A DG-IMEX Method for Two-moment Neutrino Transport: Nonlinear Solvers for Neutrino–Matter Coupling[*]

M. Paul Laiu[1] ID, Eirik Endeve[1,2] ID, Ran Chu[2] ID, J. Austin Harris[3] ID, and O. E. Bronson Messer[2,3,4] ID

[1] Multiscale Methods and Dynamics Group, Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
laiump@ornl.gov
[2] Department of Physics and Astronomy, University of Tennessee Knoxville, Knoxville, TN 37996, USA
[3] National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[4] Physics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Abstract

Neutrino–matter interactions play an important role in core-collapse supernova (CCSN) explosions, as they contribute to both lepton number and/or four-momentum exchange between neutrinos and matter and thus act as the agent for neutrino-driven explosions. Due to the multiscale nature of neutrino transport in CCSN simulations, an implicit treatment of neutrino–matter interactions is desired, which requires solutions of coupled nonlinear systems in each step of the time integration scheme. In this paper, we design and compare nonlinear iterative solvers for implicit systems with energy-coupling neutrino–matter interactions commonly used in CCSN simulations. Specifically, we consider electron neutrinos and antineutrinos, which interact with static matter configurations through the Bruenn 85 opacity set. The implicit systems arise from the discretization of a nonrelativistic two-moment model for neutrino transport, which employs the discontinuous Galerkin (DG) method for phase-space discretization and an implicit–explicit (IMEX) time integration scheme. In the context of this DG-IMEX scheme, we propose two approaches to formulate the nonlinear systems: a coupled approach and a nested approach. For each approach, the resulting systems are solved with Anderson-accelerated fixed-point iteration and Newton's method. The performance of these four iterative solvers has been compared on relaxation problems with various degrees of collisionality, as well as proto–neutron star deleptonization problems with several matter profiles adopted from spherically symmetric CCSN simulations. Numerical results suggest that the nested Anderson-accelerated fixed-point solver is more efficient than other tested solvers for solving implicit nonlinear systems with energy-coupling neutrino–matter interactions.

Unified Astronomy Thesaurus concepts: Computational methods (1965); Core-collapse supernovae (304); Radiative transfer simulations (1967); Supernova neutrinos (1666)

1. Introduction

Core-collapse supernovae (CCSNe), the explosive deaths of massive stars, are to a large extent neutrino-driven. About 99% of the gravitational potential energy released in a core-collapse event ($\sim 10^{53}$ erg) is radiated away by neutrinos, which also act as a driver for the expulsion of matter. Near the end of a massive star's life (i.e., a star with mass exceeding about 10 $M_\odot$), its iron core, which does not produce energy by nuclear burning, is held up against gravity by the pressure from degenerate electrons. However, the mass of the iron core continues to increase due to silicon burning on its surface. Once the mass of the iron core reaches about 1.4 $M_\odot$ (the Chandrasekhar limit), the electron degeneracy pressure becomes insufficient in balancing gravity, and the core collapses in on itself. The collapse proceeds until the central rest mass density exceeds the nuclear matter densities ($>10^{14}$ g cm^{-3}), when the matter equation of state (EoS) stiffens to halt the collapse and a shock wave is launched into the collapsing outer core. The outward-propagating shock wave loses energy by dissociating iron nuclei into free nucleons. Furthermore, electron capture on nucleons in the hot matter behind the shock produces copious amounts of neutrinos and antineutrinos, which can escape the system once the shock reaches sufficiently low densities (about 10^{12} g cm^{-3}). The combination of neutrino emission and iron dissociation weakens the shock, which eventually stalls at a radius of about 100–200 km from the center of the star. At this point, the region below the shock can be divided into the *cooling layer* and the *gain layer*, separated by the gain surface, where neutrino heating and cooling balance. In the cooling layer, extending from the surface of the proto–neutron star to the gain surface, there is net energy loss by neutrino emission. At lower densities, in the heating layer extending from the gain surface to the shock surface, there is net heating by neutrinos emanating from below; see the diagram in Figure 1. In the neutrino reheating explosion mechanism (Bethe & Wilson 1985), energy deposition by neutrinos in the heating layer revives the supernova shock wave to disrupt the massive star in a CCSN explosion. This basic description is supported by recent numerical simulations (e.g., Lentz et al. 2015; Melson et al. 2015; Burrows et al. 2020), but the details are more complicated. The CCSN explosion emerges from a complex interplay between between neutrino transport and hydrodynamic (or magnetohydrodynamic) processes playing out within a curved spacetime (see, e.g., Janka 2012; Burrows 2013; Hix et al. 2014; Müller 2016 for reviews).
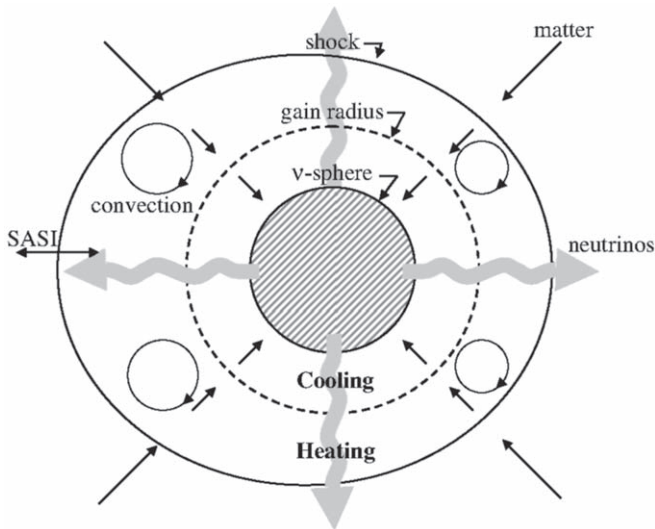

Figure 1. Schematic diagram of the neutrino reheating phase during a CCSN explosion. (This figure is originally from Bruenn et al. 2009 and reproduced here with permission from AIP Publishing.) Neutrinos emanate from the neutrinosphere—a proxy for the surface of the proto–neutron star (gray shaded region)—and the cooling layer above it. Neutrino heating mediated by neutrino–matter interactions in the gain layer between the gain radius and the shock provides energy to reinvigorate the stalled supernova shock. During shock revival, the flow below the shock may be modulated by neutrino-driven convection and the standing accretion shock instability (SASI).

Multiple interaction processes contribute to lepton number and/or four-momentum exchange between neutrinos and matter in the CCSN explosion. Neutrino emission caused by electron capture on nucleons and the inverse process of neutrino capture (absorption) dominate and contribute to both lepton number and four-momentum exchange. During stellar core collapse, electron capture on nuclei is critical to the collapse dynamics, e.g., Hix et al. (2003). Neutrino scattering on electrons and nucleons (and nuclei during collapse) is also a major contributor to the neutrino opacity. In the seminal work of Bruenn (1985), neutrino–nucleon scattering is treated as isoenergetic (the neutrino changes direction but not energy in its encounter with the nucleon), while full four-momentum exchange is considered for neutrino–electron scattering (NES). In later works (e.g., Thompson et al. 2000; Lentz et al. 2012a; Müller et al. 2012; Burrows et al. 2018), it was demonstrated that neutrino–nucleon scattering contributes to neutrino–matter thermalization in a nontrivial way. (Without modern nuclear electron-capture rates, NES also helps shape conditions in the core prior to the formation of the supernova shock; Mezzacappa & Bruenn 1993; Lentz et al. 2012a.) Neutrino pair creation and annihilation via electron–positron pairs (Bruenn 1985) or nucleon–nucleon bremsstrahlung (Hannestad & Raffelt 1998) is yet another significant process in CCSN neutrino transport. In particular, nucleon–nucleon bremsstrahlung is a dominant source of μ and τ neutrinos. Several studies (e.g., Lentz et al. 2012a; Burrows et al. 2018; Just et al. 2018) have investigated the impact of various neutrino opacities on the CCSN explosion mechanism. Although it is difficult to pin down an exact set of necessary interactions, there is a consensus view that processes that couple globally in neutrino momentum space and/or across neutrino species (e.g., NES and pair processes) must be included for a realistic description.

It is computationally expensive to include neutrino transport with satisfactory realism in simulations of CCSNe. First of all,

neutrinos interact relatively weakly with matter in the gain region, which demands a kinetic description. In addition, the hydrodynamics must be modeled without imposed spatial symmetries and with adequate resolution to capture processes that shape the explosion (see, e.g., Müller 2020). The additional requirement of including momentum space coupling neutrino–matter interactions makes realistic, large-scale simulations a computational challenge. For example, if momentum space is discretized with N_p points, a simple evaluation of the collision operator at one spatial location requires $\mathcal{O}(N_p^2)$ operations, as opposed to $\mathcal{O}(N_p)$ operations for the emission, absorption, and isoenergetic scattering operators. In current three-dimensional supernova models, the dimensionality of the neutrino transport problem is reduced by adopting one- or two-moment approaches that retain the neutrino energy dimension of momentum space (Just et al. 2015; Skinner et al. 2019; Bruenn et al. 2020), but this does not completely alleviate the computational challenge of including the critical energy-coupling interactions.

Because the neutrino mean free path can be much shorter than the spatial resolution afforded in parts of the computational domain, an implicit treatment of neutrino–matter interactions is desired. In spherically symmetric (Rampp & Janka 2002; Liebendörfer et al. 2004) and multidimensional models employing the so-called ray-by-ray approximation (Bruenn et al. 2020), the full set of transport equations is commonly solved with implicit time integration. So far, fully implicit time integration has not been the method of choice for truly multidimensional neutrino transport (but see Sumiyoshi & Yamada 2012 for an exception). Instead, implicit–explicit (IMEX) methods (Ascher et al. 1997; Pareschi & Russo 2005) have received more attention (see, e.g., Just et al. 2015; O'Connor 2015; Kuroda et al. 2016; Chu et al. 2019; Skinner et al. 2019). In the IMEX approach, collisions are treated with implicit methods, while phase-space advection is treated with explicit methods (implicit integration for the momentum space advection terms is also used; Kuroda et al. 2016). Explicit integration for phase-space advection is advantageous because it avoids solving a distributed, sparse system of nonlinear equations, and, since neutrino–matter interactions are completely local in space, the implicit part is embarrassingly parallel. Moreover, the characteristic wave speeds associated with the transport and hydrodynamics equations are not too dissimilar in relativistic systems, and it is not clear that the expected additional cost of solving the full transport equation with implicit time integration and a larger time step will pay off.

Energy-coupling neutrino–matter interactions still dominate the computational cost of IMEX-based neutrino transport schemes employing a spectral two-moment approach, where the neutrino energy domain is discretized with N_ε points. The cost per spatial point is expected to scale roughly as $\mathcal{O}(N_\varepsilon^p)$, where the power p is between 2 and 3. The lower bound ($p = 2$) is motivated by the expected cost of simply evaluating the energy-coupling operators, while the $p = 3$ scaling is due to the cost of inverting the Jacobian matrix resulting from an implicit solution algorithm based on Newton's method. In addition, the neutrino–matter interaction rates depend nonlinearly on the local thermodynamic properties of the matter, so that in a fully implicit approach (Kuroda et al. 2016), the local computational cost scales linearly with the number of iterations needed to reach convergence.

Several studies have investigated approximations that are motivated by potential gains in computational expediency. To alleviate the cost of including energy-coupling neutrino–matter interactions, Just et al. (2015) investigated the effect of evaluating the collision terms for energy-coupling interactions with matter conditions taken from the known state at high densities (above $10^{11}\,\mathrm{g\,cm^{-3}}$) and using explicit integration at lower densities and found results that were practically identical to a run with a more implicit treatment. Another approach to circumvent stiffness induced by neutrino–matter interactions is to artificially reduce the rates at high densities and use explicit time integration throughout the computational domain (Thompson et al. 2003; O'Connor 2015; Burrows et al. 2018; Just et al. 2018). We also note a simplifying approach to pair processes involving approximating pair annihilation partners by local equilibrium distributions, which essentially renders this interaction local in momentum space (O'Connor 2015). While some approximations have been shown to work well in a limited set of comparisons, they do introduce uncertainties and limit the applicability of the algorithm. Partially for these reasons, we do not follow the approach of altering the kernel or collision operator based on physical insight into the problem. Instead, we seek to develop fully implicit solvers for energy-coupling neutrino–matter interactions for CCSN simulations and related applications.

This paper details the base neutrino transport algorithms implemented in the toolkit for high-order neutrino radiation hydrodynamics (`thornado`[5]), which is being developed for simulations of CCSNe and related problems using the discontinuous Galerkin (DG) method for phase discretization. Specifically, the phase-space discretization in `thornado` is based on the *nodal* DG method (see, e.g., Hesthaven & Warburton 2008). The original DG method was developed by Reed & Hill (1973) for solving neutron transport problems. Since then, it has been extended to the Runge–Kutta DG framework for solving more general hyperbolic partial differential equations (see, e.g., Cockburn & Shu 1989, 1991, 1998, 2001; Cockburn et al. 1989, 1990 for early developments). For more recent developments of DG methods, see Shu (2016) and references therein. The DG methods are particularly attractive for transport problems, since they recover the correct asymptotic behavior in the so-called diffusion limit (Larsen & Morel 1989; Adams 2001). They can also be easily applied to problems with curvilinear coordinates, necessary when solving general relativistic problems (Teukolsky 2016). However, the DG method has so far not been applied to neutrino transport in CCSN models (but see Radice et al. 2013; Endeve et al. 2015; Chu et al. 2019 for applications in simplified settings). The DG method approximates solutions with piecewise local polynomials and tracks the evolution of coefficients associated with the polynomial expansion; thus, it is often referred to as a modal DG method. The nodal DG method uses a particular interpolating polynomial to construct the approximation, which allows it to track the evolution of nodal values at the interpolation points. This special polynomial approximation results in a simple projection operator from the target function to the polynomial space, which enables straightforward parallel implementation of the nodal DG method (Klöckner et al. 2009). The nodal DG method has been used in various applications, including solving kinetic equations (e.g., Xiong et al. 2015; Juno et al. 2018).

In this paper, we design and evaluate nonlinear solvers for neutrino–matter interactions in a two-moment model for neutrino transport within the IMEX framework. This DG-IMEX method is essentially the same as that described by Chu et al. (2019) but extended to include curvilinear spatial coordinates, multiple neutrino species, more realistic interactions, and coupling to a material background governed by a nuclear EoS. Specifically, we consider electron neutrinos and antineutrinos and develop nonlinear solvers for the opacity set of Bruenn (1985), which includes emission and absorption due to electron and neutrino capture on nucleons and nuclei, isoenergetic neutrino scattering off nucleons and nuclei, inelastic NES, and neutrino–antineutrino pair production/annihilation from electron–positron pairs. We use the SFHo EoS (Steiner et al. 2013) in the numerical experiments. The microphysics (neutrino opacities and EoS) has been tabulated by the WEAKLIB library,[6] which also provides routines for access and manipulation (e.g., interpolation and differentiation) of tabulated microphysics data.

Several nonlinear solver strategies are considered for the neutrino–matter coupling problem. As a baseline for comparison, we consider Newton's method, which can offer rapid convergence to the solution if the initial guess is sufficiently close and the objective function sufficiently regular. However, the necessity of using approximate derivatives due to tabulated opacity kernels to form the Jacobian matrix can hamper the convergence speed. In addition, the construction of the Jacobian from tabulated data and the solution of a dense linear system for each iteration is computationally expensive. We consider fixed-point iteration as an alternative to Newton's method. With fixed-point iteration, Jacobian matrix constructions and dense linear system solutions are not necessary, but the rate of convergence can be slow. We employ Anderson acceleration to improve the convergence rate of the fixed-point method. This acceleration technique was first proposed by Anderson (1965) for solving integral equations and has been used to accelerate fixed-point solutions in several applications, including solving radiation-diffusion equations (An et al. 2017), flow problems (Lott et al. 2012), nuclear reactor simulations (Hamilton et al. 2016), and a variety of nonlinear problems (Walker & Ni 2011). Anderson acceleration speeds up the convergence of standard fixed-point iterations by taking an extrapolation step based on the recent iterates that aims to minimize the residual of the new iterate. The convergence properties of Anderson acceleration were analyzed in Toth & Kelley (2015), Kelley (2018), and Evans et al. (2020), which include (i) global convergence on linear problems under the standard contraction assumption, (ii) local convergence on nonlinear problems under assumptions similar to the standard ones for local convergence of Newton's method, and (iii) an improved local convergence rate when applied to linearly converging fixed-point iterations.

Although the Anderson-accelerated fixed-point algorithm is generally faster than Newton's method in the cases investigated in this paper, we find that the computational cost associated with reevaluating the neutrino opacities in each iteration remains relatively high. This observation motivates a nested approach, where the matter quantities are updated in an outer

---

iteration loop, outside an inner iteration loop where the radiation field is iterated to convergence while the matter state is fixed. We consider two nested iteration schemes, both based on Anderson-accelerated fixed-point iteration in the outer loop. In the first case, the inner loop solve is based on Newton's method. In this nested approach, the Jacobian associated with the inner solve can be computed analytically, since the nonlinear functional is considered independent of the matter state. However, the dense linear system to be solved is of similar size as for the fully coupled Newton method. In the second case, the inner loop solve is based on Anderson-accelerated fixed-point iteration. We find that the nested iteration schemes require fewer opacity evaluations and are more efficient than the fully coupled schemes. We also find that the nested scheme with inner Newton iterations requires fewer iterations to converge for the highest mass densities than the nested scheme with inner fixed-point iterations. However, the cost per fixed-point iteration is lower than the cost of solving the dense linear system associated with each iteration in Newton's method. As a result, the nested iteration scheme with fixed-point iteration in both the inner and outer loops is the most efficient of the solvers considered.

We note that the model considered in this paper lacks the physical fidelity of current CCSN models in some key aspects. First, we consider a static fluid and adopt a nonrelativistic model. However, it is well established that both special and general relativistic effects must be included in realistic models (Bruenn et al. 2001; Lentz et al. 2012b; Müller et al. 2012). Second, we only consider electron neutrinos and antineutrinos and do not consider the nucleon–nucleon bremsstrahlung opacity (Hannestad & Raffelt 1998). This process is a dominant source for production of muon and tau neutrinos and antineutrinos (Thompson et al. 2000), which contribute significantly to the total neutrino luminosity from CCSNe. Third, neutrino–nucleon scattering is treated as isoenergetic. Still, since the work of Bruenn (1985), it has been demonstrated that, despite a relatively small energy exchange per neutrino–nucleon interaction, the relatively large cross section (when compared with NES) implies that this scattering process should be treated as inelastic (Reddy et al. 1998; Müller et al. 2012; Burrows et al. 2018). However, as the development of `thornado` matures, we intend to account for this physics and document the performance in future publications. To the best of our knowledge, the present paper documents the most advanced application of the DG method to neutrino transport.

This paper is organized as follows. In Section 2, we present the mathematical model we adopt for neutrino transport; in Section 3, we introduce the DG-IMEX scheme implemented in `thornado`; the nonlinear solvers are detailed in Section 4; and in Section 5, we present numerical results, where we compare the performance of the solvers on (1) relaxation to equilibrium and (2) proto–neutron star deleptonization. We summarize our findings and draw conclusions in Section 6.

2. Neutrino/Antineutrino Transport Equations

2.1. Boltzmann Equation and Neutrino–Matter Interactions

In nuclear astrophysics applications, neutrino transport can be modeled by the Boltzmann equation. In this paper, we consider a nonrelativistic Boltzmann equation,

$$\frac{1}{\mathrm{c}}\partial_t f + \mathcal{T}(f) = \mathcal{C}(f), \quad (1)$$

which governs the particle distribution function $f(\boldsymbol{x}, \boldsymbol{p}, t)$ that describes the density of particles at position $\boldsymbol{x} = (x^1, x^2, x^3) \in \mathbb{R}^3$ with momentum $\boldsymbol{p} = (p^1, p^2, p^3) \in \mathbb{R}^3$ at time $t \in \mathbb{R}^+$, where c is the speed of light. Adopting curvilinear phase-space coordinates, the advection operator $\mathcal{T}$ takes the form (see, e.g., Endeve et al. 2015)

$$\mathcal{T}(f) := \frac{1}{\sqrt{\gamma}}\sum_{i=1}^{3}\partial_{x^i}(\sqrt{\gamma}\,H_{\boldsymbol{x}}^i f) + \frac{1}{\sqrt{\lambda}}\sum_{i=1}^{3}\partial_{p^i}(\sqrt{\lambda}\,H_{\boldsymbol{p}}^i f), \quad (2)$$

where $H_{\boldsymbol{x}}^i f$ and $H_{\boldsymbol{p}}^i f$ denote the position space flux and momentum space flux in the corresponding ith direction, respectively, and γ, λ are the determinants of the position space and momentum space metric tensors, respectively. While the form of the advection operator in Equation (2) holds for more general (nonorthogonal) spacetime and momentum space bases (see, e.g., Cardall et al. 2013a), we restrict ourselves to orthogonal bases in this paper. The position space flux considered in this paper takes the form $H_{\boldsymbol{x}}^i f = (p^i/|\boldsymbol{p}|)f$, which is proportional to the particle propagation direction. We will restrict ourselves to spherical polar momentum coordinates, and in this case, $H_{\boldsymbol{p}}^i f$ can be obtained from Equations (A15) and (A16) in Endeve et al. (2015). The collision operator $\mathcal{C}$ models the interactions between particles and a material background and includes emission, absorption, elastic scattering on nucleons and nuclei, NES, and thermal pair processes from electron–positron creation and annihilation. This is the neutrino opacity set described in Bruenn (1985; see Table 1 therein).

In this work, we consider the transport of electron neutrinos (ν_e) and antineutrinos ($\bar{\nu}_e$), which results in the coupled equations

$$\frac{1}{\mathrm{c}}\partial_t f + \mathcal{T}(f) = \mathcal{C}(f, \bar{f}), \quad (3a)$$

$$\frac{1}{\mathrm{c}}\partial_t \bar{f} + \mathcal{T}(\bar{f}) = \bar{\mathcal{C}}(\bar{f}, f), \quad (3b)$$

where the neutrino and antineutrino distribution functions are denoted with f and $\bar{f}$, respectively, and the collision terms $\mathcal{C}$ and $\bar{\mathcal{C}}$ both depend on f and $\bar{f}$, since they include the thermal pair production and annihilation processes of neutrino–antineutrino pairs. Before giving a detail formulation of the collision terms, we first change to spherical polar momentum coordinates $(\varepsilon, \boldsymbol{\omega})$ and decompose the neutrino three-momentum as $\boldsymbol{p} = \varepsilon\,\boldsymbol{\ell}(\boldsymbol{\omega})$, where $\varepsilon \in \mathbb{R}^+$ denotes neutrino energy and the unit vector $\boldsymbol{\ell}(\boldsymbol{\omega}) \in \mathbb{R}^3$ only depends on the angular direction $\boldsymbol{\omega} \in \mathbb{S}^2$ relative to a local orthonormal basis. The momentum space volume element is then decomposed into $d\boldsymbol{p} = dV_\varepsilon\,d\boldsymbol{\omega}$, where $dV_\varepsilon := \varepsilon^2 d\varepsilon$ is the spherical shell energy volume element and $d\boldsymbol{\omega}$ is the momentum space angular element. With this notation, the particle distribution $f(\boldsymbol{x}, \boldsymbol{p}, t)$ can be written as $f(\boldsymbol{x}, \boldsymbol{\omega}, \varepsilon, t)$. At each $\boldsymbol{x}$ and t, the neutrino

collision operator then takes the form (Bruenn 1985)

$$
\begin{aligned}
\mathcal{C}(f, \bar{f}) = {}& (1 - f(\boldsymbol{\omega}, \varepsilon))\hat{\eta}(\varepsilon) - \hat{\chi}(\varepsilon)f(\boldsymbol{\omega}, \varepsilon) \\
& + \frac{\varepsilon^2}{\mathsf{c}(\mathsf{hc})^3} \int_{\mathbb{S}^2} R^{\mathrm{IS}}(\boldsymbol{\ell} \cdot \boldsymbol{\ell}', \varepsilon)f(\boldsymbol{\omega}', \varepsilon)d\boldsymbol{\omega}' \\
& - \frac{\varepsilon^2}{\mathsf{c}(\mathsf{hc})^3}f(\boldsymbol{\omega}, \varepsilon) \int_{\mathbb{S}^2} R^{\mathrm{IS}}(\boldsymbol{\ell} \cdot \boldsymbol{\ell}', \varepsilon)\, d\boldsymbol{\omega}' \\
& + \frac{1}{\mathsf{c}(\mathsf{hc})^3}(1 - f(\boldsymbol{\omega}, \varepsilon)) \int_{\mathbb{R}^+} \int_{\mathbb{S}^2} \\
& \times R^{\mathrm{IN}}(\boldsymbol{\ell} \cdot \boldsymbol{\ell}', \varepsilon, \varepsilon')f(\boldsymbol{\omega}', \varepsilon')d\boldsymbol{\omega}'\, dV_{\varepsilon'} \\
& - \frac{1}{\mathsf{c}(\mathsf{hc})^3}f(\boldsymbol{\omega}, \varepsilon) \int_{\mathbb{R}^+} \int_{\mathbb{S}^2} R^{\mathrm{OUT}} \\
& \times (\boldsymbol{\ell} \cdot \boldsymbol{\ell}', \varepsilon, \varepsilon')(1 - f(\boldsymbol{\omega}', \varepsilon'))d\boldsymbol{\omega}'\, dV_{\varepsilon'} \\
& + \frac{1}{\mathsf{c}(\mathsf{hc})^3}(1 - f(\boldsymbol{\omega}, \varepsilon)) \int_{\mathbb{R}^+} \int_{\mathbb{S}^2} \\
& \times R^{\mathrm{PR}}(\boldsymbol{\ell} \cdot \boldsymbol{\ell}', \varepsilon, \varepsilon')(1 - \bar{f}(\boldsymbol{\omega}', \varepsilon'))d\boldsymbol{\omega}'\, dV_{\varepsilon'} \\
& - \frac{1}{\mathsf{c}(\mathsf{hc})^3}f(\boldsymbol{\omega}, \varepsilon) \int_{\mathbb{R}^+} \int_{\mathbb{S}^2} \\
& \times R^{\mathrm{AN}}(\boldsymbol{\ell} \cdot \boldsymbol{\ell}', \varepsilon, \varepsilon')\bar{f}(\boldsymbol{\omega}', \varepsilon')d\boldsymbol{\omega}'\, dV_{\varepsilon'},
\end{aligned} \tag{4}
$$

with h the Planck constant, $\hat{\eta}$ the emissivity, $\hat{\chi}$ the absorption opacity, R^{IS} the elastic (isoenergetic) scattering kernel (due to scattering with nucleons and nuclei), R^{IN} and R^{OUT} the NES kernels, and R^{PR} and R^{AN} the thermal production and annihilation kernels due to pair processes. The antineutrino collision operator $\bar{\mathcal{C}}$ is defined analogously with $\hat{\eta}$, $\hat{\chi}$, and the opacity kernels R^{OP}, OP = IS, IN, OUT, PR, AN, replaced by their antineutrino counterparts.

For thermal emission and absorption, we follow the approach in, e.g., Burrows et al. (2006) and rewrite

$$
(1 - f)\hat{\eta} - \hat{\chi} f = \chi(f_0 - f), \tag{5}
$$

where the effective opacity and equilibrium distribution are defined as

$$
\chi := \hat{\eta} + \hat{\chi} \quad \text{and} \quad f_0 := \frac{\hat{\eta}}{\hat{\eta} + \hat{\chi}}, \tag{6}
$$

respectively. Similarly, the antineutrino emission and absorption term is written as $\bar{\chi}(\bar{f}_0 - \bar{f})$, with $\bar{\chi}$ and $\bar{f}_0$ defined analogously. The equilibrium distributions f_0 and $\bar{f}_0$ are given by the Fermi–Dirac distribution, which is isotropic in angle $\boldsymbol{\omega}$. Specifically,

$$
f_0(\varepsilon) = \frac{1}{e^{(\varepsilon - \mu)/\mathsf{k}_\mathsf{B}T} + 1} \quad \text{and} \quad \bar{f}_0(\varepsilon) = \frac{1}{e^{(\varepsilon - \bar\mu)/\mathsf{k}_\mathsf{B}T} + 1}, \tag{7}
$$

where k_B is the Boltzmann constant, T is the matter temperature, $\mu := \mu_{e^-} + (\mu_p - \mu_n)$ is the neutrino chemical potential, and $\bar\mu := \mu_{e^+} - (\mu_p - \mu_n)$ is the antineutrino chemical potential. Here $\mu_{e^{-(+)}}$ is the electron (positron) chemical potential, μ_n is the neutron chemical potential, and μ_p is the proton chemical potential, which are evaluated from an appropriate EoS. Since $\mu_{e^+} = -\mu_{e^-}$, we have $\bar\mu = -\mu$.

