The First Water in the Universe

Joseph Smidt[b], Brandon Wiggins[c], Francesca Samsel[a], Kristin Hoch[b], Greg Abram[a], Sam Jones[b], Alex Gagliano[d], Morgan Taylor[c]

[a] *University of Texas at Austin*
[b] *Center for Theoretical Astrophysics, Los Alamos National Laboratory*
[c] *Southern Utah University*
[d] *University of Illinois*

Abstract

Recent work in the chemistry of molecular clouds in the early universe has found reactions that may create ancient water molecules that predate our Sun. For these to take place, particular conditions of molecular density and temperature must exist. Water molecules will be destroyed by high temperatures, or when densities are insufficient to shield them from UV rays. If the environment is too cold, the reactions will not take place. Water molecules will only be created and persist in areas of the universe in which these properties are suitable, and will do so at different rates depending on local conditions.

In this work we combine a high resolution hydrodynamics simulation of the early universe closely coupled with this hydroxyl and water-producing chemistry model to determine how water molecules would be created and distributed in space and time in the early universe.

This work is enabled by the computational power of todays supercomputers and simulation technology. The complexity of the chemistry model is significantly higher than that of simple hydrodynamics, making this a computationally intensive model. Vast difference in scale of the physics involved, from the cosmological scale of the universe through the stellar scale of stars and novae to the molecular scale of chemical reactions requires that adaptive mesh refinement (AMR) techniques be used to provide resolution that varies as demanded by the physics. The visualizations presented herein will show the dynamics of the simulation as it evolves over time.

Keywords:
water, cosmology, astrochemistry, multiphysics simulations

1. Introduction

The story of the rise of life in the universe is among the most universal questions posed by scientists and poets alike. Understanding the conditions from which life might emerge is one of the fundamental pursuits of all of science.

The history of various life-giving molecules in the universe is a critical open issue. As much as half of the water in the solar system is thought to predate the Sun (Cleeves et al. 2014), which points to an ancient heritage for this substance. *How*, *where* and *when* water appears in the universe and with *what frequency* have never been answered from first principle calculations. In our submission, we reveal the first sites of water formation in the universe as calculated by new massively parallel cosmology simulations brought to life with state-of-the-art 3D visualization.

1.1. The Universe Didn't Always Have Molecules

The origin of all molecules would have taken place long before it was filled with spiral galaxies or stars with planetary systems. Shortly after the Big Bang, the universe was filled with only hydrogen and helium which collapsed under self-gravity to form pockets of gas called halos. The gas in halos would have formed the very first stars in the universe.

The first stars would have been ~40-100 times more massive than the Sun and tens of millions of times more luminous. These stars, made only from hydrogen and helium, would fuse these elements into heavier elements including carbon, oxygen, silicon and iron. They would have then died in spectacular supernova explosions which forcibly expel these heavy elements into the depths of space. These elements may have reacted with surrounding gas to form familiar chemical compounds, like carbon dioxide, methane, hydroxyl and water.

1.2. What It Takes To Make Water

Idealized one dimensional calculations reveal that water formation in the early universe may have been possible (Bialy et al. 2015). Even so, chemistry in the early universe was markedly different than interstellar chemistry today. Much of the water formed in the current cosmos is created by reactions on dust grains. Dust is made from heavier elements, but little dust would have been present in the early universe. Without dust, water forms through neutral-neutral reactions, with oxygen and hydrogen first forming hydroxyl (OH) and then reacting with hydrogen to form water.

