal Astronomical Society* 466(1):11–33 doi: 10.1093/mnras/stw3034.

Rosenberg A & J Hirschberg 2007 'V-measure: A conditional entropy-based external cluster evaluation measure' in *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)* Prague, Czech Republic: Association for Computational Linguistics pp. 410–420 URL https://www.aclweb.org/anthology/D07-1043.

Schleicher DRG, F Palla, A Ferrara, D Galli, & M Latif 2013 'Massive black hole factories: Supermassive and quasi-star formation in primordial halos' *Astronomy and Astrophysics* 558:A59 doi: 10.1051/0004-6361/201321949.

Sedov LI 1946 'Propagation of strong shock waves' *Journal of Applied Mathematics and Mechanics* 10:241–250.

Simon JD 2019 'The Faintest Dwarf Galaxies' *Annual Review of Astronomy and Astrophysics* 57:375–415 doi: 10.1146/annurev-astro-091918-104453.

Simon JD & M Geha 2007 'The Kinematics of the Ultra-faint Milky Way Satellites: Solving the Missing Satellite Problem' *The Astrophysical Journal* 670(1):313–331 doi: 10.1086/521816.

Simon JD, M Geha, QE Minor, GD Martinez, EN Kirby, JS Bullock, M Kaplinghat, LE Strigari, B Willman, PI Choi, EJ Tollerud, & J Wolf 2011 'A Complete Spectroscopic Survey of the Milky Way Satellite Segue 1: The Darkest Galaxy' *The Astrophysical Journal* 733(1):46 doi: 10.1088/0004-637X/733/1/46.

Simon JD, A Drlica-Wagner, TS Li, B Nord, M Geha, K Bechtol, E Balbinot, E Buckley-Geer, H Lin, J Marshall, B Santiago, L Strigari, M Wang, RH Wechsler, B Yanny, T Abbott, AH Bauer, GM Bernstein, E Bertin, D Brooks, DL Burke, D Capozzi, A Carnero Rosell, M Carrasco Kind, CB D'Andrea, LN da Costa, DL DePoy, S Desai, HT Diehl, S Dodelson, CE Cunha, J Estrada, AE Evrard, A Fausti Neto, E Fernandez, DA Finley, B Flaugher, J Frieman, E Gaztanaga, D Gerdes, D Gruen, RA Gruendl, K Honscheid, D James, S Kent, K Kuehn, N Kuropatkin, O Lahav, MAG Maia, M March, P Martini, CJ Miller, R Miquel, R Ogando, AK Romer, A Roodman, ES Rykoff, M Sako, E Sanchez, M Schubnell, I Sevilla, RC Smith, M Soares-Santos, F Sobreira, E Suchyta, MEC Swanson, G Tarle, J Thaler, D Tucker, V Vikram, AR Walker, W Wester, & DES Collaboration 2015 'Stellar Kinematics and Metallicities in the Ultra-faint Dwarf Galaxy Reticulum II' *The Astrophysical Journal* 808(1):95 doi: 10.1088/0004-637X/808/1/95.

Slyz AD, JEG Devriendt, G Bryan, & J Silk 2005 'Towards simulating star formation in the interstellar medium' *Monthly Notices of the Royal Astronomical Society* 356(2):737–752 doi: 10.1111/j.1365-2966.2004.08494.x URL https://doi.org/10.1111/j.1365-2966.2004.08494.x.

Smith B, S Sigurdsson, & T Abel 2008 ‘Metal cooling in simulations of cosmic structure formation’ *Monthly Notices of the Royal Astronomical Society* 385(3):1443–1454 doi: 10.1111/j.1365-2966.2008.12922.x URL https://doi.org/10.1111/j.1365-2966.2008.12922.x.

Smith B, JH Wise, BW O’Shea, ML Norman, & S Khochfar 2015 ‘The first Population II stars formed in externally enriched mini-haloes’ *Monthly Notices of the Royal Astronomical Society* 452(3):2822–2836 doi: 10.1093/mnras/stv1509 URL http://dx.doi.org/10.1093/mnras/stv1509.

