

Analysis of a Controlled Approximation for Explicit Integrations of Stiff Thermonuclear Networks

A Thesis Presented for the
Master of Science
Degree

The University of Tennessee, Knoxville

Nick Brey

May 2023

© by Nick Brey, 2023
All Rights Reserved.

Dedicated to Ed Roeske

Acknowledgments

I would like to thank Dr. Guidry for supporting me on this project, as well as Dr. Hix, Dr. Endeve and Dr. Billings for helping with the project and analysis and being on the committee. I would also like to thank the other students who worked on this project and have helped with development of the code: Adam Cole, Ashton DeRousee, Raghav Chari, Max Middlecoff and Jacob Gouge.

Additionally, I would like to thank the University of Tennessee and the Department of Physics and Astronomy for their support and funding.

Abstract

The current standard method to solve stiff coupled differential equations relies on implicit integration methods. Explicit methods are generally avoided due to the extremely small and limiting timesteps they allow when the equations are stiff. However, implicit methods are computationally expensive because of the complex calculations that need to be done at each time step. An explicit integration method can do these calculations quicker and, if allowed to take comparable timesteps to the implicit ones, would allow the entire calculation to be done faster. Previous work by Dr. Guidry, Dr. Endeve, Dr. Hix and Dr. Billings has shown that, in principle, explicit integration can take larger timesteps than normally allowed when certain approximations are used. The speed up in the calculations from implementing algebraic approximations comes at the expense of the accuracy. However, unlike other approximations typically introduced for coupling networks to hydrodynamical simulations, these approximations can be controlled by the user and allow for a quantifiable restraint on the error. The concept of a controlled approximation is introduced by providing a quantifiable way to show the trade off of accuracy for speed when using algebraic approximations for explicit integrations of stiff thermonuclear reaction networks coupled to fluid dynamics.

Table of Contents

1	Introduction and Background	1
1.1	Introduction	1
1.2	Background of Supernovae	2
2	Astrophysical Applications	5
2.0.1	Reaction Networks	7
2.1	Computational Methods	8
2.1.1	Introduction to Stiffness	10
3	Controlled Approximations	16
4	Algebraic Stabilization Using Approximation Methods	19
4.1	The Asymptotic Method	21
4.2	Partial Equilibrium	23
4.3	Adaptive Time Stepping	28
5	Results of the New C++ Code	33
5.1	Triple Alpha Network	35
5.2	CNO Cycle	36
5.2.1	CNO Main	39
5.2.2	CNO-Extended	39
5.3	Alpha Network	43
5.3.1	Constant T and ρ	43
5.3.2	Hydrodynamical Profile	46

5.4	Nova 134-Isotope	48
6	Analysis of the Approximations	53
6.1	Residuals and Uncertainties	53
6.2	Triple Alpha Analysis	55
6.3	Alpha Network Analysis	59
6.3.1	Constant Temperature and Density	59
6.3.2	Alpha Network with Hydro Profile	62
7	Conclusion	68
	Vita	73

List of Tables

2.1	Some networks used in testing the new FERN code	9
6.1	Triple Alpha Residuals	58
6.2	Alpha Network Residuals	60
6.3	Alpha Network Residuals with Hydro Profile	64

List of Figures

1.1	Periodic table showing elemental origins	4
2.1	Proton-Proton Chain	9
2.2	Isotopes and Networks	13
2.3	Forward Euler vs Backward Euler dt	14
2.4	CNO cycle	15
4.1	Triple alpha with Asy only	24
4.2	Triple Alpha with Asy+PE	29
5.1	Triple alpha summary	37
5.2	CNO main summary	40
5.3	CNO extended summary	41
5.4	CNO extended mass fraction comparison	42
5.5	Alpha summary without hydro profile	45
5.6	hydro profile used for alpha calculations	47
5.7	Alpha summary with hydro profile	49
5.8	nova-134 Hydrodynamical profile	50
5.9	nova-134 summary with hydro profile	52
6.1	The trade off in accuracy versus speed for the triple alpha calculations. . . .	57
6.2	Triple alpha mean mass fractions shown with error bars indicated by the dots.	58
6.3	Accuracy and speed trade off for the alpha network at constant T_9 and ρ . .	61
6.4	Alpha network mean mass fractions	63

6.5	Accuracy and speed trade off for the alpha network with the hydro profile of figure 5.6 giving a varying T_9 and ρ	65
6.6	Alpha network mean mass fractions with hydro profile	67

Chapter 1

Introduction and Background

1.1 Introduction

The problem of solving large, coupled reaction networks is important for many fields. The differential equations that must be solved involve timescales that may vary over orders of magnitude, causing the equations to be very stiff. A new C++ code has been developed in conjunction with this thesis work that can solve problems of this nature with a few modifications depending on the field/problem. This code is based on earlier development in references [1-6]. Our primary interest is solving these types of equations for thermonuclear reaction networks that are used in astrophysics when simulating stellar explosions, such as supernovae.

Supernovae provide insight into many different aspects of astrophysics. They play a crucial role in galactic evolution and the elemental abundances in galaxies. Additionally, some gamma-ray bursts may be linked to core collapse supernovae. Type Ia supernovae, in particular, are of great interest because they can be brighter than entire galaxies and their intrinsic brightness can be calculated from observations to yield dependable distance measurements.

Large-scale numerical simulations of supernovae are necessary to further understand these stellar explosions. The simulations of astrophysical processes require a vast amount of physics that needs to be taken in to account. The thermonuclear reaction network is just one piece of a hydrodynamical simulation for stellar explosions. Many nuclear

reactions take place in a small amount of time compared to a typical hydrodynamical time step. As the hydrodynamical evolution progresses, it uses the change in energy and isotopic abundances calculated in the thermonuclear reactions to update the state of the system. The thermonuclear calculations of the change in energy and isotopic abundance in a hydrodynamical step must be calculated quickly so as not to slow down the hydrodynamical calculation. Currently the best way to solve the differential equations describing the thermonuclear reactions uses implicit methods. The new C++ code looks to use explicit methods with some approximations in order to take advantage of the simpler solution they provide. This code can be implemented in present simulations to improve the speed/cost and with a few adaptations it could also be used for calculations/simulations in other fields.

1.2 Background of Supernovae

There are two different classifications of supernovae, Type I and Type II, based on spectra and light curves. Type I supernovae show no hydrogen in their spectra, while type II do show hydrogen. The sub classes of Type I depend on whether certain other elements are present or missing in the spectra. Type Ia shows no signs of helium in the spectra but shows a sharp silicon line as well as strong lines in the iron group of isotopes after they dim. Types Ib and Ic differ from Type Ia in their spectra and in how they explode. It is thought that Type Ib and Ic explode via core collapse, which is also how the type II supernovae explode. While neither show any hydrogen, Type Ib shows strong helium lines and Type Ic shows no helium or silicon. In the case of the Type Ib and Ic, the progenitors have likely lost their hydrogen from stellar winds or binary interactions before the collapse, leading to the absence of hydrogen in their spectra.

The different underlying physics in each type of supernovae can explain the differences in their spectra. The mechanisms for the explosion, and whether or not a supernova is Type Ia or Type Ib, Ic or II has to do with the initial mass of the progenitor. The lower mass stars ($0.6-7 M_{\odot}$) end their lives as white dwarves, anything less than $0.6 M_{\odot}$ will end up as a helium white dwarf. The white dwarf is the remnant core of a star that shed its outer layers, and is typically made up of carbon and oxygen. White dwarves have a mass in the range of $0.6 -$

$1.4 M_{\odot}$, the upper limit of the mass is known as the Chandrasekhar mass. Any star under the Chandrasekhar mass is supported against gravitational collapse by electron degeneracy pressure. Stars that end their lives with a core mass larger than the Chandrasekhar mass undergo further collapse. These stars lead to the Type Ib, Ic and type II explosions, and typically have an initial mass $> 8M_{\odot}$.

Supernovae are extremely important in the evolution of galaxies, affecting the stellar formation rate and producing/dispersing elements heavier than helium throughout the universe. Many elements heavier than carbon or oxygen up to the iron group of isotopes (Iron, Nickel, Copper, Zinc etc.) were fused in a massive star and spread through the interstellar medium via stellar winds and explosions. Core collapse supernovae are also responsible for creating elements heavier than iron through the r-process. The other possible source for the heavier, neutron-rich elements is thought to be neutron star mergers. The periodic table shown in figure 1.1, surveys the processes thought to be the origin of each element. Supernovae are also important feedback mechanisms for star formation in galaxies. In addition to spreading heavy elements into the interstellar medium, supernovae can give rise to the next generation of stars within a galaxy because the explosion can perturb nearby gas clouds and initiate their collapse. The collapsing gas clouds create proto-stars and start the cycle over again.

Supernovae have become increasingly important in the field of cosmology, especially Type Ia supernovae. Type Ia supernovae are an essential part of determining the expansion rate of the universe and exploring dark energy. The light curves of Type Ia supernovae have similar features, even though the peak luminosities may vary from case to case. The similarities in the light curves allow Type Ia explosions to be used as standardizable candles. The peak luminosity of a supernova is related to the rate of decline in the luminosity. Knowing the rate of decline in luminosity can yield the peak luminosity of the supernovae, which, in turn can give the distance. The supernovae reside in galaxies whose distance can now be found directly and with the additional knowledge of the redshift, the expansion rate of the universe can be determined. Gaining a better understanding of supernovae is vital to better understanding our universe.

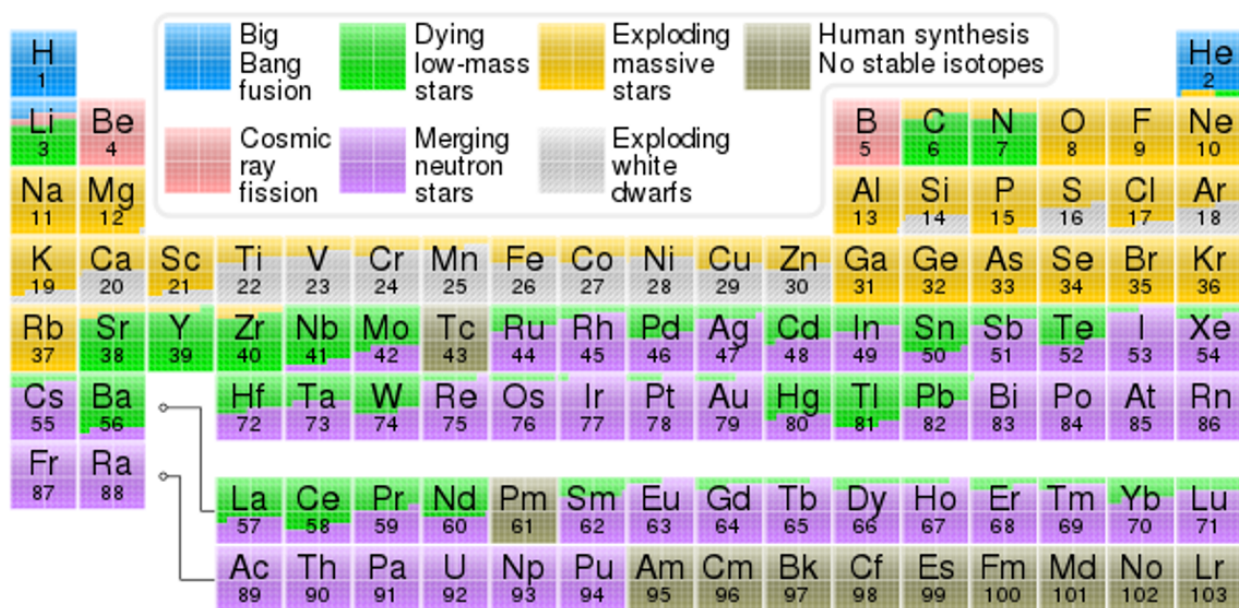


Figure 1.1: This shows the processes that are thought presently to have created the known elements of the periodic table. There is still considerable uncertainty in source details, particularly for the heavier elements, Jennifer Johnson (Ohio State University) [7]

Chapter 2

Astrophysical Applications

An important component of simulating stellar explosions such as Type Ia supernovae is evolution of the thermonuclear network of reactions that take place in the explosions. The abundances of the isotopes that are present at different times depends on the rate at which the nuclear reactions occur. The isotopes present in one reaction may also be present in others, adding to the complexity of the equations that describe the process. The abundance of a number of isotopes is characterized by an equation in the form of:

$$\frac{dY_i}{dt} = F_i^+(\mathbf{Y}) - F_i^-(\mathbf{Y}) = (f_1^+ + f_2^+ + \dots)_i - (f_1^- + f_2^- + \dots)_i = \sum_j (f_j^+ - f_j^-)_i. \quad (2.1)$$

The term Y_i is the abundance of an isotope indexed by i . The change in the abundance of an isotope is related to the flux of that isotope, F . The flux, either F_i^+ or F_i^- , is determined by how much the abundance of an isotope increases or decreases across the network, respectively. In large reaction networks, each isotope may take part in many reactions; some deplete the isotope and some produce it. The reactions where an isotope is depleted are described by the F_i^- term and the reactions that produce the isotope are described by the F_i^+ term. The coupling can be complex in large networks where each isotope takes part in several reactions and therefore, has several factors affecting the F_i^+ or F_i^- terms. The f_j^+ and f_j^- terms represent the isotopes in a given reaction group, rather than the network. The reaction groups are pairs of forward and reverse reactions that can come into equilibrium, more will be discussed on reaction groups later.

The new C++ code, named FERN (Fast Explicit Reaction Network), simulates the nuclear reactions that are present in stellar explosions that produce the Type Ia supernovae. Initially, the system is made up of primarily carbon and oxygen, as that is the makeup of a typical white dwarf. As time progresses, the elemental makeup of the system changes based on the reactions that are taking place. The abundance of an isotope depends on the reactions that contain that specific isotope. Whether or not the flux is positive or negative would partially depend on the rates of those reactions. The reaction rates play an important role in limiting the time step for explicit integration methods as well. The code reads in known rate parameters from a data file for a specific calculation and computes the reaction rates for all the reactions in the network. The rates can be found using Equation (2.2) which involves the temperature and some known constants. This is a fit formula using the REACLIB format [8],

$$R_k = \exp(p_1 + \frac{p_2}{T_9} + \frac{p_3}{T_9^{1/3}} + p_4 T_9^{1/3} + p_5 T_9 + p_6 T_9^{5/3} + p_7 \ln(T_9)). \quad (2.2)$$

The symbol T_9 is the temperature in units of a billion Kelvin, and the $p_i (i \in 1 - 7)$ values are the known constants that vary depending on the reaction. The initial state of the system is the isotopic makeup of the white dwarf, typically Carbon and Oxygen. As time progresses and the reactions create and deplete isotopes, more and more isotopes start to have non-zero abundance values and energy is released (or absorbed). When coupled to a hydrodynamical simulation, the thermonuclear calculation would be done over the interval of a hydrodynamical time step. A typical hydrodynamical time step, with certain assumptions, is on the order of 1×10^{-4} s for the simulation of Type Ia supernovae. After integration over the hydrodynamical time step, the thermonuclear code would return the abundances of each isotope and change in energy back to the hydrodynamic simulation. These rates can be related to the fluxes used in Equation (2.1) by

$$\frac{dY_i}{dt} = \sum_j \eta_j^i \lambda^i Y_j + \sum_{jk} \eta_{jk}^i \rho N_A \langle jk \rangle Y_j Y_k + \sum_{jkl} \eta_{jkl}^i \rho^2 N_A^2 \langle jkl \rangle Y_j Y_k Y_l. \quad (2.3)$$

The three terms represent the one, two and three body reactions that take place. These reactions lead to the changing abundances, $\frac{dY_i}{dt}$, or fluxes. The η terms are factors and constants that determine the sign of the term within the sum. The λ term represents the one-body decay rate, while the $\langle jk \rangle$ term is velocity-averaged two-body reaction rate, and $\langle jkl \rangle$ is the velocity-averaged three-body reaction rate. The abundance of the other species involved in a reaction are given in terms of either the Y_j, Y_k or Y_l terms, and N_A is Avogadro's number (Guidry, Stellar Processes).

2.0.1 Reaction Networks

There are a number of reaction networks that can be used in a calculation. The network sizes used to test the new code vary from as small as 3 isotope and 8 reactions to as many as 365 isotopes and 4395 reactions. The largest network contains all the isotopes and reactions common in a Type Ia supernova. The smaller networks are still significant and are subsets of reactions that are present in supernovae, but on their own may be insufficient for simulating Type Ia explosions. The smallest network, the triple alpha network, contains only isotopes of ${}^4\text{He}$, ${}^{12}\text{C}$, and ${}^{16}\text{O}$. The reactions in this network involve the creation of ${}^{12}\text{C}$ by fusing three ${}^4\text{He}$ atoms and the creation of ${}^{16}\text{O}$ by fusing a ${}^{12}\text{C}$ and ${}^4\text{He}$ atom, as well as the reverse reactions. That reaction is known as the triple-alpha process and is a key reaction that involves the formation (and depletion) of ${}^{12}\text{C}$ and ${}^{16}\text{O}$. This type of reaction is seen in red giant stars, where it converts a helium-dominated core into a carbon and oxygen white dwarf. Another small, yet important, network is the proton-proton chain. This reaction network is a multi-step process that converts protons (hydrogen ions) into a helium nucleus. This type of reaction powers stars like the Sun while in the main sequence phase of their lives.

A third network that is commonly used is the full alpha, which is a more complex and realistic network than the previous two. The alpha network spans the range of isotopes that get created in a Type Ia supernovae. These smaller networks are not very realistic since they do not contain all the isotopes and reactions that are present in supernovae, but they can be useful when having to make a necessary approximation for use in a hydrodynamical simulation. Two of the larger networks are the 150-isotope network and the 365-isotope

network. The 150 network is the minimally realistic network, as it allows for the abundances of many isotopes, and the 365 network is a more accurate picture of all the isotopes that get populated in the process. Figure (2.2) shows the isotopes that are populated in the full alpha, 150-isotope and 365-isotope networks. Various network used for testing the FERN code are shown in table 2.1.

2.1 Computational Methods

Typical calculations like the simulation of a Type Ia supernovae require large amounts of computational power. Two possible ways to run the calculations are based on different methods for solving the network of differential equations in Equation (2.1) at a specific time. One possibility is to use an explicit method for the integration, for example, a forward Euler method Equation (2.3) [9],

$$y_{n+1} = y_n + \Delta t \cdot f(y_n, t_n). \quad (2.4)$$

The terms y_n and y_{n+1} denote the current solution and the solution at the next time, respectively. The function, f depends on the current time t_n and the solution at this current time (y_n). Another possibility is an implicit method, for example, the backward Euler method (Equation 2.4). At a certain timestep, t_n , the solution at that time is given by y_n . The solution at the next time (t_{n+1}) can be described by Equation (2.3). This is the way the forward Euler method works, it uses only already-known quantities (y_n , and $f(y_n, t_n)$). While this method provides an easy way to find the solution, y_n , the time stepping must be restricted to maintain stability. The limiting time step is determined by the fastest reaction in the network. The largest time step expected in explicit integrations for this case is set by the timescale of the reaction with the fastest rate. Any time steps larger than the inverse of the fastest reaction rate would lead to instability.