Following Bruenn (1985), the neutrino scattering and pair process kernels are approximated with L-term Legendre expansions in the cosine of the scattering angle $\alpha := \boldsymbol{\ell} \cdot \boldsymbol{\ell}'$,

with expansion coefficients Φ depending on the neutrino energy. Specifically,

$$
\begin{aligned}
R^{\mathrm{IS}}(\alpha, \varepsilon) &\approx \sum_{l=0}^{L} \Phi_l^{\mathrm{IS}}(\varepsilon) P_l(\alpha) \quad \text{and} \\
R^{\mathrm{OP}}(\alpha, \varepsilon, \varepsilon') &\approx \sum_{l=0}^{L} \Phi_l^{\mathrm{OP}}(\varepsilon, \varepsilon') P_l(\alpha),
\end{aligned} \tag{8}
$$

where OP = IN, OUT, PR, and AN; P_l denotes the Legendre polynomial of degree l; and the expansion coefficients of degree l are given by

$$
\begin{aligned}
\Phi_l^{\mathrm{IS}}(\varepsilon) &= \frac{1}{C_l} \int_{-1}^{1} R^{\mathrm{IS}}(\alpha, \varepsilon) P_l(\alpha)d\alpha \quad \text{and} \\
\Phi_l^{\mathrm{OP}}(\varepsilon, \varepsilon') &= \frac{1}{C_l} \int_{-1}^{1} R^{\mathrm{OP}}(\alpha, \varepsilon, \varepsilon') P_l(\alpha)d\alpha.
\end{aligned} \tag{9}
$$

Here $C_l := \int_{-1}^{1} P_l^2(\alpha)d\alpha$, $l = 0, \ldots, L$, are normalization constants. In this work, we use $P_0(\alpha) = 1$ and $P_1(\alpha) = \alpha$; thus, the normalization constants are $C_0 = 2$ and $C_1 = \frac{2}{3}$. The antineutrino opacity kernels are approximated analogously. Due to particle conservation and detailed balance (e.g., Cernohorsky 1994), the NES and pair process kernels satisfy

$$
\begin{aligned}
R^{\mathrm{IN}}(\alpha, \varepsilon, \varepsilon') &= R^{\mathrm{OUT}}(\alpha, \varepsilon', \varepsilon), \\
R^{\mathrm{OUT}}(\alpha, \varepsilon, \varepsilon') &= R^{\mathrm{OUT}}(\alpha, \varepsilon', \varepsilon)e^{(\varepsilon - \varepsilon')/\mathsf{k}_\mathsf{B}T}, \\
R^{\mathrm{AN}}(\alpha, \varepsilon, \varepsilon') &= \bar{R}^{\mathrm{AN}}(\alpha, \varepsilon', \varepsilon), \\
R^{\mathrm{PR}}(\alpha, \varepsilon, \varepsilon') &= R^{\mathrm{AN}}(\alpha, \varepsilon, \varepsilon')e^{-(\varepsilon + \varepsilon')/\mathsf{k}_\mathsf{B}T}.
\end{aligned} \tag{10}
$$

Note that these symmetry properties are enforced in the energy space; thus, they are preserved in the angular approximation for the kernels in Equation (8).

In the following, as a first approximation, we will only include the isotropic part of the kernels (i.e., $L = 0$). More realistic treatments would include linear corrections ($L = 1$), while further corrections ($L > 1$) have been shown to result in only minor differences for NES (e.g., Smit & Cernohorsky 1996).

2.2. Angular Moment Equations

The need for high spatial resolution and unconstrained spatial dimensionality, e.g., to capture fluid dynamics in our target applications, renders direct solutions of the Boltzmann equation too expensive. However, neutrino heating rates are sensitive to the neutrino energy distribution, which demands retention of the energy dimension of momentum space. Therefore, to balance computational cost with physical fidelity, we settle for solving for a finite number of angular moments of the distribution function by adopting a two-moment model. Two-moment models are widely used to model neutrino transport in CCSNe (e.g., Just et al. 2015; O'Connor 2015; Kuroda et al. 2016; Roberts et al. 2016; Skinner et al. 2019). In the spectral two-moment model, we solve for the zeroth and first moments of the neutrino distribution function, while the second moments are obtained from a closure procedure. These moments are defined respectively as

$$
\begin{aligned}
\{\mathcal{J}, \mathcal{H}^i, \mathcal{K}^{ij}\}(\boldsymbol{x}, \varepsilon, t) &= \frac{1}{4\pi} \int_{\mathbb{S}^2} f(\boldsymbol{x}, \boldsymbol{\omega}, \varepsilon, t) \\
&\quad \times \{1, \ell^i(\boldsymbol{\omega}), \ell^i(\boldsymbol{\omega})\ell^j(\boldsymbol{\omega})\}\, d\boldsymbol{\omega},
\end{aligned} \tag{11}
$$

with moments for antineutrinos ($\bar{\mathcal{J}}$, $\bar{\mathcal{H}}^i$, and $\bar{\mathcal{K}}^{ij}$) defined analogously. Then, the zeroth moment $\mathcal{J}$ ($\bar{\mathcal{J}}$) is the spectral number density of neutrinos (antineutrinos), the first moment $\mathcal{H}^i$ ($\bar{\mathcal{H}}^i$) is the spectral number flux density of neutrinos (antineutrinos), and the second moment $\mathcal{K}^{ij}$ ($\bar{\mathcal{K}}^{ij}$) is proportional to the spectral pressure tensor of neutrinos (antineutrinos). Since the neutrino distribution function is bounded between 0 and 1, it can be shown that the moments must satisfy the bounds (Larecki & Banach 2011)

$$0 \leqslant \mathcal{J} \leqslant 1 \quad \text{and} \quad |\mathcal{H}| \leqslant (1 - \mathcal{J})\,\mathcal{J}, \qquad (12)$$

where $\mathcal{H} \coloneqq (\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3)$ and $|\mathcal{H}| = \sqrt{\mathcal{H}_i \mathcal{H}^i}$, with $\mathcal{H}_i$ associated to $\mathcal{H}^k$ via the spatial metric γ_{ik}, i.e., $\mathcal{H}_i = \gamma_{ik} \mathcal{H}^k$. (For notational convenience in this section, we sometimes use Einstein's summation convention where repeated Latin indices imply summation from 1 to 3.) Bounds equivalent to those in Equation (12) also hold for the antineutrino moments.

Taking the zeroth and first moments of Equation 3(a) leads to

$$\frac{1}{c}\partial_t \mathcal{J} + \frac{1}{4\pi}\int_{\mathbb{S}^2} \mathcal{T}(f)\, d\omega = \frac{1}{4\pi}\int_{\mathbb{S}^2} \mathcal{C}(f, \bar{f})\, d\omega, \qquad (13a)$$

$$\frac{1}{c}\partial_t \mathcal{H}_j + \frac{1}{4\pi}\int_{\mathbb{S}^2} \mathcal{T}(f)\ell_j\, d\omega = \frac{1}{4\pi}\int_{\mathbb{S}^2} \mathcal{C}(f, \bar{f})\ell_j\, d\omega, \quad (13b)$$

which, after plugging in the definitions of $\mathcal{T}$ and $\mathcal{C}$ in Equations (2) and (4), results in the moment equations for the number density and number flux,

$$\frac{1}{c}\partial_t \mathcal{J} + \frac{1}{\sqrt{\gamma}}\sum_{i=1}^{3}\partial_{x^i}(\sqrt{\gamma}\,\mathcal{H}^i) = \eta_{\text{TOT}} - \chi_{\text{TOT}}\,\mathcal{J}, \qquad (14a)$$

$$\begin{aligned}\frac{1}{c}\partial_t \mathcal{H}_j + \frac{1}{\sqrt{\gamma}}\sum_{i=1}^{3}\partial_{x^i}(\sqrt{\gamma}\,\mathcal{K}^i{}_j) &= \frac{1}{2}\,\mathcal{K}^{ik}\,\partial_{x^j}\gamma_{ik}\\ &\quad - (\chi_{\text{TOT}} + \sigma_{\text{IS}})\mathcal{H}_j,\end{aligned} \qquad (14b)$$

where we have used the facts that the position space flux $p/|p|\,f = \ell(\omega)f$ and that contributions from momentum space fluxes vanish due to boundary conditions. (An analogous set of moment equations is derived for antineutrinos.) Equation (14) holds for general, time-independent curvilinear spatial coordinates encoded in the spatial metric γ_{ij}, including the commonly used Cartesian, spherical polar, and cylindrical coordinates. The metric tensor is used to raise and lower indices on vectors and tensors; e.g., $\mathcal{H}_j = \gamma_{jk}\mathcal{H}^k$, $\mathcal{K}^i{}_j = \gamma_{jk}\mathcal{K}^{ik}$.

Remark 1. We note that Equation (14) can also be obtained from the nonrelativistic (i.e., zero fluid velocity and no gravitational fields) limit of the moment equations in Shibata et al. (2011) and Cardall et al. (2013b; see also the number conservative two-moment model discussed by Mezzacappa et al. 2020; their Equations (123) and (125)). Moreover, when including relativistic effects, one is, among other things, confronted with choosing appropriate momentum space coordinates. The most common (and perhaps most natural) choice is to use momentum space coordinates in the frame of reference of the inertial observer instantaneously comoving with the fluid (i.e., the comoving frame), as opposed to the so-called laboratory frame (see, e.g., Mihalas & Mihalas 1999). (In the absence of fluid motion and gravitational fields, which is assumed here, there is no distinction between the comoving and laboratory frames.) Specifically, the choice of comoving frame momentum coordinates provides the most straightforward framework for describing neutrino–matter interactions, and this is the choice we intend to make when including relativistic effects in the future. However, this choice complicates the advection operator associated with the moment equations, which then includes Doppler and/or gravitational frequency shift terms. The inclusion of these terms is beyond the scope of the present paper but will be considered in a future study.

On the right-hand side of Equation (14), the elastic scattering opacity is given by $\sigma_{\text{IS}} = \frac{4\pi\varepsilon^2}{c(hc)^3}\Phi_0^{\text{IS}}(\varepsilon)$, and we define the total emissivity and total opacity as

$$\begin{aligned}\eta_{\text{TOT}}(\mathcal{J}, \bar{\mathcal{J}}) &= \chi\,\mathcal{J}_0 + \eta_{\text{SC}}(\mathcal{J}) + \eta_{\text{TP}}(\bar{\mathcal{J}}) \quad \text{and}\\ \chi_{\text{TOT}}(\mathcal{J}, \bar{\mathcal{J}}) &= \chi + \chi_{\text{SC}}(\mathcal{J}) + \chi_{\text{TP}}(\bar{\mathcal{J}}),\end{aligned} \qquad (15)$$

where $\mathcal{J}_0 = \frac{1}{4\pi}\int_{\mathbb{S}^2} f_0\, d\omega = f_0$. Let $d\tilde{V}_\varepsilon \coloneqq 4\pi\varepsilon^2 d\varepsilon$ denote the scaled spherical shell energy volume element; then the opacity terms in Equation (15) are defined, respectively, as the scattering emissivity

$$\eta_{\text{SC}}(\mathcal{J}) = \frac{1}{c(hc)^3}\int_{\mathbb{R}^+} \Phi_0^{\text{IN}}(\varepsilon, \varepsilon')\mathcal{J}(\varepsilon')d\tilde{V}_{\varepsilon'}, \qquad (16)$$

the scattering opacity

$$\begin{aligned}\chi_{\text{SC}}(\mathcal{J}) = \frac{1}{c(hc)^3}\int_{\mathbb{R}^+} [\,&\Phi_0^{\text{IN}}(\varepsilon, \varepsilon')\mathcal{J}(\varepsilon')\\ &+ \Phi_0^{\text{OUT}}(\varepsilon, \varepsilon')(1 - \mathcal{J}(\varepsilon'))]\, d\tilde{V}_{\varepsilon'},\end{aligned} \qquad (17)$$

the emissivity due to thermal pair processes

$$\eta_{\text{TP}}(\bar{\mathcal{J}}) = \frac{1}{c(hc)^3}\int_{\mathbb{R}^+} \Phi_0^{\text{PR}}(\varepsilon, \varepsilon')(1 - \bar{\mathcal{J}}(\varepsilon'))d\tilde{V}_{\varepsilon'}, \qquad (18)$$

and the opacity due to thermal pair processes

$$\begin{aligned}\chi_{\text{TP}}(\bar{\mathcal{J}}) = \frac{1}{c(hc)^3}\int_{\mathbb{R}^+} [\,&\Phi_0^{\text{PR}}(\varepsilon, \varepsilon')(1 - \bar{\mathcal{J}}(\varepsilon'))\\ &+ \Phi_0^{\text{AN}}(\varepsilon, \varepsilon')\bar{\mathcal{J}}(\varepsilon')]\, d\tilde{V}_{\varepsilon'}.\end{aligned} \qquad (19)$$

We make the following remarks on Equation (14): (i) the scattering and pair process opacities depend on $\mathcal{J}$ and $\bar{\mathcal{J}}$, respectively, due to the Fermi blocking factors; (ii) since $\Phi_0^{\text{OP}} \geqslant 0$ and $\mathcal{J}, \bar{\mathcal{J}}$ are between 0 and 1, we have $\sigma_{\text{IS}}, \eta_{\text{SC}}, \chi_{\text{SC}}, \eta_{\text{TP}}$, and $\chi_{\text{TP}} \geqslant 0$; and (iii) the emissivities and opacities depend on the neutrino energy ε and local matter states (e.g., density ρ, temperature T, and electron fraction Y_e).

To close the two-moment model in Equation (14), we adopt an algebraic closure of the form (Levermore 1984)

$$\mathcal{K}^i{}_j = \frac{1}{2}[(1 - \psi)\delta^i{}_j + (3\,\psi - 1)\hat{h}^i\,\hat{h}_j\,]\,\mathcal{J}, \qquad (20)$$

where $\hat{h}^i = \mathcal{H}^i/|\mathcal{H}|$ are components of a unit vector parallel to $\mathcal{H}$, and the Eddington factor $\psi = \psi(\mathcal{J}, h)$ with $h = |\mathcal{H}|/\mathcal{J}$. We use the maximum entropy Eddington factor of Cernohorsky & Bludman (1994),

$$\psi(\mathcal{J}, h) = \frac{1}{3} + \frac{2(1 - \mathcal{J})(1 - 2\mathcal{J})}{3}\,\Theta\!\left(\frac{h}{1 - \mathcal{J}}\right), \qquad (21)$$

where the closure polynomial is given by

$$\Theta(x) = \frac{1}{5}(3 - x + 3\,x^2)x^2. \tag{22}$$

We point out that this closure is based on Fermi–Dirac statistics and suitable for designing numerical methods for the two-moment model satisfying the bounds in Equation (12) (e.g., Chu et al. 2019).

For the antineutrino transport (Equation 3(b)), a two-moment model on the antineutrino number density $\bar{\mathcal{J}}$ and number flux $\bar{\mathcal{H}}$ can be derived following similar procedure as in Equations (13)–(19). We then close the resulting two-moment model using a closure analogous to Equation (20). To simplify the notation, we denote the neutrino and antineutrino moments as $\mathcal{M} := (\mathcal{J}, \mathcal{H}, \bar{\mathcal{J}}, \bar{\mathcal{H}})$ and write the coupled two-moment models in operator form as

$$\frac{1}{\mathsf{c}}\partial t\mathcal{M} + \frac{1}{\sqrt{\gamma}}\sum_{i=1}^{3}\partial_{x^i}(\sqrt{\gamma}\,\mathcal{F}^i(\mathcal{M})) = \mathcal{G}(\mathcal{M}) + \mathcal{C}(\mathcal{M}), \tag{23}$$

where $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{C}$ are the position flux operator, geometry source operator, and collision operator, respectively. In particular,

$$\begin{aligned}\mathcal{C}(\mathcal{M}) := \ &(\eta_{\mathrm{TOT}} - \chi_{\mathrm{TOT}}\,\mathcal{J}, (\chi_{\mathrm{TOT}} + \sigma_{\mathrm{IS}})\mathcal{H}, \bar{\eta}_{\mathrm{TOT}} \\ &- \bar{\chi}_{\mathrm{TOT}}\,\bar{\mathcal{J}}, (\bar{\chi}_{\mathrm{TOT}} + \bar{\sigma}_{\mathrm{IS}})\bar{\mathcal{H}}).\end{aligned} \tag{24}$$

2.3. Coupling to the Matter Equations

Neutrino–matter interactions mediate the exchange of lepton number, momentum, and energy between matter and neutrinos. As a first approximation, we assume that the fluid remains static, and momentum exchange is ignored. Under these assumptions, the matter is described by the mass density $\rho(\boldsymbol{x})$ (fixed in time), temperature $T(\boldsymbol{x}, t)$, and electron fraction $Y_e(\boldsymbol{x}, t)$, and neutrino–matter interactions result in changes to the electron fraction,

$$\begin{aligned}\frac{1}{\mathsf{c}}\frac{dY_e}{dt} = \ &-\left(\frac{\mathsf{m_b}}{\rho}\right)\left(\frac{1}{(\mathsf{hc})^3}\right)\int_{\mathbb{R}^+}(\eta_{\mathrm{TOT}} - \chi_{\mathrm{TOT}}\,\mathcal{J}) \\ &- (\bar{\eta}_{\mathrm{TOT}} - \bar{\chi}_{\mathrm{TOT}}\,\bar{\mathcal{J}})\,d\tilde{V}_\varepsilon,\end{aligned} \tag{25}$$

and specific internal energy,

$$\begin{aligned}\frac{1}{\mathsf{c}}\frac{d\epsilon}{dt} = \ &-\left(\frac{1}{\rho}\right)\left(\frac{1}{(\mathsf{hc})^3}\right)\int_{\mathbb{R}^+}(\eta_{\mathrm{TOT}} - \chi_{\mathrm{TOT}}\,\mathcal{J}) \\ &+ (\bar{\eta}_{\mathrm{TOT}} - \bar{\chi}_{\mathrm{TOT}}\,\bar{\mathcal{J}})\varepsilon\,d\tilde{V}_\varepsilon,\end{aligned} \tag{26}$$

where $\mathsf{m_b}$ is the average baryon mass. The specific internal energy $\epsilon(\boldsymbol{x}, t)$ is defined such that $\rho\epsilon$ is the internal energy density. We note that, given any ρ and Y_e, the specific internal energy ϵ is a one-to-one (injective) function of the temperature T; i.e., one can map a given T to a unique ϵ, and vice versa. Specifically, $(\partial\epsilon/\partial T)_{\rho,Y_e} > 0$.

Together with the number density evolution equations (see Equation (14a)) for neutrinos and antineutrinos, Equations (25)

and (26) lead to the conservation of lepton number

$$\begin{aligned}&\int_D\left[\frac{\rho}{\mathsf{m_b}}\frac{dY_e}{dt} + \frac{1}{(\mathsf{hc})^3}\frac{d}{dt}\int_{\mathbb{R}^+}(\mathcal{J}(\varepsilon) - \bar{\mathcal{J}}(\varepsilon))\,d\tilde{V}_\varepsilon\right]d\boldsymbol{x} \\ &= -\oint_{\partial D}\left[\frac{\mathsf{c}}{(\mathsf{hc})^3}\int_{\mathbb{R}^+}(\mathcal{H} - \bar{\mathcal{H}})\,d\tilde{V}_\varepsilon\right]d\boldsymbol{x}\end{aligned} \tag{27}$$

and energy

$$\begin{aligned}&\int_D\left[\rho\frac{d\epsilon}{dt} + \frac{1}{(\mathsf{hc})^3}\frac{d}{dt}\int_{\mathbb{R}^+}(\mathcal{J}(\varepsilon) + \bar{\mathcal{J}}(\varepsilon))\varepsilon\,d\tilde{V}_\varepsilon\right]d\boldsymbol{x} \\ &= -\oint_{\partial D}\left[\frac{\mathsf{c}}{(\mathsf{hc})^3}\int_{\mathbb{R}^+}(\mathcal{H} + \bar{\mathcal{H}})\varepsilon d\tilde{V}_\varepsilon\right]d\boldsymbol{x},\end{aligned} \tag{28}$$

respectively. Note that $(\rho/\mathsf{m_b})Y_e = n_e$ is the electron density (technically the electron minus positron density), and that neutrinos and electrons have lepton number 1, while antineutrinos and positrons have lepton number -1. Equation (27) implies that the total lepton number in domain D (left-hand side of Equation (27)) only changes due to fluxes through the domain boundary ∂D (right-hand side of Equation (27)). A similar conservation statement holds for the total energy in D as given in Equation (28).

3. DG-IMEX Scheme for Solving the Moment Equations

3.1. Nodal DG Space and Energy Discretization

We apply the nodal DG discretization (see, e.g., Hesthaven & Warburton 2008 for an overview) to Equation (23) in space $\boldsymbol{x} \in \mathbb{R}^3$ and energy $\varepsilon \in \mathbb{R}^+$, in which a logically Cartesian mesh with coordinate aligned elements is considered. To derive the discretized equation from the nodal DG method, we first divide the computational domain $D = D_{\boldsymbol{x}} \times D_\varepsilon \subset \mathbb{R}^3 \times \mathbb{R}^+$ into a disjoint union $\mathcal{K}$ of open elements $\boldsymbol{K}$, where each element takes the form

$$\begin{aligned}\boldsymbol{K} = \ &\{(\boldsymbol{x}, \varepsilon): x^i \in K^i := (x_\mathrm{L}^i, x_\mathrm{H}^i), \\ &i = 1, 2, 3, \varepsilon \in K^\varepsilon := (\varepsilon_\mathrm{L}, \varepsilon_\mathrm{H})\},\end{aligned} \tag{29}$$

with $\Delta x^i = x_\mathrm{H}^i - x_\mathrm{L}^i$ and $\Delta\varepsilon = \varepsilon_\mathrm{H} - \varepsilon_\mathrm{L}$ denoting the side lengths of $\boldsymbol{K}$. For $i = 1, 2, 3$, the spatial surface elements in direction x^i are denoted as $\tilde{\boldsymbol{K}}^i = \times_{j\neq i}K^j$, with space variables $\tilde{\boldsymbol{x}}^i := \{x^j \in \boldsymbol{x}: j \neq i\}$ on $\tilde{\boldsymbol{K}}^i$, where $\times$ is the Cartesian product operator. We use $V_{\boldsymbol{K}}$ to denote the proper volume of the element

$$V_{\boldsymbol{K}} = \int_{\boldsymbol{K}} dV, \quad \text{where} \quad dV = \sqrt{\gamma}\,d\boldsymbol{x}\,dV_\varepsilon = \sqrt{\gamma}\,d\boldsymbol{x}\,\varepsilon^2 d\varepsilon. \tag{30}$$

We let the approximation space for the DG method, $\mathbb{V}^{N_{\boldsymbol{x}},N_\varepsilon}$, be constructed from the tensor product of one-dimensional polynomials of maximal degrees $N_{\boldsymbol{x}}$ and N_ε in space and energy, respectively. Note that functions in $\mathbb{V}^{N_{\boldsymbol{x}},N_\varepsilon}$ can be discontinuous across element interfaces. The semidiscrete DG problem is to find $\mathcal{M}_\mathrm{h} \in \mathbb{V}^{N_{\boldsymbol{x}},N_\varepsilon}$ (which approximates $\mathcal{M}$ in
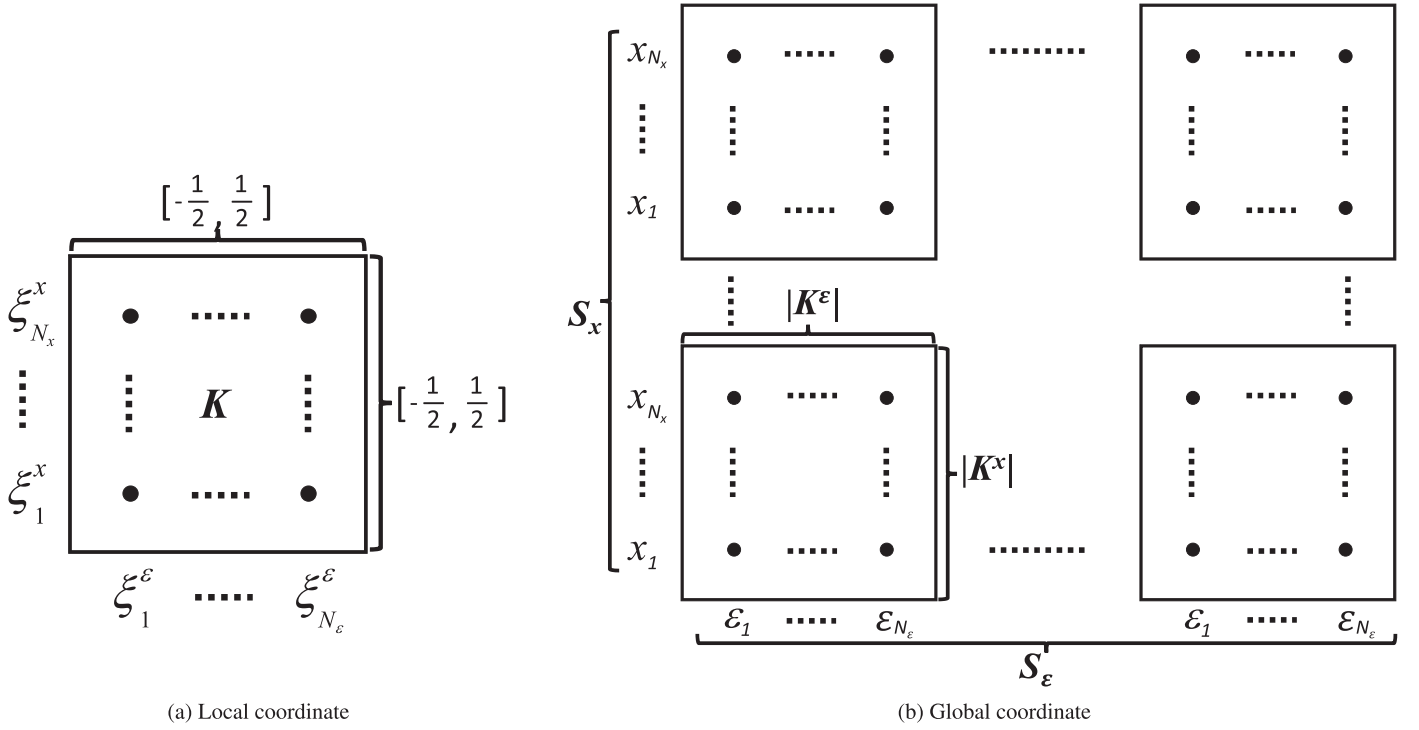
Figure 2. Example of nodal DG elements in a computational domain in $\mathbb{R} \times \mathbb{R}^+$. Figure 2(a) shows the interpolation points $\{\xi_j^i\}_{j=1}^{N_x}$ and $\{\xi_j^\varepsilon\}_{j=1}^{N_\varepsilon}$ in a nodal DG element K in the local coordinate. Figure 2(b) shows the union $\mathcal{K}$ of all elements K and the space and energy nodes—the values of the interpolation points in the global coordinate. Figure 2(b) also illustrates the collection of space nodes S_x and the collection energy nodes S_ε, which are defined in Equations (40) and (41).