Water formation in the early universe requires three basic things:

1. *Dense Gas*
   Dense gas increases the odds of particle collisions and re-

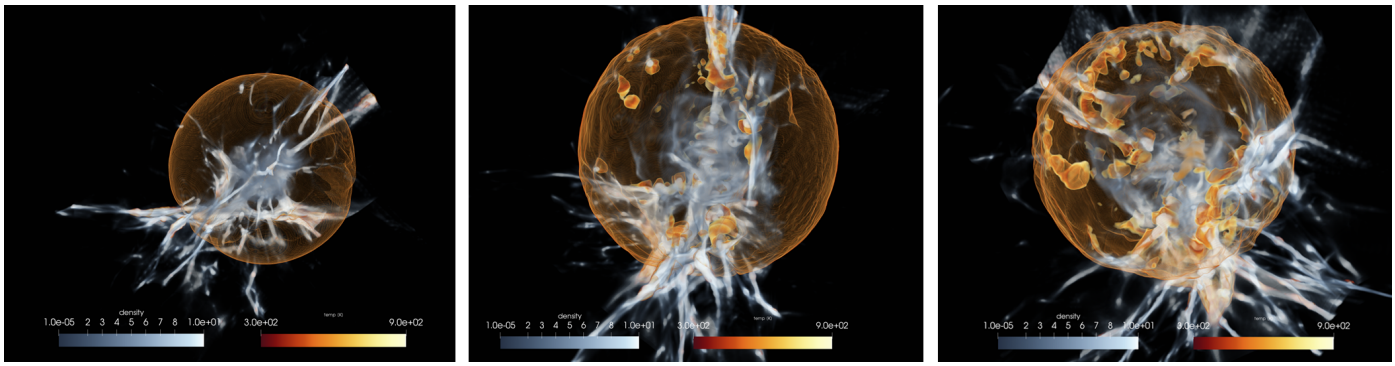

Figure 1: Multiple snapshots depicting gas density and temperature of an exploding supernova in the early universe. Chemistry in space is very sensitive heat content and density of gas and our 3-D visualizations offer new intuition into the complex process by which supernova shockwaves heat up and compress surrounding gas.

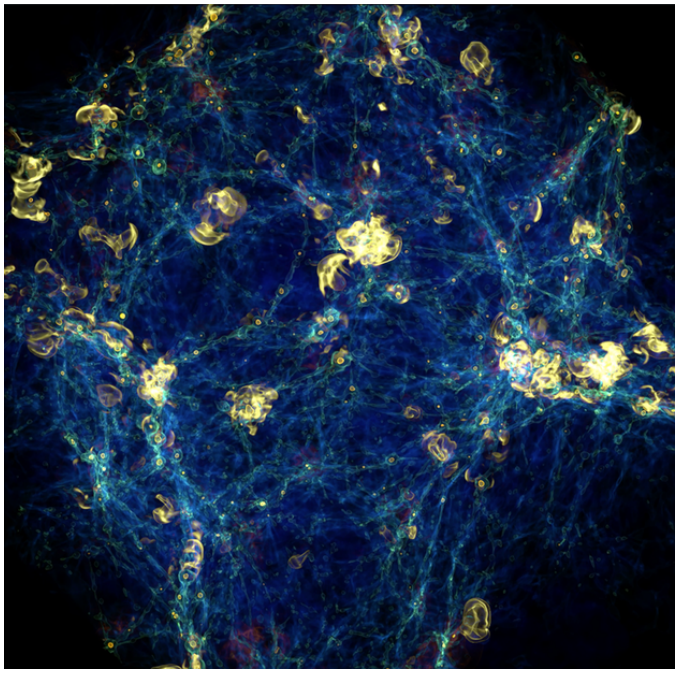

Figure 2: Most cosmology simulations only follow the collapse of dark matter and gas (blue) and the effects of stars (yellow highlights hot gas). This simulation, carried out in `Enzo` follows *only* star formation and hydrodynamics and required over one million CPU hours to complete. Following primordial molecular chemistry self-consistently between 26 molecules with over 50 reactions diminishes performance of these hero calculations by as much as a factor of 4, a crucial reason why combining chemistry with cosmology simulations has not previously been attempted.

actions and also shields water from the ultraviolet rays of stars which would dissociate it.

2. *Warm Gas Temperatures*
   Gas temperatures need to be large than ~ 300 K (100° C) to spur reactions. The gas can't be too warm, or water is destroyed.
3. *The Presence of Oxygen*
   Over 99% of the gas in the early universe would have been hydrogen and helium. Water can only form if it receives oxygen, a product only made in stars.