On the other hand, implicit methods can take much larger time steps because the backward Euler method does not introduce the instability that the forward Euler does. However, there is a tradeoff; while the time step is not restricted by stability, the solution is more complicated. The new solution in equation 2.4 [9] has the y_{n+1} appear on both sides of the equation. The solution at the next time step depends on values that are not yet known,

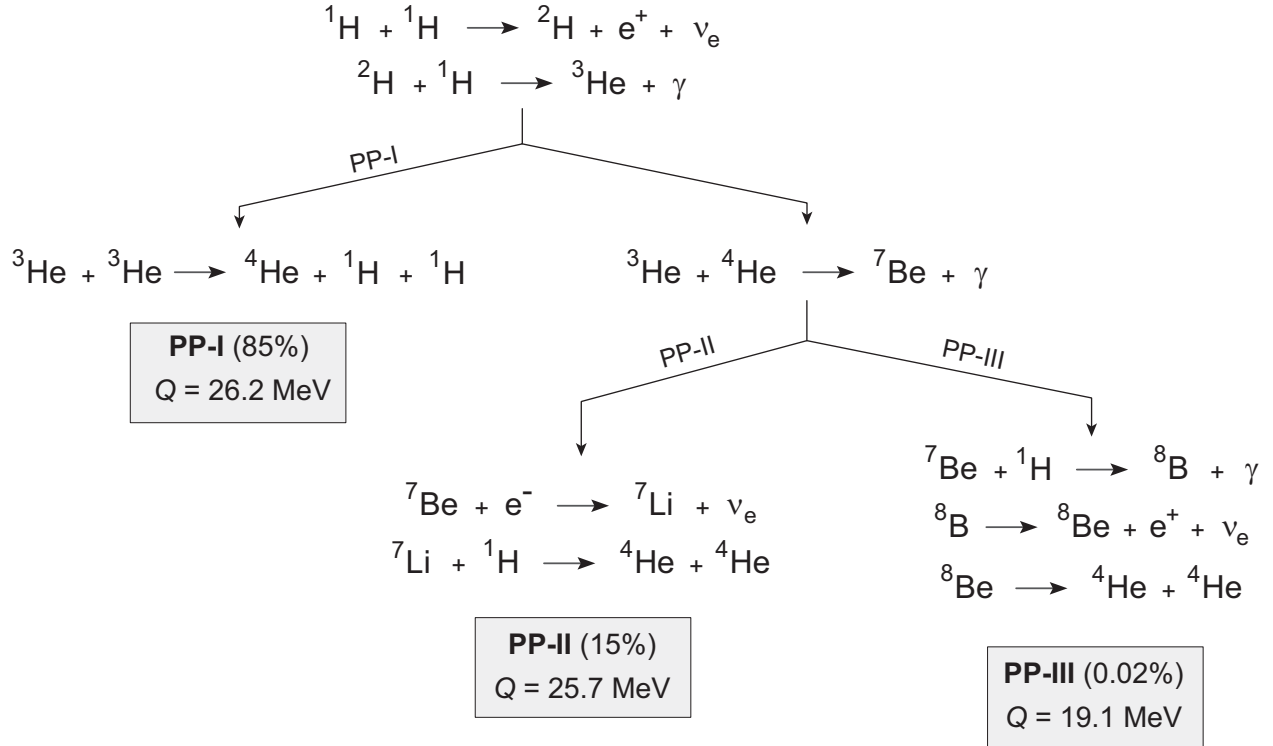


Figure 2.1: Reactions in the proton-proton chain. (Guidry, Stellar Processes)[10]

Table 2.1: Some networks used in testing the new FERN code

Network	No. of Isotopes	No. of Reactions	No. of Reaction Groups
3-Alpha	3	8	2
P-P Chain	7	28	13
Full Alpha	16	48	19
Extended CNO	16	134	42
48-Iso	48	299	132
Nova134	134	1566	693
150-Isotope	150	1604	742
268-Isotope	268	3175	1564
365-Isotope	365	4395	2165

including the solution at the next time step. The solution to the implicit method requires a solution of a matrix equation which leads to this methods being more computationally expensive for larger matrices (networks). Despite this, it can still take much larger time steps than a purely explicit method.

$$y_{n+1} = y_n + \Delta t \cdot f(y_{n+1}, t_{n+1}) \quad (2.5)$$

The difference in Δt between the forward and backward Euler methods over time (t) is shown schematically in Figure (2.3). The restrictions on the time step for the forward Euler method is caused by the instability introduced with any larger time steps. The backward Euler method does not have that limitation and continues to grow with time. When the timescale characterizations and rates for the nuclear reactions are orders of magnitude smaller than the timescale of the calculation, which is typically the case for stiff systems, the backward Euler method becomes much more favorable to use.

Both methods can be useful, depending on the calculation needing to be done, but for astrophysical simulations like these the implicit methods are generally used to get around the stiffness that limits the explicit calculation. The explicit time step that is stable for the CNO cycle is on the order of one hundred seconds, while the cycle is important for times up to 1×10^{16} seconds, or about 300 million years. It would not be feasible to run a simulation that can only take timesteps of hundreds of seconds when the integration is on the order of hundreds of millions of years; this would require about 1×10^{14} integration steps.

2.1.1 Introduction to Stiffness

In mathematical problems, stiffness has to do with numerical stability and needing to take very small step sizes to keep stability. Stiffness arises in this type of calculation where the timescales of some reactions differ from the timescales of other reactions or even the calculation as a whole by at least several orders of magnitude. The systems we are solving are considered stiff because of the need to take very small explicit time steps to maintain correct abundances, while the timescale of the calculation is many orders of magnitude greater.

The implicit methods do not have this issue of stiffness; the time steps are not limited by the instability when solving the differential equations. However, the explicit methods are able to calculate a single time step faster than implicit methods due to the simplicity of the forward Euler equation. The heavy computation used in the implicit method is generally overcome by its large time step capability, and therefore, is preferred over explicit methods. In the CNO network mentioned earlier, an implicit method could solve the problem with a few hundred time steps compared to the 1×10^{14} steps needed for the explicit method. With some approximations applied to the system to remove stiffness, the explicit integration methods have been shown to take time steps that compare to those taken by implicit methods [1]. Having comparable time stepping between the implicit and approximated explicit would mean that the explicit method would then be able to do the calculation as a whole faster, as it is able to calculate each integration step faster. In order to allow the explicit method to take larger stable time steps without compromising stability, there are certain approximations that need to be implemented to reduce the stiffness of the network. To get a better understanding how to do this, we have to look at the sources of stiffness in the calculation. The stiffness in a calculation is a property of the differential equation like the one shown in Equation (2.1). The sources of stiffness that arise from that equation are negative populations, macroscopic and microscopic equilibrium, and cyclic reactions.

Negative populations arise from the term on the left side of Equation (2.1). The Y_i term represents the abundance of each isotope and if this variable becomes negative, it would imply a negative population which is not physical and destabilizes the network. Macroscopic equilibration applies to the individual differential equations, or the $F_i^+ - F_i^-$ terms. The behavior of the F^+ and F^- terms are what dictate the evolution of the network. This is the domain of the asymptotic and quasi steady-state approximations. Microscopic equilibration acts on individual terms $(f_j^+ - f_j^-)_i$ in a given differential equation. The partial equilibrium approximation applies to this type of equilibration.

To get around the stiffness mentioned in the first three points, the new FERN code uses explicit methods with algebraic approximations that remove the stiffness and allow for time steps larger than are usual in explicit calculations. Removing the limiting reaction rates in a network by implementing algebraic approximations removes the stiffness from the differential

equations Equation (2.1) that describe the reaction network. Once the limiting reaction rate is removed, the time step is allowed to grow beyond that previous limitation. The new limit of the time step is equal dictated by a new limiting reaction rate. The purpose for using the approximations is to allow for time steps that are competitive with those of implicit methods.

The stiffness in point four arises from reactions that have a cyclic nature; an astrophysical example of such a case is the CNO cycle, whose reactions are seen in Figure (2.4). This cyclic behavior gives rise to stiffness that might not be removed by using the approximations that are implemented in the micro and macroscopic equilibrium regimes. The source of the stiffness for the CNO network lies in the β decays whose rates are only dependent on the nuclear processes for decay rather than on the temperature of the system. These β decays have characteristic constant time scales on the order of 100 seconds and at late times are the fastest reaction rates in the network. These decays hold up the rest of the network from being integrated explicitly with sufficient time steps and accuracy because of the cyclic nature of this network. To remove this stiffness from the network, we take advantage of the fact that a closed cycle requires conservation of probability, which allows for a relation to be derived between the rates of reactions and abundances. Rather than updating the abundances numerically with the previous approximations, the abundances of the isotopes that are resultant of β decays are found by using this relation between rates and abundances.

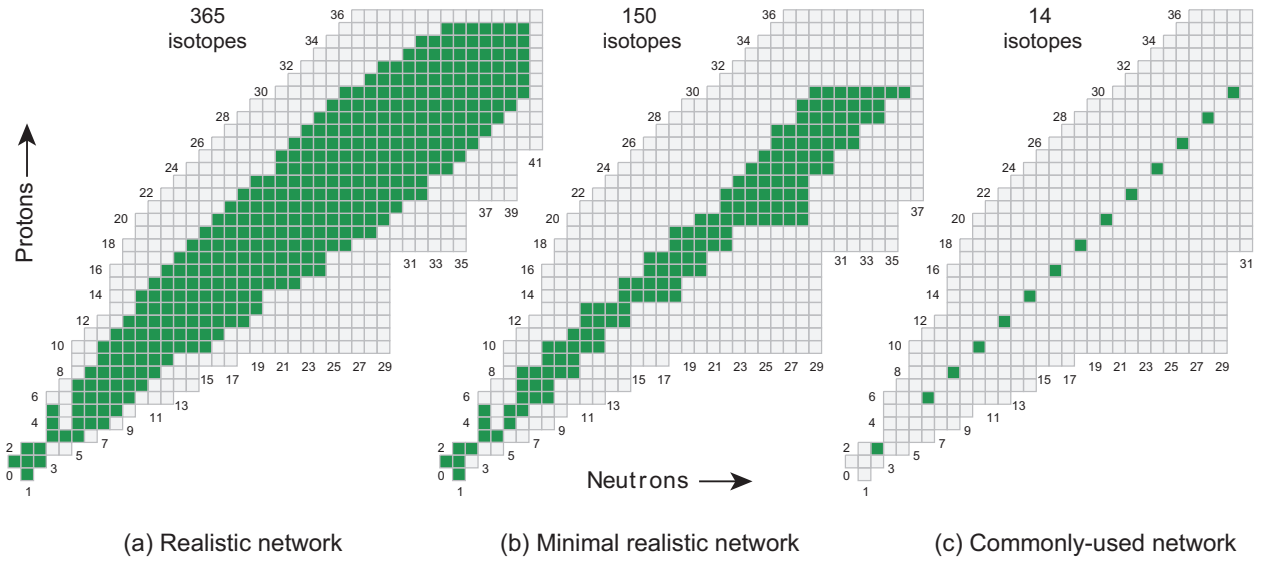


Figure 2.2: Isotopes for several networks that could be used under Type Ia supernova conditions (Guidry).

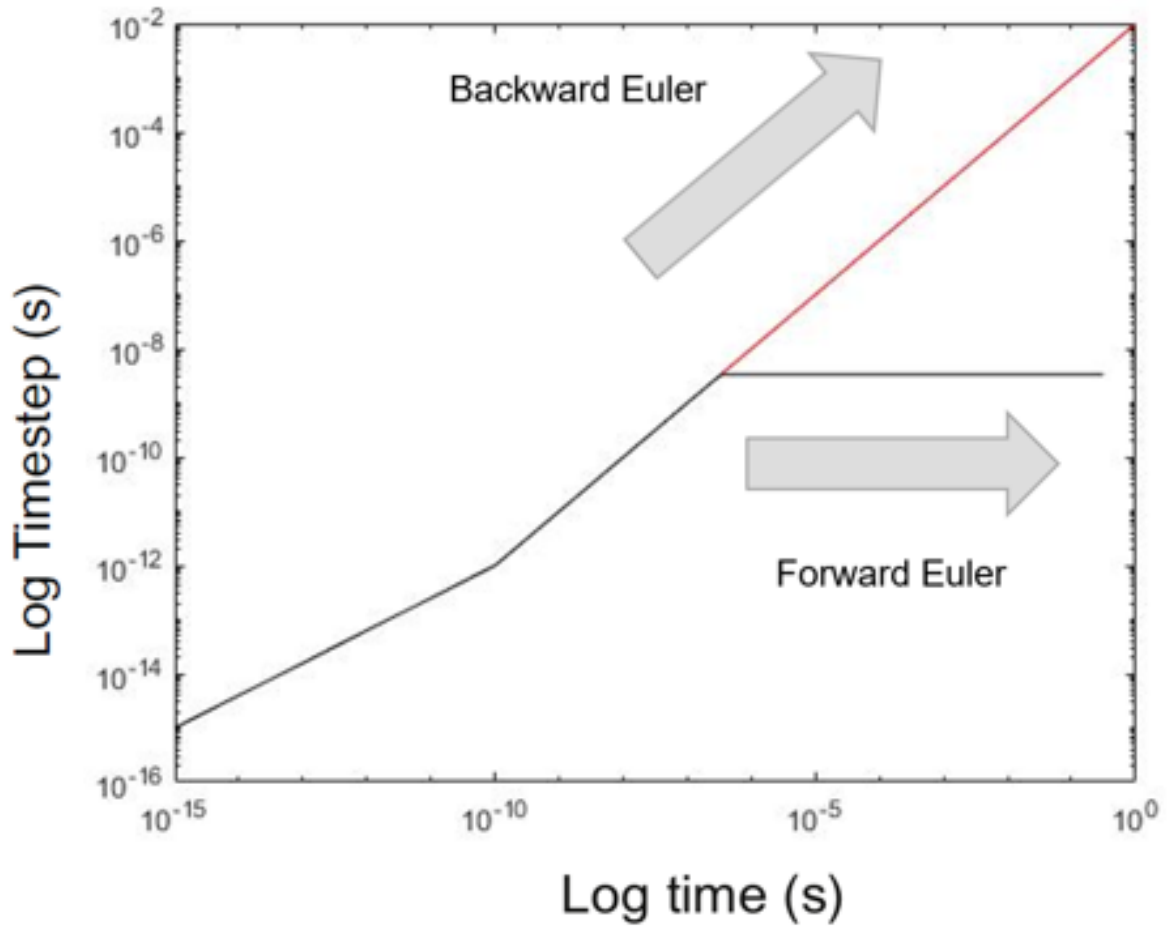


Figure 2.3: The plot shows schematically the timesteps on the y -axis and the time on the x -axis. At later times, the forward Euler method hits a regime where the stability forbids growth of the time step with time, but the backward Euler method time step continues to grow.

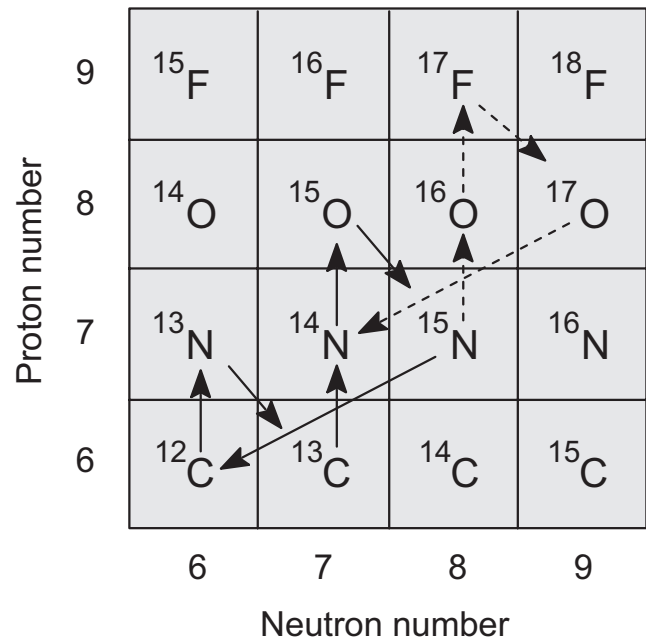
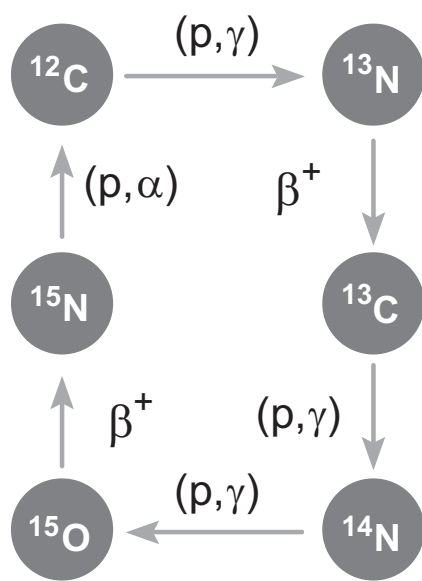


Figure 2.4: The reactions in the CNO cycle, with the β decays shown in the reaction from N13 to C13 and from O15 to N15. (Guidry, Stellar Processes)

Chapter 3

Controlled Approximations

Here we will discuss the new code as a vehicle to implement controlled approximations for use in coupling to a hydrodynamical calculation. Realistic physical systems often involve coupled environments that make the problem extremely complex. Simulations of these systems must use certain approximations to make the problem possible on computers at the present stage of hardware and software development. The simulations for Type Ia supernova are no different, one must make an approximation, but the FERN code allows the user to choose how to approximate. A vital difference between approximations is whether the errors that result from an approximation are controlled or uncontrolled.

- **Controlled Approximation:** The user has a choice of how much the parameters restrain the local error, and therefore, has control of how to limit the global error in a quantitative way.
- **Uncontrolled Approximation:** The user has a qualitative understanding of the error that is introduced by approximations, but no local quantitative control.

An overview of solving these types of large thermonuclear networks can be found in Hix and Meyer [11] and Timmes [12]. For solving the problem of extremely stiff thermonuclear networks in 3D hydrodynamical simulations of type Ia supernovae, the generally agreed-upon network that is minimally realistic is the 150-isotope network of figure 2.2. When solving large networks that couple to fluid dynamics, it has usually been necessary to truncate the thermonuclear network to that of the alpha network or smaller. This reduction in

size of the network is an example of an uncontrolled approximation, since there is no very reliable way to quantitatively estimate the error from truncating the network size locally in a given calculation. An alternative approximation is to keep the network size at $\tilde{150}$ isotopes and solve this realistic network by using controlled approximations. These controlled approximations would implement the algebraic approximations that have been shown to work in principle by Guidry [1]. Using these approximations, the user would be able to trade an acceptable amount of accuracy for speed.

In the previous chapter, it was shown that explicit methods are easier to use due to the simplicity of their solutions. The solutions only requires the knowledge of the current state of the system to compute the future state which allows the solution for a given time step to be solved much faster than by implicit methods. The implicit methods, however, are generally favored because their time stepping is not restricted by the instability of the solutions as is the case for explicit. Some implicit methods require a matrix inversion at each time step and the time required to do the matrix inversion scales poorly with network size. For larger and more realistic networks, the inversion is severely restrictive on the time used to obtain a solution. On the other hand, explicit methods do not require a matrix inversion and scale much more favorably with the size of the network. Thus, the implicit methods force approximations such as drastic reduction in the network size. This uncontrolled approximation introduces an error that is difficult to quantify because it is chosen for its ability to integrate the network implicitly more so than by its realistic physical representation of the explosion.

In typical simulations of a supernova explosion, a hydrodynamical code is coupled to thermonuclear burning and possibly neutrino transport kinetic networks. These kinetic networks may only need to be solved approximately for the full simulation if the approximations introduced can control the error. The acceptable amount of error allowed by the controlled approximations should be less than the errors and uncertainties in the whole coupled hydrodynamical-kinetic system. The thermonuclear networks and neutrino transport may be calculated later and more accurately using a hydrodynamical profile for varying the temperature and density of the system. The hydrodynamical profiles would be obtained from the more realistic simulation that has used a full network with the controlled approximations. A reasonable requirement of these approximations would restrict

the controlled error to a few percent or less for the approximated solution of the kinetic networks when coupled to the full hydrodynamical simulations. Any calculation with a strict error tolerance (much less than an order of 1% error) will slow the simulation and still not show improvement in the final results due to uncertainties that are present in the full simulation apart from the kinetic networks. Thus we propose to use approximations for explicit methods that permit larger time steps that can compete with implicit time steps, while providing a solution of large realistic networks that have an acceptable and controlled error.

Chapter 4

Algebraic Stabilization Using Approximation Methods

The drawback of standard explicit calculations is the extremely small time steps required to maintain stability in the network. FERN uses a few different methods to stabilize the network algebraically. The stabilization allows for larger time steps than purely explicit calculations could take. The goal is to be able to take time steps that are comparable with implicit ones in order to perform the whole calculation faster. There are three approximations involved, the asymptotic (ASY) developed by Guidry et al [3], quasi steady-state (QSS) developed by Guidry and Harris [4] with modifications made from Mott et al. [13-15]. The last method gets paired with either the ASY or the QSS and that is the partial equilibrium (PE) developed by Guidry et al.[2] and further elucidated by Feger [16]. The code can either use the ASY or the QSS individually and the PE can be paired with either one. This gives four possible methods to use. The pairing of ASY+PE and QSS+PE are the ones that tend to show the most improvement, but in some cases the ASY or QSS alone are good enough. The ASY and QSS approximations are used in macroscopic equilibrium, while the PE applies in microscopic equilibrium. The different combinations that may be used in a given calculation are QSS only, ASY only, QS+PE, or ASY+PE. The focus of this thesis is on the use of the asymptotic approximation and the pairing of asymptotic and partial equilibrium. Looking back at Equation (2.1), there are two different ways to group the flux terms on the right

hand side to show the micro and macroscopic terms. Equations (4.1) and (4.2) show the two different ways the terms are grouped:

$$\frac{dY_i}{dt} = F_i^+(\mathbf{Y}) - F_i^-(\mathbf{Y}) \quad (4.1)$$

$$\frac{dY_i}{dt} = \sum_j (f_j^+ - f_j^-)_i. \quad (4.2)$$

Equation (4.1) describes the fluxes in the regime of macroscopic equilibrium if the flux difference trends to zero. The sub index i , represents the i^{th} equation in the network of i isotopes. When the right-hand side tends towards 0 there is no change in isotopic abundance, $\frac{dY}{dt} \rightarrow 0$. The isotope in question is no longer changing values, it has reached an equilibrium. The other option to reach equilibrium is when the right-hand terms tend towards a constant. In this case the left-hand term ($\frac{dY}{dt}$) is not zero but maintains a constant value. This specific isotope is either increasing or decreasing at a constant rate and is assumed to be in equilibrium. The ASY approximation is used when the term tends towards zero $F_i^+ - F_i^- \rightarrow 0$ and the QSS approximation is used when the term tends towards a constant value $F_i^+ - F_i^- \rightarrow C$. In both cases that isotope is assumed to be in equilibrium. In the asymptotic approximation the i^{th} equation in Equation (2.1) that describes an isotope in equilibrium gets removed from the explicit integration of the network. It is still integrated, but not by the forward Euler method, it is approximated by the asymptotic method. The removal of these equations from the forward Euler calculation leads to a less stiff network since their removal tends to lessen the difference in the timescales for the set of remaining equations.