Equation (23)) such that (see Cockburn & Shu 2001)

$$
\begin{aligned}
&\frac{1}{\mathsf{c}}\partial_t \int_K \mathcal{M}_\mathrm{h}\, v\, dV + \sum_{i=1}^3 \int_{K^\varepsilon} \int_{\tilde{K}^i} (\sqrt{\gamma}\, \widehat{\mathcal{F}}^i(\mathcal{M}_\mathrm{h}) v|_{x_\mathrm{H}^i} \\
&- \sqrt{\gamma}\, \widehat{\mathcal{F}}^i(\mathcal{M}_\mathrm{h}) v|_{x_\mathrm{L}^i})\, d\tilde{x}^i\, dV_\varepsilon \\
&- \sum_{i=1}^3 \int_K \left(\mathcal{F}^i(\mathcal{M}_\mathrm{h}) \frac{\partial v}{\partial x^i} \right) dV = \int_K \mathcal{G}(\mathcal{M}_\mathrm{h}) v\, dV \\
&+ \int_K \mathcal{C}(\mathcal{M}_\mathrm{h}) v\, dV
\end{aligned}
\tag{31}
$$

for all $v \in \mathbb{V}^{N_x, N_\varepsilon}$ and all $K \in \mathcal{K}$. Here the numerical flux approximating the flux on the spatial surface element $\tilde{K}^i$ is denoted as $\widehat{\mathcal{F}}^i(\mathcal{M}_\mathrm{h})$. In this work, we consider the Lax–Friedrichs (LF) flux,

$$
\begin{aligned}
\widehat{\mathcal{F}}^i(\mathcal{M}_\mathrm{h})|_{x^i} &= f^{\mathrm{LF},i}(\mathcal{M}_\mathrm{h}|_{x^{i,-}}, \mathcal{M}_\mathrm{h}|_{x^{i,+}}) \\
&:= \frac{1}{2}(\, \mathcal{F}^i(\mathcal{M}_\mathrm{h}|_{x^{i,-}}) + \mathcal{F}^i(\mathcal{M}_\mathrm{h}|_{x^{i,+}}) \\
&\quad - \alpha^i(\, \mathcal{M}_\mathrm{h}|_{x^{i,+}} - \mathcal{M}_\mathrm{h}|_{x^{i,-}} \,),
\end{aligned}
\tag{32}
$$

where $\mathcal{M}_\mathrm{h}|_{x^{i,\pm}} = \lim_{\delta \to 0} \mathcal{M}_\mathrm{h}|_{x^i \pm \delta}$ are the evaluations of $\mathcal{M}_\mathrm{h}$ at the immediate right/left of x^i, which thus are functions of $(\tilde{x}^i, \varepsilon, t)$. The parameter $\alpha^i = \|\mathrm{eig}(\partial \mathcal{F}^i / \partial \mathcal{M})\|_\infty$ is the largest eigenvalue of the flux Jacobian. For massless neutrinos, which propagate at the speed of light, we can take $\alpha^i = \mathsf{c}$ (i.e., the global LF flux).

In each element K, we approximate the conserved variables $\mathcal{M}$ by

$$
\begin{aligned}
\mathcal{M}(x, \varepsilon, t) &\approx \mathcal{M}_\mathrm{h}(x, \varepsilon, t) = \sum_{n_\varepsilon = 1}^{N_\varepsilon} \sum_{\boldsymbol{n_x} = 1}^{N_x} \\
&\times \mathcal{M}_{\boldsymbol{n_x}, n_\varepsilon}(t) \ell_{\boldsymbol{n_x}}(x) \ell_{n_\varepsilon}(\varepsilon),
\end{aligned}
\tag{33}
$$

where $\{\ell_{n_\varepsilon}\}_{n_\varepsilon=1}^{N_\varepsilon}$ is chosen to be a collection of Lagrange polynomials in energy of degree N_ε, $\boldsymbol{n_x} := \{n^1, n^2, n^3\}$ is a multi-index that goes from $\boldsymbol{1} = \{1, 1, 1\}$ to $N_x = \{N_x, N_x, N_x\}$, and $\{\ell_{\boldsymbol{n_x}}\}_{\boldsymbol{n_x}=1}^{N_x}$ is the collection of multidimensional polynomials defined as $\ell_{\boldsymbol{n_x}}(x) := \prod_{i=1}^3 \ell_{n^i}(x^i)$, with $\{\ell_{n^i}\}_{n^i=1}^{N_x}$ one-dimensional Lagrange polynomials in x^i of degree N_x. It then follows from these definitions that $\{\ell_{\boldsymbol{n_x}}(x) \ell_{n_\varepsilon}(\varepsilon)\}_{\boldsymbol{n_x}=1, n_\varepsilon=1}^{N_x, N_\varepsilon}$ forms a basis for $\mathbb{V}^{N_x, N_\varepsilon}$ on K. Motivated by the numerical experiments reported in Bassi et al. (2013), we consider here the Lagrange polynomials with Gauss–Legendre interpolation points (instead of Gauss–Legendre–Lobatto points); e.g., ℓ_{n_ε} on interval K^ε is defined as

$$
\ell_{n_\varepsilon}(\varepsilon) = \ell_{N_\varepsilon, n_\varepsilon}^{\mathrm{loc}}(\xi) := \prod_{\substack{j=1 \\ j \neq n_\varepsilon}}^{N_\varepsilon} \frac{\xi - \xi_j^\varepsilon}{\xi_{n_\varepsilon}^\varepsilon - \xi_j^\varepsilon},
\tag{34}
$$

$$
\forall \xi \in I := [-1/2, 1/2], \quad n_\varepsilon = 1, \dots, N_\varepsilon.
$$

The local variable $\xi(\varepsilon) := \dfrac{\varepsilon - \varepsilon_\mathrm{C}}{\Delta \varepsilon}$ with the center $\varepsilon_\mathrm{C} :=$ $(\varepsilon_\mathrm{L} + \varepsilon_\mathrm{H})/2$, and interpolation points $\{\xi_j^\varepsilon\}_{j=1}^{N_\varepsilon}$ are given by the N_ε-point Gauss–Legendre quadrature abscissas on I. On interval K^i, ℓ_{n^i} is defined analogously as $\ell_{n^i}(x^i) = \ell_{N_x, n^i}^{\mathrm{loc}}(\xi)$, with Gauss–Legendre interpolation points $\{\xi_j^i\}_{j=1}^{N_x}$ on I. With

this choice of basis, it follows that the expansion in Equation (33) becomes a nodal representation of $\mathcal{M}_h$, i.e., $\mathcal{M}_h(x_{n_x}, \varepsilon_{n_\varepsilon}, t) = \mathcal{M}_{n_x,n_\varepsilon}(t)$, where $x_{n_x} = \{x_{n^1}^1, x_{n^2}^2, x_{n^3}^3\}$ and $\varepsilon_{n_\varepsilon}$ are the global space and energy variables corresponding to the local interpolation points $\{\xi_j^i\}_{j=1}^{N_x}$ and $\{\xi_j^\varepsilon\}_{j=1}^{N_\varepsilon}$ on element K, respectively. Figure 2 shows an example of nodal DG elements in a reduced space $x \in \mathbb{R}$ and energy $\varepsilon \in \mathbb{R}^+$ with the interpolation points in both local and global coordinates.

We then follow the standard practice and approximate Equation (31) by a semidiscrete system consisting of $(N_x)^3 N_\varepsilon$ equations, each on a nodal value $\mathcal{M}_{n_x,n_\varepsilon}$, $n_x = 1, \dots, N_x$, $n_\varepsilon = 1, \dots, N_\varepsilon$. To derive these equations, we approximate the integrals in Equation (31) using the N_x and N_ε point Gauss–Legendre quadrature rules in space and energy, respectively, with the associated weights $\{w_{n^i}\}_{n^i=1}^{N_x}$ and $\{w_{n_\varepsilon}\}_{n_\varepsilon=1}^{N_\varepsilon}$ normalized such that $\sum_{n^i=1}^{N_x} w_{n^i} = 1$ and $\sum_{n_\varepsilon=1}^{N_\varepsilon} w_{n_\varepsilon} = 1$. Specifically, for $v(x, \varepsilon) = \ell_{n_x}(x)\ell_{n_\varepsilon}(\varepsilon)$, we have

$$\partial_t \int_K \mathcal{M}_h v \, dV \approx w_{n_x,n_\varepsilon} \varepsilon_{n_\varepsilon}^2 \sqrt{\gamma}_{n_x} |K| \, \partial_t \mathcal{M}_{n_x,n_\varepsilon} \qquad (35)$$

and

$$\int_K (\mathcal{G}(\mathcal{M}_h) + \mathcal{C}(\mathcal{M}_h))v \, dV \approx w_{n_x,n_\varepsilon} \\ \times \varepsilon_{n_\varepsilon}^2 \sqrt{\gamma}_{n_x} |K|(\mathcal{G}(\mathcal{M}_h)_{n_x,n_\varepsilon} + \mathcal{C}(\mathcal{M}_h)_{n_x,n_\varepsilon}), \qquad (36)$$

where $|K| = (\prod_{i=1}^3 \Delta x^i)\Delta \varepsilon$, $w_{n_x,n_\varepsilon} = (\prod_{i=1}^3 w_{n^i})w_{n_\varepsilon}$, and γ_{n_x} denotes the value of γ at x_{n_x}. For the surface integrals and "volume terms," we denote the index $\tilde{n}_i := \{n^j \in n_x : j \neq i\}$ and obtain, e.g.,

$$\int_{K^\varepsilon} \int_{\tilde{K}^i} \sqrt{\gamma} \, \widehat{\mathcal{F}}^i(\mathcal{M}_h)v|_{x_H^i} d\tilde{x}^i \, dV_\varepsilon \approx w_{\tilde{n}_i,n_\varepsilon} \\ \times \varepsilon_{n_\varepsilon}^2 (\sqrt{\gamma}_{\tilde{n}_i} \widehat{\mathcal{F}}_{\tilde{n}_i,n_\varepsilon}^i \ell_{n^i})|_{x_H^i} |\tilde{K}^i| |K^\varepsilon| \qquad (37)$$

and

$$\int_K \mathcal{F}^i(\mathcal{M}_h)\frac{\partial v}{\partial x^i} dV \approx w_{\tilde{n}_i,n_\varepsilon} \varepsilon_{n_\varepsilon}^2 \sum_{k^i=1}^{N_x} \\ \times w_{k^i}\left(\sqrt{\gamma}_{\tilde{n}_i} \mathcal{F}_{\tilde{n}_i,n_\varepsilon}^i \frac{\partial \ell_{k^i}}{\partial x^i}\right)\Bigg|_{x_{k^i}^i} |\tilde{K}^i| |K^\varepsilon|, \qquad (38)$$

where $|K^\varepsilon| = \Delta\varepsilon$, $|\tilde{K}^i| = \prod_{j\neq i}\Delta x^j$, $w_{\tilde{n}_i,n_\varepsilon} = (\prod_{j\neq i} w_{n^j})w_{n_\varepsilon}$, $\gamma_{\tilde{n}_i}$ is the evaluation of γ at $\tilde{x}_{\tilde{n}_i}^i$, and $\widehat{\mathcal{F}}_{\tilde{n}_i,n_\varepsilon}^i$ and $\mathcal{F}_{\tilde{n}_i,n_\varepsilon}^i$ are the evaluations of $\widehat{\mathcal{F}}^i(\mathcal{M}_h)$ and $\mathcal{F}^i(\mathcal{M}_h)$ at $(\tilde{x}_{\tilde{n}_i}^i, \varepsilon_{n_\varepsilon})$, respectively. Here $\gamma_{\tilde{n}_i}$, $\widehat{\mathcal{F}}_{\tilde{n}_i,n_\varepsilon}^i$, and $\mathcal{F}_{\tilde{n}_i,n_\varepsilon}^i$ are functions of x^i.

Plugging the terms in Equations (35)–(38) into Equation (31) and dividing through by $w_{n_x,n_\varepsilon} \sqrt{\gamma}_{n_x} \varepsilon_{n_\varepsilon}^2 |K|$ leads to the semidiscrete form of the moment equations when the test function is $v = \ell_{n_x}\ell_{n_\varepsilon}$,

$$\frac{1}{c}\partial_t \mathcal{M}_{n_x,n_\varepsilon} = -\frac{1}{\sqrt{\gamma}_{n_x}}\sum_{i=1}^3 \frac{1}{w_{n^i}\Delta x^i}(\sqrt{\gamma}_{\tilde{n}_i} \widehat{\mathcal{F}}_{\tilde{n}_i,n_\varepsilon}^i \ell_{n^i})|_{x_L^i}^{x_H^i} \\ + \frac{1}{\sqrt{\gamma}_{n_x}}\sum_{i=1}^3 \frac{1}{w_{n^i}\Delta x^i}\sum_{k^i=1}^{N_x} w_{k^i}\left(\sqrt{\gamma}_{\tilde{n}_i} \mathcal{F}_{\tilde{n}_i,n_\varepsilon}^i \frac{\partial \ell_{k^i}}{\partial x^i}\right)\Bigg|_{x_{k^i}^i} \\ + \mathcal{G}(\mathcal{M}_h)_{n_x,n_\varepsilon} + \mathcal{C}(\mathcal{M}_h)_{n_x,n_\varepsilon}, \qquad (39)$$

for $n_x = 1, \dots, N_x$ and $n_\varepsilon = 1, \dots, N_\varepsilon$ in all $K \in \mathcal{K}$. Equation (39) defines the spatial and energy discretization of the moment equations and provides the basis for implementation in thornado. Before discussing the time discretization, we further simplify Equation (39) utilizing the structure of the collision operator $\mathcal{C}$. Here we introduce the collection of space and energy nodes from all elements, denoted as

$$S := \bigcup_{K \in \mathcal{K}}\{(x_{n_x}, \varepsilon_{n_\varepsilon}) \in K : n_x = 1, \dots, N_x, n_\varepsilon = 1, \dots, N_\varepsilon\}, \qquad (40)$$

and we define the spatial component of S as S_x and the energy component of S as S_ε, i.e.,

$$x \in S_x, \; \varepsilon \in S_\varepsilon \quad \text{if and only if} \quad (x, \varepsilon) \in S. \qquad (41)$$

A simplified illustration of S_x and S_ε is given in Figure 2(b), in which $x \in \mathbb{R}$. At each $x_k \in S_x$, the nodal value of $\mathcal{M}_h$ on all $\varepsilon \in S_\varepsilon$ is then denoted as $\mathcal{M}_k(t) := \{\mathcal{M}_h(x_k, \varepsilon_q, t) : \varepsilon_q \in S_\varepsilon\}$. With these notations, we write Equation (39) in the operator form in the remainder of this paper as

$$\frac{1}{c}\partial_t \mathcal{M}_k = \mathsf{F}_M(\mathcal{M}_h)_k + \mathsf{G}_M(\mathcal{M}_h)_k + \mathsf{C}_M(\mathcal{M}_k), \\ \forall k \text{ such that } x_k \in S_x, \qquad (42)$$

where F_M, G_M, and C_M denote, respectively, the discrete position space flux operator, the discrete geometry source, and the discrete collision operator. Here the collision term $\mathsf{C}_M(\mathcal{M}_k)$ depends only on $\mathcal{M}_k$ instead of the full discretized solution $\mathcal{M}_h$, since the physical interactions modeled in the collision operator (see Equation (4)) are independent of the position x while coupled in the energy domain. Specifically, let $(\mathsf{J}_k, \mathsf{H}_k, \bar{\mathsf{J}}_k, \bar{\mathsf{H}}_k)$ denote the discretized moments $\mathcal{M}_k$; then it follows from Equation (24) that

$$\mathsf{C}_M(\mathcal{M}_k) := (\eta_{\text{TOT}} - \chi_{\text{TOT}} \mathsf{J}_k, (\chi_{\text{TOT}} + \sigma_{\text{IS}})\mathsf{H}_k, \\ \bar{\eta}_{\text{TOT}} - \bar{\chi}_{\text{TOT}} \bar{\mathsf{J}}_k, (\bar{\chi}_{\text{TOT}} + \bar{\sigma}_{\text{IS}})\bar{\mathsf{H}}_k), \qquad (43)$$

where σ_{IS} and $\bar{\sigma}_{\text{IS}}$ are the values of σ_{IS} and $\bar{\sigma}_{\text{IS}}$ on the energy nodes S_ε, respectively, and η_{TOT}, χ_{TOT}, $\bar{\eta}_{\text{TOT}}$, and $\bar{\chi}_{\text{TOT}}$ are the discrete counterparts of η_{TOT}, χ_{TOT}, $\bar{\eta}_{\text{TOT}}$, and $\bar{\chi}_{\text{TOT}}$. These discrete opacities are given by replacing the energy integrals in Equations (15)–(19) with the numerical integrals using S_ε as quadrature points, for example,

$$\eta_{\text{SC}}(\mathcal{J}) = \frac{1}{c(hc)^3}\int_{\mathbb{R}^+} \Phi_0^{\text{IN}}(\varepsilon, \varepsilon')\mathcal{J}(\varepsilon')d\tilde{V}_{\varepsilon'} \\ \approx \frac{1}{c(hc)^3}\sum_{q=1}^{|S_\varepsilon|} w_q^\varepsilon \Phi_0^{\text{IN}}(\varepsilon, \varepsilon_q)\mathsf{J}_q =: \eta_{\text{SC}}(\mathsf{J}), \qquad (44)$$

where w_q^ε denotes the weight associated with energy node ε_q, and $|S_\varepsilon|$ denotes the total number of energy nodes in S_ε. Specifically, for $\varepsilon_q \in S_\varepsilon$ in some energy element K_ε,

$$w_q^\varepsilon := 4\pi\varepsilon_q^2 w_{n_\varepsilon}|K_\varepsilon|, \qquad (45)$$

where w_{n_ε} is the local Gauss–Legendre weight at ε_q in K_ε defined earlier in this section.

Finally, we apply this same nodal DG discretization on the electron fraction and specific internal energy evolution (Equations (25) and (26)). Augmenting Equation (42) to the

resulting semidiscrete equations then gives

$$\frac{1}{\mathsf{c}} \partial_t \mathcal{V}_k = \mathsf{F}(\mathcal{V}_\mathrm{h})_k + \mathsf{G}(\mathcal{V}_\mathrm{h})_k + \mathsf{C}(\mathcal{V}_k), \tag{46}$$

$$\forall \boldsymbol{k} \text{ such that } \boldsymbol{x}_k \in S_x,$$

where $\mathcal{V} := (Y_e, \epsilon, \mathcal{M})$, $\mathcal{V}_\mathrm{h} := (Y_{e,\mathrm{h}}, \epsilon_\mathrm{h}, \mathcal{M}_\mathrm{h})$ is the fully discretized version of $\mathcal{V}$, and $\mathcal{V}_k$ denotes the nodal value of $\mathcal{V}_\mathrm{h}$ at point $\boldsymbol{x}_k \in S_x$. Here the operators $\mathsf{F} := (0, 0, \mathsf{F_M})$, $\mathsf{G} := (0, 0, \mathsf{G_M})$, and C is given by the discrete version of the right-hand sides in Equations (25) and (26) and $\mathsf{C_M}$. Specifically, in C, the energy integrals in Equations (25) and (26) are evaluated using a quadrature with the energy nodes $\varepsilon_q \in S_\varepsilon$ as abscissas and weights defined in Equation (45).

Remark 2. In multidimensional CCSN simulations, the use of curvilinear coordinates, e.g., spherical polar coordinates, suffers from an excessively stringent Courant–Friedrichs–Lewy (CFL) time-step restriction due to singularities at the origin and poles (see, e.g., Müller 2020, Section 2.1.1 and references therein). Although these singularities are not an issue in the spherically symmetric CCSN models considered here, they will become a problem when extending the nodal DG scheme to multidimensional CCSN models. Several techniques have been developed to address this issue, such as mesh coarsening (Skinner et al. 2019), element averaging/merging (Asaithambi & Mahesh 2017; Müller et al. 2019), and spectral filtering (Müller et al. 2019). For future multidimensional simulations, we will consider (i) adopting one of the aforementioned techniques or (ii) using Cartesian coordinates in combination with adaptive mesh refinement.

3.2. IMEX Time Integration Scheme

An IMEX time integration scheme (Ascher et al. 1997; Pareschi & Russo 2005) is considered here for solving the two-moment model in Equation (23). When applied to transport equations with collision terms, IMEX schemes usually handle the collision term with an implicit method while applying an explicit method on the advection term (Hu et al. 2018; see also Just et al. 2015; O'Connor 2015; Kuroda et al. 2016; Skinner et al. 2019 for applications of IMEX-type schemes to neutrino transport). This approach relaxes the excessive time-step restriction from an explicit and stiff collision term and avoids the spatially coupled nonlinear solves from an implicit advection term. While the class of IMEX schemes considered in this paper is detailed in Chu et al. (2019), we include it in the following paragraph for completeness. We also stress that even though the nonlinear solution strategies given in Section 4 are motivated from the implicit part of this class of IMEX schemes, they are general enough to be used with any IMEX scheme that treats the collision term implicitly.

To perform time integration, we discretize the time interval $[t_0, t_\mathrm{f}]$ into N time steps, $t_0 = t^0 < t^1 < ... < t^N = t_\mathrm{f}$, and denote $\mathcal{V}_k(t^n)$ as $\mathcal{V}_k^n$, $n = 1, ..., N$. At each spatial node $\boldsymbol{x}_k$ and time t^n, the IMEX scheme integrates the semidiscrete Equation (46) in time via

$$\mathcal{V}_k^{(0)} = \mathcal{V}_k^n, \tag{47a}$$

$$\mathcal{V}_k^{(i)} = \sum_{j=0}^{i-1} c_{ij} \mathcal{V}_k^{(ij)} + a_{ii} \mathsf{c} \, \Delta t \, \mathsf{C}(\mathcal{V}_k^{(i)}), \quad i = 1, ..., s, \tag{47b}$$

$$\mathcal{V}_k^{n+1} = \mathcal{V}_k^{(s)}, \tag{47c}$$

where s is the number of stages, the parameters $c_{ij} \geqslant 0$, $\sum_{j=0}^{i-1} c_{ij} = 1$, $a_{ii} > 0$, and $\mathcal{V}_k^{(ij)}$ is given by an explicit update,

$$\mathcal{V}_k^{(ij)} = \mathcal{V}_k^{(j)} + \hat{c}_{ij} \, \mathsf{c} \, \Delta t \, (\mathsf{F}(\mathcal{V}_\mathrm{h}^{(j)})_k + \mathsf{G}(\mathcal{V}_\mathrm{h}^{(j)})_k), \tag{48}$$

with parameters $\hat{c}_{ij} \geqslant 0$. Thus, in each time step, the s-stage IMEX scheme requires s evaluations of the discrete flux and geometry operators F and G and s inversions of the discrete collision operator C. Here the number of evaluations of F and G is identical to the number of stages due to the fact that, by reusing values of F and G from earlier stages, each stage only requires one additional evaluation of F and G. In general, the inversion of C is the dominant cost in the IMEX scheme. In the remainder of this section, we present the details of the nonlinear system arising from inverting C.

At each stage i of the IMEX scheme, Equation 47(b) can be considered as the nonlinear system

$$\mathcal{V}_k^{(+)} = \mathcal{V}_k^{(*)} + \tau \, \mathsf{C}(\mathcal{V}_k^{(+)}), \tag{49}$$

where $\tau := a_{ii} \mathsf{c} \, \Delta t > 0$ denotes the effective time step, $\mathcal{V}_k^{(*)}$ denotes the weighted sum of explicit updates $\mathcal{V}_k^{(ij)}$, and $\mathcal{V}_k^{(+)}$ denotes the unknown nodal values to be solved.

Let $(Y_{e,k}, \epsilon_k)$ denote the nodal value of (Y_e, ϵ) at $\boldsymbol{x}_k$, and let $(\mathsf{J}_k, \mathsf{H}_k, \bar{\mathsf{J}}_k, \bar{\mathsf{H}}_k)$ denote the discrete moment $\mathcal{M}_k$, which collects the values of moments $(\mathcal{J}, \mathcal{H}, \bar{\mathcal{J}}, \bar{\mathcal{H}})$ at $\boldsymbol{x}_k$ on all energy nodes in S_ε. Equation (49) can then be considered as a nonlinear system on $\mathcal{V}_k := (Y_{e,k}, \epsilon_k, \mathsf{J}_k, \mathsf{H}_k, \bar{\mathsf{J}}_k, \bar{\mathsf{H}}_k)$. For notational simplicity, we suppress all $\boldsymbol{k}$ subscripts when denoting the nodal values of electron fraction, specific internal energy, and moments in the remainder of this paper. It follows from the definition of C that, at each $\boldsymbol{x}_k$, the resulting nonlinear system is given by

$$\mathsf{Y}_e^{(+)} = \mathsf{Y}_e^{(*)} - \tau \, \frac{\mathsf{m_b}}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(2)} (\eta_\mathrm{TOT} - \chi_\mathrm{TOT} \mathsf{J}_q^{(+)}) \\ - (\bar{\eta}_\mathrm{TOT} - \bar{\chi}_\mathrm{TOT} \bar{\mathsf{J}}_q^{(+)}), \tag{50a}$$

$$\epsilon^{(+)} = \epsilon^{(*)} - \tau \, \frac{1}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(3)} (\eta_\mathrm{TOT} - \chi_\mathrm{TOT} \mathsf{J}_q^{(+)}) \\ + (\bar{\eta}_\mathrm{TOT} - \bar{\chi}_\mathrm{TOT} \bar{\mathsf{J}}_q^{(+)}), \tag{50b}$$

$$\mathsf{J}^{(+)} = \mathsf{J}^{(*)} + \tau (\eta_\mathrm{TOT} - \chi_\mathrm{TOT} \mathsf{J}^{(+)}), \tag{50c}$$

$$\bar{\mathsf{J}}^{(+)} = \bar{\mathsf{J}}^{(*)} + \tau (\bar{\eta}_\mathrm{TOT} - \bar{\chi}_\mathrm{TOT} \bar{\mathsf{J}}^{(+)}), \tag{50d}$$

$$\mathsf{H}^{(+)} = \mathsf{H}^{(*)} - \tau (\chi_\mathrm{TOT} + \sigma_\mathrm{IS}) \mathsf{H}^{(+)}, \tag{50e}$$

$$\bar{\mathsf{H}}^{(+)} = \bar{\mathsf{H}}^{(*)} - \tau (\bar{\chi}_\mathrm{TOT} + \bar{\sigma}_\mathrm{IS}) \bar{\mathsf{H}}^{(+)}. \tag{50f}$$

In Equations 50(a) and (b), a quadrature is used to evaluate the energy integrals in Equations (25) and (26), as described when defining C in Equation (46). Here $(\mathsf{J}_q, \bar{\mathsf{J}}_q)$ denotes the value of $(\mathsf{J}, \bar{\mathsf{J}})$ at energy node $\varepsilon_q \in S_\varepsilon$. The physical constants are absorbed into the weights, i.e.,

$$w_q^{(2)} := \frac{1}{(\mathsf{hc})^3} w_q^\varepsilon \quad \text{and} \quad w_q^{(3)} := \frac{1}{(\mathsf{hc})^3} \varepsilon_q w_q^\varepsilon, \tag{51}$$

with w_q^ε defined in Equation (45).

