

## An adaptive mesh, GPU-accelerated, and error minimized special relativistic hydrodynamics code

Po-Hsun Tseng ,<sup>1</sup> Hsi-Yu Schive ,<sup>1, 2, 3, 4</sup> and Tzihong Chiueh ,<sup>1, 2, 3</sup>

<sup>1</sup> Institute of Astrophysics, National Taiwan University, Taipei 10617, Taiwan

<sup>2</sup> Department of Physics, National Taiwan University, Taipei 10617, Taiwan

<sup>3</sup> Center for Theoretical Physics, National Taiwan University, Taipei 10617, Taiwan

<sup>4</sup> Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan

### ABSTRACT

We present a new special relativistic hydrodynamics (SRHD) code capable of handling coexisting ultra-relativistically hot and non-relativistically cold gases. We achieve this by designing a new algorithm for conversion between primitive and conserved variables in the SRHD solver, which incorporates a realistic ideal-gas equation of state covering both the relativistic and non-relativistic regimes. The code can handle problems involving a Lorentz factor as high as  $10^6$  and optimally avoid the catastrophic cancellation. In addition, we have integrated this new SRHD solver into the code GAMER (<https://github.com/gamer-project/gamer>) to support adaptive mesh refinement and hybrid OpenMP/MPI/GPU parallelization. It achieves a peak performance of  $7 \times 10^7$  cell updates per second on a single Tesla P100 GPU and scales well to 2048 GPUs. We apply this code to two interesting astrophysical applications: (a) an asymmetric explosion source on the relativistic blast wave and (b) the flow acceleration and limb-brightening of relativistic jets.

*Keywords:* processes—relativistic shocks. galaxies—jets. methods: numerical. software—development, simulations.

### 1. INTRODUCTION

Many high energy astrophysical problems involve relativistic flows. The problems include, for example, collimated jets in active galactic nuclei (AGN) (Chiueh et al. 1991; Li et al. 1992; Blandford et al. 2019), collapsar models of long-duration gamma-ray bursts (Woosley 1993), magnetized relativistic winds and nebulae from pulsars (Kennel & Coroniti 1984a; Kennel & Coroniti 1984b; Chiueh 1989; Chiueh et al. 1998), and mildly relativistic wide-angle outflows in neutron star mergers (Mooley et al. 2018b; Mooley et al. 2018a; Ghirlanda et al. 2019; Fong et al. 2019). The full scope of these problems generally involves substantial temperature changes between jets (winds) and ambient gases. For this reason, the pioneering works of Taub (1948), Mathews (1971), and Mignone et al. (2005) suggested Taub-Mathews equation of state (TM EoS) that approximates the exact EoS (Synge 1957) for ultra-

relativistically hot (high- $T$  hereafter) gases coexisting with non-relativistically cold (low- $T$  hereafter) gases.

In addition, Noble et al. (2006) first compared the accuracy of several schemes for recovering primitive variables in the Riemann problems by means of self-checking tests (see Appendix B for details). Mignone & McKinney (2007) further proposed an inversion scheme for an arbitrary EoS and suggested that directly evolving the reduced energy density (i.e. the energy density subtracting the rest mass energy density from the total energy density) can avoid catastrophic cancellation in the non-relativistic limit. However, very few studies have systematically investigated how serious the catastrophic cancellation bears upon simulation results. This is partially due to the lack of exact solutions with which numerical results can be compared.

In this paper, we propose a new numerical scheme for conversion between primitive and conserved variables in the presence of both high- $T$  and low- $T$  gases. The new scheme is carefully tailored to avoid catastrophic cancellation. To verify its accuracy, we numerically derive the exact solutions of two relativistic Riemann problems with the TM EoS and compare with the simulation results. It demonstrates that our new special

Corresponding author: Po-Hsun Tseng  
zengbs@gmail.com

relativistic hydrodynamics (SRHD) code can minimize numerical errors compared with conventional methods.

We have integrated this new SRHD solver into the code GAMER (Schive et al. 2010; Schive et al. 2018) to facilitate GPU acceleration and adaptive mesh refinement (AMR). This new code, GAMER-SR, yields good weak and strong scalings using up to 2048 GPUs on **Piz-Daint**, the supercomputer at the Swiss National Supercomputing Centre (CSCS). Finally, we present two astrophysical applications, an asymmetric explosion and self-accelerating jets, to demonstrate the capability of this new code in extreme conditions. All simulation data are analysed and visualized using the package `yt` (Turk et al. 2011).