Equation (4.2) represents the equation for microscopic equilibrium, the regime of the PE approximation. Here the $(f_j^+ - f_j^-)_i$ terms represent individual reactions. When the flux of any group of forward-reverse reactions is approximately zero, those reactions can be removed from the equation (total flux is set to 0 and the reaction won't contribute to the equation). Individual reactions can come into equilibrium even if the differential equation is not macroscopically in equilibrium. This approximation reduces the stiffness in the equations by removing the terms with the fastest rates. Unlike the asymptotic

approximation the partial equilibrium does not remove equations from the network. Rather, it removes the stiffest terms in the equations. This lessens the difference in the timescales as well. Combining the partial equilibrium method with either the asymptotic or quasi steady-state removes stiffness in the calculation in two different ways, which allows for larger time steps.

4.1 The Asymptotic Method

The asymptotic approximation is used in macroscopic equilibrium and removes whole differential equations from the numerical integration of the network. The equation in macroscopic equilibrium is still integrated, but it is replaced by a stable algebraic solution. The asymptotic approximation applies to the differential equation whose right-hand side tends towards zero. The left-hand term, therefore, also tends towards 0, $\frac{dY}{dt} \rightarrow 0$. This means the abundance of an isotope is very slowly changing in this case.

The differential equation in Equation (2.)1 can be written in the form of Equation (4.3),

$$\frac{dy}{dt} = F^+(t) - k(t) \cdot Y(t), \quad (4.3)$$

where Y is the abundance of an isotope, F^+ is the increasing flux terms and the $k(t)$ term is the rate parameters. The abundance of an isotope changes with time depending on the rate at which it is produced and depleted. When it is in equilibrium, the rate at which this isotope is created F^+ is roughly equal to the rate at which it is depleted F^- . The depletion of an isotope results in the F^- term is proportional to the abundance, Y , times a rate factor, k (Equation 4.4).

$$F_i^- = (k_1^i + k_2^i + k_3^i + \dots + k_j^i) \cdot Y_i \quad \text{or} \quad k^i = \frac{F_i^-}{Y_i}. \quad (4.4)$$

With the subscript i denoting a specific isotope in the network. The rate factors k_j^i describe the reaction rates where the isotope in question is on the left side of the reaction equation. In the forward reactions, isotopes on the left side are depleted and isotopes on the right side are created. These processes depend on the conditions (temperature and density) as well as the present abundances. The timescale of these rates is given by a characteristic time, τ . This

characteristic time $\tau = \frac{1}{k^i}$, can vary over many orders of magnitude, leading to stiffness. The characteristic time is the timescale for the total depletion rate, k , of any specific isotope Y_i . Using the re-written form of equation 2.1 that is seen in Equation (4.3) and by substituting Equation (4.4) into that form, the abundances in a given time step, n , can be written as

$$y_i(t_n) = \frac{F_i^+(t_n)}{k^i(t_n)} - \frac{1}{k^i(t_n)} \frac{dy_i}{dt}. \quad (4.5)$$

For the asymptotic approximation to be valid $F_i^+ \sim F_i^-$, so that the change in abundance for an isotope is also near 0. Once this is the case, the approximation is used to calculate the abundances, y_i at a given time step as seen in

$$y_n = \frac{1}{1 + k_n \Delta t} (y_{n-1} + F_n^+ \Delta t) \quad (4.6)$$

To determine whether an isotope's abundance is updated by this approximation or by the standard forward Euler method, we define a critical value, κ . This critical value is compared to $k^i \Delta t$ for each isotope, and if κ is set to be 1.0, then the two possible outcomes are:

1. $\kappa > k^i \Delta t$
2. $\kappa \leq k^i \Delta t$

If the value of $k \Delta t$ for a specific isotope is greater than κ (or 1.0) then it will be updated with the approximation equation in 4.6. Otherwise, if $k \Delta t$ is less than κ (or 1.0) then that isotope will have its abundance updated by the forward Euler method since the forward Euler time step is stable for that value of Δt .

The asymptotic approximation removes whole differential equations from the network that would be solved explicitly and instead replaces those calculations with algebraically stable solutions. The removal of those equations from the explicit solution removes stiffness from the network and allows the timestep to grow past the previous explicit limit. Figure 4.1 shows the implementation of the asymptotic approximation used on the triple alpha network. The dotted red line is the typical limit of the time step with the forward Euler method. The blue and green lines are the time step with the ASY+PE and ASY only approximations applied to the calculation. The ASY+PE calculation has ideal time stepping for the purposes

of showing that these approximation methods allow for much larger time steps than typical explicit methods.

4.2 Partial Equilibrium

The partial equilibrium approximations is applied when microscopic equilibrium conditions hold. This method is only used paired with the asymptotic or the quasi steady-state method. Back in equation 4.2, the right hand side looks like $(f_j^+ - f_j^-)_i$. These terms represent individual terms in the differential equation shown in equation 4.1 that the asymptotic approximation deals with. The partial equilibrium looks more in depth at the differential equations and the source terms $(f_j^\pm)$ with individual reactions. This approximation method is applied to reaction groups in order to remove the stiffest parts of the fluxes in the differential equations in equation 4.1.

A reaction group generally consists of forward and reverse reactions (if present) for reactions that produce or deplete any given isotope. For example, in the alpha network, one of the nineteen reaction groups contains the forward and reverse reactions that create and deplete O16 and is listed below:

- $He4 + C12 \rightarrow O16$
- $He4 + C12 \rightarrow O16$
- $O16 \rightarrow He4 + C12$
- $O16 \rightarrow He4 + C12$

It may be important to note that the difference between the two forward (and two reverse) reactions is based on the REACLIB rate parameters for each, which give two different reaction rate components contributing to the same reaction. A reaction group is considered to be in equilibrium once every isotope within that reaction group has reached equilibrium. For now, equilibration for an isotope can simply be put as whether or not that isotope has reached its equilibrium abundance. The method for determining and implementing equilibrium abundances is discussed later in this section. A reaction group with every

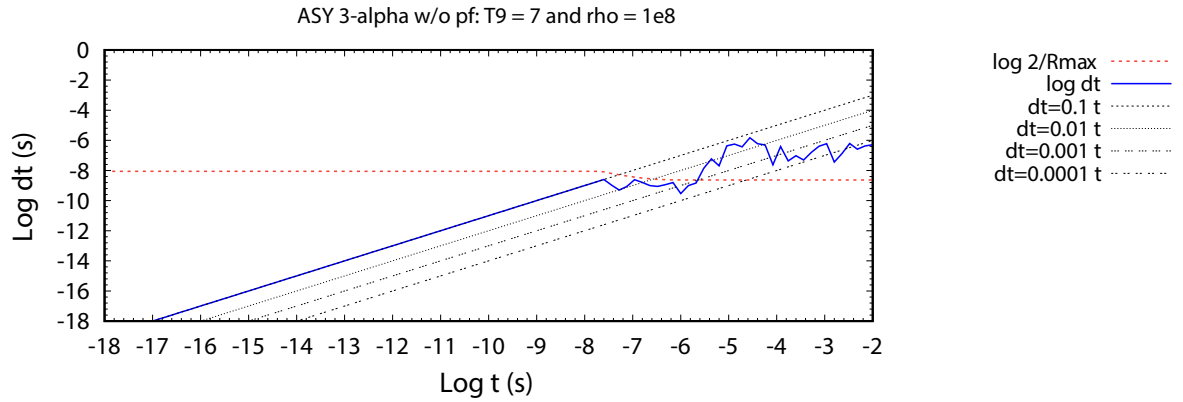


Figure 4.1: This figure shows the time step as a function of time. The red dotted line is the typical forward Euler maximum time step. A purely explicit time step would be limited to stay at or under the red line but removing some stiffness with the asymptotic approximation allows for a time step that can go above the forward Euler value (solid blue curve).

isotope in equilibrium would imply that the forward (i.e. $He4 + C12 \rightarrow O16$) and reverse ($O16 \rightarrow He4 + C12$) reactions occur at the same rate. Therefore, the isotopes within the reaction group are also produced and depleted at the same rate causing the total flux to be zero for the reaction group. Once a reaction is in approximately equilibrium, the fluxes for the participating reactions are replaced with zero. The methods for determining approximately equilibrium will also be discussed later in this section.

The source terms for an example reaction $A + B \rightarrow C + D$ can be written in terms of the rate parameters of the reaction and the abundances of the participating reactions ($f_{ab \rightarrow cd} = \pm(k_f y_a y_b - k_r y_c y_d)$). The k term is either the forward or reverse reaction rate parameter. The source term, $(f_{ab \rightarrow cd})$, will appear in the equations that include the species a, b, c and d . The abundances of a, b, c and d are also dependent on other reactions where those species are involved, so the timescale will vary between each one. Removing this one term will lessen the stiffness for multiple differential equations and therefore, reduce the stiffness in the network as a whole.

To get a better picture of the process of removing the stiffness, equation 4.7 shows the change in abundance in isotope a with the flux and source terms.

$$\frac{dy_a}{dt} = F_a^+ - F_a^- = k_r \cdot y_c \cdot y_d - k_f \cdot y_a \cdot y_b = F_a^+ - \frac{y_a}{\tau_a} \quad (4.7)$$

The flux total abundance and fluxes for the isotope, a , depend on all the reactions that it partakes in. Not all the reactions have the same rate, or timescale though. The τ term comes in as the characteristic timescale for this particular reaction. The timescale depends on the abundance of the isotopes pairing(s) in the reaction and the forward (or reverse rate). For example,

$$\tau_a = \frac{1}{k_f y_b} \quad \tau_b = \frac{1}{k_f y_a} \quad (4.8)$$

In the example with $A + B \rightarrow C + D$, there are then four different characteristic timescales present. To work down to one scale we use conserved scalars and progress variables that are described in Guidry et al. [2]. The result leads to the following equation for y_a from the example above.

$$\frac{dy_a}{dt} = k_r \cdot y_c \cdot y_d - k_f \cdot y_a \cdot y_b \quad (4.9)$$

This can also be written in the form of another differential equation,

$$\frac{dy_a}{dt} = ay_a^2 + by_a + c. \quad (4.10)$$

The coefficients a, b and c are related to the rate parameters and the conserved scalar quantities, c_1, c_2 , and c_3 . The conserved scalar quantities are a sum (or difference) of initial abundances in the reactions. They are given by the relations below.

- $c_1 = y_a^0 - y_b^0$
- $c_2 = y_a^0 + y_c^0$
- $c_3 = y_a^0 + y_d^0$

The y^0 terms denote the initial abundances. The coefficients can then be shown with these terms and the forward and reverse rate parameters.

- $a = k_r - k_f$
- $b = -k_r(c_2 + c_3) + k_f \cdot c_1$
- $c = k_r \cdot c_2 \cdot c_3$

This reaction can be described by the differential equation above (Equation 4.9) and even a 3-body reaction like the first example can be approximated by this sort of equation. When the forward and reverse rates are equal, the coefficient, a , is zero. The solution to Equation (4.10) when $a = 0$ yields the quasi steady-state solution which is not discussed here. When the coefficient, a , is not zero, we look where the discriminant of the quadratic equation: $q = 4ac - b^2$ is negative. The solution $q = 0$ is the trivial solution and when $q < 0$, the equilibrium solution as t approaches infinity is:

$$y_a = y_a^{eq} = -\frac{1}{2a}(b + (\sqrt{-q})) \quad (4.11)$$

Once y_a has been found, the other abundances, b, c and d can be found by solving for them from the relations above that use c_1, c_2 , and c_3 . The other abundances then become:

1. $y_b(t) = y_a(t) - c_1$
2. $y_c(t) = c_2 - y_a(t)$
3. $y_d(t) = c_3 - y_a(t)$

The rate at which this reaction, $A + B \rightarrow C + D$, evolves towards the equilibrium value in Equation (4.11) is no longer determined by four separate timescales. The single timescale used to determine the rate of evolution to equilibrium is given by Equation (4.12).

$$\tau = \frac{1}{\sqrt{-q}} \quad (4.12)$$

Now that the reaction can be described by one timescale, it is simpler to work with and remove the stiffness from the network. Once the reaction is approximately in equilibrium the single timescale allow the removal of all the source terms in the reaction which reduces the stiffness of a number of differential equations being solved numerically.

It was mentioned above that the partial equilibrium can be applied to the source terms approximately in equilibrium. This approximation comes from the variable ϵ . This ϵ can be seen as a critical value in the sense that κ was in the asymptotic method. The ϵ can be set to any desired value depending on how close to equilibrium is desired for a specific calculation. The ϵ term is compared to a percent difference between the current abundance of each isotope participating in the partial equilibrium and its equilibrium abundance that was found in Equation (4.11). Equation (4.13) shows the comparison between this difference and the tolerance (ϵ) term. When the percent difference is greater than ϵ , equilibrium has not been reached. When the percent difference is less than ϵ , then equilibrium has been approximately reached:

$$\left| \frac{y_i(t) - y_i^{eq}}{y_i^{eq}} \right| < \epsilon_i. \quad (4.13)$$

This comparison is done on an isotope by isotope basis.

The partial equilibrium removes stiffness from the source terms of possibly multiple differential equations that are describing isotopic abundance. When paired with either asymptotic or quasi steady-state the stiffness in the network is greatly reduced and the time steps that can be taken while remaining stable grow far beyond the traditional explicit ones and can become comparable to implicit time steps. Figure 4.2 shows the time steps that can be taken when the partial equilibrium approximation is invoked with the asymptotic approximation. The asymptotic method only was able to go above the explicit limit, but fell below 0.01% of t . With the partial equilibrium and asymptotic methods, the time stepping does not go below 1% of t except for a small dip around $t = 1 \times 10^{-7}$. The time stepping at later times is much more significant since the x -axis is $\log(t)$ and at these times dt is linear at 10% of t . This aligns with typical implicit calculations that also take time steps typically around 1% to 10% of the time.

4.3 Adaptive Time Stepping

This time step function is adaptive, meaning it can increase or decrease the value of Δt depending on the estimated error of the network at any given time. The sum of the mass fraction for all isotopes (denoted as sumX in FERN) should be equal to 1.0 ($\sum_i X_i = 1.0$) in order to conserve mass. There is some deviation from 1.0 in the calculations because we are using approximations to calculate the abundances. Just how much sumX differs from 1.0 is dependent upon a tolerance set by the user. The tolerance condition controls how much the sumX value is allowed to change between consecutive time steps. The difference in the sumX of consecutive time steps need to be less than the tolerance value to continue progression,

$$|\text{sumX}_{\text{current}} - \text{sumX}_{\text{previous}}| < \text{Tolerance}. \quad (4.14)$$

The value of Δt can affect the calculated abundances as shown back in Equation (4.6) and those abundances end up relating to the mass fractions by multiplying the abundances and the mass number for each isotope

$$X = Y \cdot A. \quad (4.15)$$

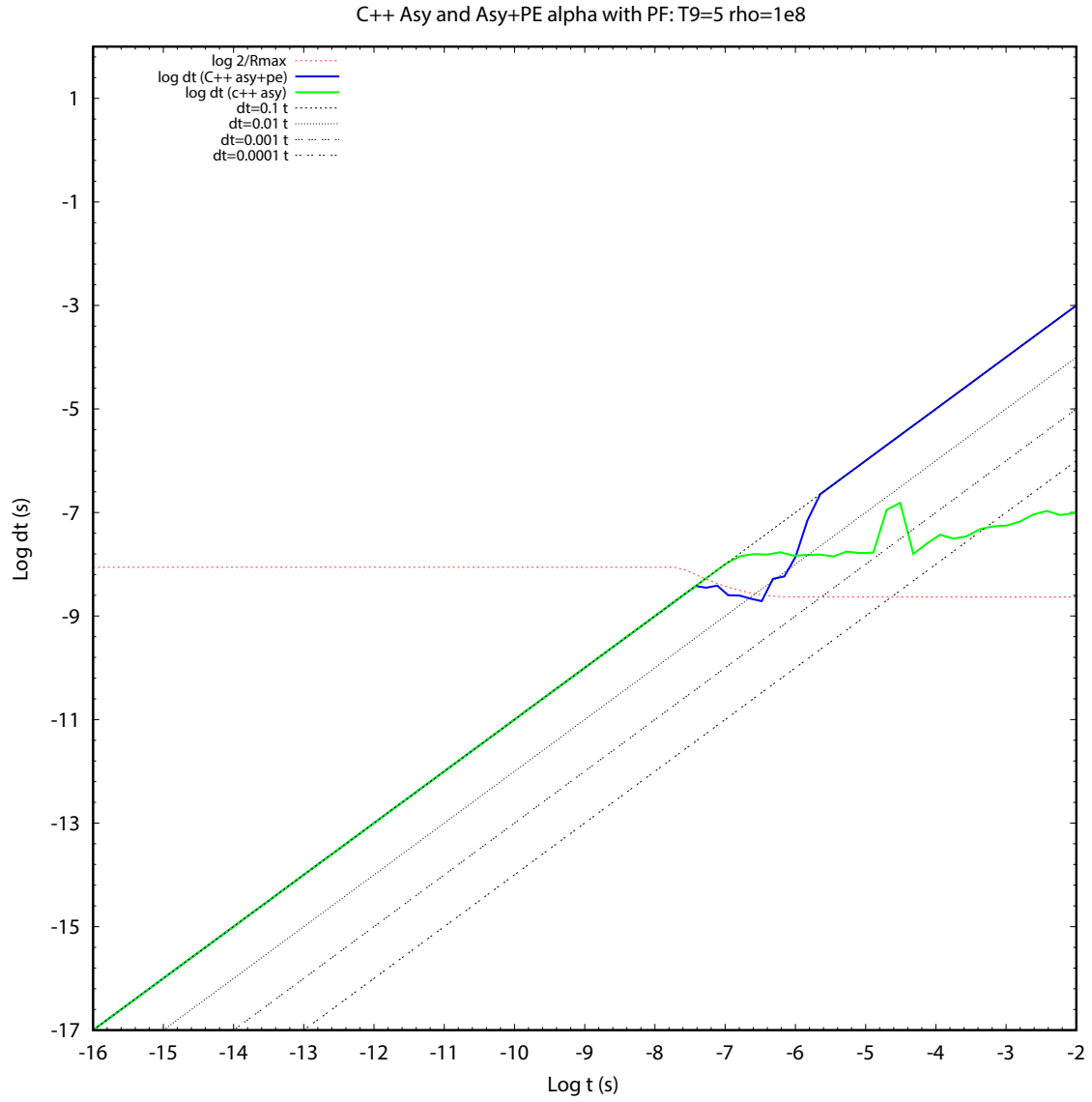


Figure 4.2: This figure shows the timesteps taken with the asymptotic and asymptotic plus partial equilibrium approximations on the triple alpha network in blue. The green line is the asymptotic only and the red line is the explicit limit.

Allowing the time step to be able to grow or decrease depending on the current status of the error is an important implementation in FERN. Allowing the time step to grow, if possible, while maintaining desired accuracy provides more efficient time stepping. On the other hand, restricting the time step in order to keep accuracy allows the final calculation to be as accurate as the user desires. The tolerance conditions in FERN are labeled as `massTol` within the code. There is an `Asy` and `PE` variable for `massTol` allowing for two different tolerance parameters for the two different regimes of the approximations.

The partial equilibrium has its own set of tolerance values in addition to the mass tolerance. The condition of whether an isotope enters partial equilibrium or not is shown back in Equation (4.13). The value of ϵ is the tolerance that determines how close to true equilibrium an isotope needs to be in order to be approximated with the PE approximation. A smaller value of ϵ would require an isotope to be closer to true equilibrium than a higher value. The second tolerance, denoted as `deviousMax` in FERN, is used to determine whether an isotope that has entered partial equilibrium stays in it or not. To stay in equilibrium, the difference in the equilibrium ratios of isotopes in a reaction group and the ratio of the reverse and forward rate parameters for a specific reaction need to be less than `deviousMax`,

$$\frac{Y_i^{eq}}{Y_j^{eq}} - \frac{k_{rev}}{k_{for}} < deviousMax. \quad (4.16)$$

If that difference is greater than the value set for `deviousMax`, then the isotope is removed from equilibrium. Both ratios depend on the isotopes that are participating in a specific reaction. A larger difference between these two ratios let the isotope stay in partial equilibrium even if it has swayed from its true equilibrium value in Equation (4.13). A lower tolerance for this difference would restrict how much an isotope can deviate from its true equilibrium value and remove it from partial equilibrium once it exceeds the tolerance.