We take a two-step approach to solve Equation (50), which first solves the fully coupled nonlinear system in Equations 50(a)–(d); plug the solution $(Y_e^{(+)}, \epsilon^{(+)}, \mathsf{J}^{(+)}, \bar{\mathsf{J}}^{(+)})$ into Equations 50(e)–(f) to compute $(\chi_\mathrm{TOT}, \sigma_\mathrm{IS}, \bar{\chi}_\mathrm{TOT}, \bar{\sigma}_\mathrm{IS})$;

and then update $(\mathbf{H}^{(+)}, \bar{\mathbf{H}}^{(+)})$. We note that, once $(\chi_{\mathrm{TOT}}, \sigma_{\mathrm{IS}}, \bar{\chi}_{\mathrm{TOT}}, \bar{\sigma}_{\mathrm{IS}})$ are known, solving Equations 50(e)–(f) is straightforward. Thus, we focus on the solution procedure of the coupled system in Equations 50(a)–(d), where the opacities $(\eta_{\mathrm{TOT}}, \chi_{\mathrm{TOT}}, \bar{\eta}_{\mathrm{TOT}}, \bar{\chi}_{\mathrm{TOT}})$ are functions of $(Y_e, \epsilon, J, \bar{J})$. Specifically, while the opacities are written explicitly as functions of $(J, \bar{J})$ in Equations (15)–(19), they also depend on the matter state (Y_e, ϵ) through the opacity kernels Φ_0^{IN}, Φ_0^{OUT}, Φ_0^{PR}, and Φ_0^{AN} in Equations (16)–(19). In a fully implicit approach, these opacities need to be updated in the solution procedure of Equations 50(a)–(d) in order to remain consistent with $(Y_e^{(+)}, \epsilon^{(+)}, J^{(+)}, \bar{J}^{(+)})$.

We proceed to investigate various approaches to solve the coupled nonlinear system in Equations 50(a)–(d).

4. Nonlinear Solution Strategies

In this section, we discuss two approaches for solving the system in Equations 50(a)–(d), which couple the evolution of the matter states (Y_e, ϵ) to the neutrino and antineutrino spectral distributions $(J, \bar{J})$. To start, we first rewrite Equations 50(a)–(d) as

$$Y_e^{(+)} = Y_e^{(*)} - \frac{m_b}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(2)} (J_q^{(+)} - \bar{J}_q^{(+)})$$
$$+ \frac{m_b}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(2)} (J_q^{(*)} - \bar{J}_q^{(*)}), \quad (52a)$$

$$\epsilon^{(+)} = \epsilon^{(*)} - \frac{1}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(3)} (J_q^{(+)} + \bar{J}_q^{(+)})$$
$$+ \frac{1}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(3)} (J_q^{(*)} + \bar{J}_q^{(*)}), \quad (52b)$$

$$J^{(+)} = J^{(*)} + \tau \, \eta_{\mathrm{TOT}}(Y_e^{(+)}, \epsilon^{(+)}, J^{(+)}, \bar{J}^{(+)})$$
$$- \tau \, \chi_{\mathrm{TOT}}(Y_e^{(+)}, \epsilon^{(+)}, J^{(+)}, \bar{J}^{(+)}) \, J^{(+)}, \quad (52c)$$

$$\bar{J}^{(+)} = \bar{J}^{(*)} + \tau \, \bar{\eta}_{\mathrm{TOT}}(Y_e^{(+)}, \epsilon^{(+)}, J^{(+)}, \bar{J}^{(+)})$$
$$- \tau \, \bar{\chi}_{\mathrm{TOT}}(Y_e^{(+)}, \epsilon^{(+)}, J^{(+)}, \bar{J}^{(+)}) \, \bar{J}^{(+)}, \quad (52d)$$

with unknowns $Y_e^{(+)} \in \mathbb{R}$, $\epsilon^{(+)} \in \mathbb{R}$, $J^{(+)} \in \mathbb{R}^{|S_\varepsilon|}$, and $\bar{J}^{(+)} \in \mathbb{R}^{|S_\varepsilon|}$. Here Equations 52(a) and (b) are derived by substituting Equations 50(c) and (d) into the right-hand sides of Equations 50(a) and (b) to remove the explicit dependency on opacities, and Equations 52(c) and (d) are identical to Equations 50(c) and (d), with an explicit expression of the dependency of opacities $(\eta_{\mathrm{TOT}}, \chi_{\mathrm{TOT}}, \bar{\eta}_{\mathrm{TOT}}, \bar{\chi}_{\mathrm{TOT}})$ on matter states (through opacity kernels) and neutrino (antineutrino) distributions. Specifically,

$$\eta_{\mathrm{TOT}}(Y_e, \epsilon, J, \bar{J}) = \chi \, J_0 + \eta_{\mathrm{SC}}(J) + \eta_{\mathrm{TP}}(\bar{J}) \quad \text{and}$$
$$\chi_{\mathrm{TOT}}(Y_e, \epsilon, J, \bar{J}) = \chi + \chi_{\mathrm{SC}}(J) + \chi_{\mathrm{TP}}(\bar{J}), \quad (53)$$

where the discrete opacities χ, χ_{OP}, and η_{OP} are computed as in Equation (44), using opacity kernels Φ^{OP} evaluated at (Y_e, ϵ), and the Fermi–Dirac distribution J_0 evaluated at chemical potential $\mu(Y_e, \epsilon)$ and matter temperature $T(Y_e, \epsilon)$.

We propose two approaches for solving the nonlinear system in Equation (52): a coupled approach and a nested approach. The former directly considers Equation (52) as a fully coupled system, while the latter formulates Equation (52) as a nested

system with Equations 52(a) and (b) in the outer layer and Equations 52(c) and (d) in the inner. Opacity kernel evaluations, i.e., evaluating Φ^{OP} at given (Y_e, ϵ), are needed when solving Equations 52(a) and (b). Since the tabulated opacity kernels are used, evaluating Φ^{OP} requires opacity table interpolations, which are the dominant cost in solving Equation (52). The nested approach aims to reduce the number of opacity kernel evaluations by giving a better prediction of $(J, \bar{J})$ through the inner solver on Equations 52(c) and (d).

We consider a fixed-point iteration method with Anderson acceleration and Newton's method as the nonlinear system solvers in both the coupled and nested approach. When solving the systems considered here, fixed-point methods are often more attractive than Newton's method because they (1) do not require the Jacobian matrix, which can be difficult to compute accurately with tabulated opacities, and (2) avoid inversion of dense linear systems. However, the rate of convergence can be slower for fixed-point methods than that of Newton-based methods. The performance of these two types of solvers on systems arising from each approach is compared in the numerical results reported in Section 5. In the following subsections, we state the coupled fixed-point algorithm (Section 4.1), the coupled Newton's method (Section 4.2), the nested fixed-point algorithm (Section 4.3), and the nested Newton's method (Section 4.4).

4.1. Coupled Fixed-point Algorithm

To simplify the notation, we denote the matter states as $\boldsymbol{u} := (Y_e^{(+)}, \epsilon^{(+)})$, the discretized neutrino and antineutrino distributions as $\boldsymbol{\mathcal{U}} := (J^{(+)}, \bar{J}^{(+)})$, and the collection of all of the unknowns as $\boldsymbol{U} := (\boldsymbol{u}, \boldsymbol{\mathcal{U}})$ in the remainder of the paper. To formulate the system in Equation (52) as a fixed-point problem, we write it as

$$\boldsymbol{U} = \boldsymbol{G}(\boldsymbol{U}) := \begin{pmatrix} \mathbf{g}(\boldsymbol{\mathcal{U}}) \\ \boldsymbol{\mathcal{G}}(\boldsymbol{u}, \boldsymbol{\mathcal{U}}) \end{pmatrix}, \quad (54)$$

where

$$\mathbf{g}(\boldsymbol{\mathcal{U}}) := \begin{pmatrix} -\dfrac{m_b}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(2)} (J_q^{(+)} - \bar{J}_q^{(+)}) + C_{Y_e} \\ -\dfrac{1}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(3)} (J_q^{(+)} + \bar{J}_q^{(+)}) + C_\epsilon \end{pmatrix} \quad (55)$$

with $C_{Y_e} = Y_e^{(*)} + \dfrac{m_b}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(2)} (J_q^{(*)} - \bar{J}_q^{(*)})$, $C_\epsilon = \epsilon^{(*)} + \dfrac{1}{\rho} \sum_{q=1}^{|S_\varepsilon|} w_q^{(3)} (J_q^{(*)} + \bar{J}_q^{(*)})$, and

$$\boldsymbol{\mathcal{G}}(\boldsymbol{u}, \boldsymbol{\mathcal{U}}) := \begin{pmatrix} (J^{(*)} + \tau \, \eta_{\mathrm{TOT}}(\boldsymbol{u}, \boldsymbol{\mathcal{U}})) / (1 + \tau \, \chi_{\mathrm{TOT}}(\boldsymbol{u}, \boldsymbol{\mathcal{U}})) \\ (\bar{J}^{(*)} + \tau \, \bar{\eta}_{\mathrm{TOT}}(\boldsymbol{u}, \boldsymbol{\mathcal{U}})) / (1 + \tau \, \bar{\chi}_{\mathrm{TOT}}(\boldsymbol{u}, \boldsymbol{\mathcal{U}})) \end{pmatrix}. \quad (56)$$

Here $\boldsymbol{\mathcal{G}}$ is an equivalent form of Equations 52(c) and (d), which ensures that $\boldsymbol{G}$ is a contraction map, i.e., the Lipschitz constant of $\boldsymbol{G}$ is strictly less than 1.

The coupled fixed-point algorithm considers Equation (54) as a fixed-point problem with unknowns $\boldsymbol{U}$; e.g., applying the Picard iteration on Equation (54) leads to

$$\boldsymbol{U}^{[k+1]} = \boldsymbol{G}(\boldsymbol{U}^{[k]}), \quad (57)$$

where $U^{[k]}$ denotes the kth iterate of the unknowns U, starting from an initial guess $U^{[0]}$. When G is a contraction mapping, the Picard iteration guarantees that, as $k \to \infty$, the iterate $U^{[k]}$ converges to $U = (\mathsf{Y}_e^{(+)}, \epsilon^{(+)}, \mathsf{J}^{(+)}, \bar{\mathsf{J}}^{(+)})$, the solution to Equation (52). Here the opacities ($\eta, \chi, \bar{\eta}, \bar{\chi}$) in G are updated at each iteration k using $U^{[k]}$ and thus are consistent with the solution. While the Picard iteration guarantees convergence when G is a contraction, the convergence could be slow. To achieve faster convergence, we implement Anderson acceleration (Anderson 1965; Walker & Ni 2011) to solve Equation (54). Anderson acceleration utilizes information from previous iterations to update the unknowns, which is expected to give faster convergence than the Picard iteration but at a cost of additional memory usage. Specifically, in iteration k, Anderson acceleration on the coupled problem first computes the residual

$$r^{[k]} := G(U^{[k]}) - U^{[k]}, \tag{58}$$

then solves a least-squares problem $\alpha^* := argmin_{\alpha \in \mathbb{R}^{m_k+1}} \{\|\sum_{i=0}^{m_k}\alpha_i \, r^{[k-i]}\|_2^2 : \sum_{i=0}^{m_k}\alpha_i = 1\}$ with $m_k := \min\{m, k\}$, and finally updates

$$U^{[k+1]} = \sum_{i=0}^{m_k} \alpha_i^* \, G(U^{[k-i]}). \tag{59}$$

Here the truncation parameter $m \geqslant 0$ is an integer that indicates the "memory" of Anderson acceleration, i.e., the maximum number of residuals kept in memory. When $m = 0$, the solver reduces to Picard iteration. For $m > 0$, Anderson acceleration updates U using a linear combination of the last m_k iterates that leads to the minimum residual. In the numerical tests in Section 5, we use $m = 2$, which we have found to significantly reduce the number of iterations when compared to Picard. For $m = 2$, the additional memory required for Anderson acceleration is small, since each implicit solve is local in space. In addition, the least-squares problem for α_i^* is small and can be written out explicitly or solved using LAPACK's DGELS.

4.2. Coupled Newton's Method

The other solver we considered for the nonlinear coupled system in Equation (52) is Newton's method, which formulates Equation (52) as a root-finding problem,

$$\begin{pmatrix} f(u, \mathcal{U}) \\ \mathcal{F}(u, \mathcal{U}) \end{pmatrix} =: \mathbf{F}(U) = \mathbf{0}, \tag{60}$$

where

$$
f(u, \mathcal{U}) := \begin{pmatrix} f_{\mathsf{Y}_e}(u, \mathcal{U}) \\ f_\epsilon(u, \mathcal{U}) \end{pmatrix}
$$
$$
:= \begin{pmatrix} \mathsf{Y}_e^{(+)} + \dfrac{\mathsf{m_b}}{\rho} \displaystyle\sum_{q=1}^{|S_\varepsilon|} w_q^{(2)}(\mathsf{J}_q^{(+)} - \bar{\mathsf{J}}_q^{(+)}) - C_{\mathsf{Y}_e} \\[2ex] \epsilon^{(+)} + \dfrac{1}{\rho} \displaystyle\sum_{q=1}^{|S_\varepsilon|} w_q^{(3)}(\mathsf{J}_q^{(+)} + \bar{\mathsf{J}}_q^{(+)}) - C_\epsilon \end{pmatrix}
\tag{61}
$$

with constants C_{Y_e} and C_ϵ defined as in Equation (55), and

$$
\mathcal{F}(u, \mathcal{U}) := \begin{pmatrix} \mathcal{F}_\mathsf{J}(u, \mathcal{U}) \\ \mathcal{F}_{\bar{\mathsf{J}}}(u, \mathcal{U}) \end{pmatrix}
$$
$$
:= \begin{pmatrix} (1 + \tau \, \chi_{\text{TOT}}(u, \mathcal{U}))\mathsf{J}^{(+)} - (\mathsf{J}^{(*)} + \tau \, \eta_{\text{TOT}}(u, \mathcal{U})) \\ (1 + \tau \, \bar{\chi}_{\text{TOT}}(u, \mathcal{U}))\bar{\mathsf{J}}^{(+)} - (\bar{\mathsf{J}}^{(*)} + \tau \, \bar{\eta}_{\text{TOT}}(u, \mathcal{U})) \end{pmatrix}.
\tag{62}
$$

Applying Newton's method to solve Equation (60) leads to the following update of U in iteration k,

$$U^{[k+1]} = U^{[k]} + \delta U^{[k]}, \tag{63}$$

where the Newton step is given by

$$\delta U^{[k]} = -J(U^{[k]})^{-1}\,\mathbf{F}(U^{[k]}), \quad \text{with} \quad J(U^{[k]}) := \mathbf{F}'(U^{[k]}). \tag{64}$$

Equation (64) shows that two key components are needed in Newton's method: evaluating the Jacobian J and solving the linear system for the Newton step. For the coupled problem in Equation (60), Jacobian evaluation at a given $\hat{U}$ requires computing $\dfrac{\partial f}{\partial u}$, $\dfrac{\partial f}{\partial \mathcal{U}}$, $\dfrac{\partial \mathcal{F}}{\partial u}$, and $\dfrac{\partial \mathcal{F}}{\partial \mathcal{U}}$ at $(\hat{u}, \hat{\mathcal{U}}) =: \hat{U}$. From Equation (61), it is clear that evaluating $\dfrac{\partial f}{\partial u}$ and $\dfrac{\partial f}{\partial \mathcal{U}}$ at $(\hat{u}, \hat{\mathcal{U}})$ is straightforward with minimal cost. However, it follows from Equation (62) that evaluating $\dfrac{\partial \mathcal{F}}{\partial u}$ and $\dfrac{\partial \mathcal{F}}{\partial \mathcal{U}}$ at $(\hat{u}, \hat{\mathcal{U}})$ requires gradients of the opacities ($\eta_{\text{TOT}}, \chi_{\text{TOT}}, \bar{\eta}_{\text{TOT}}, \bar{\chi}_{\text{TOT}}$) with respect to the matter state u and the neutrino and antineutrino number densities $\mathcal{U}$, respectively. In particular, Equations (15)–(19) and (44) imply that $\dfrac{\partial \Phi^{\text{OP}}}{\partial u}$, which is the gradient of opacity kernels with respect to the matter state, is involved in the computation of $\dfrac{\partial \mathcal{F}}{\partial u}(\hat{u}, \hat{\mathcal{U}})$. As discussed earlier, tabulated opacity kernels are considered in this paper. Thus, we can only obtain an approximate $\dfrac{\partial \Phi^{\text{OP}}}{\partial u}$ from the tabulated quantities. Further, when the opacity kernels are not tabulated in terms of u, the gradient $\dfrac{\partial \Phi^{\text{OP}}}{\partial u}$ has to be approximated using the chain rule. The detailed calculations of these kernel derivatives are given in the Appendix. Once the approximate Jacobian is obtained, the linear solve in Equation (64) is performed via LAPACK's DGESV.

4.3. Nested Fixed-point Algorithm

The next approach we consider is a nested algorithm, which formulates Equation (54) as a nested fixed-point problem with two layers,

$$u = \mathbf{g}(\hat{\mathcal{U}}(u)), \tag{65a}$$

$$\hat{\mathcal{U}}(u) = \mathcal{G}(u, \hat{\mathcal{U}}(u)), \tag{65b}$$

where the outer layer, Equation 65(a), is a fixed-point problem on the matter states u, and the inner layer, Equation 65(b), is on the distributions $\hat{\mathcal{U}}$ for fixed matter states u. These two problems are nested in the sense that evaluating the right-hand side of Equation 65(a) at a given u requires solving

Equation 65(b). For example, applying the Picard iteration to both Equations 65(a) and (b) gives the iterative scheme

$$\boldsymbol{u}^{[k+1]} = \mathbf{g}(\hat{\boldsymbol{\mathcal{U}}}(\boldsymbol{u}^{[k]})), \qquad (66a)$$

where $\hat{\boldsymbol{\mathcal{U}}}(\boldsymbol{u}^{[k]}) = \boldsymbol{\mathcal{U}}^{[k,*]}$, the limit point of the inner Picard iteration

$$\boldsymbol{\mathcal{U}}^{[k,\ell+1]} = \boldsymbol{\mathcal{G}}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k,\ell]}). \qquad (66b)$$

In practice, we use Anderson acceleration with $m = 2$, as described in Section 4.1, to accelerate both the outer and inner solves separately.

The nested approach was considered in Laiu et al. (2020) for relaxing the nonlinear coupling between the electric field and electron concentration when solving implicit systems for semiconductor models. Here the nested approach is motivated by the fact that in solving Equation (54), the most costly part is evaluating the opacity kernels Φ at a given matter state $\boldsymbol{u}$, which is performed in $\boldsymbol{\mathcal{G}}$ whenever $\boldsymbol{u}$ is updated. Therefore, while the coupled approach seems simple and straightforward, the nested structure in Equation (65) justifies the additional complexity by reducing the number of updates (on $\boldsymbol{u}$) in Equation 65(a) via a more accurate distribution update given by solving Equation 65(b) at the current matter state. Note that the matter state $\boldsymbol{u}$ is fixed in the solution procedure of the inner problem in Equation 65(b), which does not require opacity kernel evaluations and results in much cheaper inner iterations.

4.4. Nested Newton's Method

The nested Newton's method formulates the inner layer of the nested system in Equation (65) as a root-finding problem, resulting in

$$\boldsymbol{u} = \mathbf{g}(\hat{\boldsymbol{\mathcal{U}}}(\boldsymbol{u})), \qquad (67a)$$

$$\boldsymbol{\mathcal{F}}(\boldsymbol{u}, \hat{\boldsymbol{\mathcal{U}}}(\boldsymbol{u})) = \mathbf{0}, \qquad (67b)$$

where the outer layer in Equation 67(a) is still a fixed-point problem on $\boldsymbol{u}$, and the inner layer in Equation 67(b) is a root-finding problem on $\hat{\boldsymbol{\mathcal{U}}}$ for fixed $\boldsymbol{u}$, with $\boldsymbol{\mathcal{F}}$ defined in Equation (62). Here Equation 67(a) is solved using Anderson acceleration, and, whenever the right-hand side of Equation 67(a) is evaluated at some given $\boldsymbol{u}$, the inner problem in Equation 67(b) is solved via Newton's method to obtain $\hat{\boldsymbol{\mathcal{U}}}(\boldsymbol{u})$ that is used to evaluate $\mathbf{g}(\hat{\boldsymbol{\mathcal{U}}}(\boldsymbol{u}))$. This nested solver is identical to the nested fixed-point algorithm in Section 4.3, except that the inner problem is solved via Newton's method instead of Anderson acceleration. Specifically, let $\boldsymbol{u}^{[k]}$ be the kth iterate in the outer layer, then $\hat{\boldsymbol{\mathcal{U}}}(\boldsymbol{u}^{[k]}) = \boldsymbol{\mathcal{U}}^{[k,*]}$, which is the limit point of the Newton iterate

$$\boldsymbol{\mathcal{U}}^{[k,\ell+1]} = \boldsymbol{\mathcal{U}}^{[k,\ell]} + \delta\boldsymbol{\mathcal{U}}^{[k,\ell]}, \quad \text{with}$$
$$\delta\boldsymbol{\mathcal{U}}^{[k,\ell]} = -\left[\frac{\partial\boldsymbol{\mathcal{F}}}{\partial\boldsymbol{\mathcal{U}}}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k,\ell]})\right]^{-1} \boldsymbol{\mathcal{F}}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k,\ell]}). \qquad (68)$$

Since the matter state $\boldsymbol{u}$ is fixed in the inner iterations, Equations (15)–(19) and (44) imply that the Jacobian $\dfrac{\partial\boldsymbol{\mathcal{F}}}{\partial\boldsymbol{\mathcal{U}}}$ can be calculated with no additional opacity kernel evaluation, which justifies this nested approach. The reason we choose not to formulate the outer layer as a root-finding problem and solve

it with Newton's method is to avoid the costly opacity kernel gradient approximation discussed in Section 4.2.

5. Numerical Experiments

The four iterative solvers introduced in Section 4 are compared in this section. First, to investigate the iterative solvers in isolation, we report on results obtained on relaxation problems under conditions expected in CCSNe. Then we compare the iterative solvers in the context of the IMEX scheme in Equations (47) and (48) on proto–neutron star deleptonization problems using matter conditions from spherically symmetric CCSN simulations at various times after core bounce.

5.1. Implementation Details

In the numerical tests discussed in this section, the DG-IMEX scheme and the nonlinear solvers are implemented following the specifics below, unless otherwise noted.

1. DG scheme—We consider problems with one spatial dimension (imposing spherical symmetry). For the relaxation problems in Section 5.2, we solve the space-homogeneous problem in Equation (49) for a single spatial element. For the proto–neutron star deleptonization problems in Section 5.3, the spatial domain $r \in [0, 300]$ km is divided into 128 geometrically progressing elements, with the first element of size $\Delta x = 1$ km and the last element of size $\Delta x \approx 4.54$ km. In both tests, the energy domain covering $\varepsilon \in [0, 300]$ MeV is divided into 16 geometrically progressing elements, where the first element has $\Delta\varepsilon = 4$ MeV and the last element has $\Delta\varepsilon \approx 50$ MeV. The spatial and energy DG elements considered here are linear ($k = 1$).

2. IMEX scheme—In the proto–neutron star deleptonization problem in Section 5.3, we use the IMEX scheme in Equations (47) and (48) with two stages ($s = 2$). When written in the so-called Shu–Osher form (as in Equations (47) and (48)), the coefficients are given by Chu et al. (2019):

$$\begin{bmatrix} c_{10} & \\ c_{20} & c_{21} \end{bmatrix} = \begin{bmatrix} 1 & \\ 1/2 & 1/2 \end{bmatrix}, \quad \begin{bmatrix} \hat{c}_{10} & \\ \hat{c}_{20} & \hat{c}_{21} \end{bmatrix} = \begin{bmatrix} 1 & \\ 0 & 1 \end{bmatrix}, \quad \text{and}$$
$$\begin{bmatrix} a_{11} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}. \qquad (69)$$

This scheme consists of two evaluations of the explicit part and two implicit solves. We use the realizability-enforcing limiter in Chu et al. (2019) to enforce realizable moments (see Equation (12)) after each stage.

3. The EoS and neutrino opacity tables—In all tests, we use a tabulated version of the SFHo EoS (Steiner et al. 2013). Thermodynamic (dependent) variables are tabulated as a function of mass density, temperature, and electron fraction (ρ, T, and Y_e). The EoS table covers the ranges $\rho \in [1.66 \times 10^3, 3.16 \times 10^{15}]$ g cm^{-3} using $N_\rho = 185$ points (logarithmically spaced to achieve about 15 points per decade), $T \in [1.16 \times 10^9, 1.84 \times 10^{12}]$ K using $N_T = 81$ points (logarithmically spaced to achieve about 25 points per decade), and $Y_e \in [0.01, 0.6]$ using $N_{Y_e} = 30$ points (linearly spaced). The neutrino opacities

are taken from Bruenn (1985), with all input thermo-dynamic quantities computed with the SFHo EoS. The absorption and scattering opacities (χ and σ_{IS}) are tabulated in terms of the neutrino energy ε in addition to ρ, T, and Y_e (using the same resolution as the EoS table). The neutrino energy range covers $\varepsilon \in [0.1, 300]$ MeV using $N_\varepsilon = 40$ logarithmically spaced points. The NES and pair creation and annihilation kernels are tabulated in terms of neutrino energy pairs ε and ε' (using the same points as for χ and σ_{IS}), T (using the same points as in the EoS table), and the degeneracy parameter $\eta = \mu_e/(\mathsf{k}_{\mathrm{B}}T) \in [1 \times 10^{-3}, \ 2.5 \times 10^{3}]$ using $N_\eta = 60$ logarithmically spaced points. To evaluate dependent variables from the table following, e.g., Mezzacappa & Messer (1999), we use bilinear interpolation (or the higher-dimensional equivalent), while derivatives with respect to any of the independent variables are computed by taking the derivative of the interpolation formula. When interpolating the opacity kernels, we enforce the symmetries in Equation (10).

4. Nonlinear solvers—The four nonlinear solvers are implemented following the description in Section 4, with one exception that, in the implementation, the effective emission and absorption opacity χ in Equation (15) is lagged. In other words, when solving Equation (52), χ is evaluated at the starting matter state ($Y_e^{(*)}$, $\epsilon^{(*)}$) and is not being updated in the solution procedure. We choose to lag χ in the nonlinear solvers to simplify the Jacobian calculation in Newton's method. Since χ is usually varying slowly in time, lagging χ has minimal impact on the solution accuracy. For the fixed-point solvers, χ can be updated at each iteration at a minor additional cost. As mentioned in Section 4, we choose the truncation parameter in Anderson acceleration to be $m = 2$, unless otherwise specified. In this case, the least-squares problem for determining α^* in Equation (59) becomes an inversion of a 3×3 matrix and is thus solved analytically. A numerical justification of this choice of m is given in Section 5.2. In Newton's method, the Jacobian matrix is constructed using the derivatives given in the Appendix.

We also note that Jacobian-free Newton–Krylov (JFNK) methods, where the Newton step is computed by solving an approximate Newton system with a Krylov solver, are not well suited for this problem (see, e.g., Knoll & Keyes 2004 for a comprehensive survey of these methods). The reason is that, in the JFNK methods, one evaluation of the opacity kernels is needed in every Krylov iteration to approximate the Jacobian matrix. Thus, the JFNK methods require several opacity evaluations per Newton iteration, while the standard Newton's method only requires one, which makes the JFNK methods more expensive for solving these problems, where the opacity evaluation is a dominant computational cost.

5. Nonlinear solver initial guess—When solving the coupled nonlinear system in Equation (52), a natural choice of initial guess for the unknowns ($Y_e^{(+)}$, $\epsilon^{(+)}$, $\mathsf{J}^{(+)}$, $\bar{\mathsf{J}}^{(+)}$) is ($Y_e^{(*)}$, $\epsilon^{(*)}$, $\mathsf{J}^{(*)}$, $\bar{\mathsf{J}}^{(*)}$), the weighted sum from explicit steps defined in Equation (49). In this work, we add a "presolve" step that aims to provide a better starting point for the iterative solvers and speed up the computation. Specifically, for a given ($Y_e^{(*)}$, $\epsilon^{(*)}$, $\mathsf{J}^{(*)}$, $\bar{\mathsf{J}}^{(*)}$), the presolve step solves a

subsystem of Equation (52), which is obtained by setting the opacities η_{SC}, η_{TP}, χ_{SC}, and χ_{TP} in Equation (53) to be zero, i.e., with only emission, absorption, and isoenergetic scattering. The solution of this simplified system then serves as the initial guess for ($Y_e^{(+)}$, $\epsilon^{(+)}$, $\mathsf{J}^{(+)}$, $\bar{\mathsf{J}}^{(+)}$) in the nonlinear solvers for the system in Equation (52). This presolve step is computationally inexpensive (no opacity table interpolations are needed, since the emission and absorption opacity χ is lagged as discussed in the previous paragraph) while giving a reasonable initial guess for the full system. We observe that, without the presolve step, the coupled and nested Newton's methods in Sections 4.2 and 4.4 could diverge if the initial guess ($Y_e^{(*)}$, $\epsilon^{(*)}$, $\mathsf{J}^{(*)}$, $\bar{\mathsf{J}}^{(*)}$) is too far away from the solution.