Smith BD, MJ Turk, S Sigurdsson, BW O’Shea, & ML Norman 2009 ‘Three modes of metal-enriched star formation in the early universe’ *The Astrophysical Journal* 691(1):441 URL http://stacks.iop.org/0004-637X/691/i=1/a=441.

Spencer ME, M Mateo, MG Walker, EW Olszewski, AW McConnachie, EN Kirby, & A Koch 2017 ‘The Binary Fraction of Stars in Dwarf Galaxies: The Case of Leo II’ *The Astronomical Journal* 153(6):254 doi: 10.3847/1538-3881/aa6d51.

Stacy A, TH Greif, & V Bromm 2010 ‘The first stars: formation of binaries and small multiple systems’ *Monthly Notices of the Royal Astronomical Society* 403(1):45–60 doi: 10.1111/j.1365-2966.2009.16113.x.

Stacy A, V Bromm, & AT Lee 2016 ‘Building up the Population III initial mass function from cosmological initial conditions’ *Monthly Notices of the Royal Astronomical Society* 462(2):1307–1328 doi: 10.1093/mnras/stw1728.

Szegedy C, W Liu, Y Jia, P Sermanet, S Reed, D Anguelov, D Erhan, V Vanhoucke, & A Rabinovich 2014 ‘Going Deeper with Convolutions’ *arXiv e-prints* :arXiv:1409.4842.

Taylor GI 1950 ‘The formation of a blast wave by a very intense explosion i. theoretical discussion’ *Proceedings of the Royal Society of London Series A Mathematical and Physical Sciences* 201(1065):159–174.

The Planck Collaboration, PAR Ade, N Aghanim, C Armitage-Caplan, M Arnaud, M Ashdown, F Atrio-Barandela, J Aumont, C Baccigalupi, AJ Banday, & et al 2014 ‘Planck 2013 results. xvi. cosmological parameters’ *Astronomy & Astrophysics* 571:A16 doi: 10.1051/0004-6361/201321591 URL http://dx.doi.org/10.1051/0004-6361/201321591.

Thornton K, M Gaudlitz, HT Janka, & M Steinmetz 1998 ‘Energy input and mass redistribution by supernovae in the interstellar medium’ *The Astrophysical Journal* 500(1):95–119 doi: 10.1086/305704 URL https://doi.org/10.1086/305704.

Tompson J, R Goroshin, A Jain, Y LeCun, & C Bregler 2014 ‘Efficient Object Localization Using Convolutional Networks’ *arXiv e-prints* :arXiv:1411.4280.

Torrealba G, SE Koposov, V Belokurov, & M Irwin 2016a ‘The feeble giant. Discovery of a large and diffuse Milky Way dwarf galaxy in the constellation of Crater’ *Monthly Notices of the Royal Astronomical Society* 459(3):2370–2378 doi: 10.1093/mnras/stw733.

Torrealba G, SE Koposov, V Belokurov, M Irwin, M Collins, M Spencer, R Ibata, M Mateo, A Bonaca, & P Jethwa 2016b ‘At the survey limits: discovery of the Aquarius 2 dwarf galaxy in the VST ATLAS and the SDSS data’ *Monthly Notices of the Royal Astronomical Society* 463(1):712–722 doi: 10.1093/mnras/

stw2051.

Torrealba G, V Belokurov, SE Koposov, K Bechtol, A Drlica-Wagner, KAG Olsen, AK Vivas, B Yanny, P Jethwa, AR Walker, TS Li, S Allam, BC Conn, C Gallart, RA Gruendl, DJ James, MD Johnson, K Kuehn, N Kuropatkin, NF Martin, D Martinez-Delgado, DL Nidever, NED Noël, JD Simon, GS Stringfellow, & DL Tucker 2018 'Discovery of two neighbouring satellites in the Carina constellation with MagLiteS' *Monthly Notices of the Royal Astronomical Society* 475(4):5085–5097 doi: 10.1093/mnras/sty170.