Excerpts from the FERN code's time step function are provided to allow for more detail on how the adaptive time step function works. The function starts off by storing the values of `sumX` and `dt` from the previous time step and computing a Forward Euler value and an `Asy` value for the current step.

```
if(fastestCurrentRate > 0){
```

```

dt_FE = sf/fastestCurrentRate;
} else {
dt_FE = 0.00001*t;
}

```

```

dt_EA = computeTimeStep_EA(dtLast);
if(dt_EA < dt_FE) choice2 = 1;
dt = min(dt_FE,dt_EA);

```

The time step using the Asy method begins by updating the abundances (and mass fractions) with the current value of dt and comparing the current sumX to the previous sumX (this difference is labeled as diffX in the FERN code). If the value of diffX satisfies the tolerance condition, the time step is stored and passed on. For values of dt that lead to $\text{diffX} > \text{massTol}$, the code enters a while loop to decrease dt until it the tolerance condition of $\text{diffX} < \text{massTol}$ is satisfied.

```

while( diffX > massTol AND iterations <= maxit)[
dtt_0 = dtt;
dtt *= downbumper;
dt = dtt;
updatePopulations(dtt);
diffX = abs(sumX-sumXlast);
iterations ++;
totalIterations ++;
]
diffXfinal = diffX;

```

The current value of the time step (dtt) is stored in the variable dtt_0 before it is modified. The downbumper, currently set to 0.7, decreases the time step to be at 70% its original value and stores it in the variable, dt . The abundances are updated via the Asy method with the new time step value, and diffX is recalculated. The code will break out of the while loop once the value of diffX has satisfied the condition $\text{diffX} > \text{massTol}$, or if it has iterated the

decrease in dt over 100 times. Once an acceptable value for dt has been calculated, a trial time step is computed to start of the next iteration step.

```
double maxupdt = 2.0;
double maxdowndt = 0.5;
double dt_MAXFAC = 0.1;
dtmin = min(dt_new, maxupdt*dt_old);
dt_new = max( dtmin, maxdowndt*dt_old );
// Don't let dt exceed dt_MAXFAC * t for accuracy reasons
dt_new = min(dt_new, dt_MAXFAC*t);
// Return the new trial timestep that will be the starting point
// for the next integration step.
```

The trial time step is kept within the range of 0.5 - 2.0x the current value and is capped off at a value of 10% t to ensure accuracy. The trial time step will be the starting point for the next value of dt . At the next iteration, when the code computes the Asy time step, this value will be what is initially used before any modifications are made to it.

The adaptive time stepping allows for a controlled approximation based on how tight or loose the tolerance conditions are. The tighter the tolerance conditions, the less error that is allowed with each calculation. The other side of the coin is loosening the tolerance conditions to trade of accuracy for speed. The value of the time step relates to the tolerances based on the value of the sum of the mass fractions with a certain Δt .

Chapter 5

Results of the New C++ Code

The new C++ code, named FERN++, allows for an approximation that is controlled by the user, to trade off the accuracy of a calculation versus speed. The end of the last chapter discussed the parameters controlling the accuracies as well as the time step by determining how much error was allowed. To run a calculation that prioritizes accuracy, the user would run the calculation with tight tolerances; to prioritize speed, run a calculation with looser tolerances.

When coupled with a hydrodynamical code, the thermonuclear code would calculate the changes in the abundances and the energies in a single hydrodynamical time step. The values that the thermonuclear calculations return is seen in plots of the differential energy release and the mass fraction evolution over the course of the calculation. Other useful plots that show how these approximation methods works are the plot with the time step versus time and the fraction of isotopes (ASY) or reaction groups (PE) that are in equilibrium.

Examining the differential energy shows the net amount of energy that was released or absorbed per time step. Since the network contains both forward and reverse reactions, endothermic and exothermic reactions are present to add or remove energy from the system. The sum of all the energy produced/absorbed in every reaction at a single time step would yield the net change in energy for that period of time. The differential energies can then be related to the mass fractions by the Q value of the reactions and the flux of the isotopes. The Q value is the energy difference, in terms of mass, between the reactants and products of a given reaction. When an isotope is in or close to equilibrium, its flux is approximated to be

zero, causing the differential energies to tend towards zero. In this case, a fully equilibrated network will show no changes in energy and the differential energy will be zero. Alternatively, another way to calculate the differential energies more accurately is to use the difference in mass of the network before and after an iteration. The conservation of energy allows the differential energy to be calculated by multiplying the abundance of an isotope by its mass excess. Therefore, the differential energies should be as accurate as the mass fractions. In addition to the differential energies, the change in the abundances are important quantities to analyze with the thermonuclear network. The mass fraction plot displays the mass fraction evolution of each isotope in the network. The mass fractions are related to the abundances by the atomic mass of a specific isotope, therefore, the behavior of each isotope's abundance can be inferred from the mass fraction plots. Since these two measurables are returned to the hydrodynamical code, the resultant plots are used to judge the accuracy of the calculations.

The effectiveness of the approximations can be seen in plots of the time step growing with time, Δt vs t , and the fractions in equilibrium. An ideal time step would be within about 1-10% of the time, this would yield time stepping that is comparable to implicit methods. The fraction in equilibrium plot shows the percentage of isotopes in equilibrium for the asymptotic approximation as well as the reaction groups that are in approximate equilibrium that satisfy the requirement to use the partial equilibrium approximation. Pairing this with the time step plot allows a better look at how the approximations let the time step grow past the typical Forward Euler limit. These two plots display the speed up that occurs with the trade off in accuracy

The user will have control on exactly how they trade off accuracy for speed. The networks below have had several calculations done with different tolerance conditions ranging from accurate to fast. There are four cases that have been used for the purpose of showing the controlled approximation in action; Reference calculation, Accurate calculation, Intermediate calculation, and Fast calculation. The reference calculations have been run with an asymptotic approximation with very strict tolerances ($diffX = 1 \times 10^{-9}$ or 1×10^{-10} in most cases). Earlier comparisons with implicit calculations and with Forward Euler with (tiny) stable time steps indicate that the reference calculations can be considered exact. The accurate cases add the partial equilibrium approximation or loosen the tolerances, in most

cases it is a combination of both. The partial equilibrium tolerance in this calculation is stricter to make sure the isotopes are close to true equilibrium before being approximated so as to maintain high accuracy when compared to the reference calculations. The medium cases loosen the tolerances a bit more, but still show some restraint on the time step. This case can be seen as the best compromise between the accurate and fast cases. The fast cases have even looser tolerance conditions to allow for more efficient time stepping. These calculations are still restricted enough to keep the sum of the mass fractions equal to 1.0 within a desired tolerance of 1×10^{-4} to 1×10^{-6} so as not to obtain a final result that is unphysical while keeping an accuracy comparable to implicit calculations. The fast cases show how well the approximations work by cutting down on the number of integration steps, usually by a few to several orders of magnitude. Just how much error is introduced to attain this speed is discussed below and looked at more in depth in the next chapter.

5.1 Triple Alpha Network

The triple alpha network contains the reactions that make up the triple alpha process. This process converts three helium nuclei (or alpha particles) into ^{12}C . The fusion of two helium nuclei form an unstable isotope of ^8Be , which has a half-life on the order of 1×10^{-17} seconds. In order to form the ^{12}C isotope, another alpha particle must fuse with the ^8Be isotope before it decays. Another product of this process is ^{16}O , which is created from the fusion of ^{12}C and an alpha particle. This fusion process only occurs at high temperatures in the core of a star that contains a sufficient amount of helium. Enough helium must be present in order to have a high probability of fusing the helium nucleus with the highly unstable ^8Be isotope. This reaction occurs in stars that have reached their helium burning phase and have temperatures on the order of $T = 1 \times 10^8\text{K}$.

The results of this network that are shown in figure 5.1 are the calculations done with the fast case. The calculations done with loose tolerance conditions still match up well with the reference calculations represented by the dashed lines. The slight dip seen in the differential energy for the fast calculations coincides with the implementation of the partial equilibrium approximation. The timing of the approximations match up with the growth in the time

step above the forward Euler limit, confirming the principle of these methods. The ASY+PE approximation completes the calculation in 691 steps, while the strict asymptotic reference case took about 28,000 steps. The approximations introduced a speed up in the calculation by a factor of about 40.

Full equilibration is reached at a time of about $\log(t) = -4$ from interpolating the C++ fraction in equilibrium lines on the middle plot depicting $\log[dt(s)]$. The full equilibration of the network is reflected in the differential energy plot as well as the mass fraction plot. The differential energy plot plateaus around this time about the value of $\frac{dE}{dt} = 0$, as expected from a network that has reached 100% equilibration. The mass fractions plateau as well around their expected equilibrium values.

5.2 CNO Cycle

The CNO cycle is the reaction that powers the stars with a mass $M_\star > 1.5M_\odot$ because stars of those masses have high enough core temperatures for the CNO cycle to run efficiently. Following the reactions depicted in Figure 2.4, the CNO cycle ends up converting four hydrogen atoms into one helium atom. The CNO cycle can be present in a thermonuclear explosion of a star as a nova with the accreted hydrogen forming an envelope around the carbon oxygen white dwarf. The CNO reactions that occur can use this new source of hydrogen to trigger the thermonuclear reactions that cause steep rise in temperature.

The time scale for this calculation under stellar conditions is a few hundred million years, $t_{\text{final}} = 1 \times 10^{16}\text{s}$. Invoking the ASY+PE approximations at times when Δt is of order 1×10^{14} or 1×10^{15} , causes obvious errors when calculating the mass fractions for ^{13}C and ^{15}N . The time steps at these late times are too large to resolve the β decays which lead to error in the resulting isotopes. This type of stiffness is not removed from either the ASY or PE approximation methods, so a different solution is required.

Chapter 2 had briefly mentioned that at late times, the reactions with the fastest rate are the β decay from $^{13}\text{N} \rightarrow ^{13}\text{C}$ or from $^{15}\text{O} \rightarrow ^{15}\text{N}$ which has a time scale of $\sim 100\text{s}$. The rates of the β decays are independent of the temperature and density, unlike the other reactions in the network, which are extremely temperature dependent. In more massive main

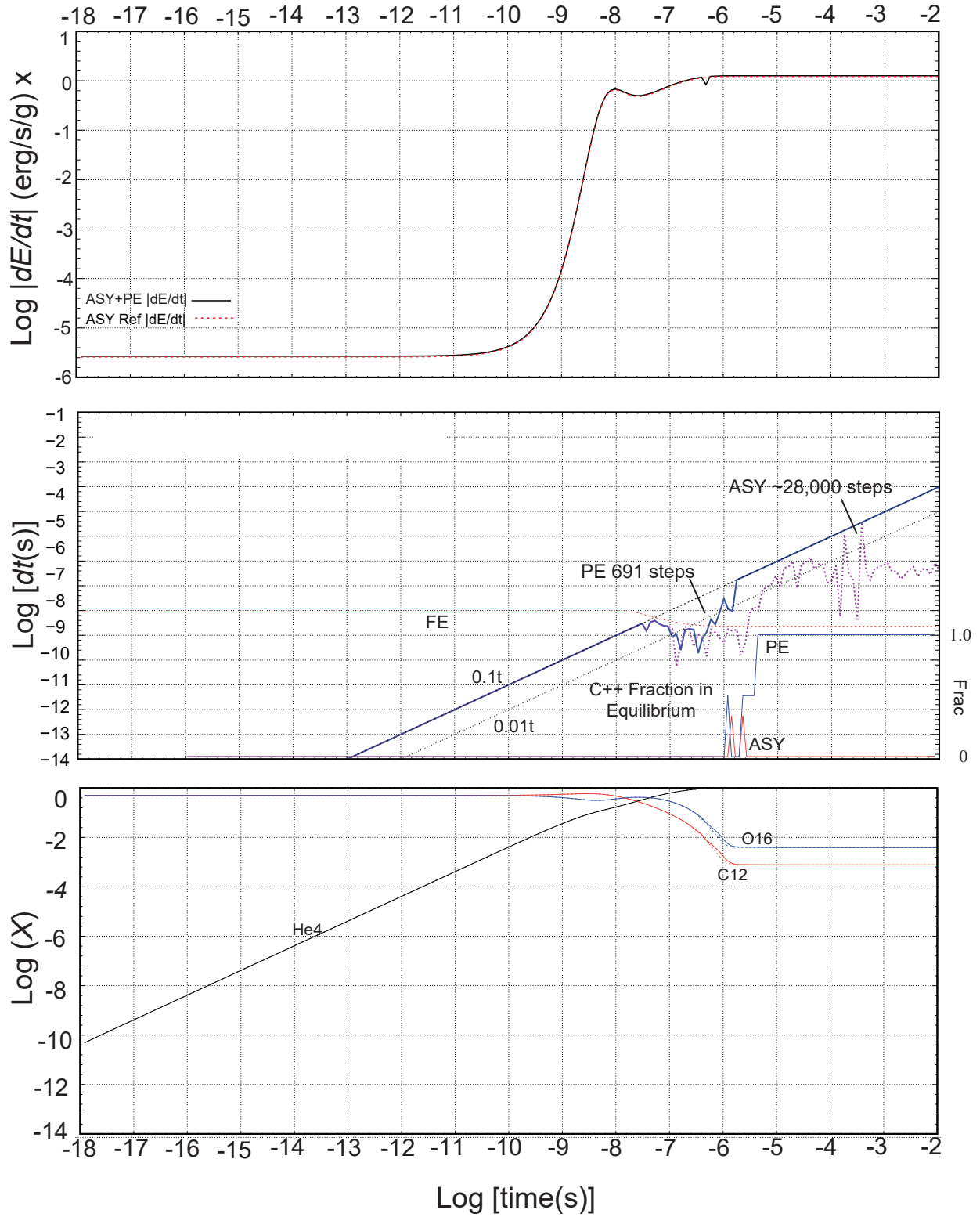


Figure 5.1: Summary of Triple Alpha results with constant $T_9=7$ and $\rho = 1e8$. The calculation done with ASY+PE is shown in solid lines, while the exact (reference) calculations are shown as dashed lines. The dotted line labelled FE indicates the largest stable forward Euler time step.

sequence stars where the CNO cycle is dominant, the reaction rates for other species in the network are orders of magnitude slower than the β decays. The algebraic approximations applied to these reactions allow the timescales to end up orders of magnitude larger than the β decays at later times in a calculation. Invoking the ASY+PE approximations at late times yields a Δt that is of order 1×10^{14} or 1×10^{15} , which is 13-14 orders of magnitude larger than the fastest timescale in the cycle. The large difference between these two timescales leads to obvious errors when calculating the mass fractions for ^{13}C and ^{15}N . To correct these errors and still take competitive time steps, the stiffness in this cyclic reaction must be removed. The partial equilibrium method is unable to remove this stiffness because these reactions do not come into forward-reverse equilibrium under typical stellar conditions that allow for the CNO cycle to be effective. This is particularly the case for the β decays such as $^{15}\text{N} \rightarrow e^+ + \nu_e$ because the inverse reaction requires the presence of positrons and electron neutrinos, which are not abundant enough in CNO burning conditions.

However, once a steady-state is reached, the closed cyclic nature of the CNO cycle requires that particle number be conserved in the flow. This requirement implies a relationship between the abundances of the species in the cycle and the reaction rates. In the case of the CNO cycle, once the hydrogen is depleted, there is no more helium being created and the cycle reaches a steady state. Therefore, the β decays can be essentially removed from the network and the abundances of ^{13}C and ^{15}N would be approximated based on this condition. The calculation of ^{13}C and ^{15}N can be seen in the equations 5.1 and 5.2,

$$Y_{C13} = \frac{\text{Rate}_{\text{p+C12} \rightarrow \text{N13} + \gamma}}{\text{Rate}_{\text{p+C13} \rightarrow \text{N14} + \gamma}} \cdot Y_{N14}, \quad (5.1)$$

$$Y_{N15} = \frac{\text{Rate}_{\text{p+N14} \rightarrow \text{O15} + \gamma}}{\text{Rate}_{\text{p+N15} \rightarrow \text{C12} + \alpha}} \cdot Y_{N14}. \quad (5.2)$$

Approximating the mass fractions of ^{13}C and ^{15}N from known values of rates and the abundance of more correct values removes the cyclic stiffness in the network and reduces the error in the approximation to a much more acceptable value, as will be seen in the extended CNO cycle results.

5.2.1 CNO Main

The CNO calculations are done up to a final time of $t = 1 \times 10^{16}$ seconds to cover the typical burning time for the CNO cycle. The asymptotic approximation alone works well enough in the main CNO network. The PE approximation was kept but appears to have not been invoked by the code. With the ASY alone, the speed up of a looser tolerance is about 8.5 times faster than the same calculation with a stricter tolerance. The mass fractions in the fast calculation show some deviation from the exact values at late times as hydrogen is being depleted, figure 5.2. While entirely possible the error seen is a consequence of the approximations, a likely explanation is that the cyclic stiffness described above is causing incorrect abundances to be calculated at large time steps coinciding with hydrogen depletion.

5.2.2 CNO-Extended

The extended CNO cycle used contains 15 isotopes with 91 reactions. The summary of this calculation, seen in figure 5.3, shows the effectiveness of the cyclic stabilization (CS) approximation used for this network. The ASY+PE approximations already allowed for time steps of the order 1×10^{16} , but now the β decay time scales have been fully removed, thus eliminating the smallest time scales from the network. The isotopes that are populated via β decays (^{13}C and ^{15}N) are now able to be calculated more accurately when these larger time steps are taken. Without the implementation of the cyclic stabilization approximation, these isotopes had mass fractions that were orders of magnitude off from the correct values. The comparison of the mass fraction calculation with and without the cyclic stabilization is shown in figure 5.4.

Equilibration in the network is reached at hydrogen depletion, leading to the steady state of reactions occurring in the cycle. The time step plots shown in figure 5.4, show the approximation continues to take time steps within 1-10% of t , in the same way the ASY+PE calculation does. The mass fraction plots show that the ^{13}C and ^{15}N isotopes are calculated closer to the exact values using the ASY+PE+CS methods than just ASY+PE.

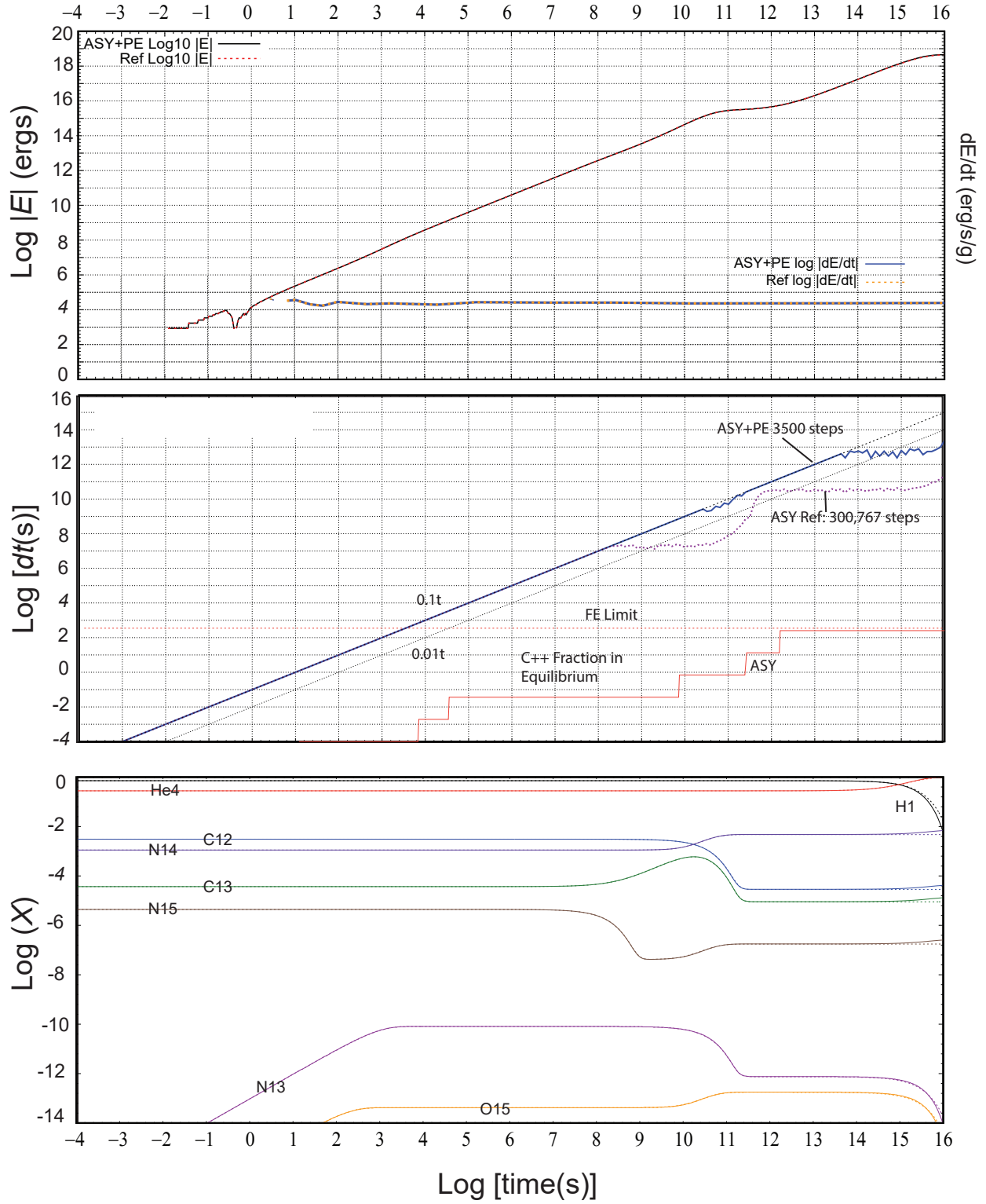


Figure 5.2: Summary of the CNO main network with $T_9 = 0.02$ and $\rho = 1 \text{e}3 \text{ g/cm}^3$.