6. Nonlinear solver convergence criteria—We set the convergence criteria for the iterative solvers based on the relative residual of the system in Equation (52) at the current iterate. Specifically, for the coupled solvers in Sections 4.1 and 4.2, the convergence criteria are

$$|f_{Y_e}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k]})| \leqslant \mathtt{tol}\ |Y_e^{[0]}|,$$
$$|f_\epsilon(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k]})| \leqslant \mathtt{tol}\ |\epsilon^{[0]}|, \tag{70a}$$

$$\|\boldsymbol{\mathcal{F}}_{\mathsf{J}}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k]})\| \leqslant \mathtt{tol}\ \|\mathsf{J}^{[0]}\|,$$
$$\|\boldsymbol{\mathcal{F}}_{\bar{\mathsf{J}}}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k]})\| \leqslant \mathtt{tol}\ \|\bar{\mathsf{J}}^{[0]}\|, \tag{70b}$$

where $\mathtt{tol} > 0$ is a constant relative tolerance, and (f_{Y_e}, f_ϵ, $\boldsymbol{\mathcal{F}}_{\mathsf{J}}$, $\boldsymbol{\mathcal{F}}_{\bar{\mathsf{J}}}$) are the residuals as defined in Equations (61) and (62). As for the nested solvers in Sections 4.3 and 4.4, the outer layer (Equations 65(a) and 67(a)) uses the convergence criteria in Equation 70(a), while the convergence criteria in the inner layer (Equations 65(b) and 67(b)) are given by

$$\|\boldsymbol{\mathcal{F}}_{\mathsf{J}}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k,\ell]})\| \leqslant \mathtt{tol}\ \|\mathsf{J}^{[k,0]}\|,$$
$$\|\boldsymbol{\mathcal{F}}_{\bar{\mathsf{J}}}(\boldsymbol{u}^{[k]}, \boldsymbol{\mathcal{U}}^{[k,\ell]})\| \leqslant \mathtt{tol}\ \|\bar{\mathsf{J}}^{[k,0]}\|. \tag{71}$$

When all convergence criteria are satisfied, the solvers return the current iterate as the solution. In all numerical experiments, we choose the norm to be the discrete L^2 norm on the energy domain, i.e., $\|\mathsf{J}\| := (\sum_{q=1}^{|S_\varepsilon|} w_q^{(2)} \mathsf{J}_q^2)^{1/2}$, and we set the relative tolerance to be $\mathtt{tol} = 10^{-8}$.

Remark 3. In the coupled fixed-point and the two nested solvers, the basic version of Anderson acceleration outlined in Section 4.1 is implemented. It is known (see, e.g., Walker & Ni 2011 and references therein) that the iterations in Anderson acceleration may suffer from "stagnation" when the least-squares problem is ill conditioned. Due to the choice of small truncation parameter (e.g., $m = 2$) and the contractive property of the collision operator, we do not encounter the stagnation issue in any of the numerical tests presented in this section. Stagnation may potentially become a practical concern when applying Anderson acceleration to solve more complicated systems, e.g., fully coupled neutrino radiation hydrodynamics. A common approach to mitigate the stagnation issue is to control the condition number of the least-squares problems. This can be achieved by (i) solving a proper reformulation of the least-squares problems using QR factorization (Ni & Walker 2010), (ii) modifying the truncation parameter m adaptively (Yang et al. 2009), and/or (iii) regularizing the least-squares problems (Scieur et al. 2020). Alternatively, one
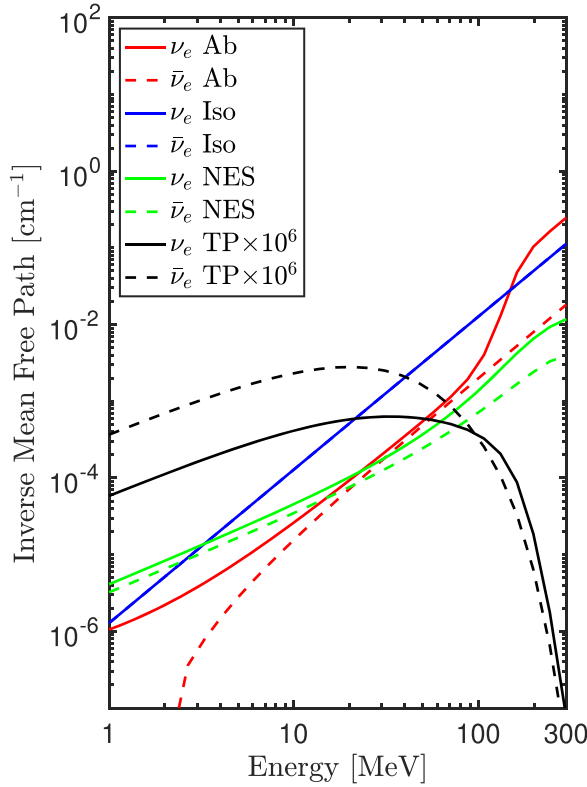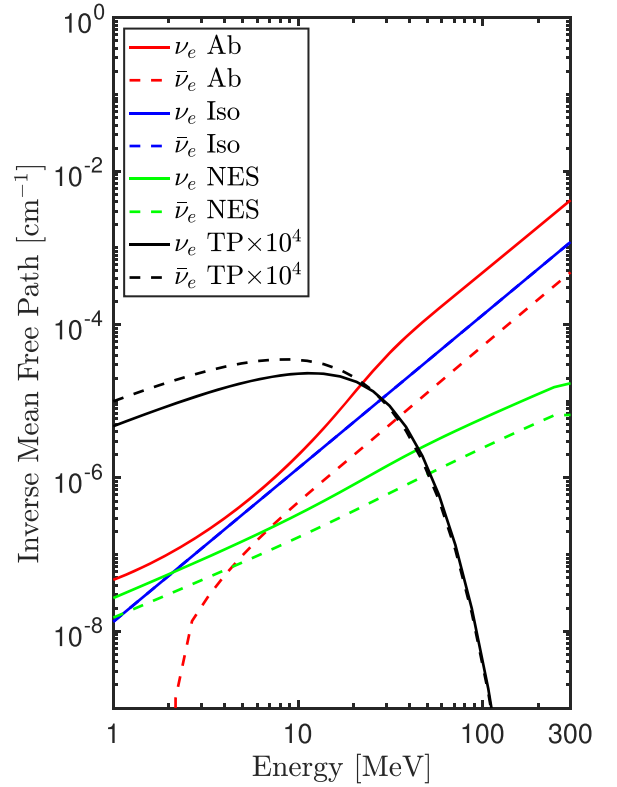
(a) $\rho = 1.084 \times 10^{14}$ g cm^{-3}, $T = 1.845 \times 10^{11}$ K, $Y_e = 0.2728$

(b) $\rho = 1.032 \times 10^{12}$ g cm^{-3}, $T = 8.806 \times 10^{10}$ K, $Y_e = 0.1347$

Figure 3. Neutrino opacities for high (left panel) and moderate (right panel) mass density. In each panel, for neutrinos (solid lines) and antineutrinos (dashed lines), we plot, vs. neutrino energy ε, the absorption opacity (χ and $\bar{\chi}$; red lines), the elastic (isoenergetic) scattering opacity (σ_{IS} and $\bar{\sigma}_{\mathrm{IS}}$; blue lines), the NES opacity (χ_{SC} and $\bar{\chi}_{\mathrm{SC}}$; green lines), and the opacity due to pair creation and annihilation (χ_{TP} and $\bar{\chi}_{\mathrm{TP}}$; black lines). The NES and pair creation and annihilation opacities were computed with the neutrino and antineutrino number densities set to zero; see Equations (17) and (19).

may consider the recently proposed globally convergent variant of Anderson acceleration (Zhang et al. 2020), in which ill conditioning is handled via regularization and global convergence is guaranteed using safeguarding steps.

5.2. Relaxation Problem

The first class of test problems we consider here is the *relaxation* problem, where the neutrino and antineutrino transport Equations 3(a) and (b) are solved with only the collision terms considered, i.e., the space-homogeneous case where the advection terms $\mathcal{T}(f)$ and $\mathcal{T}(\bar{f})$ are zero. In these relaxation problems, the collision operator *relaxes* the distributions f and $\bar{f}$ to the equilibrium Fermi–Dirac distributions f_0 and $\bar{f}_0$ in Equation (7), respectively, as time evolves. Due to the lack of advection terms, there is no spatial coupling. Thus, these problems can be solved independently in space, which makes them ideal test cases for the nonlinear collision system solvers considered in this paper. In this setup, the semidiscrete moment and matter equations reduce from Equation (46) to

$$\frac{1}{\mathrm{c}}\partial_t \boldsymbol{\mathcal{V}}_{\boldsymbol{k}} = \mathbf{C}(\boldsymbol{\mathcal{V}}_{\boldsymbol{k}}), \quad \forall \boldsymbol{k} \text{ such that } \boldsymbol{x}_{\boldsymbol{k}} \in S_{\boldsymbol{x}}. \quad (72)$$

For the relaxation problems, we discretize Equation (72) in time with the backward Euler method. At each time step, this time discretization results in a coupled system that takes the form of Equation (50) with the effective time step $\tau = \Delta t$. To solve this system, we apply the nonlinear solvers in Section 4

to the subsystem in Equation (52) (with $\tau = \Delta t$) to obtain $(Y_e^{(+)}, \epsilon^{(+)}, \mathsf{J}^{(+)}, \bar{\mathsf{J}}^{(+)})$, which are then used to compute $(\mathbf{H}^{(+)}, \bar{\mathbf{H}}^{(+)})$. We start with trivial initial states for $\mathbf{H}$ and $\bar{\mathbf{H}}$, which remain unchanged with time.

We test the nonlinear solvers on the relaxation problem in Equation (72) with two initial matter states that present problems with different degrees of collisionality. The first state, which represents the high-density, strongly collisional region inside a proto–neutron star, is sampled at radius $r^{(1)} = 9.756$ km from the center of a collapsed stellar core, and the second state, which represents the lower-density regions around the surface of a proto–neutron star with relatively weaker collisionality, is sampled at radius $r^{(2)} = 39.52$ km. We obtained the matter states at these two locations from a spherically symmetric CCSN simulation 50 ms after core bounce (Liebendörfer et al. 2005; results from the VERTEX code using a $15\,M_\odot$ progenitor). Specifically, the matter state at $r^{(1)}$ is given by $\rho^{(1)} = 1.084 \times 10^{14}$ g cm^{-3}, $T^{(1)} = 1.845 \times 10^{11}$ K, and $Y_e^{(1)} = 0.2728$; and the matter state at $r^{(2)}$ is $\rho^{(2)} = 1.032 \times 10^{12}$ g cm^{-3}, $T^{(2)} = 8.806 \times 10^{10}$ K, and $Y_e^{(2)} = 0.1347$. The inverse mean free paths associated with the neutrino opacities for these matter states are plotted versus neutrino energy in Figure 3. Considering the isoenergetic scattering opacity (which is equal for neutrinos and antineutrinos and increases with neutrino energy as ε^2), the mean free path $\lambda_{\mathrm{IS}} = \sigma_{\mathrm{IS}}^{-1}$ varies from about 10 to about 10^{-4} km in the high collisional state (left panel) in the energy range $\varepsilon \in [1,$
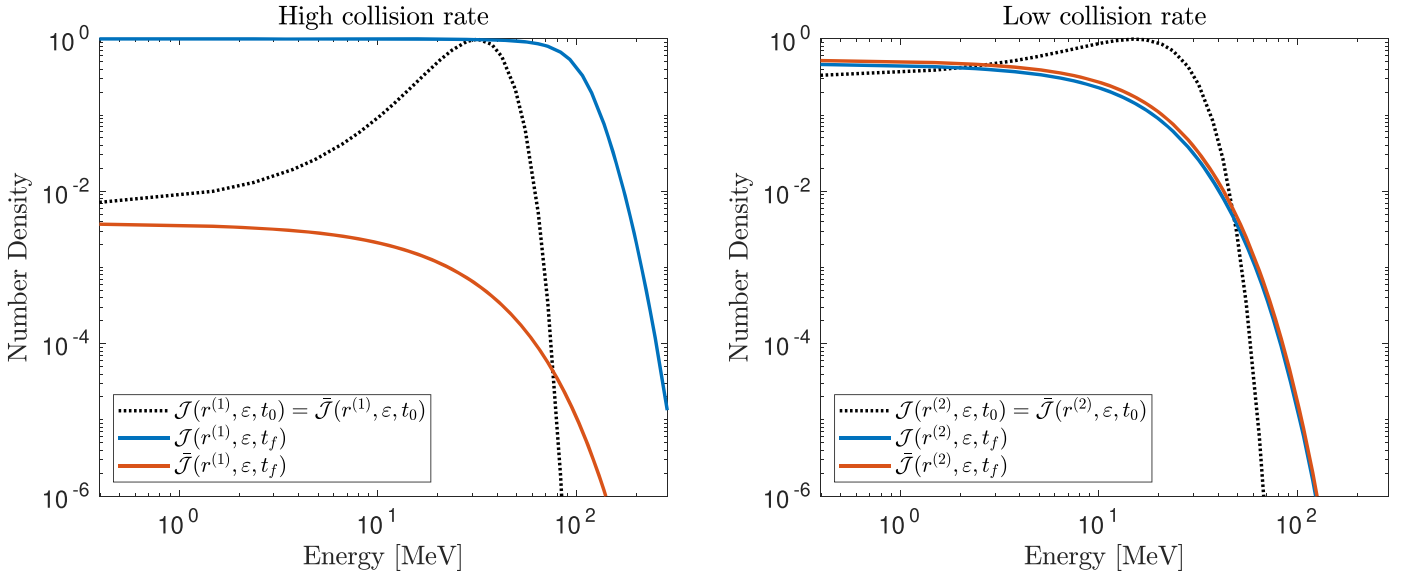
(a) Number density $\mathcal{J}(r^{(1)}, \varepsilon, t)$ at $t = 0$ and $t = t_f = 0.5$ ms. At $r^{(1)}$, with initial matter state $\rho = 1.084 \times 10^{14}$ g cm^{-3}, $T = 1.845 \times 10^{11}$ K, $Y_e = 0.2728$.

(b) Number density $\mathcal{J}(r^{(2)}, \varepsilon, t)$ at $t = 0$ and $t = t_f = 30$ ms. At $r^{(2)}$, with initial matter state $\rho = 1.032 \times 10^{12}$ g cm^{-3}, $T = 8.806 \times 10^{10}$ K, $Y_e = 0.1347$.

Figure 4. Initial and final number densities vs. neutrino energy for relaxation problems with high (left panel) and low (right panel) collision rates. Each plot shows the initial number density given in Equation (73) for both neutrinos and antineutrinos (black dashed line), the final neutrino number density (blue solid line), and the final antineutrino number density (red solid line).

300] MeV. In the low collisional state, the scattering mean free path varies from $\lambda_{IS} \approx 10^3$ to 10^{-2} km. Correspondingly, the collision time $\tau_{IS} = \lambda_{IS}/c$ varies from about 3×10^{-2} to 3×10^{-7} ms in the high collisional state case and from about 3 to 3×10^{-5} ms in the low collisional state case.

We run the simulations from initial time $t_0 = 0$ to final time $t_f = 0.5$ ms for the high-collision case and from $t_0 = 0$ to final time $t_f = 30$ ms for the low-collision case. This is to guarantee that the distributions have relaxed to the equilibrium distributions by the end of the simulations. Each simulation is solved on a single spatial element (the relaxation problem is space-homogeneous) with 16 geometrically progressing energy elements that divide the energy domain [0, 300] MeV, where the first element has $\Delta \varepsilon = 4$ MeV and the last element has $\Delta \varepsilon \approx 50$ MeV. In the simulations, the initial matter states take the electron fraction $Y_{e,0} = Y_e^{(i)}$ and specific internal energy ϵ_0 calculated at $(\rho^{(i)}, T^{(i)}, Y_e^{(i)})$ from the EoS with $i = 1$ and 2. The initial neutrino and antineutrino moments are given by

$$\mathcal{J}_0(r^{(i)}, \varepsilon) = \bar{\mathcal{J}}_0(r^{(i)}, \varepsilon) = 0.99$$
$$\times \exp\left(-\frac{(\varepsilon - 2k_B T^{(i)})^2}{2 \times 10^2}\right) \text{ and} \quad (73)$$
$$\mathcal{H}_0(r^{(i)}, \varepsilon) = \bar{\mathcal{H}}_0(r^{(i)}, \varepsilon) = 0$$

for $i = 1$ and 2, which were chosen such that the initial moments are away from the expected final equilibrium distributions while making sure the initial moments are realizable, i.e., satisfy Equation (12).

Figure 4 illustrates the initial and final equilibrium electron neutrino and antineutrino number densities for the two relaxation problems. Here the initial number densities are given in Equation (73), and the final number densities were generated by solving Equation (72) using the nodal DG and backward Euler discretization, together with the proposed nonlinear solvers. We validated these results by comparing them to number densities from simulations on finer energy-temporal meshes, in which no

noticeable differences were observed. The number densities for the high-collision case are shown in Figure 4(a), and the ones for the low-collision case are shown in Figure 4(b). For the high-collision case, the temperature is 1.845×10^{11} K at $t_0 = 0$ and increases by 7.3% to 1.979×10^{11} K at $t_f = 0.5$ ms, while the electron fraction drops by 14.0% from 0.2728 to 0.2347. For the low-collision case, the temperature rises by 18.8% from 8.806×10^{10} K at $t_0 = 0$ to 1.046×10^{11} K at $t_f = 30$ ms; meanwhile, the electron fraction increases by 2.2% from 0.1347 to 0.1376. From these plots, we observe that, in the high-collision case, the neutrino number density $\mathcal{J}$ is at least 2 orders of magnitude higher than the antineutrino number density $\bar{\mathcal{J}}$ at the final equilibrium; thus, they are further away from the initial densities than the ones in the low-collision case. This is one of the reasons that the high collision rate problems are more challenging than the low collision rate problems in the earlier stage, as discussed in the following paragraphs.

The conservation of lepton number and energy (see Equations (27) and (28), respectively) in the relaxation tests is shown in Figure 5. Since the relaxation problem is space-homogeneous, the right-hand sides of Equations (27) and (28) are both zero. It then follows from the nonlinear system formulation in Equation (52) that the lepton number and energy are conserved if and only if Equations 52(a) and (b) are satisfied. Thus, the lepton number and energy are expected to be conserved up to the nonlinear solver tolerance at each point on the spacetime grid, which is confirmed in the results reported in Figure 5. In addition, we also observe from Figure 5 that, in both the high and low collision rate cases, tightening the nonlinear solver tolerance from 10^{-8} to 10^{-12} indeed improves the conservation results.

Figure 6 shows iteration counts versus time for each nonlinear solver on the two relaxation problems using various time-step sizes: $\Delta t = 10^{-4}$, 10^{-3}, and 10^{-2} ms. These time-step sizes are motivated by the fact that in the context of the IMEX scheme in Equations (47) and (48), the maximum stable time step is $\Delta t = C_{CFL} \times 3 \times 10^{-3}(\Delta x/1 \text{ km})$ ms, where
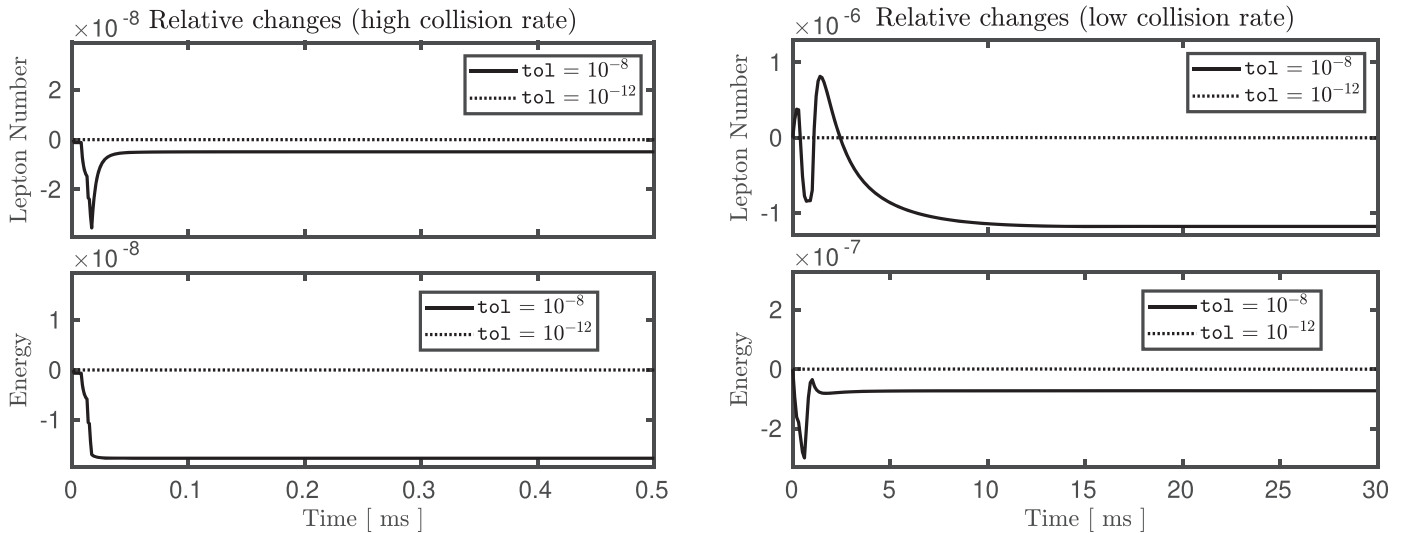
(a) Changes in lepton number and energy relative to their initial values in the relaxation simulation at $r^{(1)}$ from $t = 0$ to $t = t_f = 0.5$ ms. The initial matter state is $\rho = 1.084 \times 10^{14}$ g cm^{-3}, $T = 1.845 \times 10^{11}$ K, $Y_e = 0.2728$.

(b) Changes in lepton number and energy relative to their initial values in the relaxation simulation at $r^{(2)}$ from $t = 0$ to $t = t_f = 30$ ms. The initial matter state is $\rho = 1.032 \times 10^{12}$ g cm^{-3}, $T = 8.806 \times 10^{10}$ K, $Y_e = 0.1347$.

Figure 5. Lepton number and energy conservation (see Equations (27) and (28), respectively) in relaxation problems with high (left column) and low (right column) collision rates. Each plot shows the relative changes in lepton number and energy in the relaxation simulations with the nonlinear solver tolerance `tol` set to 10^{-8} (black solid line) and 10^{-12} (black dashed line). Tightening the tolerance from 10^{-8} to 10^{-12} leads to approximately 4 orders of magnitude smaller relative changes in both lepton number and energy.

$C_{\text{CFL}} \lesssim 1$ is the CFL number. For a spatial resolution $\Delta x = \mathcal{O}(1 \text{ km})$, our chosen time steps bracket what is typically used in CCSN simulations. The results for the high-collision problem are shown in Figure 6(a), while the ones for the low-collision problem are shown in Figure 6(b). In these figures, the top plot shows the "outer" iteration counts for each solver, while the bottom plot shows the averaged "inner" iteration counts for the two nested solvers. Here the outer iteration counts represent the number of iterations needed for solving Equations (54), (60), 65(a), and 67(a) in the coupled AA (Anderson acceleration), coupled Newton, nested AA, and Nested newton solvers, respectively; the inner iteration counts are the number of iterations needed for solving Equations 65(b) and 67(b) in the nested AA and nested Newton solvers, respectively. Since the inner Equations 65(b) and 67(b) are solved in every outer iteration, the reported inner iteration counts in Figures 6(a) and (b) are averaged over the number of times that Equations 65(b) and 67(b) were solved; i.e., the total number of inner iterations taken in the solver is the product of the outer iteration count and the averaged inner iteration count.

Before comparing the solvers, we first observe from the results that the implicit system in Equation (52) in the high collision rate case indeed requires more iterations to reach convergence than the same system does in the low collision rate case. Also, the overall iteration count grows as Δt increases. These observations agree with our expectation, since increasing Δt effectively increases the collision rates, and higher collision rates result in stronger coupling of the number densities, which makes the system in Equation (52) harder to solve.

For the nonlinear solver performance, we observe that the coupled Newton solver requires more iterations than the coupled AA solver on harder problems, e.g., problems with higher collision rates, larger time steps, or at an earlier stage (further away from equilibrium); on easier problems, the

coupled Newton solver converges in fewer iterations than the coupled AA solver. As expected, the nested solvers indeed reduce the number of outer iterations when compared to the coupled solvers, presumably by providing a more accurate update of the neutrino and antineutrino number densities from the inner iteration. The nested AA and nested Newton solvers share nearly identical outer iteration counts. This is due to the fact that both nested solvers use Anderson acceleration in the outer layer (Equations 65(a) and 67(a)), which takes number densities $\hat{\mathcal{U}}(\boldsymbol{u})$ from the inner layers (Equations 65(b) and 67(b)) of the two solvers. Since the problems in the inner layers are equivalent, the solutions $\hat{\mathcal{U}}(\boldsymbol{u})$ are identical up to the residual tolerance. The nested AA solver generally requires more (inner) iterations to converge than the nested Newton solver does, especially on harder problems. We also note that the comparison of iteration counts here does not fully reflect the performance of the solvers in terms of computational time. In particular, as discussed in Section 4.3, opacity kernel evaluations make outer iterations much more computationally expensive than inner iterations. In addition, the iterations in Anderson acceleration are computationally cheaper than the ones in Newton's method, since they do not require constructing and inverting the Jacobian matrix. We defer the comparison of computational times to Section 5.3, where a more realistic test problem is considered.