Torrey P, M Vogelsberger, F Marinacci, R Pakmor, V Springel, D Nelson, J Naiman, A Pillepich, S Genel, R Weinberger, & L Hernquist 2019 'The evolution of the mass–metallicity relation and its scatter in IllustrisTNG' *Monthly Notices of the Royal Astronomical Society* 484(4):5587–5607 doi: 10.1093/mnras/stz243 URL https://doi.org/10.1093/mnras/stz243.

Tumlinson J, MS Peeples, & JK Werk 2017 'The circumgalactic medium' *Annual Review of Astronomy and Astrophysics* 55(1):389–432 doi: 10.1146/annurev-astro-091916-055240 URL https://doi.org/10.1146/annurev-astro-091916-055240.

Turk MJ, BD Smith, JS Oishi, S Skory, SW Skillman, T Abel, & ML Norman 2011 'yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data' *The Astrophysical Journal* 192:9 doi: 10.1088/0067-0049/192/1/9.

Vogelsberger M, S Genel, D Sijacki, P Torrey, V Springel, & L Hernquist 2013 'A model for cosmological simulations of galaxy formation physics' *Monthly Notices of the Royal Astronomical Society* 436(4):3031–3067 doi: 10.1093/mnras/stt1789.

Vogelsberger M, V Genel Sand Springel, P Torrey, D Sijacki, D Xu, G Snyder, S Bird, D Nelson, & L Hernquist 2014 'Properties of galaxies reproduced by a hydrodynamic simulation' *Nature* 509 doi: 10.1038/nature13316 URL https://doi.org/10.1038/nature13316.

Vogelsberger M, F Marinacci, P Torrey, & E Puchwein 2020 'Cosmological simulations of galaxy formation' *Nature Reviews Physics* 2(1):42–66 doi: 10.1038/s42254-019-0127-2.

Walker MG, M Mateo, EW Olszewski, J Peñarrubia, NW Evans, & G Gilmore 2009a 'A Universal Mass Profile for Dwarf Spheroidal Galaxies?' *The Astrophysical Journal* 704(2):1274–1287 doi: 10.1088/0004-637X/704/2/1274.

Walker MG, M Mateo, EW Olszewski, B Sen, & M Woodroofe 2009b 'Clean Kinematic Samples in Dwarf Spheroidals: An Algorithm for Evaluating Membership and Estimating Distribution Parameters When Contamination is Present' *The Astronomical Journal* 137(2):3109–3138 doi: 10.1088/0004-6256/137/2/3109.

Walker MG, M Mateo, EW Olszewski, S Koposov, V Belokurov, P Jethwa, DL Nidever, V Bonnivard, I Bailey John I, EF Bell, & SR Loebman 2016 'Magellan/M2FS Spectroscopy of Tucana 2 and Grus 1' *The Astrophysical Journal* 819(1):53 doi: 10.3847/0004-637X/819/1/53.

Wells AI & ML Norman 2021 'Predicting Localized Primordial Star Formation with Deep Convolutional Neural Networks' *The Astrophysical Journal Supplement Series* 254(2):41 doi: 10.3847/1538-4365/abfa17.

Wells AI & ML Norman 2022 'Connecting Primordial Star-forming Regions and Second-generation Star Formation in the Phoenix Simulations' *The Astrophysical Journal* 932(1):71 doi: 10.3847/1538-4357/ac6c87.

Welsh L, R Cooke, & M Fumagalli 2019 'Modelling the chemical enrichment of Population III supernovae: the origin of the metals in near-pristine gas clouds' *Monthly Notices of the Royal Astronomical Society* 487(3):3363–3376 doi: 10.1093/mnras/stz1526 URL https://doi.org/10.1093/mnras/stz1526.

Welsh L, R Cooke, & M Fumagalli 2020 'The stochastic enrichment of Population II stars' *Monthly Notices of the Royal Astronomical Society* 500(4):5214–5228 doi: 10.1093/mnras/staa3342 URL https://doi.org/10.1093/mnras/staa3342.

Whalen D, BW O'Shea, J Smidt, & ML Norman 2008a 'How the first stars regulated local star formation. i. radiative feedback' *The Astrophysical Journal* 679(2):925–941 doi: 10.1086/587731 URL https://doi.org/10.1086%2F587731.