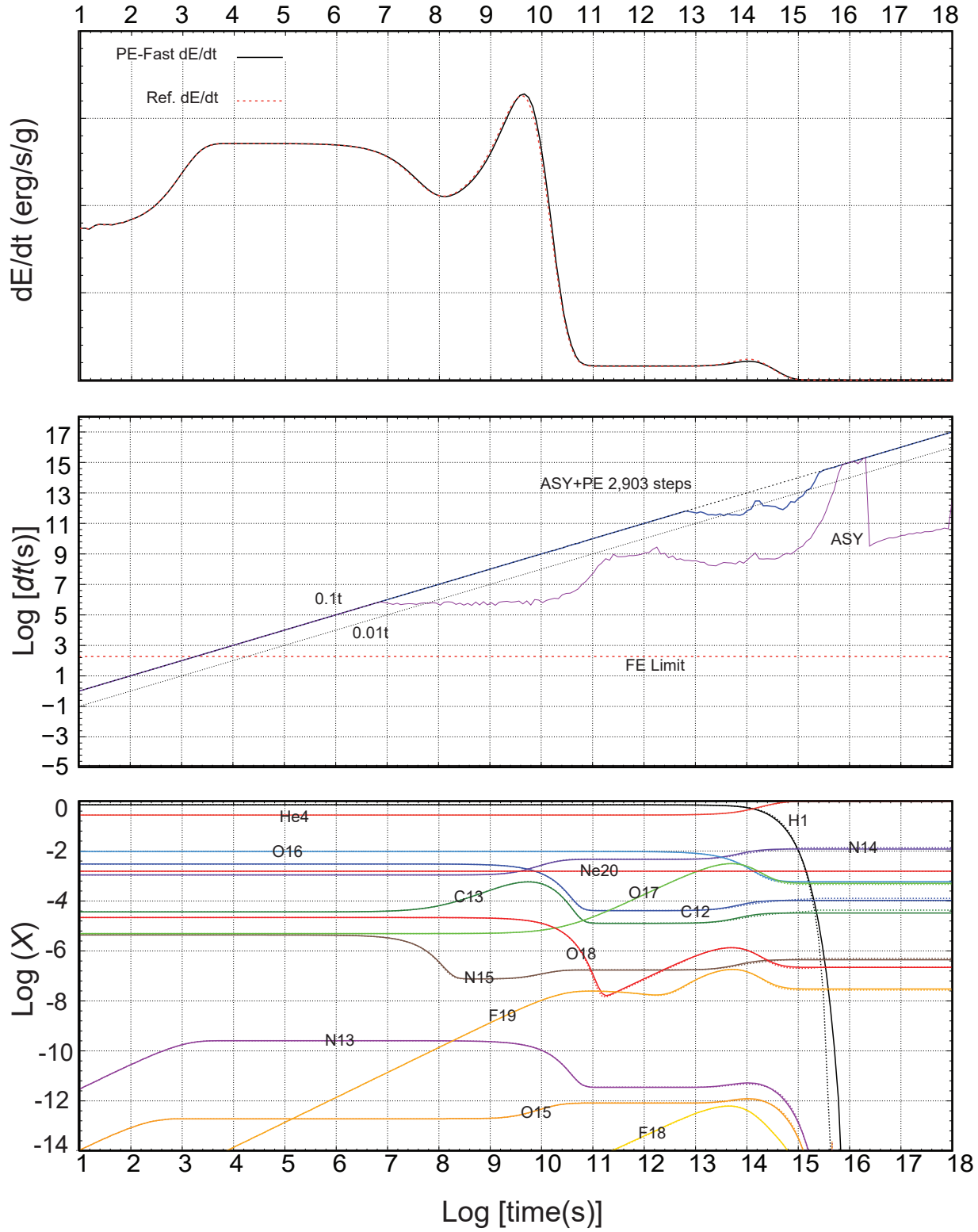


Figure 5.3: Summary of the CNO extended network with $T_9 = 0.025$ and $\rho = 100 \text{ g/cm}^3$.

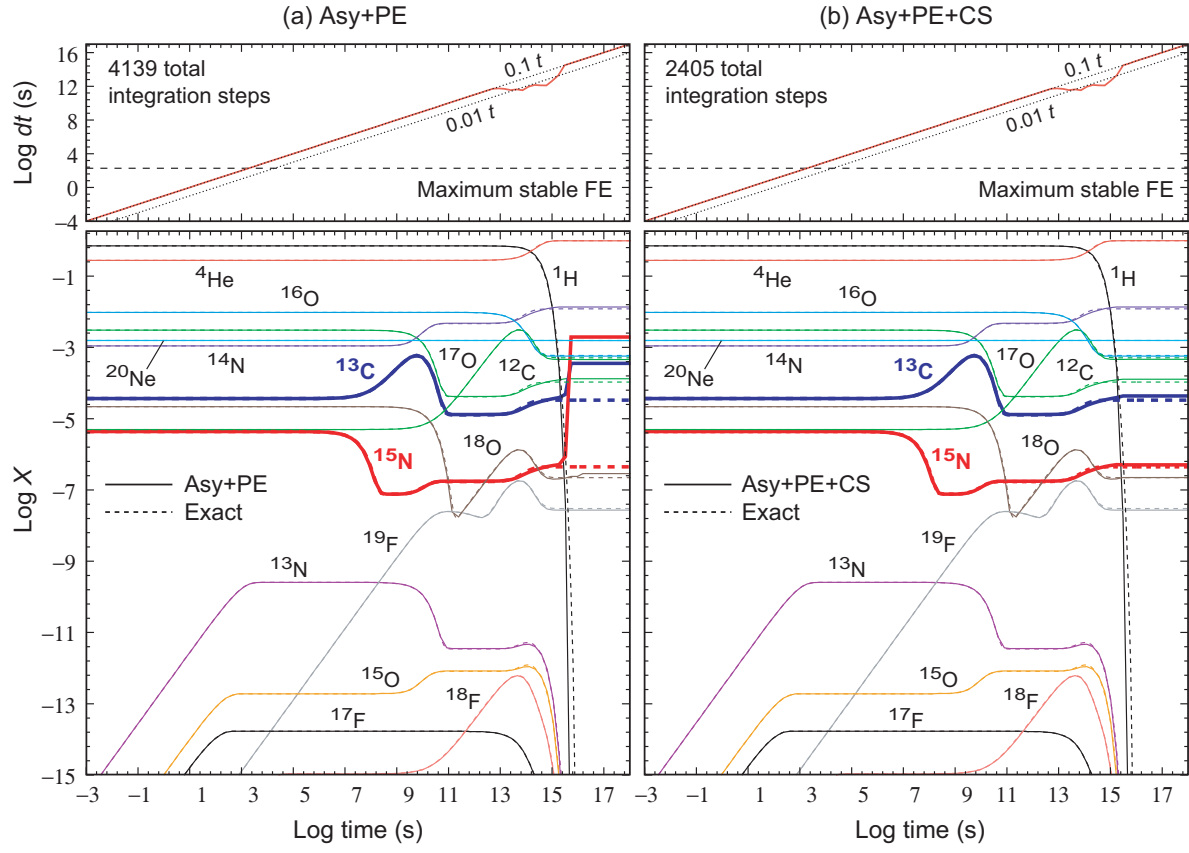


Figure 5.4: Mass fractions in the extended CNO with constant $T_9 = 0.025$ and $\rho = 100 \text{ g/cm}^3$ with (right) and without (left) the implementation of the cyclic stabilization approximation method.

5.3 Alpha Network

Everything up to this point has been calculated with the assumption of constant temperature and density. A more realistic simulation takes in to account the change in temperature and density as the network evolves. The code has the capability to use the constant temperature and density as well as to read and use hydrodynamical profiles that govern the evolution of temperature (T_9) and density (ρ) over the course of a specific calculation. The reaction rates are dependent on the temperature and as the reactions progress the energy of the system changes, consequently leading to further changes in the temperature. For a Type Ia simulation the density initially does not change since the star is made up of degenerate matter. Eventually the density does begin to change over time leading to a change in the probability of collisions. There are different hydro profiles used depending on the calculation being done. One of the more commonly used hydro profiles for our calculations was developed by Viktor Chupryna [17]. This hydro profile consists of an almost constant density, but a very sharp rise in the temperature.

5.3.1 Constant T and ρ

The alpha network (or alpha-16) is an often-used network that contains a minimal number of isotopes up to iron. The initial abundances of carbon and oxygen use fusion with alpha particles to climb up to iron. There are a few reactions that involve fusing carbon or oxygen with a heavier isotope to get an iron group isotope instead of just using alpha particle fusion to step up to those heavier isotopes.

The results for the calculation done at constant temperature ($T_9=7$) and density ($\rho = 1 \times 10^8 \text{ g/cm}$) are shown in figure 5.5. Aside from a slight discrepancy between the fast and exact calculations for the differential energies, there is no noticeable deviation between the two calculations. Even the discrepancy seen is not very significant in the big picture. This error occurs close to $\frac{dE}{dt} = 0$, and is orders of magnitude less than the peak differential energies.

The error present in the differential energies is not correlated with any error seen in the mass fractions at that time. The errors that are present in the mass fractions occur after

the approximations are applied to the network. The approximations begin removing the stiffness in the network as Δt reaches the forward Euler limit. The time step grows beyond this limit with the aid of the asymptotic approximation. The errors seen in the mass fractions occur in a time range of $\log(t) = -6 \rightarrow -4$. This time frame coincides with the rise of the number of reaction groups entering partial equilibrium. As was the case for the triple alpha network, once the network has been fully equilibrated, the mass fractions are constant at their equilibrium values.

The time stepping in the calculation stays within the ideal range of 1-10% of t , only dipping below 10% during the same time range mentioned earlier where the fraction of PE rises. The asymptotic reference calculation took just over 11 million steps, while the fast calculation required only 1,095 yielding a speed up by a factor of 10,000.

The order at which the reaction groups enter equilibrium are of some interest as well. There are nineteen reaction groups in the alpha network, all of which contain both forward and reverse reactions. In the accurate calculation the first three groups to come into equilibrium are RG 9, RG 8, and RG 12. Reaction group 9 contains the forward (and reverse) reaction of ${}^4\text{He} + {}^{28}\text{Si} \rightarrow {}^{32}\text{S}$, this reaction group enters equilibrium at $t = 2.08 \times 10^{-8}\text{s}$. The second group to enter equilibrium is RG 8. This group contains the forward (and reverse) reaction of ${}^4\text{He} + {}^{24}\text{Mg} \rightarrow {}^{28}\text{Si}$ and enters equilibrium at $t = 2.44 \times 10^{-8}\text{s}$. The third group to reach equilibrium is RG 12, which contains the forward (and reverse) reaction for ${}^4\text{He} + {}^{40}\text{Ca} \rightarrow {}^{44}\text{Ti}$. This group enters equilibrium at $t = 1.57 \times 10^{-7}\text{s}$.

The medium and fast calculations have the same first three reactions groups reaching equilibrium: RG9, RG 10, and RG8. Reaction groups 8 and 9 were both among the first groups in the accurate case, but RG 10 is new for these two calculations. Reaction Group 10 contains the forward reaction of ${}^4\text{He} + {}^{32}\text{S} \rightarrow {}^{36}\text{Ar}$. In the medium calculation RG 9 enters equilibrium at $t = 7.36 \times 10^{-9}\text{s}$, RG 10 enters at $t = 2.37 \times 10^{-8}$ and RG 8 enters at $t = 2.52 \times 10^{-8}\text{s}$. In the fast calculation RG 9 enters equilibrium at $t = 6.54 \times 10^{-9}\text{s}$, RG 10 enters at $t = 1.86 \times 10^{-8}\text{s}$ and RG 8 enters at $t = 2.48 \times 10^{-8}\text{s}$.

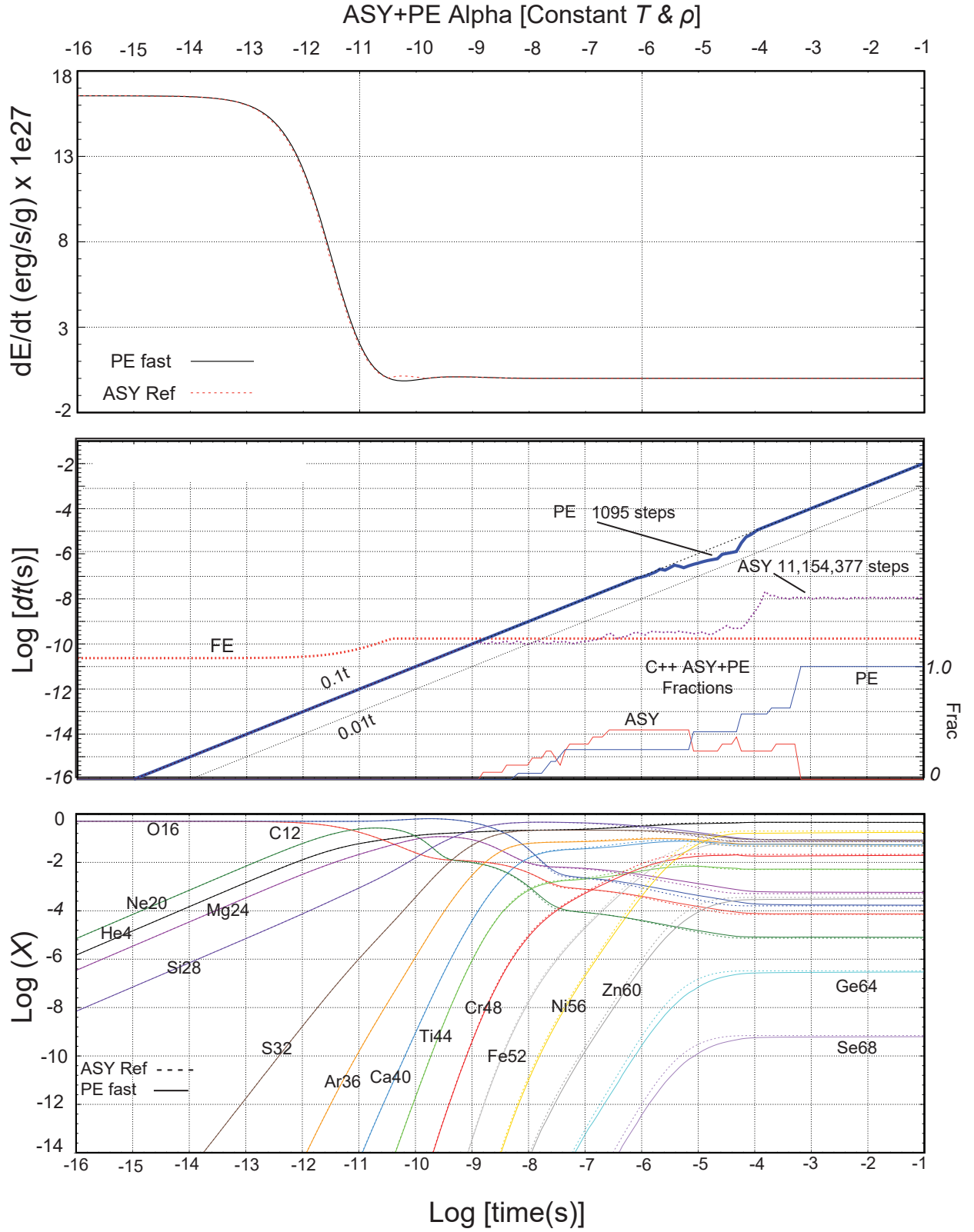


Figure 5.5: Summary of the alpha network at constant $T_9 = 7$ and $\rho = 1e8 \text{ g/cm}^3$.

5.3.2 Hydrodynamical Profile

The alpha network, being an often-used network in supernovae simulations, is one that is ideal for more realistic calculations using a hydrodynamical profile. One set of temperature and density profiles tested can be seen in figure 5.6. The density remains constant at $\rho = 1 \times 10^8$ g/cm³, while the temperature has a steep increase from $T_9 = 2$ to about $T_9 = 6.6$. This sharp rise in temperature drastically affects the reactions as seen by the sudden changes in the mass fractions in figure 5.7.

The differential energies do not appear as accurate as previous calculations, this is due to the change in scale for the y -axis. The usual axis is a linear scale, however, this calculation only shows a sharp peak in the differential energies and not much detail. The use of the log scale for this, allows more detail to be shown. The deviations seen in the fast calculation occur later in time when the energies are several orders of magnitude below the peak energy release. The error seen in the differential energies does not contribute significantly to the total calculation. The peak of $\frac{dE}{dt}$ has a value of about 1×10^{25} ergs/g/s, while the errors starts at about 3-4 orders of magnitude below and that gap keeps increasing.

The error seen in the mass fractions occurs in the same time frame that the differential energy errors, both coincide with the rise in reaction groups reaching partial equilibrium conditions. The approximations, while introducing noticeable error, keeps the time step growing above the forward Euler limit. With some exception, the time step is optimal and maintains a value between 1-10% of t for nearly the entire calculation.

The first reaction groups to enter equilibrium using a hydrodynamical profile are a bit different than when using constant temperature and density. The first three groups to reach equilibrium in the accurate case are RG 6, RG 9, and RG 12. Reaction group 6 contains two forward (and two reverse) reactions for ${}^4\text{He} + {}^{20}\text{Ne} \rightarrow {}^{24}\text{Mg}$. This group reached equilibrium at $t = 7.583 \times 10^{-6}$ s. Reaction groups 9 and 12 were both mentioned before in the case of constant temperature and density, RG 9 contains ${}^4\text{He} + {}^{28}\text{Si} \rightarrow {}^{32}\text{S}$ and RG 12 contains ${}^4\text{He} + {}^{40}\text{Ca} \rightarrow {}^{44}\text{Ti}$. Those groups came into equilibrium at $t = 7.589 \times 10^{-6}$ s and $t = 7.603 \times 10^{-6}$ s respectively.