Next, we explore the effect of the truncation parameter m on the convergence of Anderson acceleration by comparing the iteration count for the coupled AA solver with different values of m on both the high and low collision rate relaxation problems considered in the previous test. In this comparison, the time-step size is fixed to $\Delta t = 10^{-3}$ ms, and m varies from zero to 4, where $m = 0$ resembles the simple Picard iteration. From the results reported in Figure 7, we observe that, in these problems, Anderson acceleration does converge faster as the value of m increases, while the marginal benefit becomes

(a) High collision rate, $\rho = 1.084 \times 10^{14}$ g cm^{-3}, $T = 1.845 \times 10^{11}$ K, $Y_e = 0.2728$

(b) Low collision rate, $\rho = 1.032 \times 10^{12}$ g cm^{-3}, $T = 8.806 \times 10^{10}$ K, $Y_e = 0.1347$
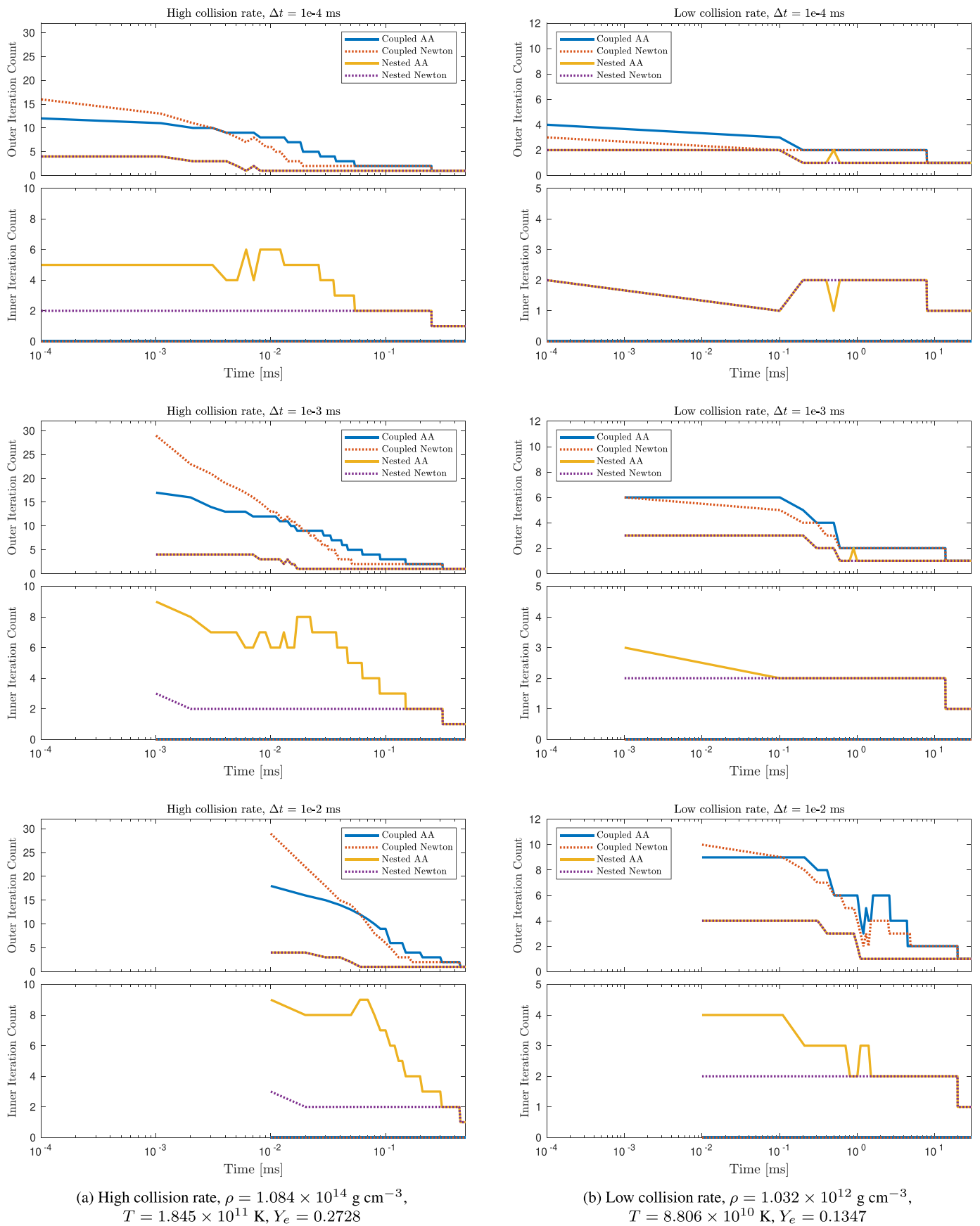
Figure 6. Iteration counts of the nonlinear solvers on relaxation problems. The iteration counts of the coupled AA, coupled Newton, nested AA, and nested Newton solvers on relaxation problems in the high (left column) and low (right column) collision rate regimes with time-step sizes varying from 10^{-4} to 10^{-2} ms.

(a) High collision rate, $\rho = 1.084 \times 10^{14}$ g cm^{-3}, $T = 1.845 \times 10^{11}$ K, $Y_e = 0.2728$

(b) Low collision rate, $\rho = 1.032 \times 10^{12}$ g cm^{-3}, $T = 8.806 \times 10^{10}$ K, $Y_e = 0.1347$
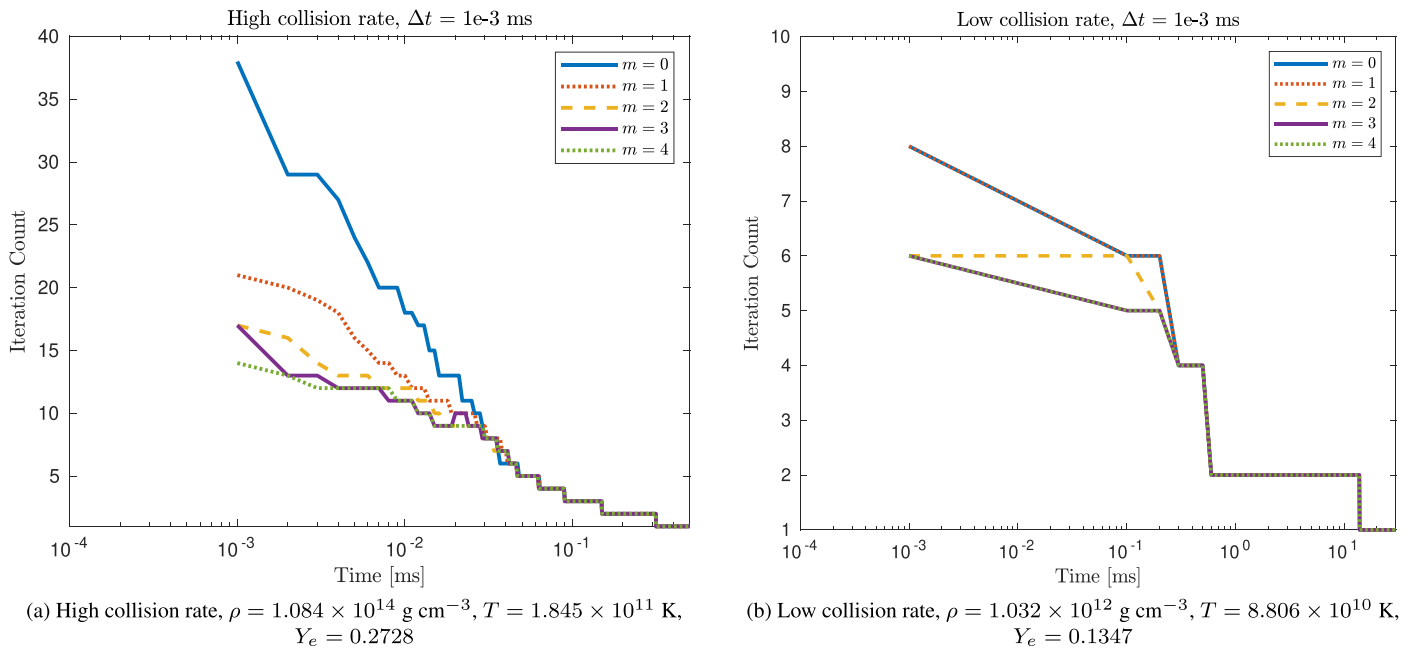
Figure 7. Iteration counts of the coupled AA solver with truncation parameter $m = 0, \ldots, 4$ on relaxation problems.

insignificant for $m \geqslant 2$. This result justifies our choice of $m = 2$ in the numerical tests throughout the paper, since higher values of m lead to larger memory footprints and more expensive least-squares solves in Anderson acceleration with minimal improvement in the iteration counts.

We next investigate how early termination of the iterative solvers affects solution accuracy. We choose to test the nested AA solver with early termination of the outer loop and compare the resulting solution to a converged reference solution at the final time. Here the solution process is terminated either when the convergence criteria (Equation 70(a)) are satisfied or when the outer iteration count for solving Equation 65(a) reaches a preset maximum (MaxIter). We test the solver on the relaxation problem with a low collision rate and time step $\Delta t = 10^{-1}$ ms, which is much larger than the usual time step for a stable explicit scheme. The choice of a large time step makes the effect of early termination more pronounced. Figure 8 reports the iteration counts for MaxIter = 1 and 2 on the test problem, along with the electron neutrino and antineutrino (energy-integrated) number densities, temperatures, and electron fractions at the final time. These results are compared to a fully converged reference solution (MaxIter = 100), where the nested fixed-point solver converges well before the nominal maximal outer iteration is reached, as shown in Figure 8(a). Here the presolve step discussed in Section 5.1 is turned off. We note that when setting MaxIter = 1, the nested AA solver (without presolve) resembles an approach used in earlier works such as Just et al. (2015), where the radiation quantities are updated using opacities computed from the lagged matter states; i.e., matter states at time t^n are used to update the radiation quantities at t^{n+1}. The results in Figure 8 show that when MaxIter = 1, while the relative differences in the earlier time steps are rather significant (from 30% to 0.5%), the solution still converges to an identical (up to the solver tolerance) equilibrium at the final time. However, when MaxIter = 2, it is clear that the solution converges to a different equilibrium. Indeed, it can be seen from Figure 8(e) that the lepton number and energy are not conserved in the solution with MaxIter = 2, while they are conserved

up to the solver tolerance in the fully converged solution (MaxIter = 100). When MaxIter = 2, the changes in lepton number and energy lead to a different equilibrium. As for the solution with MaxIter = 1, despite the fact that the solution process of the nonlinear system (Equation (52)) is terminated early, the lepton number and energy are actually conserved in exact arithmetic. This is because, when MaxIter = 1, the early terminated nested iterates satisfy Equations 52(a) and (b) exactly, which enforces lepton number and energy convergences and leads to the correct equilibrium. However, the early terminated solutions generally do not satisfy Equations 52(c) and (d) and result in rather inaccurate solutions in the transient state. In Figure 9, we repeat the test but with the presolve step turned on. Here we also include the option MaxIter = 0 in the comparison, in which the radiation and matter quantities are only updated in the presolve step, i.e., with the NES and pair processes ignored. From Figure 9, it can be observed that even with MaxIter = 0, the presolve step gives a fairly accurate solution at the final time, as the lepton number and energy are conserved up to the solver tolerance in the presolve step. When the maximum iteration is allowed to be higher, the presolve step improves the solution accuracy, at least in the earlier stage. We observe that while the solution is still inaccurate when MaxIter = 2, it is much closer to the reference solution when the presolve is turned on.

5.3. Deleptonization Problem

We further investigate and compare the performance of the nonlinear solvers in a more realistic setting with a proto–neutron star deleptonization problem, using initial matter profiles from spherically symmetric CCSN simulations. In this test, we solve the full moment equations presented in Section 2.2 using the DG phase-space discretization in Section 3.1 and the IMEX time integration scheme in Section 3.2.

For this test, we adopt matter profiles from Liebendörfer et al. (2005). Specifically, we use profiles for mass density, temperature, and electron fraction obtained with the VERTEX code using a $15\,M_\odot$ progenitor from Woosley & Weaver (1995; model G15 in Liebendörfer et al. 2005). Other thermodynamics
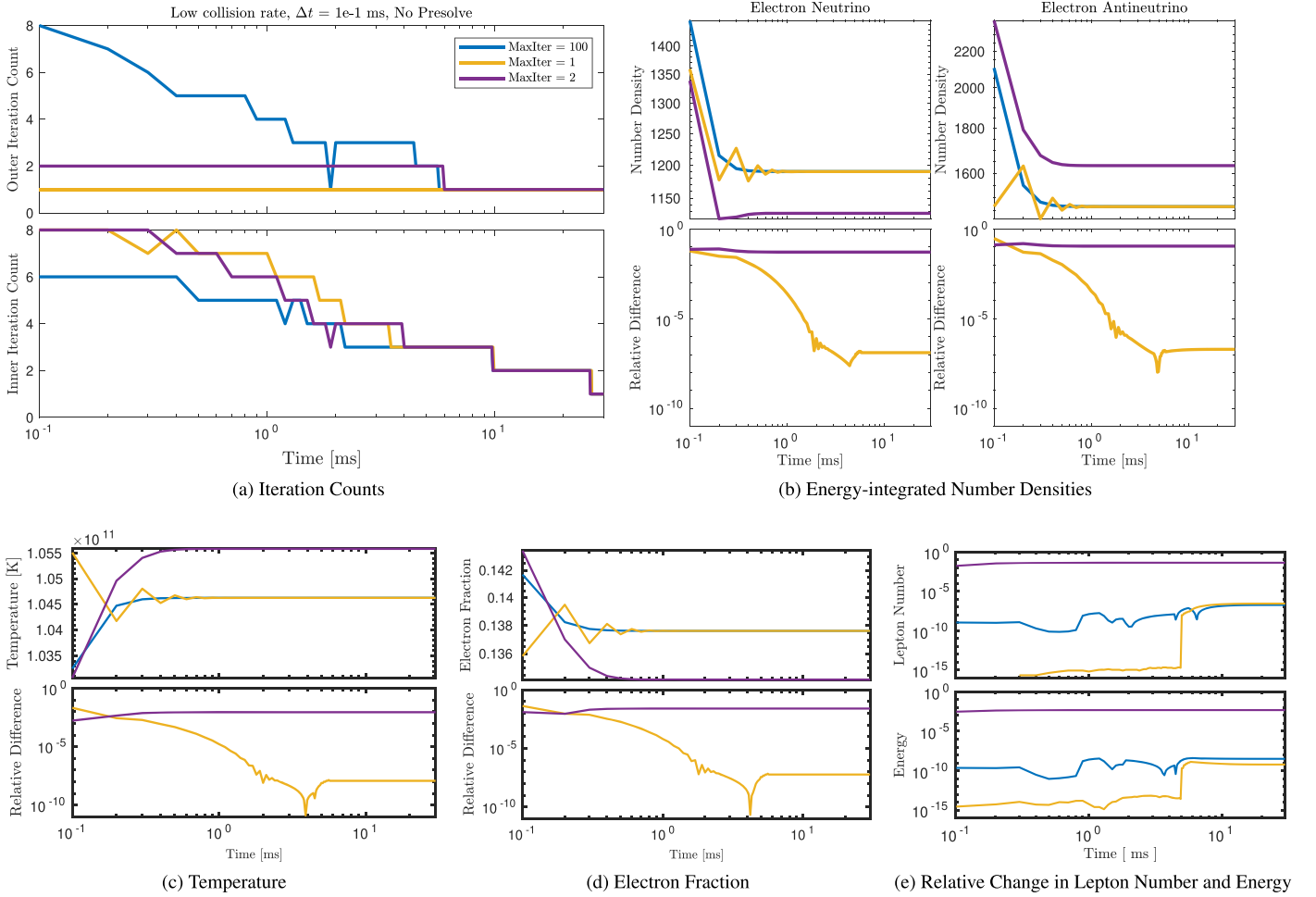
Figure 8. Iteration counts, energy-integrated number densities, temperature, electron fraction, and lepton number and energy conservation results for the nested AA solver on the relaxation problem with low collision rate $\rho = 1.032 \times 10^{12}$ g cm^{-3}, $T = 8.806 \times 10^{10}$ K, $Y_e = 0.1347$. The solver is applied without the presolve step and run to various maximum outer iterations (MaxIter) with time step $\Delta t = 10^{-1}$ ms.

quantities (e.g., internal energy and electron, proton, and neutron chemical potentials) are obtained from the tabulated SFHo EoS (Steiner et al. 2013). To investigate the sensitivity to conditions encountered over an extended period covering the neutrino heating phase, we run the comparison on profiles taken at $t = 50$, 100, 150, and 250 ms after core bounce. The initial matter profiles are plotted in Figure 10.

Since we do not have the radiation quantities from Liebendörfer et al. (2005), to initialize the radiation field, we adopt the analytical distribution function from the homogeneous sphere test, f_{HS} (e.g., Smit et al. 1997), which is a solution to the steady-state transport problem of radiation emanating from a sphere whose constant absorption opacity and emissivity inside a radius R_0 are χ_0 and $\chi_0 f_0$, respectively,

$$f_{HS}[f_0, \chi_0, R_0](r, \mu) = f_0(1 - e^{-\chi_0 \, s[R_0](r,\mu)}), \tag{74}$$

where f_0 is an isotropic equilibrium distribution,

$$s[R_0](r, \mu)$$
$$= \begin{cases} r\mu + R_0 \, g[R_0](r, \mu) & \text{if} \quad r < R_0, \ \mu \in [-1, +1], \\ 2 R_0 \, g[R_0](r, \mu) & \text{if} \quad r \geqslant R_0, \ \mu \in [(1 - (R_0/r)^2)^{1/2}, \\ 0 & \text{otherwise}, \end{cases} \tag{75}$$

and $g[R_0](r, \mu) = [1 - (r/R_0)^2(1 - \mu^2)]^{1/2}$.

To adopt the homogeneous sphere distribution to the current setting, we first estimate the energy-dependent neutrinosphere radius $R_\nu(\varepsilon)$,

$$\int_{R_\nu(\varepsilon)}^{\infty} \chi(\varepsilon, r) dr = \frac{2}{3}, \tag{76}$$

where χ is the absorption opacity. (Neutrinosphere radii for the various initial profiles used here are plotted versus neutrino energy in the lower right panel of Figure 10.) The neutrino distribution function is then set to

$$f(r, \varepsilon, \mu)$$
$$= \begin{cases} f_{HS}[f_0(r, \varepsilon), \chi(r, \varepsilon), R_\nu(\varepsilon)](r, \mu), & r < R_\nu(\varepsilon) \\ f_{HS}[f_0(R_\nu(\varepsilon)), \chi(R_\nu(\varepsilon)), R_\nu(\varepsilon)](r, \mu), & r \geqslant R_\nu(\varepsilon), \end{cases} \tag{77}$$

where $f_0(r, \varepsilon)$ is taken to be the Fermi–Dirac distribution in Equation (7). Finally, the initial moments are computed as

$$\{\mathcal{J}, \mathcal{H}\}(r, \varepsilon) = \frac{1}{2} \int_{-1}^{1} f(r, \varepsilon, \mu) \mu^{\{0,1\}} \, d\mu. \tag{78}$$

(The same procedure is adopted to initialize the antineutrinos.)

The evolution of the electron fraction in the deleptonization problem is illustrated in Figure 11, where we plot the electron
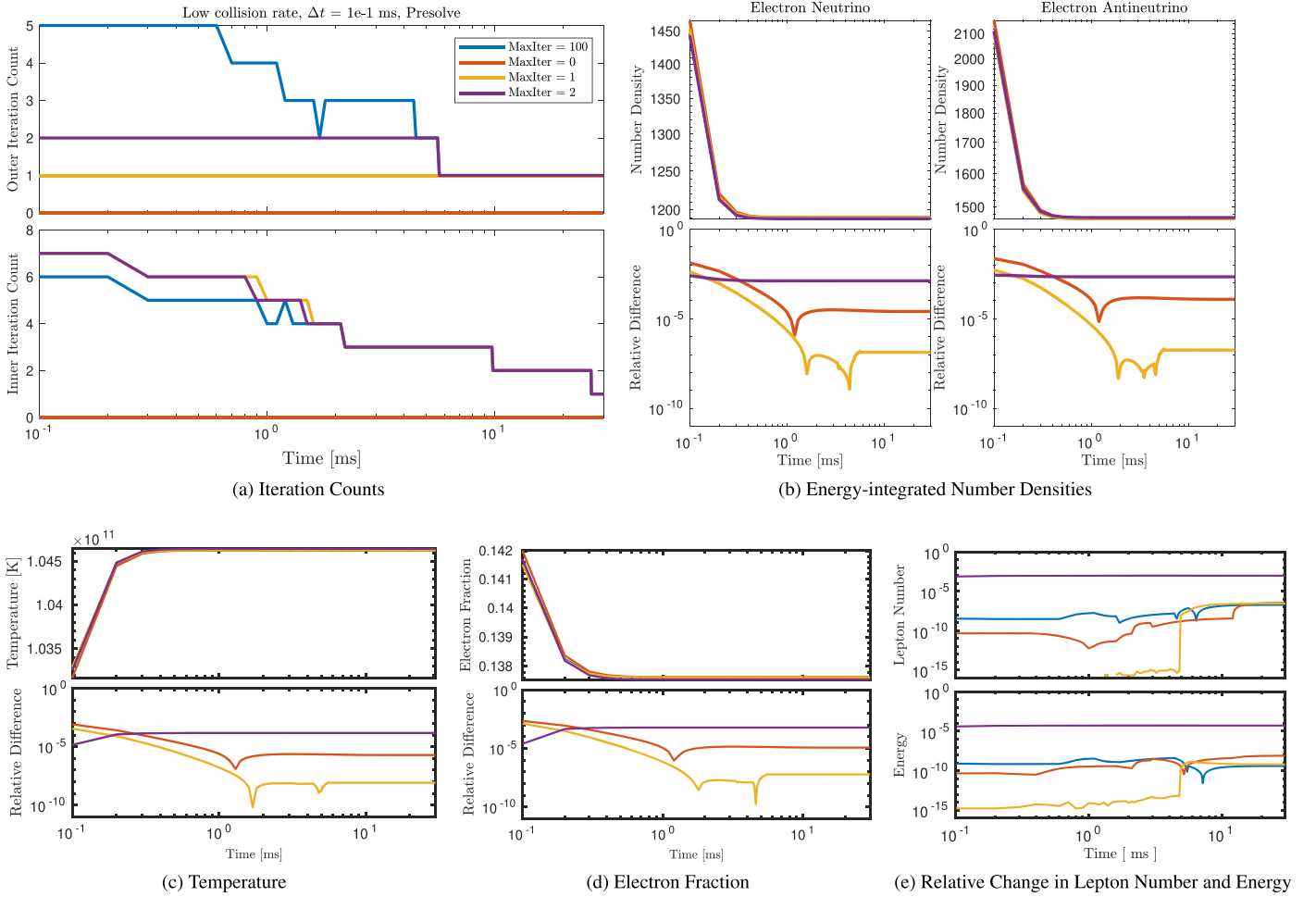
(a) Iteration Counts

(b) Energy-integrated Number Densities

(c) Temperature

(d) Electron Fraction

(e) Relative Change in Lepton Number and Energy

Figure 9. Iteration counts, energy-integrated number densities, temperature, electron fraction, and lepton number and energy conservation results for the nested AA solver on the relaxation problem with low collision rate $\rho = 1.032 \times 10^{12}$ g cm^{-3}, $T = 8.806 \times 10^{10}$ K, $Y_e = 0.1347$. The solver is applied with the presolve step and run to various maximum outer iterations (MaxIter) with time step $\Delta t = 10^{-1}$ ms.

fraction versus mass density over 10 ms of evolution starting from the 100 ms profile in Figure 10 (dashed lines). Here the evolution of the electron fraction was generated from the DG-IMEX scheme with the proposed nonlinear solvers as described in Section 5.1. We validated the result by comparing it against solutions from simulations on finer spatial energy–temporal meshes, in which no noticeable differences were observed. For higher densities ($\rho \gtrsim 3 \times 10^{12}$ g cm^{-3}), neutrinos are effectively trapped, and the electron fraction remains largely unchanged. For lower densities, neutrinos (created by electron capture on protons) are not trapped and escape the computational domain, which results in a lowering of the electron fraction (deleptonization). Figure 12 shows conservation of lepton number and energy over 5 ms of evolution in the deleptonization problem starting from the 100 ms postbounce matter profile. Here the lepton number and energy both comprise two parts: (i) the interior lepton number and energy in the computation domain, i.e., the time integrals of the left-hand sides in Equations (27) and (28), respectively, and (ii) the accumulated outflow lepton number and energy at the boundary, which are, respectively, the negations of the right-hand sides in Equations (27) and (28), integrated in time. The conservation results of the lepton number/energy, as well as the evolution of the interior lepton number/energy and the accumulated outflow lepton number/energy, are illustrated in

Figure 12(a). The evolution of the individual components of the interior lepton number/energy, including matter lepton number (electron number), neutrino lepton number, internal energy, and neutrino energy, are shown in Figure 12(b).

For each profile, the deleptonization problems are simulated from the profile time (50, 100, 150, and 250 ms after core bounce) to 5 ms after the profile time, which are referred to as the initial time $t_0 = 0$ and the final time $t_f = 5$ ms in the remainder of the paper. The IMEX time integration scheme discussed in Section 5.1 is used in the simulations, where the time step Δt is determined by the stability requirement for the explicit advection part. These problems are solved in the spatial domain $r \in [0, 300]$ km and energy domain $[0, 300]$ MeV, which are divided into 128 and 16 geometrically progressing elements, respectively. Here the first spatial element has $\Delta x = 1$ km, the last spatial element has $\Delta x \approx 4.54$ km, the first energy element has $\Delta \varepsilon = 4$ MeV, and the last energy element has $\Delta \varepsilon \approx 50$ MeV.

Figure 13 shows the iteration counts of the nonlinear solvers on the deleptonization problem for various profiles. The results for profiles taken at 50, 100, 150, and 250 ms after core bounce are illustrated in Figures 13(a)–(d), respectively. As in Figure 6, the top plot in these figures shows the "outer" iteration counts of each solver, while the bottom plot shows the averaged "inner" iteration counts of the nested solvers, where the inner

(a) Mass Density

(b) Temperature

(c) Electron Fraction
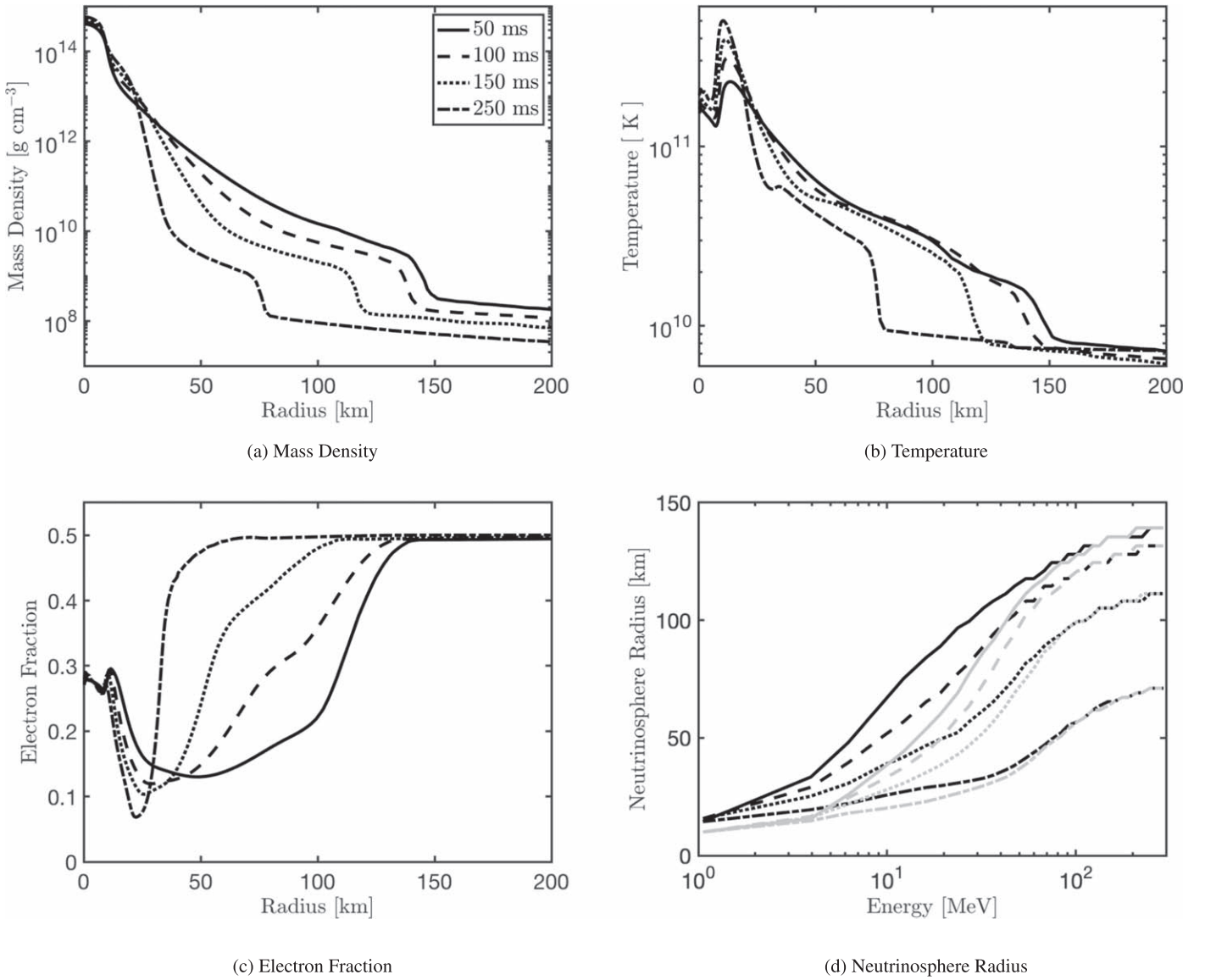
(d) Neutrinosphere Radius

Figure 10. Initial matter profiles used in the deleptonization problem, taken from Liebendörfer et al. (2005). Plotted vs. radius are mass density (upper left panel), temperature (upper right panel), and electron fraction (lower left panel) for various postbounce times: 50 (solid), 100 (dashed), 150 (dotted), and 250 (dashed–dotted) ms. The neutrinosphere radii, defined in Equation (76), for the respective profiles—for electron neutrinos (black) and electron antineutrinos (gray)—are also plotted vs. neutrino energy (lower right panel).

iteration counts are averaged over the number of times that the inner Equations 65(b) or 67(b) were solved. In addition, here both the outer and inner iteration counts are averaged over time (from t_0 to t_f), and the averaged iteration counts are plotted against the mass density, which corresponds to spatial locations for each profile. (We have found that the number of iterations at a given location varies little from t_0 to t_f.)