How or *if* these conditions actually arise from natural scenarios in a cosmological context is an open question, requiring HPC and state-of-the-art visualization to reveal.

1.3. Current Computational Cosmology

Full cosmological simulations are fairly common. In such calculations, a simulation box representing a sizeable volume of the universe is run which follows the collapse of dark matter into a system of filaments with hydrogen and helium collecting in halos at filament nodes. The effects of stars are simulated, which heat or remove gas from halos by violent supernova explosions. Usually, a small set of chemical reactions is carried out to model the cooling and collapse of the gas. This minimal chemistry is carried out explicitly by software libraries like `Grackle`. Cosmological simulations require extensive use of adaptive mesh refinement, the art of subdividing cells to gain more resolution, to resolve both large cosmic scales and the small scales where stars form.

While water is formed through a relatively simple reaction chain, predicting the *amount* and *distribution* of water in the early universe requires the simulation of a much larger reaction network. For example, in circumstances where water may have formed, carbon monoxide, a frequent molecule in dense gas may have also formed in sufficient abundance to starve water production of oxygen.

The situation is complicated, however, by the *sheer number* of reactions which carbon, oxygen and hydrogen might undergo in interstellar gas, an issue which rendered previous attempts to calculate water formation in cosmological simulations cost-prohibitive. Additionally, large reaction networks like ours cannot be solved by direct updates to variables with ready-made equations. Chemical species must instead be updated by inverting a mid-sized matrix with a size determined by the number of molecular species in the simulation. This matrix must be inverted for each cell, as many as hundreds of times in a given computer cycle depending on the conditions in each cell.

The computational cosmology research collaboration at the Center for Theoretical Astrophysics at Los Alamos National Laboratory (LANL) have previously carried out demanding calculations on numerous phenomena in the early universe including simulating the growth of ancient massive black holes in the early universe using high resolution cosmological simulations

(e.g. Smidt et al. 2018; Smidt, Wiggins & Johnson 2016), predicting the merger sites of binary neutron stars on the cosmic web (Wiggins et al. 2018) and calculating the brightness and character of light from ancient supernovae (e.g. Smidt et al. 2014; Chatzopoulos et al. 2015; Mesler et al. 2014). LANL's cutting-edge position in cosmological simulations along with its expertise in supernovae and multi-physics calculations render it an ideal institution to address the challenging task of simulating the appearance of the first water in the cosmos.

2. Method

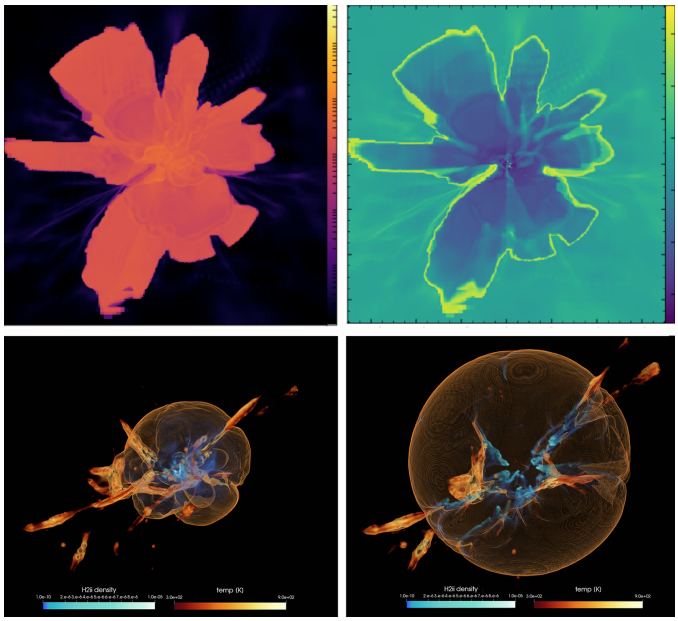

Figure 3: *Top panels:* right - 2D slice plot of temperature of a star dying in a supernova explosion created in `yt`, left - slice of H_2II, hydroxyl, rendered in `yt`. *Bottom panels:* Temperature and H_2II volume-rendered together at early and late stages of the supernova. Normal use of cosmology visualization tools such as *yt* usually result in 2-D plots of a single quantity such as temperature. Visualization methods, shown in the lower panels, allow cosmologists to draw rapid connections of relationship *between* physical quantities and the underlying chemistry by allowing multiple quantities to share the same 3D space. In the lower plots, the creation of H_2II is seen to be the direct result of the supernova shock wave.