This paper is organized as follows. We introduce the equation of state and our new scheme for conversion between primitive and conserved variables in Section 2. In Section 3, we describe numerical methods, including the AMR structure, GPU acceleration, flexible time-steps, and correction of unphysical results. In Sections 4 and 5, we conduct numerical experiments to demonstrate the accuracy in both the non-relativistic (NR) and ultra-relativistic (UR) limits, the performance scalability, as well as the limitation of GAMER-SR. Finally, we present two astrophysical applications in Section 6 and draw the conclusion in Section 7.

Note that the speed of light and the Boltzmann constant are hard-coded to 1 in GAMER-SR. However, these physical constants are retained in this paper, except in Appendices, for dimensional consistency.

## 2. RELATIVISTIC HYDRODYNAMIC EQUATIONS

### 2.1. Relativistic hydrodynamics

Mass and energy-momentum conservation laws of a special relativistic ideal fluid follow

$$\partial_\nu(\rho U^\nu) = 0, \quad (1a)$$

$$\partial_\nu T^{\mu\nu} = 0, \quad (1b)$$

where

$$T^{\mu\nu} = \rho h U^\mu U^\nu / c^2 + p \eta^{\mu\nu}. \quad (2)$$

$\rho$  and  $p$  are the proper mass density and the pressure,  $U^\mu$  the four-velocity,  $\eta^{\mu\nu}$  the metric tensor of Minkowski space, and  $c$  the speed of light.  $h$  is the specific enthalpy, related to the specific thermal energy  $\epsilon$  by

$$h = c^2 + \epsilon + \frac{p}{\rho}. \quad (3)$$

An equation of state,  $h(\rho, p)$ , is required to close Equation (1) and will be discussed in Section 2.2. Throughout this paper, lower-case Greek indices run from 0 to 3, Latin ones from 1 to 3, and the Einstein

summation convention is used, except when stated otherwise.

Equation (1) can be rewritten into a convenient conservative form for numerical integration:

$$\partial_t D + \partial_j(DU^j/\gamma) = 0, \quad (4a)$$

$$\partial_t M^i + \partial_j(M^i U^j/\gamma + p \delta^{ij}) = 0, \quad (4b)$$

$$\partial_t E + \partial_j(M^j c^2) = 0, \quad (4c)$$

where  $\gamma$  is the Lorentz factor, and  $\delta^{ij}$  is the Kronecker delta notation.

The five conserved quantities  $D$ ,  $M^i$ , and  $E$  are the mass density, the momentum densities, and the total energy density, respectively. All conserved variables are related to primitive variables  $(\rho, U^i, p)$  through

$$D = \rho \gamma, \quad (5a)$$

$$M^i = D h U^i / c^2, \quad (5b)$$

$$E = D h y - p. \quad (5c)$$

Nevertheless, Mignone & McKinney (2007) suggest evolving the reduced energy density,

$$\tilde{E} := E - D c^2, \quad (6)$$

instead of the total energy density; otherwise, extraction of a tiny thermal energy for a cold gas from the total energy will lead to catastrophic cancellation. An intuitive approach is to subtract Equation (4a) from Equation (4c) so that we can obtain a new energy equation. However, the new energy flux,  $(M^j - D U^j / \gamma) c^2$ , also suffers from catastrophic cancellation in the NR limit. An appropriate new energy flux avoiding such a problem is  $(\tilde{E} + p) U^j / \gamma$ , which is mathematically equivalent to  $(M^j - D U^j / \gamma) c^2$ . The reduced energy equation for numerical integration can thus be cast into

$$\partial_t \tilde{E} + \partial_j((\tilde{E} + p) U^j / \gamma) = 0, \quad (7)$$

which is to replace Equation (4c).

Moreover, solving the Lorentz factor  $\gamma$  as three-velocity ( $v = \sqrt{v^i v_i}$ ) approaches  $c$  can seriously suffer from catastrophic cancellation when using  $\gamma = 1/\sqrt{1 - v^i v_i/c^2}$ . Therefore, we explicitly adopt four-velocities ( $U^i$ ) instead of three-velocity ( $v^i$ ) for numerical computations and solve the Lorentz factor in terms of four-velocities as

$$\gamma = \sqrt{1 + U^i U_i / c^2}, \quad (8)$$

by which significant digits in  $\gamma$  can be kept when  $\gamma \gg 1$ .