From Figure 13, we observe that the simulations with all four profiles give consistent results; the nested solvers require fewer outer iterations than the coupled ones do, and the nested AA solver requires more inner iteration to converge than the nested Newton solver does, especially for harder problems (at higher mass density). This observation also agrees with the results reported in Section 5.2 on the relaxation problems. Computational times for these simulations are reported in Table 1, where tests 1, 5, 9, and 13 correspond to simulations shown in Figure 13(a), tests 2, 6, 10, and 14 correspond to simulations shown in Figure 13(b), and so on. In Tables 1 and 2, we report

the total computation time t_{Tot} as well as the detailed timing measurements in each simulation, such as the computational time spent on (i) solving the nonlinear system in Equation (52) in the implicit step (t_{Im}); (ii) the opacity evaluations/ interpolations when solving Equation (52) (t_{Op}); (iii) linear algebra operations, such as the least-squares solve in Anderson acceleration or the assembly and inversion of Jacobian matrices in Newton's method (t_{LA}); (iv) the presolve step (t_{Ps}); (v) the explicit update of the advection term (t_{Ex}); and (vi) the positivity limiter (t_{PL}).[7] The reported timing results in Tables 1 and 2 are all linearly scaled such that the highest reported computational time (16,085 s from test 2) is scaled to 100,

_____

[7] Here the positivity limiters are applied after each explicit and implicit update to enforce the realizability of the moments. The computation time for the positivity limiter (t_{PL}) reported in Tables 1 and 2 is relatively large (compared to t_{Ex}), and we expect that t_{PL} could be further reduced by a more sophisticated implementation. In any case, it does not affect the observations we made in this paper.
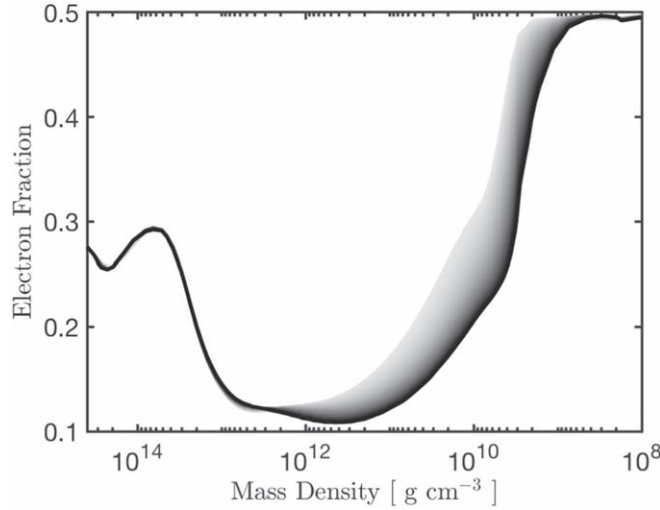
Figure 11. Electron fraction vs. mass density over 10 ms of evolution from the initial state given by the 100 ms postbounce matter profile in Figure 10. The time evolution is indicated by the gray scale, which goes from lightest (initial state) to darkest (final state).
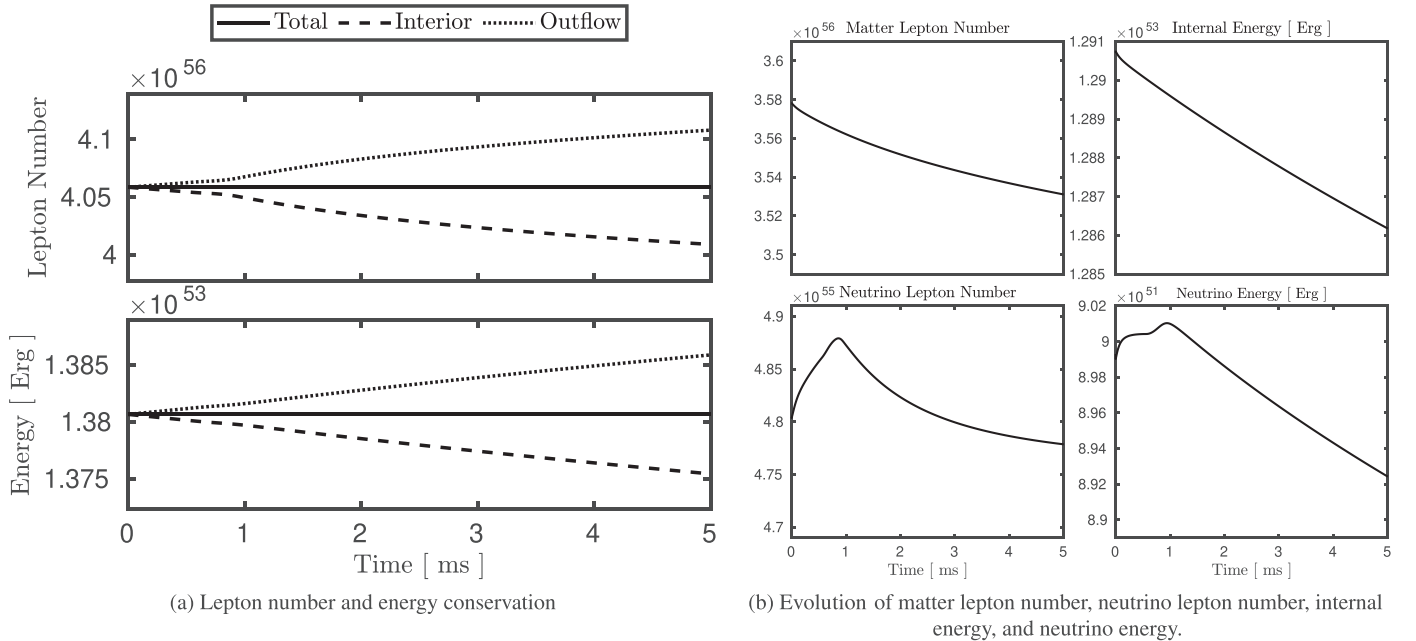

(a) Lepton number and energy conservation

(b) Evolution of matter lepton number, neutrino lepton number, internal energy, and neutrino energy.

Figure 12. Conservation in the deleptonization problem. The left panel shows the lepton number and energy (see Equations (27) and (28)) over 5 ms of evolution starting from the 100 ms postbounce matter profile. Here the interior lepton number/energy in the computational domain are plotted with black dashed lines, and the accumulated outflow lepton number/energy are plotted with black dotted lines. Note that the accumulated outflows (black dotted lines) are shifted up by the value of the initial lepton number/energy for better illustration. The conserved lepton number and energy are plotted with black solid lines. The right panels show the evolution of the components of the interior lepton number and energy, including matter lepton number (electron number), neutrino lepton number, internal energy, and neutrino energy.

which corresponds to the total computation time when the coupled Newton solver is used to solve the deleptonization problem with the 100 ms profile. For each simulation, we also record the solver configurations, such as the value of the truncation parameter m in Anderson acceleration, the maximum allowed iteration (MaxIter), and whether the presolve step is performed or not. We also note that these reported data are not mutually exclusive, e.g., $t_{Tot} \approx t_{Im} + t_{Ex} + t_{PL}$ and $t_{Im} \approx t_{Op} + t_{LA} + t_{Ps}$. Figure 14 provides a column chart that visualizes the results with the 100 ms profile reported in Table 1. There is no qualitative difference between the results from problems with different profiles. The column chart in

Figure 14 confirms that the majority of the computational time in these simulations is spent on opacity evaluation/interpolation, which is proportional to the outer iteration counts. It also shows that the nested solvers indeed speed up the computations by taking inner iterations to reduce the number of outer iterations. Further, we observe that the lower computational cost on the linear algebra operations (t_{LA}) makes the nested AA solver outperform the nested Newton solver in terms of the total computation time by about 8%, despite the higher inner iteration counts. These results indicate that the nested AA solver leads to the least computation time for the deleptonization problems among all of the tested solvers.
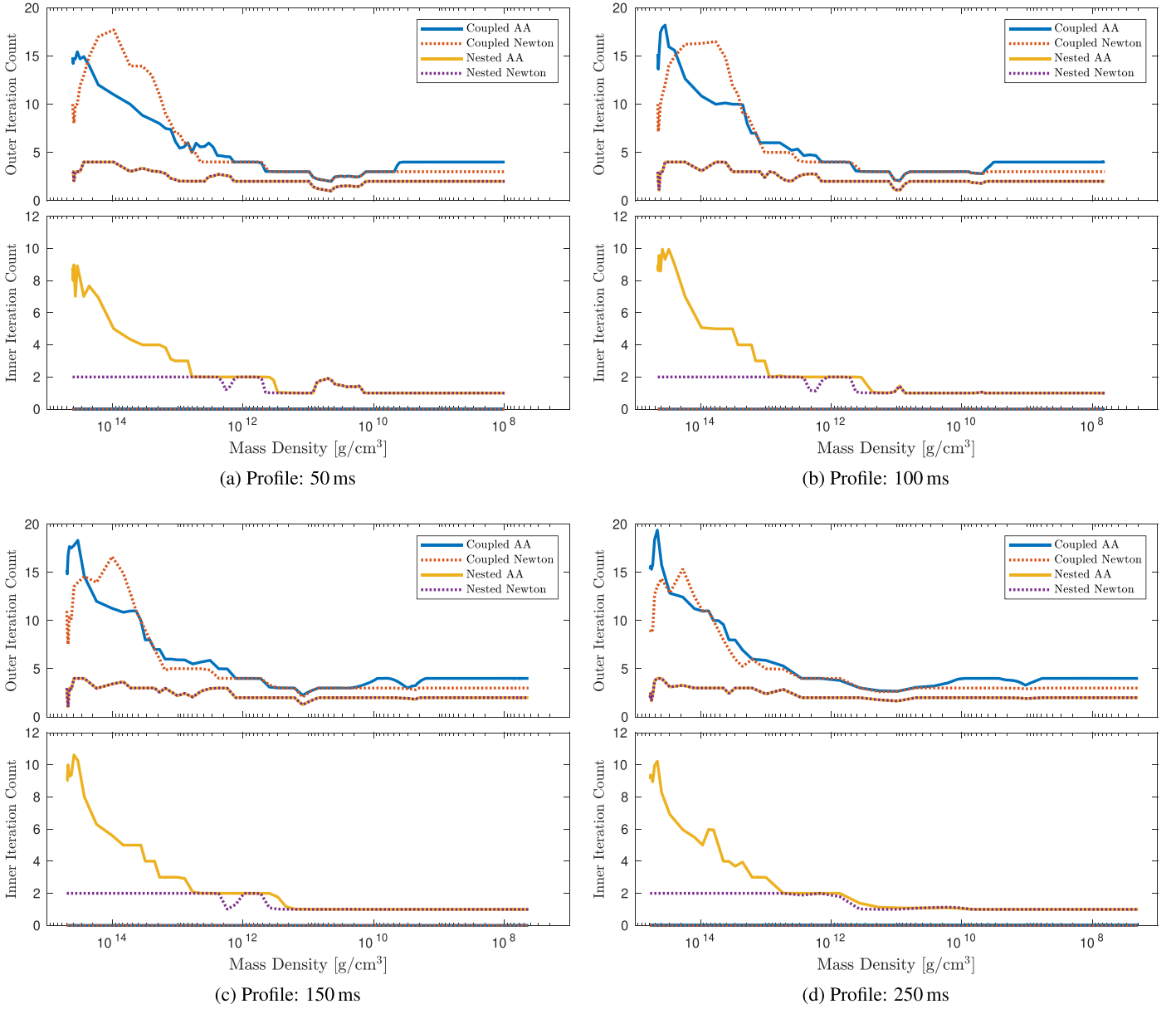
Figure 13. Time-averaged iteration counts of the nonlinear solvers on deleptonization problems. Shown are the time-averaged iteration counts of the coupled AA, coupled Newton, nested AA, and nested Newton solvers on deleptonization problems with various profiles.

To further analyze the performance of the nested AA solver, we experiment with different solver parameters on the deleptonization problem using the 100 ms profile. Specifically, we tested the nested AA solver with the Anderson acceleration truncation parameter set to $m = 0$, 1, and 2 for solving the outer system in Equation 65(a). This experiment is to verify the benefit of Anderson acceleration on solving the smaller outer system in Equation 65(a), which has only two unknowns. In addition, for each choice of m, we initialize the solver with and without the presolve step introduced in Section 5.1, which helps us assess the effect of that step in a more realistic setting. The resulting iteration counts are shown in Figure 15, which are averaged over the time from t_0 to t_f as in Figure 13. The computation times are reported in Table 2 (tests 17–21). From these results, we observe that moving from Picard iteration ($m = 0$) to Anderson acceleration ($m > 0$) does reduce the outer iteration count and thus

improves the computation time by around 18%, especially for harder problems (at higher mass density). However, unlike the results for the coupled AA solver reported in Figure 7, increasing the truncation parameter from $m = 1$ to 2 does not lead to any observable reduction in either the iteration counts or computation time. This result is not unexpected, since Anderson acceleration is applied here to solve the outer system in Equation 65(a) with only two unknowns, while the results reported in Figure 7 are from solving the fully coupled system in Equation (54). Another observation from Figure 15 is that there is no clear benefit in applying the presolve step on this problem. The presolve step does slightly reduce the iteration counts; however, the additional cost of the presolve step wipes out the gains from fewer iterations. We suspect that the diminished benefit of the presolve step is due to the fact that in the deleptonization problem, the explicit time-step restriction from the advection term limits the stiffness of the
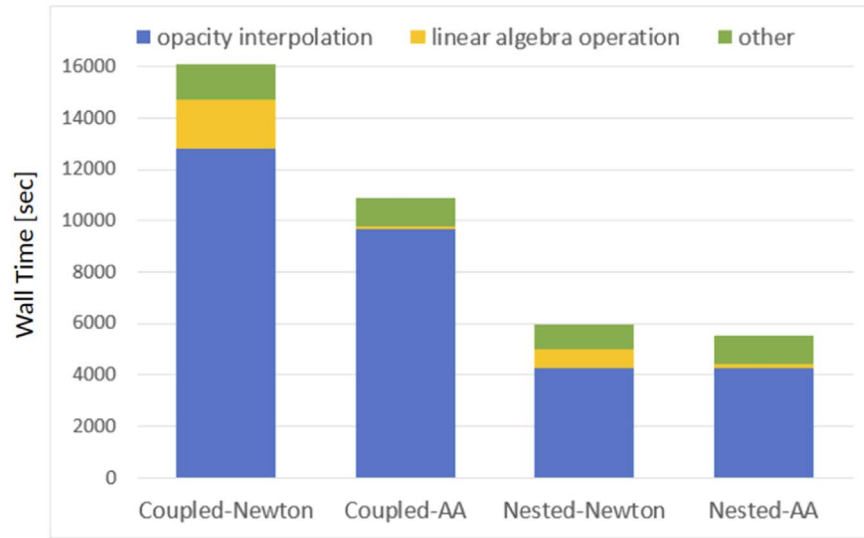
Figure 14. Computation time breakdown for the nonlinear solvers on the deleptonization problem with a 100 ms profile. The opacity interpolation time (t_{Op} in Tables 1 and 2) and the dense linear algebra operation time (t_{LA} in Tables 1 and 2) are the blue and yellow sections of the columns, respectively. The green sections represent the time spent on all other computations.

Table 1
Overview of Nonlinear Solver Timing Results for the Deleptonization Problem

No.	Solver	Profile	t_{Tot}	t_{Im}	t_{Op}	t_{LA}	t_{Ps}	t_{Ex}	t_{PL}	m	Presolve	MaxIter
1	Coupled Newton	50 ms	98.8	96.8	78.6	11.7	2.7	0.8	1.0	⋯	Yes	100
2	Coupled Newton	100 ms	100	97.9	79.5	11.9	2.7	0.8	1.0	⋯	Yes	100
3	Coupled Newton	150 ms	97.5	95.4	77.5	11.5	2.7	0.8	1.0	⋯	Yes	100
4	Coupled Newton	250 ms	92.9	90.8	73.8	10.8	2.7	0.8	1.0	⋯	Yes	100
5	Coupled AA	50 ms	65.2	63.2	57.6	0.9	2.7	0.8	1.0	2	Yes	100
6	Coupled AA	100 ms	67.8	65.8	60.0	1.0	2.7	0.8	1.0	2	Yes	100
7	Coupled AA	150 ms	69.0	67.0	61.1	1.0	2.7	0.8	1.0	2	Yes	100
8	Coupled AA	250 ms	67.5	65.4	59.7	1.0	2.6	0.8	1.0	2	Yes	100
9	Nested Newton	50 ms	36.6	34.5	26.0	4.8	2.7	0.8	1.0	2	Yes	100
10	Nested Newton	100 ms	36.9	34.9	26.4	4.7	2.7	0.8	1.0	2	Yes	100
11	Nested Newton	150 ms	36.4	34.3	26.1	4.5	2.7	0.8	1.0	2	Yes	100
12	Nested Newton	250 ms	35.7	33.7	25.8	4.3	2.6	0.8	1.0	2	Yes	100
13	Nested AA	50 ms	33.7	31.6	26.0	0.9	2.7	0.8	1.0	2	Yes	100
14	Nested AA	100 ms	34.3	32.2	26.5	0.9	2.7	0.8	1.0	2	Yes	100
15	Nested AA	150 ms	33.8	31.7	26.0	0.9	2.7	0.8	1.0	2	Yes	100
16	Nested AA	250 ms	33.1	31.0	25.5	0.8	2.6	0.8	1.0	2	Yes	100

Notes. The detailed computational time is reported for the results shown in Figure 13. Here the four iterative solvers are compared on four initial matter profiles with identical solver parameters. The results are linearly scaled so that the highest measurement (boxed; 16,085 s) is scaled to 100.

t_{Tot}: total simulation time; t_{Im}: implicit solution time; t_{Op}: opacity interpolation time;

t_{LA}: dense linear algebra time; t_{Ps}: initial presolve time; t_{Ex}: explicit update time;

t_{PL}: positivity limiter time; m: Anderson acceleration truncation parameter;

Presolve: whether the presolve step is performed in solver initialization;

MaxIter: maximum allowed (outer) iteration for (nested) iterative solvers;

Relations: $t_{Tot} \approx t_{Im} + t_{Ex} + t_{PL}$, $t_{Im} \approx t_{Op} + t_{LA} + t_{Ps}$.

implicit problem and forces the implicit update to be small, which makes the effect of presolve insignificant. For deleptonization problems, larger implicit time steps could potentially be achieved by techniques such as subcycling the

explicit steps or by adopting the general multirate framework proposed by Sandu (2019). We expect to see a greater impact of the presolve step on these problems with larger time steps, as is observed in the relaxation tests.
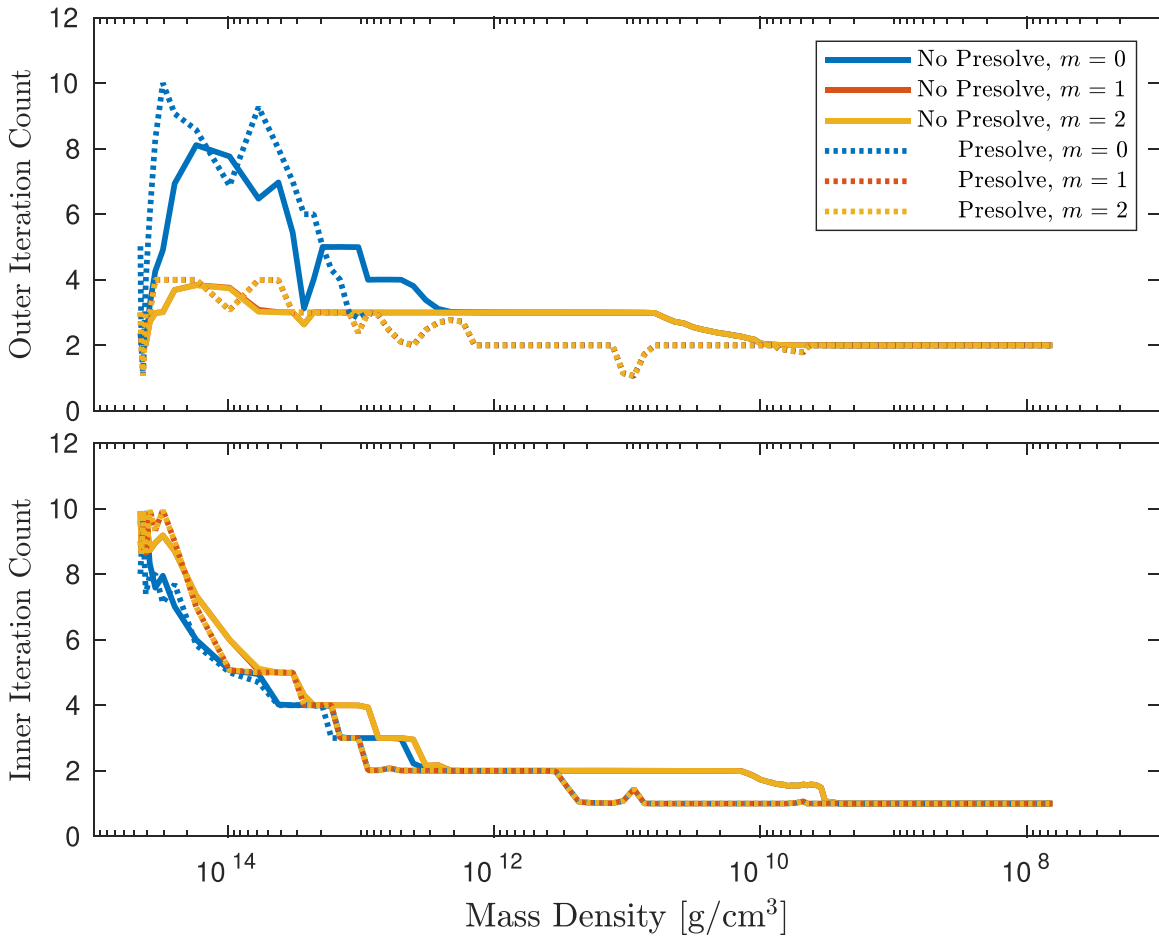
Figure 15. Time-averaged iteration counts for the nested AA algorithm with various solver configurations on the deleptonization problem with a 100 ms profile. Results for the nested AA algorithm without/with presolve and with truncation parameter values $m = 0$, 1, and 2 in the outer loop are reported. Here the iteration counts for $m = 2$ overlap with the corresponding results for $m = 1$, which implies that there is no benefit in moving from $m = 1$ to 2 when solving the outer loop problem.

Finally, we investigate the effect of early termination for the nested AA solver on the deleptonization problems. In these tests, the solver configurations are identical to the ones reported in Figures 8 and 9 for the relaxation problems; e.g., the outer loop of the nested AA solver is terminated early by restricting the maximum number of outer iterations (MaxIter). Here we test the solver on the deleptonization problem with the 100 ms initial matter profile for MaxIter = 1, 2, and 100 and the presolve step discussed in Section 5.1 turned off. We then repeat the test for MaxIter = 0, 1, 2, and 100 with the presolve step turned on. The results without and with the presolve step are shown in Figures 16 and 17, respectively, where the iteration counts are reported along with the electron neutrino and antineutrino (energy-integrated) number densities, temperatures, and electron fractions at the final time $t = 5$ ms. Here the fully converged solutions (MaxIter = 100) are considered as reference solutions, where the nested solver converges well before the nominal maximal outer iteration is reached, as shown in Figures 16(a) and 17(a). As mentioned in the earlier subsection, the nested AA solver with MaxIter = 1 and the presolve step turned off is effectively lagging the opacities by computing them from the matter states at the previous time step while updating the radiation quantities at the current time step

(see, e.g., Just et al. 2015). From Figure 16, we observe that the early terminated solutions are in good agreement with the reference solution. However, allowing two outer iterations does not give a better solution than the one with only a single outer iteration, which is possibly due to the issue of lepton number and energy conservation for early terminated solutions, as discussed in Section 5.2. Another potential reason is that Anderson acceleration does not guarantee monotone decreasing of the residuals (see Toth & Kelley 2015; Kelley 2018; Pollock & Rebholz 2019; Evans et al. 2020 for convergence analysis of AA). The behavior of the residuals from Anderson acceleration alternately increasing and decreasing was observed in Pollock & Rebholz (2019), and a potential explanation is given from Pollock & Rebholz (2019), Theorem 4.5. Similar results can be observed in Figure 17, where the presolve step is in effect. Here the comparison includes the case that MaxIter = 0, where both the radiation and matter quantities are solely updated in the presolve step, with the NES and pair processes omitted. The relative difference in the solution with MaxIter = 0 is mostly around 10^{-2}–10^{-3}; however, the difference could go up to 10^1 for the energy-integrated antineutrino number density in the high mass density region. We also note that the presolve step reduces the difference in the solution with MaxIter = 2 by a

Table 2
Overview of Nonlinear Solver Timing Results for the Deleptonization Problem

#	Solver	Profile	$t_{\rm Tot}$	$t_{\rm Im}$	$t_{\rm Op}$	$t_{\rm LA}$	$t_{\rm Ps}$	$t_{\rm Ex}$	$t_{\rm PL}$	m	Presolve	MaxIter
14	Nested AA	100 ms	100	94.0	77.1	2.6	7.8	2.4	2.9	2	Yes	100
17	Nested AA	100 ms	117.8	111.9	96.1	2.0	⋯	2.4	2.9	0	No	100
18	Nested AA	100 ms	99.9	94.0	81.5	1.5	⋯	2.4	2.9	1	No	100
19	Nested AA	100 ms	99.7	93.8	81.3	1.5	⋯	2.4	2.9	2	No	100
20	Nested AA	100 ms	118.2	112.3	91.4	3.3	7.7	2.4	2.9	0	Yes	100
21	Nested AA	100 ms	99.3	93.4	76.7	2.6	7.7	2.4	2.9	1	Yes	100
22	Nested AA	100 ms	47.2	41.2	34.0	0.8	⋯	2.4	2.9	2	No	1
23	Nested AA	100 ms	83.9	77.9	67.2	1.2	⋯	2.5	2.9	2	No	2
24	Nested AA	100 ms	100.9	94.9	82.3	1.5	⋯	2.4	2.9	2	No	100
25	Nested AA	100 ms	16.9	10.8	3.7	1.2	7.6	2.5	2.9	2	Yes	0
26	Nested AA	100 ms	54.2	48.2	36.2	1.9	7.7	2.4	2.9	2	Yes	1
27	Nested AA	100 ms	89.1	83.0	68.2	2.2	7.8	2.4	2.9	2	Yes	2

Notes. The detailed computational time is reported for the results shown in Figures 15–17. Here the nested AA solver is tested on the 100 ms profile with various solver parameters. The results are linearly scaled so that the total time in test 14 (boxed; 5518 s) is scaled to 100.