Whalen D, B van Veelen, BW O'Shea, & ML Norman 2008b 'The destruction of cosmological minihalos by primordial supernovae' *The Astrophysical Journal* 682(1):49–67 doi: 10.1086/589643 URL https://doi.org/10.1086%2F589643.

Wheeler C, PF Hopkins, AB Pace, S Garrison-Kimmel, M Boylan-Kolchin, A Wetzel, JS Bullock, D Kereš, CAe Faucher-Gigu` ere, & E Quataert 2019 'Be it therefore resolved: cosmological simulations of dwarf galaxies with 30 solar mass resolution' *Monthly Notices of the Royal Astronomical Society* 490(3):4447–4463 doi: 10.1093/mnras/stz2887 URL https://doi.org/10.1093/mnras/stz2887.

Willman B & J Strader 2012 '"Galaxy," Defined' *The Astronomical Journal* 144(3):76 doi: 10.1088/0004-6256/144/3/76.

Willman B, M Geha, J Strader, LE Strigari, JD Simon, E Kirby, N Ho, & A Warres 2011 'Willman 1—A Probable Dwarf Galaxy with an Irregular Kinematic Distribution' *The Astronomical Journal* 142(4):128 doi: 10.1088/0004-6256/142/4/128.

Wise JH & T Abel 2011 'enzo+moray: radiation hydrodynamics adaptive mesh refinement simulations with adaptive ray tracing' *Monthly Notices of the Royal Astronomical Society* 414(4):3458–3491 doi: 10.1111/j.1365-2966.2011.18646.x URL +http://dx.doi.org/10.1111/j.1365-2966.2011.18646.x.

Wise JH & R Cen 2009 'Ionizing Photon Escape Fractions From High-Redshift Dwarf Galaxies' *The Astrophysical Journal* 693(1):984–999 doi: 10.1088/0004-637X/693/1/984.

Wise JH, T Abel, MJ Turk, ML Norman, & BD Smith 2012 'The birth of a galaxy - II. The role of radiation pressure' *Monthly Notices of the Royal Astronomical Society* 427(1):311–326 doi: 10.1111/j.1365-2966.2012.21809.x.

Wise JH, MJ Turk, ML Norman, & T Abel 2012 'The birth of a galaxy: Primordial metal enrichment and stellar populations' *The Astrophysical Journal* 745(1):50 URL http://stacks.iop.org/0004-637X/745/i=1/a=50.

Woosley SE & A Heger 2015 *The Deaths of Very Massive Stars* Cham: Springer International Publishing chapter 7 pp. 199–225 doi: 10.1007/978-3-319-09596-7_7 URL https://doi.org/10.1007/978-3-319-09596-7_7.

Xu H, JH Wise, & ML Norman 2013 ‘POPULATION III STARS AND REMNANTS IN HIGH-REDSHIFT GALAXIES’ *The Astrophysical Journal* 773(2):83 doi: 10.1088/0004-637x/773/2/83 URL https://doi.org/10.1088/0004-637x/773/2/83.

Xu H, JH Wise, ML Norman, K Ahn, & BW O’Shea 2016 ‘Galaxy properties and uv escape fractions during the epoch of reionization: Results from the renaissance simulations’ *The Astrophysical Journal* 833(1):84 URL http://stacks.iop.org/0004-637X/833/i=1/a=84.

Zhang X, Y Wang, W Zhang, Y Sun, S He, G Contardo, F Villaescusa-Navarro, & S Ho 2019 ‘From Dark Matter to Galaxies with Convolutional Networks’ *arXiv e-prints* :arXiv:1902.05965.

Zhang Z, C Wu, S Coleman, & D Kerr 2020 ‘Dense-inception u-net for medical image segmentation’ *Computer Methods and Programs in Biomedicine* 192:105395 doi: https://doi.org/10.1016/j.cmpb.2020.105395 URL http://www.sciencedirect.com/science/article/pii/S0169260719307904.

Zhu JY, T Park, P Isola, & AA Efros 2017 ‘Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks’ *arXiv e-prints* :arXiv:1703.10593.