We combined the capabilities of *three* codes, all running in-line, to create the simulations in this submission, `Enzo`, `Grackle` and `WaterNet`. Our results are visualized with ParaView. One of these codes is our stand-alone molecular chemistry solver, `WaterNet`, which was written by the authors over a period of two years to address this very problem. We describe the importance of each code base below.

2.1. Enzo

For cosmology calculations, we use `Enzo` (Bryan et al. 2014), a cosmological adaptive mesh refinement (AMR) code. `Enzo` follows the collapse and self-gravity of the dark matter cosmic web, solves the equations of hydrodynamics for gas falling into potential wells and models subsequent star formation. As stars are produced, `Enzo` uses a ray-tracing radiative transfer scheme to model the effects of light on surround gas and also models the production of shock waves as stars die in supernova explosions.

Our simulation root grids vary from 128^3 to 512^3 with 4 subgrids and 14 levels of refinement. Each simulation requires $\sim 100,000$ CPU hours and hundreds of dumps, each of which is ~ 10 GB.

2.2. Grackle

`Grackle` (Smith et al. 2017) is a cooling library which solves simple chemistry for hydrogen, helium to assess how halos in *Enzo*'s simulation should collapse. Molecular hydrogen emits infrared light which allows halos to glow and cool.

`Grackle` solves what cosmologist call *primordial chemistry*. It's a simple type of chemistry which involves only hydrogen and helium and electrons. It is distinct from *molecular chemistry*, which contains many more compounds with atoms other than hydrogen and helium. `Grackle`'s implementation of primordial chemistry is an semi-explicit solver, meaning that variables are *updated directly* with a hard-coded formula in a strategic order. This simple implementation is *inexpensive* but would have severe limitations in accurately modeling a complex reaction network. We describe the implementation of our molecular chemistry network below.

2.3. WaterNet

Our non-equilibrium chemical network evolves abundances of H, H^+, H^-, H_2, e, D, D^+, HD, O, OH, H_2O, O_2, O^+, OH^+, H_2O^+, H_3O^+, O_2^-, C^+, C, CH, CH_2, CH_3, CH_4, CO, CO^+ and CO_2 with 61 reactions between the species. `Enzo` passes chemical abundance, temperature and density information to

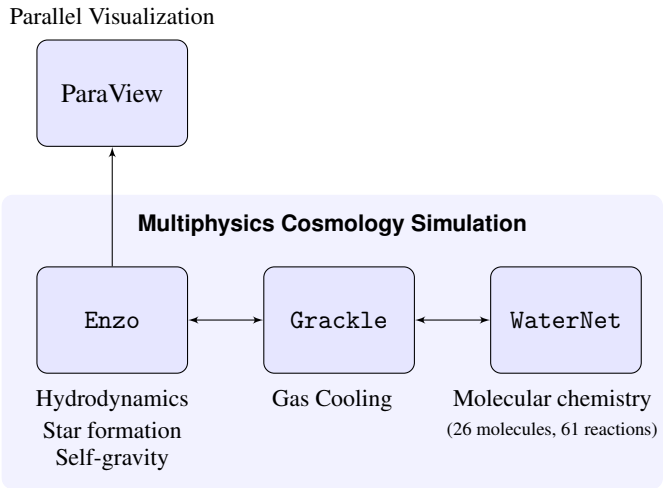

Figure 4: Schematic of relation between codes. `Enzo`, a parallel cosmology code, uses `Grackle`, a cooling library which informs how gas collapses to form structures including stars. `WaterNet`, expands the chemistry solve to some 26 different molecules and treats 61 chemical reactions. Our results are visualized by ParaView.