$t_{\rm Tot}$: total simulation time; $t_{\rm Im}$: implicit solution time; $t_{\rm Op}$: opacity interpolation time;

$t_{\rm LA}$: dense linear algebra time; $t_{\rm Ps}$: initial presolve time; $t_{\rm Ex}$: explicit update time;

$t_{\rm PL}$: positivity limiter time; m: Anderson acceleration truncation parameter;

Presolve: whether the presolve step is performed in solver initialization;

MaxIter: maximum allowed (outer) iteration for (nested) iterative solvers;

Relations: $t_{\rm Tot} \approx t_{\rm Im} + t_{\rm Ex} + t_{\rm PL}$, $t_{\rm Im} \approx t_{\rm Op} + t_{\rm LA} + t_{\rm Ps}$.

few orders of magnitude in the low mass density region. The reason is that, in the low-density region, the nonlinear solve usually converges within two iterations with the presolve step. The results in Figures 16 and 17 suggest that, for problems requiring few iterations, limiting MaxIter to 1 can give sufficiently accurate solutions while reducing the computation time by roughly a factor of 2 from the fully converged case, as shown in the results for tests 22–27 in Table 2.

6. Summary and Discussion

We have investigated several iterative solvers for nonlinear systems arising from the discretization of a nonrelativistic two-moment model for neutrino transport with opacities from Bruenn (1985), coupled with static matter configurations. Specifically, we have incorporated the nonlinear solvers in a DG-IMEX scheme, as implemented in the toolkit for high-order neutrino radiation hydrodynamics (thornado). Within the IMEX time integration scheme currently adopted in thornado, updating the neutrino transport and matter equations requires solving a coupled nonlinear system on the radiation moments and matter states (internal energy and electron fraction). We have considered two approaches to solve the nonlinear system: a coupled approach that directly solves the fully coupled system and a nested approach that formulates the nonlinear system as a nested system with the outer system governing the matter states and the inner system governing the neutrino number densities. The nested approach is introduced to reduce the number of opacity evaluations/interpolations required in the solution procedure and thus becomes more efficient than the coupled approach. Two iterative solvers—the

Anderson-accelerated fixed-point solver and Newton's method —are implemented for both the coupled and nested approaches. We have tested the four solvers on relaxation problems with various collision rates and time steps, as well as on proto–neutron star deleptonization problems with postbounce matter profiles from spherically symmetric CCSN simulations.

The numerical results confirm that both nested solvers indeed require fewer iterations to converge (and thus less computational time) than the coupled solvers, due to the lower number of opacity interpolations performed in the solution procedure. The nested Anderson acceleration solver requires more inner iterations to converge but, due to the low cost per iteration, less computation time than the nested Newton's method, which is a consequence of the heavier dense linear algebra operations in Newton's method. In addition to the advantage in computation time, another benefit for using solvers based on Anderson acceleration over Newton's method is the simplicity of implementation, particularly for solving problems in which the derivatives are not readily available, such as the coupled nonlinear systems in CCSN simulations considered in this paper. For the test problems considered in this paper, we also observe that forcing the nested Anderson acceleration solver to terminate after the first outer iteration could lead to reasonably accurate results. This observation confirms that, on these problems, solving the nonlinear coupled system in the implicit step using lagged opacity kernels from the previous time step could give a sufficiently accurate solution, which has also been observed by others (e.g., Just et al. 2015).

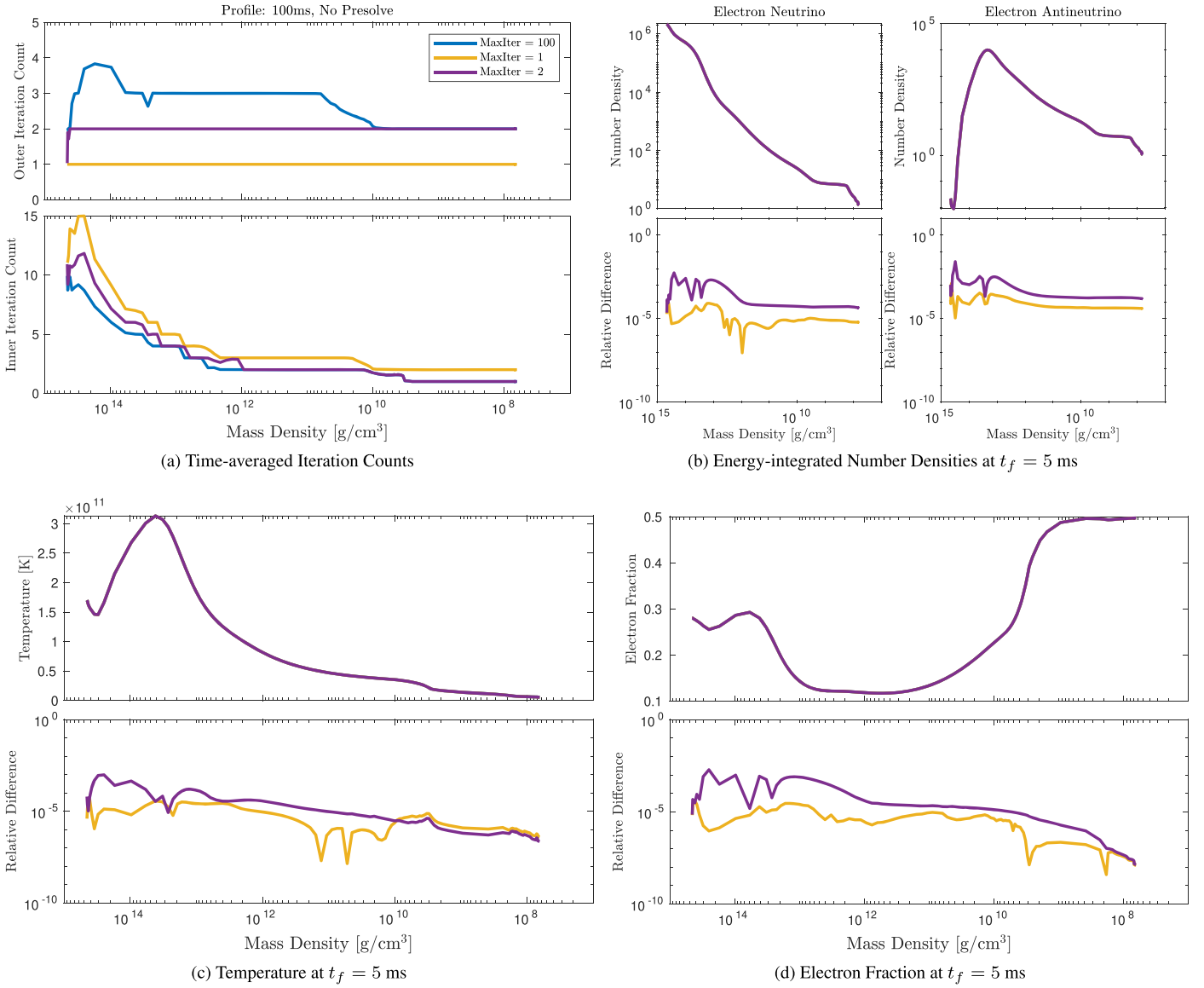Moving forward, we will continue to expand on the capabilities in thornado and incorporate the more

(a) Time-averaged Iteration Counts

(b) Energy-integrated Number Densities at $t_f = 5$ ms

(c) Temperature at $t_f = 5$ ms

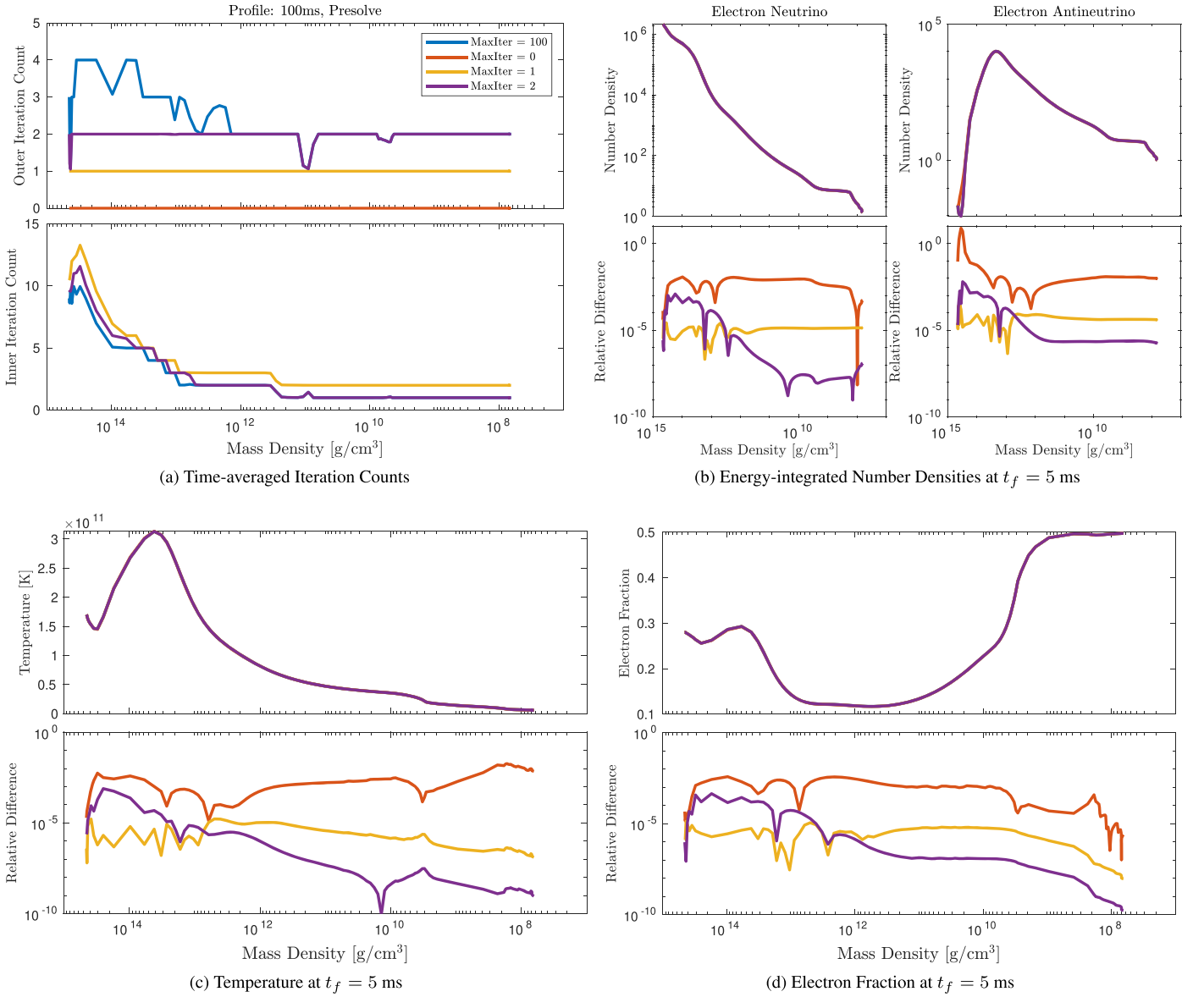(d) Electron Fraction at $t_f = 5$ ms

Figure 16. Time-averaged iteration counts and final-time ($t_f = 5$ ms) solutions for the nested AA solver without the presolve step, running to various maximum outer iterations (MaxIter). The blue line (MaxIter = 100) is considered as a reference.

comprehensive physics needed for realistic CCSN models. Next steps toward this goal include (i) incorporating special and general relativistic effects into the neutrino transport model; (ii) including muon and tau neutrinos; (iii) updating the opacity set to include, e.g., modern electron-capture rates, bremsstrahlung, and inelastic scattering on nucleons; (iv) coupling the neutrino transport equations with fluid equations to self-consistently model neutrino radiation hydrodynamics;

(v) porting the nonlinear solvers to modern hardware architectures (e.g., GPUs) and further analyzing implementation performance; and (vi) comparing thornado to other well-established CCSN simulation codes, such as AGILE-BOLTZTRAN (Liebendörfer et al. 2004), following the approaches in Just et al. (2015) and O'Connor et al. (2018). We are making progress in these directions and plan to report on the results in future publications.

(a) Time-averaged Iteration Counts

(b) Energy-integrated Number Densities at $t_f = 5$ ms

(c) Temperature at $t_f = 5$ ms

(d) Electron Fraction at $t_f = 5$ ms

Figure 17. Time-averaged iteration counts and final-time ($t_f = 5$ ms) solutions for the nested AA solver with the presolve step, running to various maximum outer iterations (MaxIter). The blue line (MaxIter = 100) is considered as a reference.

Appendix
Kernel Derivatives

The NES and pair kernels in Equations (16)–(19) are tabulated in terms of temperature T and degeneracy parameter $\eta = \mu_e/kT$; i.e., for given ε and ε', they are functions of the form

$$\Phi = \Phi(T, \eta). \qquad (A1)$$

However, for the coupled Newton's method in Section 4.2, we need kernel derivatives with respect to internal energy ϵ and

electron fraction Y_e. The electron chemical potential, along with all other quantities given by the EoS, is tabulated in terms of ρ, T, and Y_e. On the one hand, the variation of the kernel in Equation (A1) is

$$
\begin{aligned}
d\Phi &= \left(\frac{\partial \Phi}{\partial T}\right)_\eta dT + \left(\frac{\partial \Phi}{\partial \eta}\right)_T d\eta \\
&= \frac{1}{kT}\left(\frac{\partial \Phi}{\partial \eta}\right)_T \left(\frac{\partial \mu_e}{\partial \rho}\right)_{T,Y_e} d\rho + \left\{\left(\frac{\partial \Phi}{\partial T}\right)_\eta \right. \\
&\quad \left. + \left(\frac{\partial \Phi}{\partial \eta}\right)_T \left[\frac{1}{kT}\left(\frac{\partial \mu_e}{\partial T}\right)_{\rho,Y_e} - \frac{\eta}{T}\right]\right\} dT \\
&\quad + \frac{1}{kT}\left(\frac{\partial \Phi}{\partial \eta}\right)_T \left(\frac{\partial \mu_e}{\partial Y_e}\right)_{\rho,T} dY_e,
\end{aligned}
\tag{A2}
$$

where we used

$$
\begin{aligned}
d\eta &= \left(\frac{\partial \eta}{\partial \rho}\right)_{T,Y_e} d\rho + \left(\frac{\partial \eta}{\partial T}\right)_{\rho,Y_e} dT + \left(\frac{\partial \eta}{\partial Y_e}\right)_{\rho,T} dY_e \\
&= \frac{1}{kT}\left(\frac{\partial \mu_e}{\partial \rho}\right)_{T,Y_e} d\rho + \left[\frac{1}{kT}\left(\frac{\partial \mu_e}{\partial T}\right)_{\rho,Y_e}\right. \\
&\quad \left. - \frac{\eta}{T}\right] dT + \frac{1}{kT}\left(\frac{\partial \mu_e}{\partial Y_e}\right)_{\rho,T} dY_e.
\end{aligned}
\tag{A3}
$$

On the other hand, by considering the kernel as a function of ϵ and Y_e, the variation is

$$
\begin{aligned}
d\Phi &= \left(\frac{\partial \Phi}{\partial \epsilon}\right)_{Y_e} d\epsilon + \left(\frac{\partial \Phi}{\partial Y_e}\right)_\epsilon dY_e \\
&= \left(\frac{\partial \Phi}{\partial \epsilon}\right)_{Y_e}\left(\frac{\partial \epsilon}{\partial \rho}\right)_{T,Y_e} d\rho + \left(\frac{\partial \Phi}{\partial \epsilon}\right)_{Y_e}\left(\frac{\partial \epsilon}{\partial T}\right)_{\rho,Y_e} dT \\
&\quad + \left[\left(\frac{\partial \Phi}{\partial \epsilon}\right)_{Y_e}\left(\frac{\partial \epsilon}{\partial Y_e}\right)_{\rho,T} + \left(\frac{\partial \Phi}{\partial Y_e}\right)_\epsilon\right] dY_e,
\end{aligned}
\tag{A4}
$$

where we used

$$
d\epsilon = \left(\frac{\partial \epsilon}{\partial \rho}\right)_{T,Y_e} d\rho + \left(\frac{\partial \epsilon}{\partial T}\right)_{\rho,Y_e} dT + \left(\frac{\partial \epsilon}{\partial Y_e}\right)_{\rho,T} dY_e.
\tag{A5}
$$

Comparing Equations (A2) and (A4), we have

$$
\left(\frac{\partial \Phi}{\partial T}\right)_\eta + \left[\frac{1}{kT}\left(\frac{\partial \mu_e}{\partial T}\right)_{\rho,Y_e} - \frac{\eta}{T}\right]\left(\frac{\partial \Phi}{\partial \eta}\right)_T = \left(\frac{\partial \epsilon}{\partial T}\right)_{\rho,Y_e}\left(\frac{\partial \Phi}{\partial \epsilon}\right)_{Y_e},
\tag{A6}
$$

$$
\frac{1}{kT}\left(\frac{\partial \mu_e}{\partial Y_e}\right)_{\rho,T}\left(\frac{\partial \Phi}{\partial \eta}\right)_T = \left(\frac{\partial \epsilon}{\partial Y_e}\right)_{\rho,T}\left(\frac{\partial \Phi}{\partial \epsilon}\right)_{Y_e} + \left(\frac{\partial \Phi}{\partial Y_e}\right)_\epsilon.
\tag{A7}
$$

Solving for $(\partial \Phi/\partial \epsilon)_{Y_e}$ and $(\partial \Phi/\partial Y_e)_\epsilon$, we obtain the derivatives we need for Newton's method in terms of derivatives that can be computed directly from the opacity

and EoS tables,

$$
\begin{aligned}
\left(\frac{\partial \Phi}{\partial \epsilon}\right)_{Y_e} &= \left[\frac{1}{kT}\left(\frac{\partial \mu_e}{\partial T}\right)_{\rho,Y_e} - \frac{\eta}{T}\right] \bigg/ \left(\frac{\partial \epsilon}{\partial T}\right)_{\rho,Y_e}\left(\frac{\partial \Phi}{\partial \eta}\right)_T \\
&\quad + \left(\frac{\partial \epsilon}{\partial T}\right)_{\rho,Y_e}^{-1}\left(\frac{\partial \Phi}{\partial T}\right)_\eta,
\end{aligned}
\tag{A8}
$$

$$
\begin{aligned}
\left(\frac{\partial \Phi}{\partial Y_e}\right)_\epsilon &= \left\{\frac{1}{kT}\left(\frac{\partial \mu_e}{\partial Y_e}\right)_{\rho,T} - \left(\frac{\partial \epsilon}{\partial Y_e}\right)_{\rho,T}\left[\frac{1}{kT}\left(\frac{\partial \mu_e}{\partial T}\right)_{\rho,Y_e}\right.\right. \\
&\quad \left.\left. - \frac{\eta}{T}\right] \bigg/ \left(\frac{\partial \epsilon}{\partial T}\right)_{\rho,Y_e}\right\}\left(\frac{\partial \Phi}{\partial \eta}\right)_T \\
&\quad - \left\{\left(\frac{\partial \epsilon}{\partial Y_e}\right)_{\rho,T} \bigg/ \left(\frac{\partial \epsilon}{\partial T}\right)_{\rho,Y_e}\right\}\left(\frac{\partial \Phi}{\partial T}\right)_\eta.
\end{aligned}
\tag{A9}
$$

As discussed in Section 5.1 (and following, e.g., Mezzacappa & Messer 1999), tabulated quantities are evaluated using bilinear or trilinear interpolation, while derivatives are estimated by direct differentiation of the respective interpolation formulae.

ORCID iDs

M. Paul Laiu ⓘ https://orcid.org/0000-0002-2215-6968
Eirik Endeve ⓘ https://orcid.org/0000-0003-1251-9507
Ran Chu ⓘ https://orcid.org/0000-0001-7951-049X
J. Austin Harris ⓘ https://orcid.org/0000-0003-3023-7140
O. E. Bronson Messer ⓘ https://orcid.org/0000-0002-5358-5415

References

Adams, M. L. 2001, NSE, 137, 298
An, H., Jia, X., & Walker, H. F. 2017, JCoPh, 347, 1
Anderson, D. G. 1965, J. ACM, 12, 547
Asaithambi, R., & Mahesh, K. 2017, JCoPh, 341, 377
Ascher, U., Ruuth, S., & Spiteri, R. 1997, ApNM, 25, 151
Bassi, F., Franchina, N., Ghidoni, A., & Rebay, S. 2013, IJNMF, 71, 1322
Bethe, H. A., & Wilson, J. R. 1985, ApJ, 295, 14
Bruenn, S., Mezzacappa, A., Hix, W., et al. 2009, JPhCS, 180, 012018
Bruenn, S. W. 1985, ApJS, 58, 771
Bruenn, S. W., Blondin, J. M., Hix, W. R., et al. 2020, ApJS, 248, 11
Bruenn, S. W., De Nisco, K. R., & Mezzacappa, A. 2001, ApJ, 560, 326
Burrows, A. 2013, RvMP, 85, 245
Burrows, A., Radice, D., Vartanyan, D., et al. 2020, MNRAS, 491, 2715
Burrows, A., Reddy, S., & Thompson, T. A. 2006, NuPhA, 777, 356
Burrows, A., Vartanyan, D., Dolence, J. C., Skinner, M. A., & Radice, D. 2018, SSRv, 214, 33
Cardall, C. Y., Endeve, E., & Mezzacappa, A. 2013a, PhRvD, 88, 023011
Cardall, C. Y., Endeve, E., & Mezzacappa, A. 2013b, PhRvD, 87, 103004
Cernohorsky, J. 1994, ApJ, 433, 247
Cernohorsky, J., & Bludman, S. A. 1994, ApJ, 433, 250
Chu, R., Endeve, E., Hauck, C. D., & Mezzacappa, A. 2019, JCoPh, 389, 62
Cockburn, B., Hou, S., & Shu, C.-W. 1990, MaCom, 54, 545
Cockburn, B., Lin, S.-Y., & Shu, C.-W. 1989, JCoPh, 84, 90
Cockburn, B., & Shu, C.-W. 1989, MaCom, 52, 411
Cockburn, B., & Shu, C.-W. 1991, MMNA, 25, 337
Cockburn, B., & Shu, C.-W. 1998, JCoPh, 141, 199
Cockburn, B., & Shu, C.-W. 2001, JSCom, 16, 173
Endeve, E., Hauck, C. D., Xing, Y., & Mezzacappa, A. 2015, JCoPh, 287, 151
Evans, C., Pollock, S., Rebholz, L. G., & Xiao, M. 2020, SJNA, 58, 788
Hamilton, S., Berrill, M., Clarno, K., et al. 2016, JCoPh, 311, 241
Hannestad, S., & Raffelt, G. 1998, ApJ, 507, 339
Hesthaven, J. S., & Warburton, T. 2008, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis and Applications (Berlin: Springer)
Hix, W. R., Lentz, E. J., Endeve, E., et al. 2014, AIPA, 4, 041013

Hix, W. R., Messer, O. E. B., Mezzacappa, A., et al. 2003, PhRvL, 91, 201102
Hu, J., Shu, R., & Zhang, X. 2018, SJNA, 56, 942
Janka, H.-T. 2012, ARNPS, 62, 407
Juno, J., Hakim, A., TenBarge, J., Shi, E., & Dorland, W. 2018, JCoPh, 353, 110
Just, O., Bollig, R., Janka, H. T., et al. 2018, MNRAS, 481, 4786
Just, O., Obergaulinger, M., & Janka, H.-T. 2015, MNRAS, 453, 3386
Kelley, C. 2018, AcNum, 27, 207
Klöckner, A., Warburton, T., Bridge, J., & Hesthaven, J. S. 2009, JCoPh, 228, 7863
Knoll, D. A., & Keyes, D. E. 2004, JCoPh, 193, 357
Kuroda, T., Takiwaki, T., & Kotake, K. 2016, ApJS, 222, 20
Laiu, M. P., Chen, Z., & Hauck, C. D. 2020, JCoPh, 417, 109567
Larecki, W., & Banach, Z. 2011, JQSRT, 112, 2486
Larsen, E. W., & Morel, J. E. 1989, JCoPh, 83, 212
Lentz, E. J., Bruenn, S. W., Hix, W. R., et al. 2015, ApJL, 807, L31
Lentz, E. J., Mezzacappa, A., Messer, O. E. B., Hix, W. R., & Bruenn, S. W. 2012a, ApJ, 760, 94
Lentz, E. J., Mezzacappa, A., Messer, O. E. B., et al. 2012b, ApJ, 747, 73
Levermore, C. D. 1984, JQSRT, 31, 149
Liebendörfer, M., Messer, O. E. B., Mezzacappa, A., et al. 2004, ApJS, 150, 263
Liebendörfer, M., Rampp, M., Janka, H.-T., & Mezzacappa, A. 2005, ApJ, 620, 840
Lott, P., Walker, H., Woodward, C., & Yang, U. 2012, AdWR, 38, 92
Melson, T., Janka, H.-T., & Marek, A. 2015, ApJL, 801, L24
Mezzacappa, A., & Bruenn, S. W. 1993, ApJ, 410, 740
Mezzacappa, A., Endeve, E., Messer, O. E. B., & Bruenn, S. W. 2020, LRCA, 6, 4
Mezzacappa, A., & Messer, O. 1999, JCoAM, 109, 281
Mihalas, D., & Mihalas, B. W. 1999, Foundations of Radiation Hydrodynamics (New York: Dover)
Müller, B. 2016, PASA, 33, e048
Müller, B. 2020, LRCA, 6, 3
Müller, B., Janka, H.-T., & Marek, A. 2012, ApJ, 756, 84
Müller, B., Tauris, T. M., Heger, A., et al. 2019, MNRAS, 484, 3307
Ni, P., & Walker, H. 2010, A Linearly Constrained Least-squares Problem in Electronic Structure Computations, Tech. Rep. MS-1-13-46, Worcester Polytechnic Institute
O'Connor, E. 2015, ApJS, 219, 24
O'Connor, E., Bollig, R., Burrows, A., et al. 2018, JPhG, 45, 104001
Pareschi, L., & Russo, G. 2005, JSCom, 25, 129
Pollock, S., & Rebholz, L. 2019, Anderson Acceleration for Contractive and Noncontractive Operators, arXiv:1909.04638
Radice, D., Abdikamalov, E., Rezzolla, L., & Ott, C. D. 2013, JCoPh, 242, 648
Rampp, M., & Janka, H. T. 2002, A&A, 396, 361
Reddy, S., Prakash, M., & Lattimer, J. M. 1998, PhRvD, 58, 013009
Reed, W., & Hill, T. 1973, Triangular Mesh Methods for the Neutron Transport Equation, Tech. Rep. LA-UR-73-479, Los Alamos National Laboratory
Roberts, L. F., Ott, C. D., Haas, R., et al. 2016, ApJ, 831, 98
Sandu, A. 2019, SJNA, 57, 2300
Scieur, D., d'Aspremont, A., & Bach, F. 2020, Math. Program, 179, 47
Shibata, M., Kiuchi, K., Sekiguchi, Y., & Suwa, Y. 2011, PThPh, 125, 1255
Shu, C.-W. 2016, JCoPh, 316, 598
Skinner, M. A., Dolence, J. C., Burrows, A., Radice, D., & Vartanyan, D. 2019, ApJS, 241, 7
Smit, J. M., & Cernohorsky, J. 1996, A&A, 311, 347
Smit, J. M., Cernohorsky, J., & Dullemond, C. P. 1997, A&A, 325, 203
Steiner, A. W., Hempel, M., & Fischer, T. 2013, ApJ, 774, 17
Sumiyoshi, K., & Yamada, S. 2012, ApJS, 199, 17
Teukolsky, S. A. 2016, JCoPh, 312, 333
Thompson, T. A., Burrows, A., & Horvath, J. E. 2000, PhRvC, 62, 035802
Thompson, T. A., Burrows, A., & Pinto, P. A. 2003, ApJ, 592, 434
Toth, A., & Kelley, C. 2015, SJNA, 53, 805
Walker, H., & Ni, P. 2011, SJNA, 49, 1715
Woosley, S. E., & Weaver, T. A. 1995, ApJS, 101, 181
Xiong, T., Jang, J., Fengyan, L., & Qiu, J.-M. 2015, JCoPh, 284, 70
Yang, C., Meza, J. C., Lee, B., & Wang, L.-W. 2009, ACM Trans. Math. Softw., 36, 10
Zhang, J., O'Donoghue, B., & Boyd, S. 2020, SIAM J. Optim., 30, 3170