ViktorProfile_400.inp

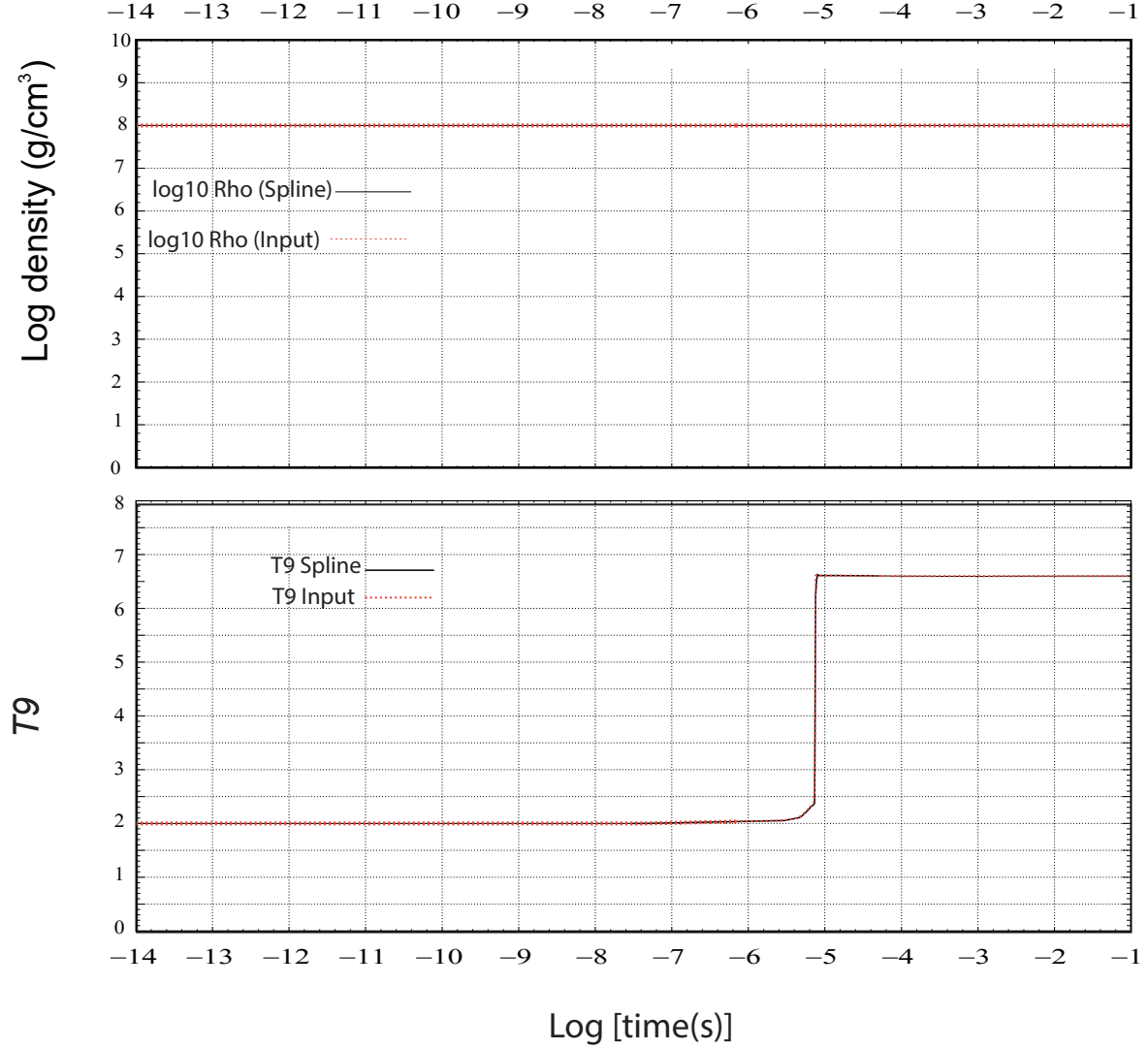


Figure 5.6: Hydrodynamical profiles used for the alpha network with varying T_9 and near constant density. The dotted curves marked Input are the profile table that was read in. The solid curves marked Spline are a cubic spline interpolation from that table used in the calculation. The profile is a result of work done by Viktor Chupryna.

In the medium calculation, the first groups in equilibrium are RG 6, RG 9, and RG 8. Reaction group 6 enters equilibrium at $t = 7.583 \times 10^{-6}$ s, while RG9 enters shortly after that at $t = 7.589 \times 10^{-6}$ s again. Reaction group 8 enters equilibrium third, ahead of RG12, at a time of $t = 7.6001 \times 10^{-6}$ s, RG12 enters equilibrium right after at $t = 7.6049 \times 10^{-6}$ s. The first groups to reach equilibrium in the fast calculation are RG 9, RG 10 and RG 12. Reaction group 9 enters equilibrium at $t = 7.58 \times 10^{-6}$ s, RG 10 enters equilibrium at $t = 7.59 \times 10^{-6}$ s, and RG 12 enters equilibrium at $t = 7.60 \times 10^{-6}$ s.

5.4 Nova 134-Isotope

The nova 134 network initially contains various isotopes of helium, carbon and oxygen as well as trace amounts ($Y = 1 \times 10^{-7}$ to 1×10^{-9}) of other elements such as fluorine and neon and magnesium. This calculation also involves a hydro profile with a steep increase in temperature followed by a rapid decline. The density in this calculation does not remain constant as it did in the Type Ia alpha network; rather the density has a sharp drop that occurs at the time of the temperature spike, as can be seen in figure 5.8.

The number of curves shown in the plot of the mass fractions versus time (Figure 5.9) is too high to be able to reasonably resolve the differences between the exact and fast calculations in most cases (though examination of individual curves shows good agreement). A more useful overall indicator of accuracy in this case would be the plot showing the differential energy. Since the values of $\frac{dE}{dt}$ are what has the most influence on the hydrodynamical code, and because the fluxes in abundances relate to the differential energies, we can infer the accuracy of the calculation from this plot. Both the total energy release and the differential energy are shown in the top graph. While the total energy releases do not show any error, there is a slight deviation in the fast calculation from the exact in the differential energy. This error appears at the same time the rapid rise in temperature occurs and causes a number of reactions to come in to play, just before the time, $\log(t) = 2$. The differential energies shown in this calculation were calculated using the Q values rather than the mass excess, which leads to the errors seen in figure 5.9. At this same time, the time step undergoes a drop below 1% of t , but quickly recovers and manages to maintain optimal

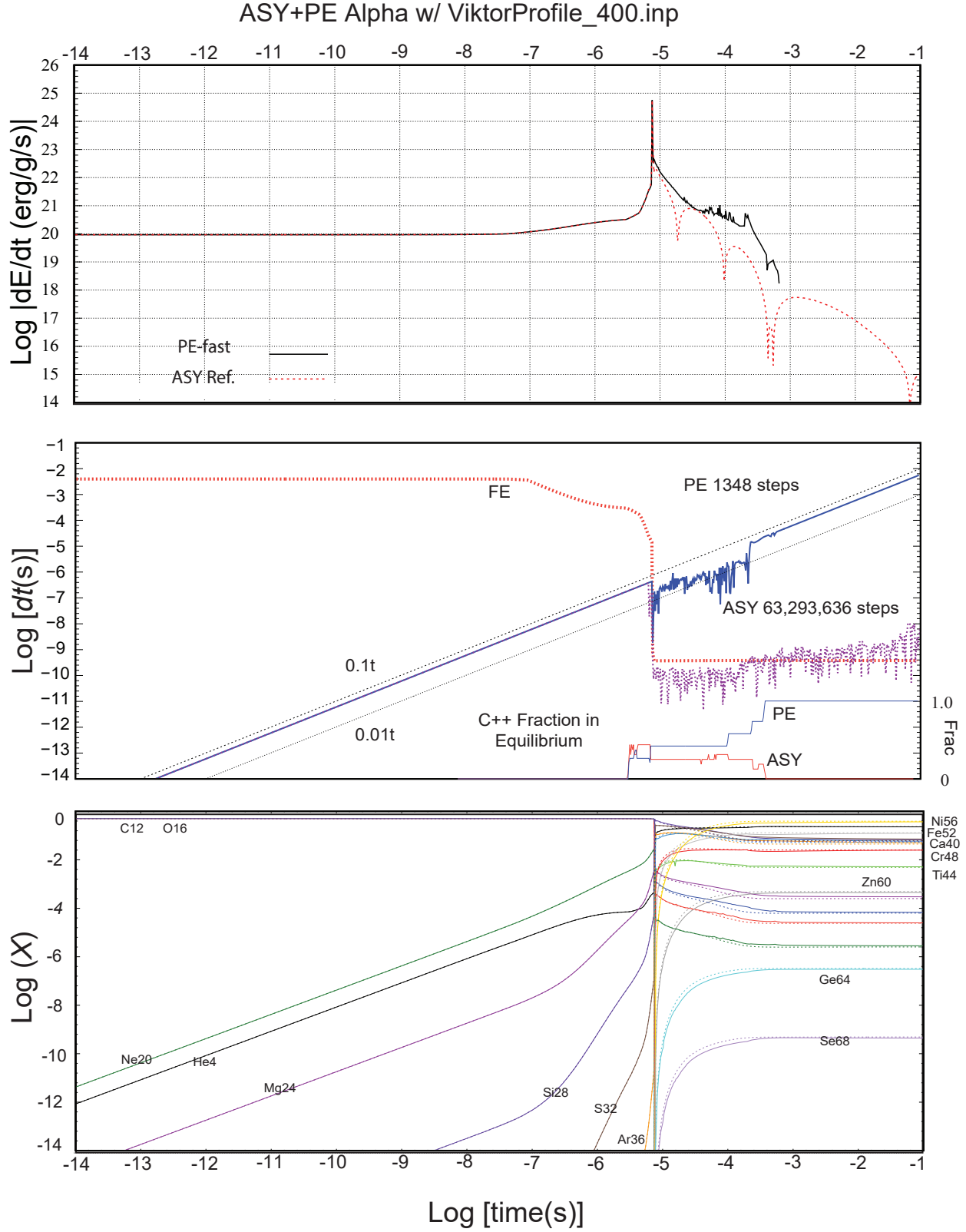


Figure 5.7: Summary of the alpha network with varying T_9 and ρ using the hydrodynamical profile *ViktorProfile400.inp*.

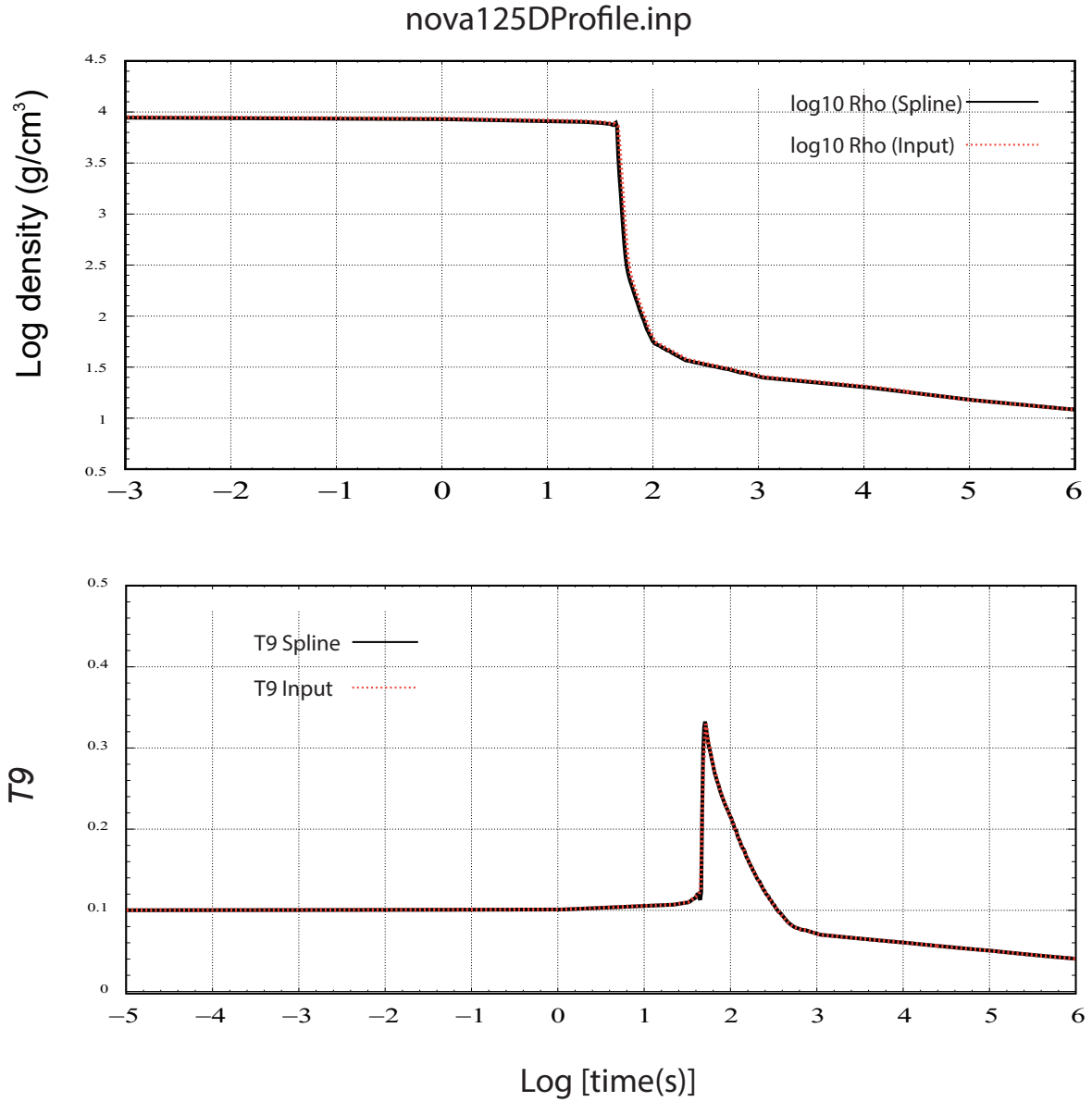


Figure 5.8: Hydrodynamical profile for the nova-134 calculation done with the *125DProfile400.inp*.

time stepping for the rest of the calculation. The asymptotic approximation alone in this network is able to cut down on the number of steps taken from nearly 41,000 for the correct case to just 617 in the fast asymptotic case. The time steps taken in the fast case are able to outdo the forward Euler steps by eight orders of magnitude over much of the calculation.

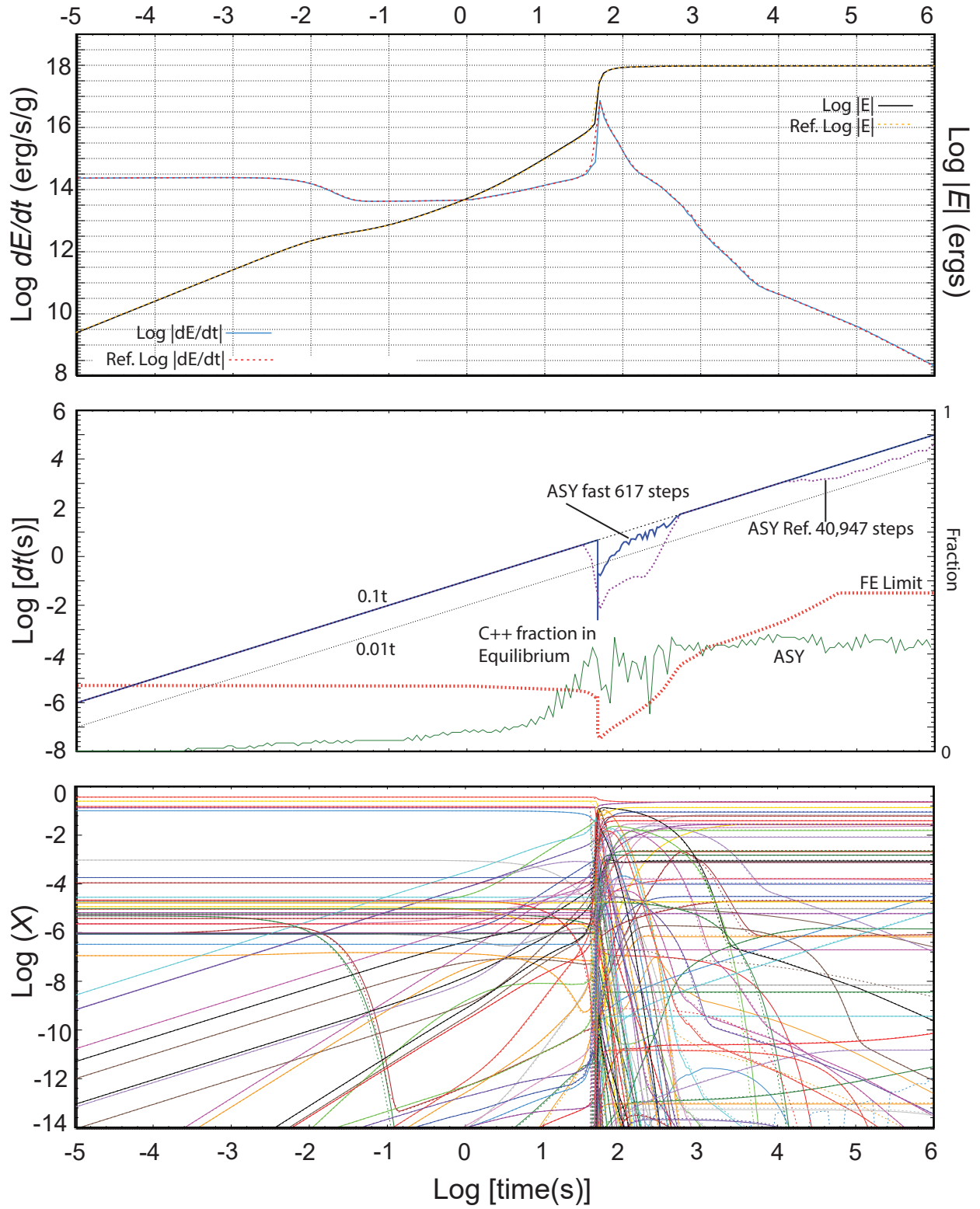


Figure 5.9: Summary of the nova-134 network with varying T_9 and ρ using the hydrodynamical profile *125DProfile400.inp*.

Chapter 6

Analysis of the Approximations

The previous chapter presented how the trade off in accuracy yielded a much faster calculation but did not quantify any of the errors introduced from using the approximations. The time stepping graphs were able to demonstrate how much faster the approximations can solve the network, while the mass fractions gave a visual representation of the accuracy. This chapter will dive into an analysis of the different calculations that used the ASY+PE approximations in order to quantify the error more precisely. The residuals (absolute errors) are calculated for each isotope in the network along with any uncertainties between the calculations made. In addition to the analysis, further discussion will be made to clarify the idea of a controlled approximation.

6.1 Residuals and Uncertainties

The errors in the calculations that used the Asy+PE approximation methods are seen in the residuals for each isotope. Both the accurate and fast mass fraction calculations have been compared to the mass fractions of the reference calculation on an isotope by isotope basis. The residuals are simply the difference between the calculated mass fraction and the correct value in the reference calculation, $\text{Residuals} = |X_{\text{calc}} - X_{\text{ref}}|$. There are two ways in which the residuals are used for analysis in this paper. The first way is on an isotope by isotope basis. Looking at the individual isotope's residuals can show the behavior of the accuracy over the course of the calculation and the maximum value of the residual gives

an idea of the degree of accuracy with which that isotope has been calculated. The second method uses the residuals to find the root mean square of the error at a specific time. The total error in the network is found by summing the residuals of each isotope at each output time. The sum of the residuals at each output time allows us to track the total error in the network over the calculation. We find at which output time the error across the network is the greatest, and use the residuals of the mass fraction for each isotope at that specific output time to calculate the root mean square (RMS) of the error. The RMS of the error provides a description of the accuracy of the FERN calculation with a single value shown as,

$$RMS = \sqrt{\sum_i |X_{\text{calc},i} - X_{\text{exact},i}|^2}, \quad (6.1)$$

where the index, i , represents the i^{th} isotope in the network. The root mean square of the error represents what the average deviation is between the mass fractions of the exact and FERN calculations. This method is meant to give a general sense of the accuracy across the network as a whole in a given calculation. Note that this method of error analysis is best suited for a description of the accuracy in the network as a whole rather than accuracy of individual isotopes within the network. The weaker isotopes (generally of order $X < 1 \times 10^{-4}$) have residuals that are very small compared to the stronger isotopes ($X > 1 \times 10^{-4}$). The residuals are generally two orders of magnitude (or more) smaller than the value of the mass fractions, so weaker isotopes have much smaller residuals. The root mean square may be best thought of as the average error in the stronger isotopes, which make up a nearly the entire mass of the system (99% of the mass). The best sense of accuracy for individual isotopes would lie within the residuals for each isotope.

The other important statistic used is the uncertainty in the calculation of the mass fractions. The three cases used for this analysis were the reference, accurate and fast calculations. The uncertainty here refers to the variation of a specific calculation from the mean value of those three sampled calculations. This allows us to place error bars on the mean mass fraction for any/all isotopes. The uncertainty for each isotope between the

three calculations is shown by equation 6.1, [18].

$$U = \sqrt{\frac{\sum_i (X_i - \mu)^2}{n(n-1)}} \quad (6.2)$$

The X_i term is the mass fraction for a specific calculation and μ is the mean of the three calculations mentioned earlier. The term in the denominator is the number of calculations used, so $n = 3$. This has been done for each isotope in the networks which used ASY+PE. The uncertainties are consistently low for each network, but there is still interest in looking at it in more detail. The values of the uncertainty are mostly relevant in the stronger isotopes when both the Asy and PE approximations are in effect. The uncertainties are used as error bars when plotting the mean mass fractions over time. These error bars represent the range of values in mass fractions an isotope may have when running a calculation using tolerances between the accurate and fast calculations.

6.2 Triple Alpha Analysis

The triple alpha network showed few errors in its mass fraction plot as well as the differential energy plot. This result is not too surprising as it is a simple network with only three isotopes and eight reactions. The residuals in the accurate and fast cases allow a more detailed look at the error that is introduced by the approximations. In terms of the error in the network as a whole, the time at which the maximum error occurs varies between the accurate and fast cases. The time the maximum error was reached in the accurate case was $t = 1.20 \times 10^{-6}$ s and the time for the fast case was $t = 3.98 \times 10^{-7}$ s. It happens to be at these times that each isotope had its maximum residual value too. While not too surprising in a small network, this is not always the case in larger networks. Table 6.1 shows the maximum value of the absolute error obtained by each isotope in the course of the calculations. Because the time of maximum error in the network occurred with the maximum error in each isotope, there is no column showing an isotope's contribution to the total error in the network. The last row

in the network provides the root mean square of the error, which quantifies the accuracy of the calculation.