WaterNet afterwhich it solves the rate equations

$$F_i = \frac{dY_i}{dt} = \sum_{j,k} k_{jk} Y_j Y_k + \sum_l k_l Y_l - Y_i \left[\sum_m k_{im} Y_m - \sum_n k_n \right] \quad (1)$$

for each species in our network. Our network is taken from Omukai et al. (2005).

The inclusion of 26 addition species inside of Enzo posed a serious demand on memory. Enzo normally tracks 10-20 fields per cell and our implementation of WaterNet within Enzo ups this number to 40 or 50. To avoid memory issues, we used the Grizzy cluster at LANL which possesses 125 GB of ram per node, the most of any HPC platform at LANL for open research.

Reaction rates between various chemical reactions in our network can differ by up to 5 orders of magnitude. To avoid instability in the solver, equation (1) must be solved *implicitly.* This means we need to solve a linear system that requires a 26×26 matrix inversion in each cell. This needs to occur at least *three* times every cycle[1].

Often, however, the network must be sub-cycled to avoid the growth of unacceptable errors. If the network fails a convergence criterion after a matrix inversion, it is then inverted twice over half the period of time, after which a check is performed again. The code is allowed to subdivide Δt one hundred times for every time step taken by Enzo to achieve greater accuracy. The molecular chemistry calculation by WaterNet is far and away our performance bottleneck, consuming up to three or four times the amount of time as all the physics in Enzo and Grackle combined.

2.4. Visualization Aiding the Science and Simulation

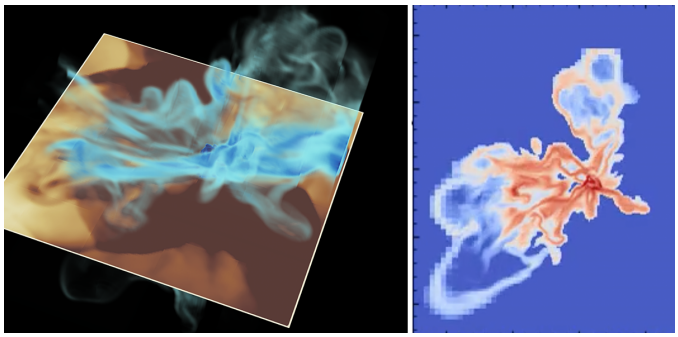

Figure 5: *Left:* Water density volume rendering with particle density slice plane. *Right:* yt water density plot. 3-D visualizations reveal dramatic multi-dimensional asymmetries in the predicted distribution of newly-formed water in the aftermath of the supernova, evidence of water production's complex dependence on availability of precursor molecules and gas temperatures and densities.

Visualizations in computational cosmology are generally 2D slices as seen in the right-hand panel of Figure 5. This has a number of limitations; volume-rendering inside of yt is cumbersome, multiple variables cannot be plotted simultaneously, and data cannot be plotted interactively. ParaView allows us to create three dimensional images with multiple variables. It also allows data to be plotted interactively, allowing us to explore data quickly and efficiently. Using yt for these simulations, a single slice plot of a single time step took 13 minutes to create. Given the size of the dataset and the time required for even a single plot, the vast majority of the data would go uninvestigated with yt. Paraview allows us to explore the data, find features which are of interest, and generate high quality images in far less time than yt.

Because of the complex interactions between multiple variables, it is essential that we be able to see how those variable interact. The most intuitive way of doing this is to plot them simultaneously, so that the relationships can be clearly seen rather than requiring looking between multiple plots.

Visualization of Enzo data in ParaView requires that the data be converted from the $\sim$ 10 GB time steps in native Enzo format to vti image data, which is smaller and easier to visualize. This conversion was also run on LANL HPC systems, taking $\sim 1,000$ CPU hours, making what would have been an intensive process locally into a trivial process with HPC systems. Likewise, the ParaView visualizations were run on HPC systems, allowing us to create images from hundreds of time steps quickly.

2.5. Verification and Validation via Visualization

Visualization allows results to be verified throughout the process of writing and testing code. In adding the molecular chemistry to Enzo, it is essential to ensure that the chemistry is integrating properly into the code. By comparing visualizations with and without the molecular chemistry, we can see how the chemistry is affecting the behavior of the code. It is also essential to ensure that the molecular chemistry matches with the theoretical understanding. Visualizations allowed us to discover and fix a problem with the molecular chemistry which allowed water to form in areas where UV radiation was too intense for water formation, indicating a problem in the code. The ability to visualize interactively and with multiple variables simultaneously makes this process even easier because the relationships between variables are readily apparent and a full three dimensional image can be examined in a few minutes, rather than the hours it would take to generate the relevant slices in yt.

3. The Formation of Water

Supernova shock waves distribute heavy elements including carbon and oxygen. They also compress and heat surrounding gas which spurs reactions. Water formation is expected to form around shock fronts (Bialy et al. 2015). But the relationship between a supernova and its environment is complex, requiring the aide of full 3D, interactive visualization to truly understand.

Because of the complexity of the scenario and the chemical reactions which result, we first simulate a star born in isolation to gain intuition for the phenomenon.

[1]The reason our matrix is inverted three times in a given time step per cell is motivated by checking convergence. Results from a single matrix solve over the entire time step are compared with two matrix solves over half the time step. If the difference in solutions is about some tolerance, additional solves are carried out

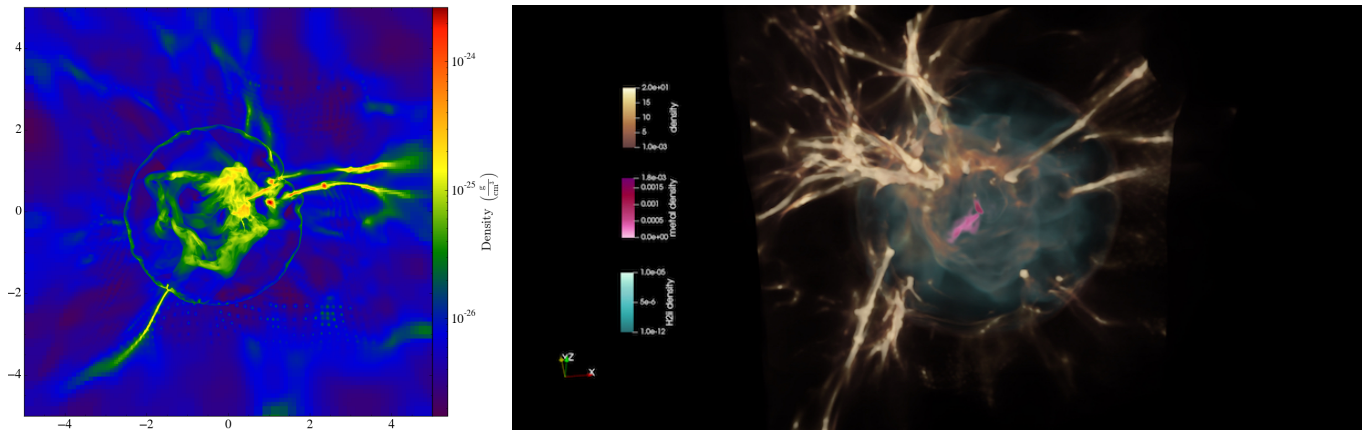

Figure 6: Side-by-side comparison of a 2D slice in a single variable, density (left), and a carefully colored 3D volume renders (right) reveals complexity and drama in cosmological supernova feedback which is completely hidden in the 2D approach. The full 3D volume render conclusively reveals the destruction of the entire halo and multiple shock fronts from previous supernovae, features not obvious from a 2D slice. Plotting density, metal density and H_2II density together further shows the causal relationships between variables, in this case the creation of H_2II and ejection of metals within the shock wave of an expanding supernova.