The isotope with the largest contribution to the error in both the accurate and fast calculations is helium, while oxygen is not much lower. The error increases by a factor of 64 in the case of helium when going from the accurate to fast calculations, oxygen showed a factor of 67 increase, while carbon showed an increase of 66 times. While these values appear large, the reality is they don't actually introduce a significant amount of error in the network. The true mass fraction of helium at this time is $X[\text{He}] = 0.994418$, the accurate case gives a value of $X[\text{He}] = 0.994278$, differing only in the fourth decimal. This would yield a percent error of 0.01%. The fast case has the maximum error at a slightly different time where the true value of $X[\text{He}] = 0.931644$ and the fast calculation of $X[\text{He}] = 0.922678$, which yields a percent error of 0.96%.

Figure 6.1, shows a summarized analysis of the results for the accurate, medium and fast calculations. The mass fractions show the root mean square of the error at the time of the maximum error in the network as a quantitative measure of accuracy. The time step plots show the number of steps that each calculation took to integrate the network as a quantitative measure of speed. The controlled approximation is able to show the trade off in accuracy versus speed quantitatively now, rather than just qualitatively. While the error has increased between the accurate case to the fast case, as seen in the residuals and root mean square of the error, it is still well within the range of desired error for the approximations made to the kinetic network.

The uncertainties have been calculated for each isotope as well and the mean of each mass fraction between the reference, accurate, and fast calculations can give a range of values for each isotope. The uncertainties would allow the mass fractions to have a range of values at any time based on the type of calculation being done. Figure 6.2 shows the plot of the mass fractions with the uncertainties calculated for each isotope acting as the error bars. The error bars represent the range of possible values for each isotope's mass fraction in a calculation with tolerance conditions that fall between the accurate and fast cases. The uncertainties in the triple alpha network are just as small as the absolute errors, with helium having the largest uncertainty ($U_{\text{He}} = 0.003, U_{\text{C}} = 0.0006, U_{\text{O}} = 0.0024$).

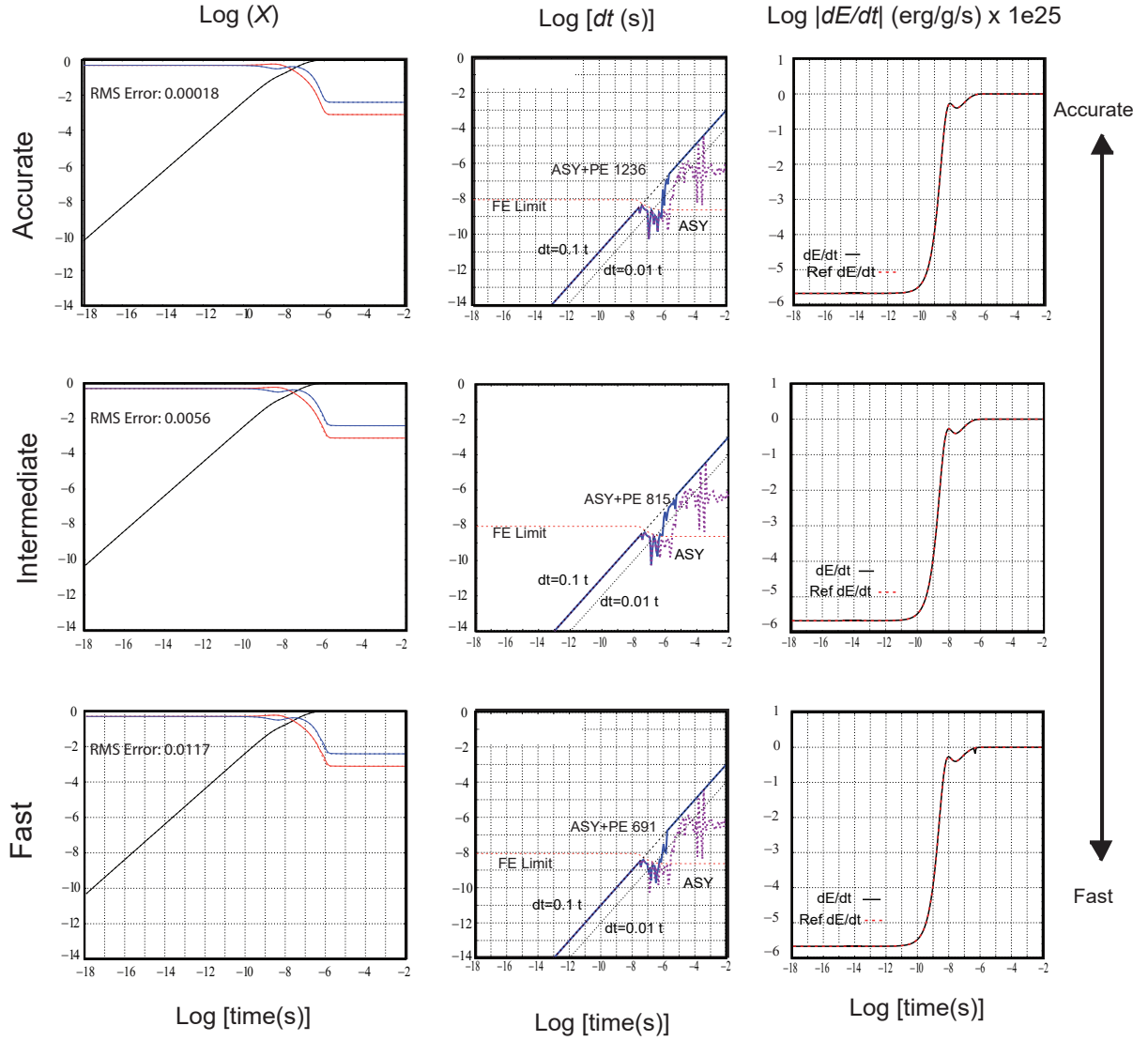


Figure 6.1: The trade off in accuracy versus speed for the triple alpha calculations.

Table 6.1: Triple Alpha Residuals

Isotope	Max. Error (Acc.)	Max. Error (Fast)
He-4	1.40×10^{-4}	8.97×10^{-3}
C-12	2.59×10^{-5}	1.72×10^{-3}
O-16	1.08×10^{-4}	7.33×10^{-3}
RMS Err.	1.8×10^{-4}	1.17×10^{-2}
Steps	1236	691

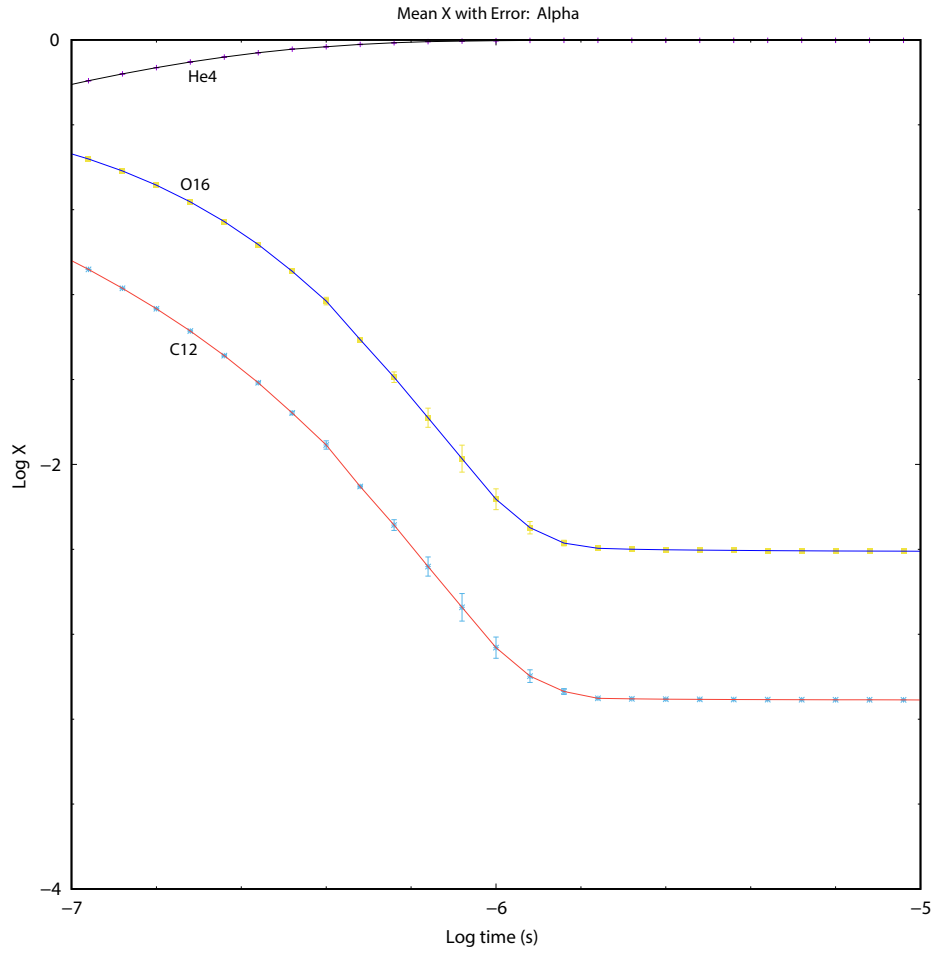


Figure 6.2: Triple alpha mean mass fractions shown with error bars indicated by the dots.

6.3 Alpha Network Analysis

6.3.1 Constant Temperature and Density

The alpha network had more noticeable errors between the fast and the reference calculations on the mass fraction plot. Both calculations had the same isotope contribute the most to the error at the time where the error in the network was the greatest. The accurate case had the peak error in the network at $t = 1.82 \times 10^{-5}$ s and the fast had its peak at $t = 1.5 \times 10^{-5}$ s. In both cases, ^{56}Ni contributed the most. There are more isotopes in the fast calculation that had contributed almost as much, but ^{56}Ni is still the largest contributor. Table 6.2 shows the residuals of each isotope in two different situations. The first grouping shows each isotope's maximum absolute error that it reached in the calculation for both the accurate and fast cases. The second grouping shows each isotope's absolute error at the time when the total error in the network was at a maximum for both the accurate and fast calculations. The last row again, shows the root mean square of the error for the two calculations

Figure 6.3 shows the summarized analysis of the alpha network calculations again for the accurate, medium, and fast calculations. The root mean squares of the error show that the accuracy gets worse as the time stepping becomes better, but the RMS error in the fast calculation is still an acceptable amount for the kinetic networks to return to the hydrodynamical code. The trade off in accuracy gives a factor of about 4 in the RMS of the error, while the fast calculation takes a factor of 20 fewer steps in the calculation. The RMS values still suggest that the error introduced by the approximations is an acceptable amount when compared to the other errors expected in the hydrodynamical calculation.

The uncertainties for each isotope were again calculated for the alpha network between the reference, accurate and fast calculations. The same few isotopes that contributed the most to the error in the fast calculations (^{56}Ni , ^{28}Si , ^{32}S , and ^4He) are the ones that yield the highest uncertainties as well. The plot of the mass fractions with error bars shown in figure 6.4 is only of these isotopes with the addition of ^{52}Fe . The remaining isotopes in the network had uncertainties that were too small to show anything of significance on the plot. The largest uncertainty is seen in the isotope of ^{56}Ni with $U_{\text{Ni}} = 0.0140$ and ^{28}Si followed

Table 6.2: Alpha Network Residuals

Isotope	Max. Error (Acc.)	Max. Error (Fast)	Error at $t_{\max}$ (Acc.)	Error at $t_{\max}$ (fast)
He-4	6.31×10^{-3}	2.90×10^{-2}	5.99×10^{-3}	2.55×10^{-2}
C-12	1.54×10^{-5}	2.03×10^{-4}	1.00×10^{-5}	3.20×10^{-5}
O-16	2.49×10^{-4}	9.23×10^{-4}	2.40×10^{-5}	1.11×10^{-4}
Ne-20	1.89×10^{-5}	1.79×10^{-4}	1.00×10^{-6}	3.00×10^{-6}
Mg-24	7.35×10^{-5}	4.88×10^{-4}	7.00×10^{-5}	3.53×10^{-4}
Si-28	8.50×10^{-3}	4.12×10^{-2}	8.44×10^{-3}	4.01×10^{-2}
S-32	6.54×10^{-3}	2.90×10^{-2}	6.54×10^{-3}	2.90×10^{-2}
Ar-36	3.35×10^{-3}	1.34×10^{-2}	3.35×10^{-3}	1.29×10^{-2}
Ca-40	2.55×10^{-3}	9.66×10^{-3}	2.42×10^{-3}	6.97×10^{-3}
Ti-44	1.88×10^{-4}	1.35×10^{-3}	1.69×10^{-4}	2.54×10^{-4}
Cr-48	5.99×10^{-4}	4.86×10^{-3}	1.51×10^{-4}	2.76×10^{-3}
Fe-52	3.49×10^{-3}	1.69×10^{-2}	3.39×10^{-3}	1.69×10^{-2}
Ni-56	1.12×10^{-2}	4.63×10^{-2}	1.12×10^{-2}	4.46×10^{-2}
Zn-60	2.33×10^{-5}	9.44×10^{-5}	2.30×10^{-5}	8.80×10^{-5}
Ge-64	2.41×10^{-8}	9.76×10^{-8}	$< 1 \times 10^{-7}$	$< 1 \times 10^{-7}$
Se-68	5.59×10^{-11}	2.24×10^{-10}	$< 1 \times 10^{-7}$	$< 1 \times 10^{-7}$
RMS Err.	1.75×10^{-2}	7.45×10^{-2}	-	-
Steps	22,473	1,095	-	-

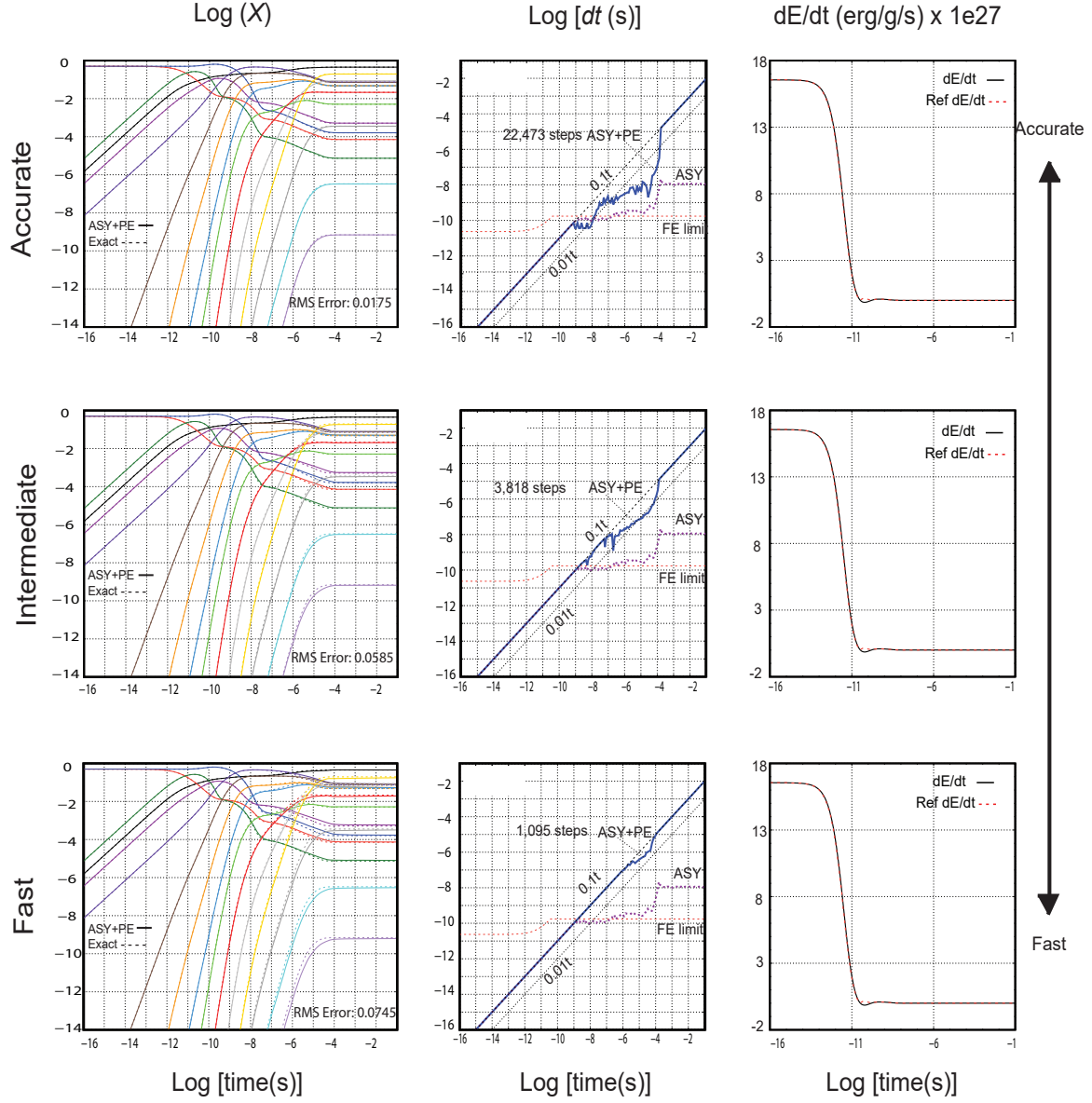


Figure 6.3: Accuracy and speed trade off for the alpha network at constant T_9 and ρ

closely with $U_{Si} = 0.0126$. The remaining notable uncertainties are Helium, $U_{He} = 0.0091$, Sulfur $U_S = 0.0088$ and Iron $U_{Fe} = 0.0052$.

6.3.2 Alpha Network with Hydro Profile

When using hydro profiles, the tolerances can be as strict as needed to accurately calculate the abundances and energies. The approximations used for the kinetic networks in the hydrodynamical code provide a fast and reasonably accurate way to retrieve a temperature and density profile in a given zone. The use of the approximations for the kinetic network can be used with the hydrodynamical simulation to obtain a temperature and density profile. This hydro profile can then be used offline in conjunction with a more accurate calculation of the kinetic network in order to achieve the desired accuracy in the results for the kinetic network. The hydro profile's use for this calculation show that the algebraic approximations and the FERN code can solve stiff networks that involve varying temperatures and densities.

The alpha calculation with a hydrodynamical profile showed some noticeable error in the fast calculation. Not too surprisingly, ^{56}Ni had the largest absolute error with the hydro profile added to the calculation and it contributed the most to the total error when the error in the network was at a maximum. The times for the largest error for the accurate and fast calculations are at $t = 8.55 \times 10^{-5}\text{s}$ and $t = 9.63 \times 10^{-5}\text{s}$ respectively. Table 6.3 shows the residuals of each isotope in the same way table 6.2 did. The first grouping shows each isotope's maximum absolute error that it reached in the calculation for both the accurate and fast cases. The second grouping shows each isotope's absolute error at the time when the total error in the network was at a maximum for both calculations.

Figure 6.5 shows the summarized analysis of the alpha network calculations again for the accurate, medium, and fast calculations. The tradeoff of accuracy in favor of speed is much more apparent in this calculation, but the root mean square of the errors suggests that not much error is introduced from the approximations. The steps taken has decreased by a factor of about 20 while the RMS has increased by less than a factor of 3. The RMS error reveals the accuracy in the network with these approximations is still sufficient when paired with a hydrodynamical calculation.