3.1. The First Star

We simulate a star 40 times the mass of the Sun forming in relative isolation a few 100 million years after the Big Bang. This is a "first star", born from gas made of only hydrogen and helium. This massive star only lives a few million years before exploding as an energetic supernova, showering heavy elements, including carbon and oxygen, into the surrounding medium.

The composition of the supernova ejecta are originally composed of only carbon and oxygen and other heavy elements. But as the supernova expands and cools, meaningful structure in chemical abundances is apparent in the supernova remnant in full 3D volume-rendered visualizations which helps scientist understand the chemical processes taking place. Different elements appear in various regions depending upon local gas properties like temperature and density. The shock wave *first* ionizes the molecular hydrogen and heating the gas, which reacts with oxygen to form hydroxyl (OH). Later, very small amounts of H_2O are created within compressed gas in the supernova shock fronts as the hydroxyl reacts with ambient hydrogen. Carbon also reacts with the hydrogen to form trace methane along shock boundaries.

3D visualizations reveal the extent of the complexity and turbulent structure of diffused hot gas in the wake of the supernova shock hinted at in traditional 2D plots (see Figure 7). The supernova shock wave heats and compresses surrounding gas, a complicated interaction which develops strongly asymmetric features. This heat energy later spurs the creation of molecules.

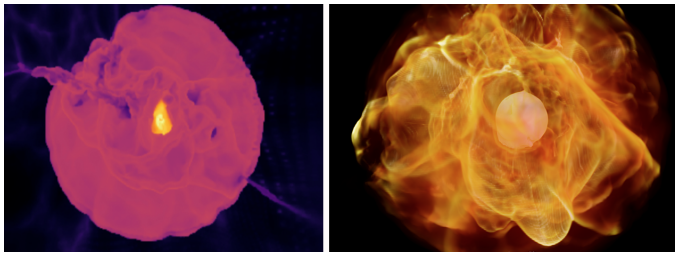

Figure 7: *Left:* temperature as rendered in a 2D slice plot in `yt`. *Right:* interactive 3D volume renderings of temperature in the wake of the supernova shock wave reveal this seething, turbulent environment in unprecedented detail.

Water is not created in appreciable abundances from enrichment by supernova in isolation. The shock wave becomes ever less dense, allowing ultra violet light for neighbor stars to break water into hydroxyl and hydrogen.

3.2. A 10 million solar mass halo

Does the story of water formation change when supernovae no longer take place in isolation? We simulated a large cosmological box at low resolution to find regions in our simulated cosmos where a halo with the mass of around 10 million Suns formed. We then restarted the simulation from the beginning with a smaller, high resolution box centered on the location of such a halo. The halo emerges hundreds of millions of years later as a merger of several large filaments. 3-D visualization reveals the complex branching structure of neighboring filaments.

In this massive halo, stars are born with such rapidity that the halo is destroyed by a series of supernova explosions which expel their heavy elements (termed *metals* by astronomers) into cosmological voids. 3D visualization reveals the extent the destruction the halo, a state difficult to discern from 2D slices. The turbulent structure of ionized H_2II behind the shock front reveal a history of multiple supernova explosions. Much of the oxygen and carbon produced by the star will not go on to make molecules because supernova explosions drive most of the gas from the halo. Water formation requires *dense* regions of gas to form.

3D visualizations shows how shock waves directly influence the chemistry in cosmological halos. As shock waves propagate outwards H_2, a molecule abundant within the halo, is immediately heated and ionized. Detailed visualization reveals the complex turbulent gas structure behind the shock which might serve to mix molecules are spur reactions. Because blast waves in our simulations start out as spheres, any anisotropy

in the supernova remnant comes from hydrodynamic interactions with the cosmic web. These insights give LANL scientists clues that an early-universe supernova's placement in the cosmic web strongly affects the turbulent mixing of heavy elements and chemical compounds in the wake of supernova shocks.

3.3. A 100 million solar mass halo

The story changes again with a 100 million solar mass halo. In these scenarios, the gravity from the halo is so strong the supernovae cannot expel the gas though they appear with equal frequency.