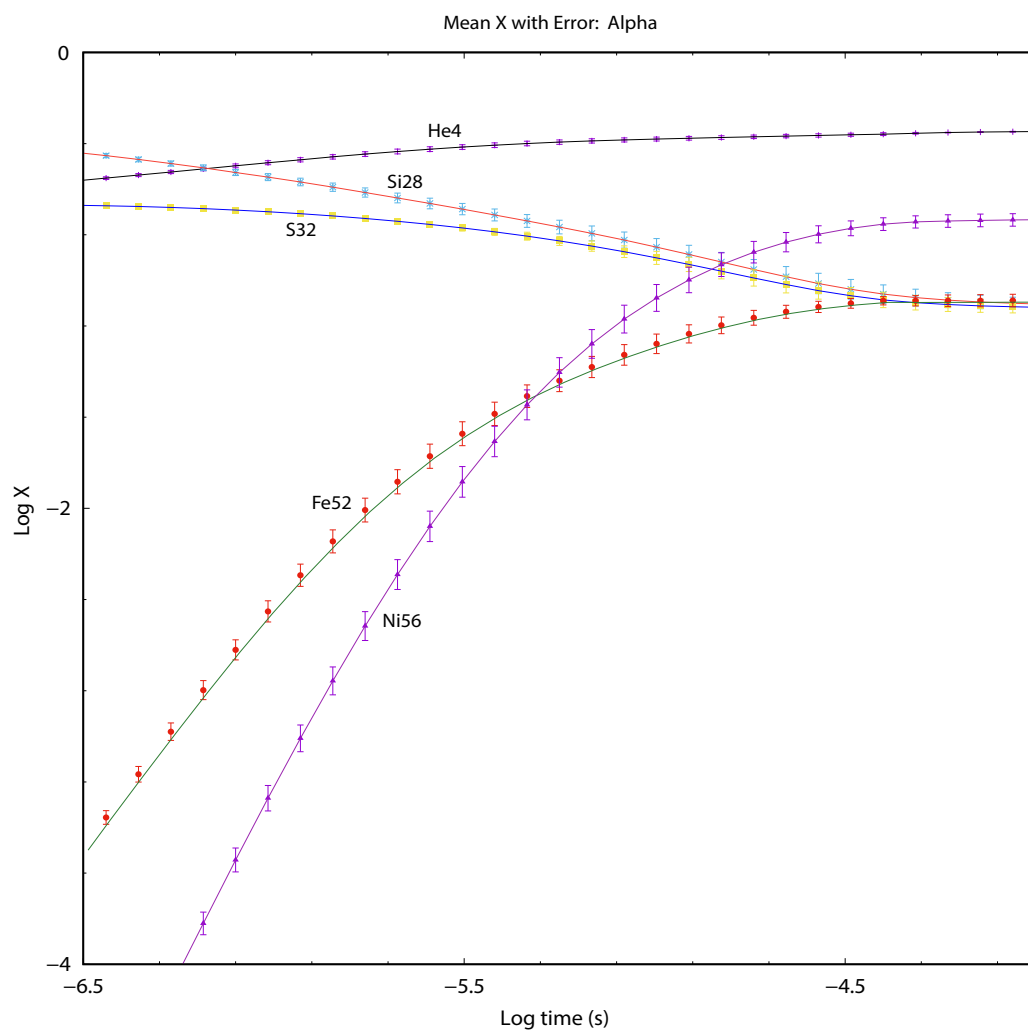


Figure 6.4: Alpha network mean mass fractions shown with error bars indicated by the dots. The 5 isotopes shown had high enough uncertainties where showing error bars was reasonable. For the remaining isotopes not shown, the uncertainties were too low for error bars to show any additional information.

Table 6.3: Alpha Network Residuals with Hydro Profile

Isotope	Max. Error (Acc.)	Max. Error (Fast)	Error at $t_{\max}$ (Acc)	Error at $t_{\max}$ (Fast)
He-4	6.38×10^{-3}	1.21×10^{-2}	4.16×10^{-3}	1.04×10^{-2}
C-12	3.68×10^{-5}	1.32×10^{-3}	8.00×10^{-6}	1.20×10^{-5}
O-16	1.51×10^{-4}	4.15×10^{-3}	2.20×10^{-5}	5.80×10^{-5}
Ne-20	4.02×10^{-6}	1.21×10^{-3}	1.00×10^{-6}	1.00×10^{-6}
Mg-24	2.83×10^{-4}	2.62×10^{-2}	7.70×10^{-5}	2.04×10^{-4}
Si-28	2.43×10^{-2}	3.77×10^{-2}	1.41×10^{-2}	3.70×10^{-2}
S-32	1.26×10^{-2}	3.23×10^{-2}	1.26×10^{-2}	3.23×10^{-2}
Ar-36	7.69×10^{-3}	1.75×10^{-2}	7.11×10^{-3}	1.72×10^{-2}
Ca-40	1.04×10^{-2}	1.90×10^{-2}	6.19×10^{-3}	1.36×10^{-2}
Ti-44	8.28×10^{-4}	4.25×10^{-3}	4.08×10^{-4}	8.26×10^{-4}
Cr-48	7.21×10^{-4}	5.65×10^{-3}	1.88×10^{-4}	1.15×10^{-3}
Fe-52	6.71×10^{-3}	2.14×10^{-2}	6.34×10^{-3}	1.77×10^{-2}
Ni-56	2.85×10^{-2}	7.04×10^{-2}	2.85×10^{-2}	7.04×10^{-2}
Zn-60	4.09×10^{-5}	1.01×10^{-4}	3.90×10^{-5}	9.80×10^{-5}
Ge-64	3.21×10^{-8}	7.97×10^{-8}	$< 1 \times 10^{-7}$	$< 1 \times 10^{-7}$
Se-68	5.21×10^{-11}	1.33×10^{-10}	$< 1 \times 10^{-7}$	$< 1 \times 10^{-7}$
RMS Err.	3.63×10^{-2}	9.09×10^{-2}	-	-
Steps	25,371	1,348	-	-

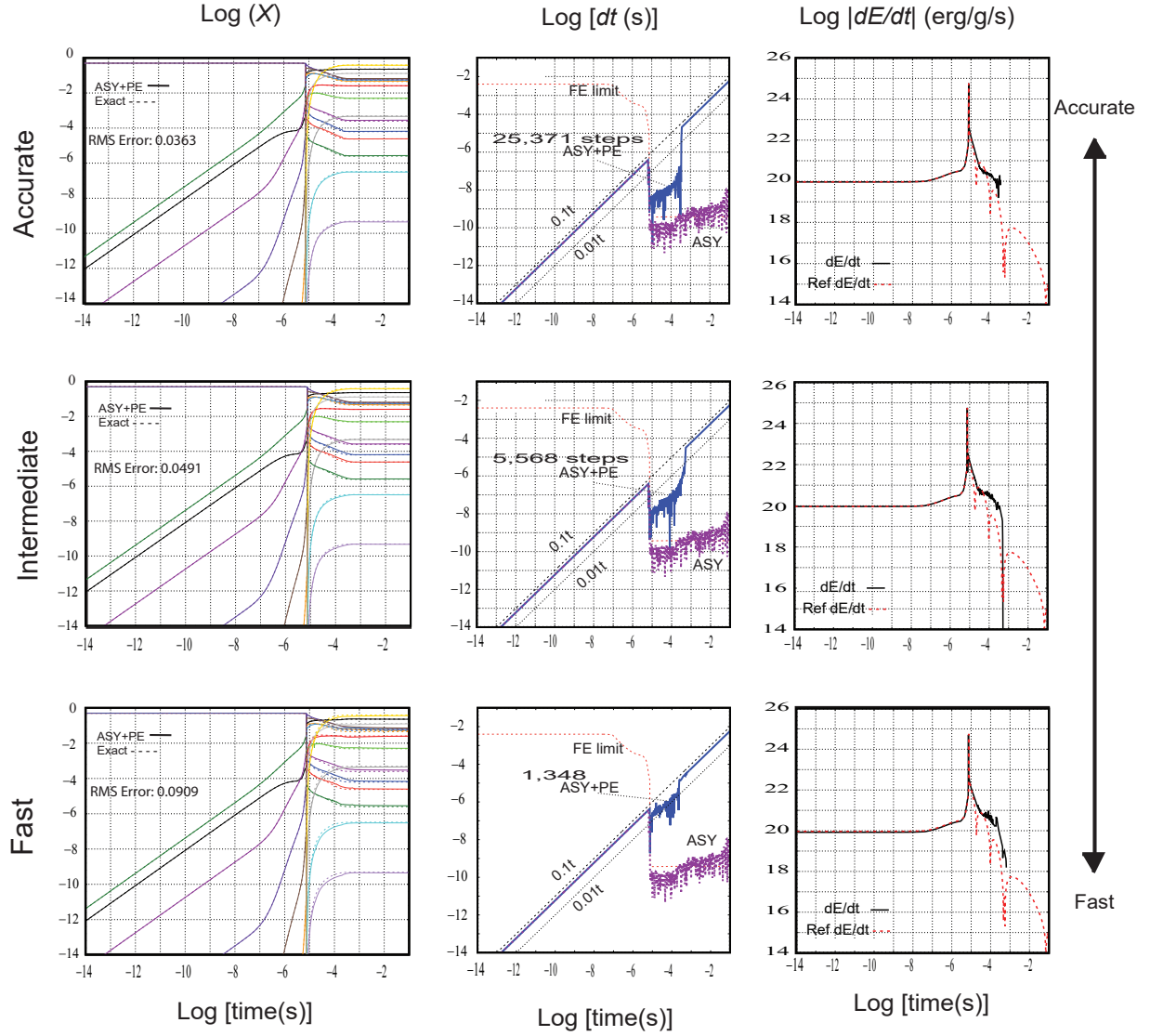


Figure 6.5: Accuracy and speed trade off for the alpha network with the hydro profile of figure 5.6 giving a varying T_9 and ρ .

The three isotopes ^{28}Si , ^{32}S , and ^{56}Ni all have the highest uncertainties. This is not surprising based on how high the absolute errors were for those isotopes. Additionally, ^{24}Mg and ^{52}Fe had uncertainties close to that of ^{32}S . The largest uncertainty is seen in ^{56}Ni with $U_{\text{Ni}} = 0.0205$ and in ^{28}Si with $U_{\text{Si}} = 0.0110$. The remaining three isotopes of significance had uncertainties of $U_{\text{S}} = 0.0094$, $U_{\text{Mg}} = 0.0087$, and $U_{\text{Fe}} = 0.0069$. These five isotopes have the mass fractions plotted with their uncertainties in the error bars in figure 6.6. Calcium and Argon have both been added to the plot since their uncertainties were large enough for error bars to be visible in the plot: $U_{\text{Ca}} = 0.006$ and $U_{\text{Ar}} = 0.0052$.

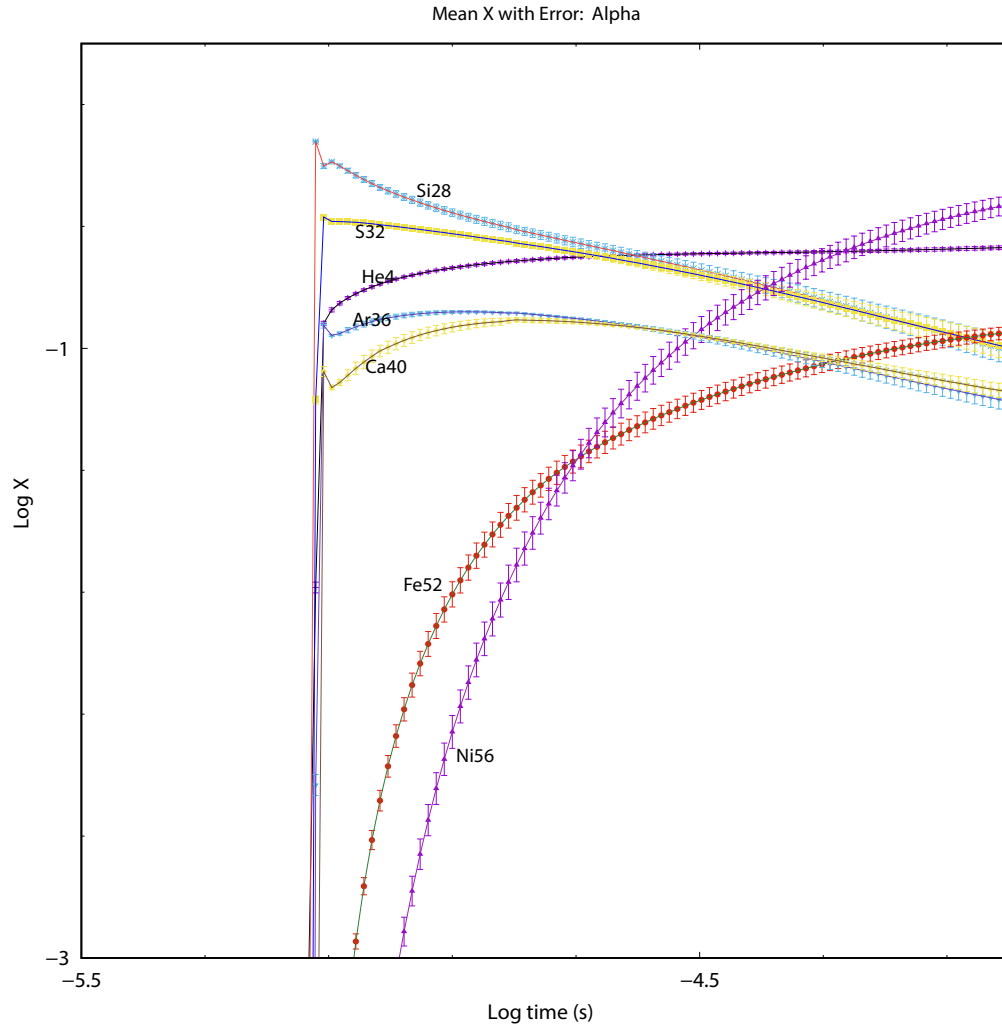


Figure 6.6: Alpha network calculation with the Viktor hydro profile in fig 5.6 of 7 isotopes plotted with error bars. The isotopes not shown did not have an uncertainty high enough to add any new information.

Chapter 7

Conclusion

The new FERN code has improved upon previous implementations of explicit algebraic methods with the addition of a newer adaptive time stepping algorithm and a new method for removing stiffness. The new time stepping algorithm has shown it can re-create previous results obtained using the algebraic approximations. FERN has been able to show correct results using Asy+PE on networks as large as the alpha network and the extended CNO network, and has been able to run the nova-134 network where only Asy is needed. The full calculation of a minimally realistic network, like the 150-isotope, still needs to be done and with the addition of the new time stepping algorithm it is more attainable. However, there is still work to be done to ensure the partial equilibrium works correctly on larger networks that contain a number of singlet reaction groups if default reaction libraries are used.

The new method for removing stiffness in cyclic reactions has proven effective but can still be optimized further. The stiffness that has been introduced in the CNO cycle by the β decays has been successfully removed and the populations of the C13 and N15 isotopes have been approximated with another algebraic method. In order to examine this method in the context of a controlled approximation requires an analysis for the calculations of this network similar to that of chapter 6. Having been a very recent contribution to this project, the analysis of the network is not included in this thesis.

The controlled approximation allows the user to have a quantitative way to control the speed and error in a calculation. The speed is quantitatively shown by the number of steps taken in the integration, while the error is shown in different ways. The residuals of the error

show exactly how much each isotope's mass fraction deviates from its exact value when the explicit approximation methods are invoked. The residuals have been used in two different ways to give an estimate of the accuracy. The first way is looking at the each isotope's maximum residual in the entire calculation. The second way is looking at each isotope's contribution to the total error at the time in the calculation where the error across the network is at a maximum. Both ways measure the accuracy of the calculation. The isotope by isotope analysis provides a more in depth look at the behavior of each isotope over the calculation while looking at the maximum error gives the error of the network as a whole.

The root mean square of the error expands upon the use of the residuals. When the error in the network is at a maximum, the error contributed by each isotope is used to find the root mean square of the error. This yields a measure of the accuracy of the entire network with just one value, similar to how the number of steps is used to give the speed in a single value. The values obtained for the root mean square are more useful for the isotopes that contribute the most to the mass of the network (99%) since the RMS is used to describe the network as a whole.

The Uncertainty provides further analysis in to the error between three of the calculations used. The uncertainty is able to show the accuracy for any network on an isotope by isotope basis as well. The important piece for the uncertainty lies in the mean values of the mass fractions between the three calculations: reference, accurate, and fast. The mean mass fraction for each isotope can represent an expected value for the mass fraction when using tolerances between the strict reference case and the loose, fast case. When the mean mass fractions are plotted, the uncertainty is shown in the error bars. The error bars on the plots with the mean mass fractions are the range of values which the isotopes may obtain depending on the tolerances used. Any tolerances that are between the reference and fast would yield mass fractions within the error bars. The uncertainties may be used to determine how accurate any isotope is calculated and whether or not the level of accuracy is acceptable.

These analyses have been done for networks which have been successfully tested with the Asy+PE approximations. This analysis is still possible and useful for those networks that only require the Asy approximations, such as the nova-134 network. The results from these analyses show how accurate the calculations are when speed in the calculation is prioritized.

The analysis done for each network by looking at the residuals and uncertainties of each isotope is a new way to judge the effectiveness of the approximations. The increase in speed for the calculations has been shown in previous works, but the accuracy has not been quantified in the same way. With the addition of the network analysis, both the speed and accuracy are now quantifiable, leading to the possibility of a robust, quantitatively controlled approximation for coupling realistic networks to fluid dynamics. Further analysis can be done, especially once the partial equilibrium has been successfully tested for larger networks. Additionally, as mentioned before the CNO-extended network will need to be analyzed to quantify how well the new approximation method for stiffness in cycles works. So far, the triple alpha and alpha networks have shown promising results in the ability of FERN to run calculations prioritizing either speed or accuracy. Having a quantifiable trade off in error for speed will let users choose the desired accuracy and speed for calculations of kinetic networks that couple to realistic fluid dynamic simulations.

References

- [1] Guidry, M. (2012). Algebraic stabilization of explicit numerical integration for extremely stiff reaction networks. *Journal of Computational Physics*, 231(16), 5266–5288. <https://doi.org/10.1016/j.jcp.2012.04.026>
- [2] Guidry, M. W., Billings, J. J., & Hix, W. R. (2013). Explicit integration of extremely stiff reaction networks: partial equilibrium methods. *Computational Science & Discovery*, 6(1), 015003. <https://doi.org/10.1088/1749-4699/6/1/015003>
- [3] Guidry, M. W., Budiardja, R., Feger, E., Billings, J. J., Hix, W. R., Messer, O. E. B., Roche, K. J., McMahon, E., & He, M. (2013). Explicit integration of extremely stiff reaction networks: asymptotic methods. *Computational Science & Discovery*, 6(1), 015001. <https://doi.org/10.1088/1749-4699/6/1/015001>
- [4] Guidry, M. W., & Harris, J. A. (2013). Explicit integration of extremely stiff reaction networks: quasi-steady-state methods. *Computational Science & Discovery*, 6(1), 015002. <https://doi.org/10.1088/1749-4699/6/1/015002>
- [5] Brock, B., Belt, A., Billings, J. J., & Guidry, M. (2015). Explicit integration with GPU acceleration for large kinetic networks. *Journal of Computational Physics*, 302, 591–602. <https://doi.org/10.1016/j.jcp.2015.09.013>
- [6] Haidar, A., Brock, B., Tomov, S., Guidry, M., Billings, J. J., Shyles, D., & Dongarra, J. (2016). Performance analysis and acceleration of explicit integration for large kinetic networks using batched GPU computations. 2016 IEEE High Performance Extreme Computing Conference (HPEC). <https://doi.org/10.1109/hpec.2016.7761605>
- [7] Johnson, Jennifer. (Ohio State University)
- [8] Rauscher, T., & Thielemann, F.-K. (2000). Astrophysical Reaction Rates From Statistical Model Calculations. *Atomic Data and Nuclear Data Tables*, 75(1-2), 1–351. [https://doi.org/10.1016/S0375-9606\(00\)00001-1](https://doi.org/10.1016/S0375-9606(00)00001-1)

[org/10.1006/adnd.2000.0834](https://doi.org/10.1006/adnd.2000.0834)

- [9] Forward and Backward Euler Methods. (n.d.). Web.mit.edu. https://web.mit.edu/10.001/Web/Course_Notes/Differential_Equations_Notes/node3.html
- [10] Guidry, M. W. (2019). Stars and stellar processes. Cambridge University Press.
- [11] Hix, W. R., & Meyer, B. S. (2006). Thermonuclear kinetics in astrophysics. Nuclear Physics A, 777, 188–207. [https : //doi.org/10.1016/j.nuclphysa.2004.10.009](https://doi.org/10.1016/j.nuclphysa.2004.10.009)
- [12] Timmes, F. X. (1999). Integration of Nuclear Reaction Networks for Stellar Hydrodynamics. The Astrophysical Journal Supplement Series, 124(1), 241–263. <https://doi.org/10.1086/313257>
- [13] D.R. Mott, E.S. Oran, B. van Leer, Differential equations of reaction kinetics, J. Comp. Phys. 164 (2000) 407–428.
- [14] D.R. Mott, New Quasi-Steady-State and Partial-Equilibrium Methods for Integrating Chemically Reacting Systems, doctoral thesis, University of Michigan, 1999.
- [15] D. Mott, E. Oran, B. van Leer, Identifying and Imposing Partial Equilibrium in Chemically Reacting Systems, in: AIAA-2003-667, 41st Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 2003.
- [16] Feger, Elisha Don, "Evaluating Explicit Methods for Solving Astrophysical Nuclear Reaction Networks." PhD diss., University of Tennessee, 2011. https://trace.tennessee.edu/utk_graddiss/1048
- [17] Chupryna, Viktor, "Explicit Methods in the Nuclear Burning Problem for Supernova Ia Models." PhD diss., University of Tennessee, 2008. https://trace.tennessee.edu/utk_graddiss/484
- [18] Uncertainty Formula — Calculation (Examples with Excel Template). (2019, July 14). EDUCBA. <https://www.educba.com/uncertainty-formula/>

Vita

Nick Brey was born in Evanston, Illinois in May of 1996. He received his undergraduate degree from the University of Tennessee in Physics/Astronomy. In 2019 he joined the University of Tennessee graduate program in Physics and Astronomy. During his time in graduate school he competed on the Men's Track and Field team and worked as a Graduate Teaching Assistant for the undergraduate physics labs. He graduated in the Spring of 2023 and begins work with the Earth and Planetary Science Department to earn a doctoral degree in the Spring of 2023 as well.