Appreciable water formation directly after a supernova would require a scenario in which a supernova takes place in proximity to very dense region of gas. A fortuitously placed supernova could compress and heat this gas to temperatures conducive to water formation and deposit the ingredients of water. On the left side of Figure 5 water density is rendered in blue while the gas density is shown on the brown slicing plane enabling scientists to observe the relationships between variables as the simulation unfolds. On the right is the yt plot of water density, the means used to observe formations before we developed the 3D visualizations.

Finding such a halo requires one to be able to fly through and manipulate data in 3D, a prospect impossible with current cosmology visualization tools which require hours to visualize a single volume-rendered frame. Additionally, the optimal conditions for water formation are a confluence of various factors and volume rendering multiple variables simulataneously allow us to find regions of high gas temperature and density quickly. Navigating the complex 3D filamentary structure of the halo with 2D slice plots would be time prohibitive. Parallel, real-time visualization carried out on LANL HPC platform Grizzly permits us to find optimal regions for water formation in our hypothetical halo. Clumps of gas likely to form a second generation of heavy-element enriched stars are the most likely locations for abundant water in the early universe.

Acknowledgments

This work was funded by the National Nuclear Security Administration (NNSA) Advanced Simulation and Computing (ASC) Program. We particularly want to acknowledge LANL ASC Production Visualization Project. We thank John Patchett, Los Alamos National Laboratory and Annie Bares, University of Texas at Austin. We thank Steven Meredith, Southern Utah University, for narrating our video submission. Calculations were carried out on Turquoise computing platforms at LANL, including Wolf, Pinto and Grizzly on Institutional Computing (IC) Allocations. LA-UR-18-27663.

References

Bialy, S., & Sternberg, A. 2015, Monthly Notices of the Royal Astronomical Society, 450, 4424

Bialy, S., Sternberg, A., Loeb, A. 2015, Astrophysical Journal Letters, 804, L29

Bryan, G. Norman, M., O'Shea B. et al. 2014, Astrophysical Journal Supplements, 211, 19

Chatzopoulos, E., van Rossum, D. R., Craig, W. J., Whalen, D., Smidt, J., Wiggins, B. 2015, Astrophysical Journal, 799, 18

Cleeves, L. I., Bergin, E. A., Alexander, C. M., Du, F., Graniger, D., Oberg, K. I., Harries, T. J. 2014, Science, 345, 1590

Hopkins, P. 2015. Monthly Notices of the Royal Astronomical Society, 450, 53

Imprellizzeri, V., Mckean, J. P., Castangia, P., Roy, A. L., Henkel, C., Brunthaler, A., Wucknitz, O. 2008. Nature, 456, 927

Li, X. 2005, Toms, 31, 302

Mesler, R. A., Whalen, D. J., Smidt, J., Fryer, C. L., Lloyd-Ronning, N. M., Pihlstrom, Y. M., 2014, 787, 91

Omukai, K., Tsuribe, T., Schneider, R., Ferrara, A. 2005, Astrophysical Journal, 626, 627

Smidt, J., Whalen, D. J., Johnson, J. L., Surace, M., Li, H., 2018, Astrophysical Journal, 865, 126

Smidt, J., Wiggins, B. K., Johnson, J. L., 2016, Astrophysical Journal Letters, 829, L6

Smidt, J., Whalen, D., Wiggins, B. K., Even, W., Johnson, J. L., Fryer, C. L., 2014, Astrophysical Journal, 797, 97

Smith, B. D., Bryan, G. L., Glover, S. C. O. et al. 2017, Monthly Notices of the Royal Astronomical Society, 466, 2217

Visser, R., van Dishoeck, E. F., & Black, J. H. 2009, Astronomy and Astrophysics, 503, 323

Wiggins, B. K., Fryer, C. L., Smidt, J. M., Hartmann, D., Lloyd-Ronning, N., Belcynski, C., 2018, Astrophysical Journal, 865, 27

Wolcott-Green, J., Haiman, Z., Bryan, G. L. 2011. Monthly Notices of the Royal Astronomical Society, 418, 838