In addition, unlike the three-velocity bounded by  $c$ , four-velocity  $U^i$  has no upper limit and therefore can greatly reduce the risk of having  $v > c$  due to numerical errors.

## 2.2. Equations of state

GAMER-SR supports two kinds of EoS, the Taub-Mathews EoS (TM; [Taub 1948](#), [Mathews 1971](#), [Mignone et al. 2005](#)) and the polytropic EoS with a constant ratio of specific heats  $\Gamma$ . Assuming an ideal fluid in local thermal equilibrium and obeying the non-degenerate Maxwell-Jüttner statistics ([Jüttner 1911](#)), the exact EoS ([Synge 1957](#)) derived from the kinetic theory of relativistic gases is given by

$$\frac{h_{\text{exact}}}{c^2} = \frac{K_3(mc^2/k_B T)}{K_2(mc^2/k_B T)}, \quad (9)$$

where  $k_B$  and  $T$  are the Boltzmann constant and temperature, respectively, and  $K_n$  the  $n$ -th order modified Bessel function of the second kind. However, direct use of Equation (9) is computationally inefficient because the evaluation of Bessel function is numerically expensive.

Alternatively, the TM EoS is an approximation of Equation (9) and given by

$$\frac{h_{\text{TM}}}{c^2} = 2.5 \left( \frac{k_B T}{mc^2} \right) + \sqrt{2.25 \left( \frac{k_B T}{mc^2} \right)^2 + 1}. \quad (10)$$

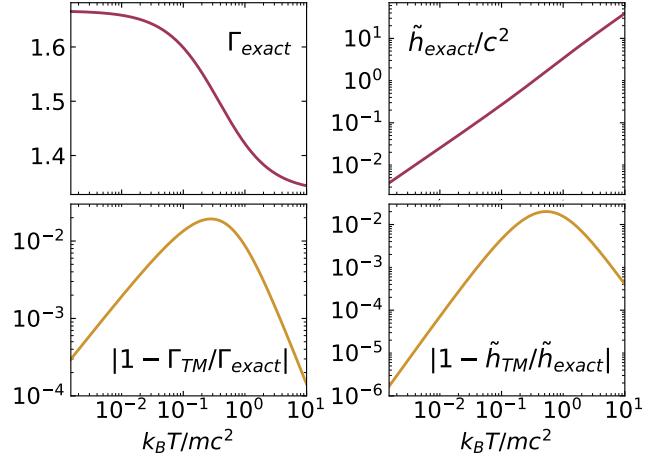
The effective  $\Gamma$  can be found by equating Equation (9) or Equation (10) to the polytropic EoS,

$$\frac{h_\Gamma}{c^2} = 1 + \frac{\Gamma}{\Gamma - 1} \left( \frac{k_B T}{mc^2} \right), \quad (11)$$

and solving  $\Gamma$  for the exact or TM EoS, respectively. As depicted in Figure 1, the maximum relative errors  $|1 - \Gamma_{\text{TM}}/\Gamma_{\text{exact}}|$  and  $|1 - \tilde{h}_{\text{TM}}/\tilde{h}_{\text{exact}}|$  are found to be only 1.9 and 2.0 per cent, respectively. In addition, Equation (10) approaches Equation (9) in both high- and low- $T$  limits. Detailed comparisons between Equation (9) and Equation (10) have been presented previously ([Mignone et al. 2005](#); [Ryu et al. 2006](#); [Mignone & McKinney 2007](#)) and we do not repeat here.

On the other hand, the polytropic EoS has the advantage of simplicity and therefore has been used in many SRHD codes, such as FLASH ([Fryxell et al. 2000](#)), CAFE ([Lora-Clavijo et al. 2015](#)), and XTREEM-FV ([Núñez-de la Rosa & Munz 2016](#)). However, the polytropic EoS cannot handle the case where relativistic gases and non-relativistic gases coexist, primarily because the ratio of specific heats depends sensitively on temperature when  $k_B T \sim mc^2$  (see the upper left panel in Figure 1). Moreover, the polytropic EoS with a non-relativistic  $\Gamma = 5/3$  and a relativistic  $\Gamma = 4/3$  does not satisfy the Taub's fundamental inequality for ideal gases ([Taub 1948](#))

$$\left[ \frac{h}{c^2} - \left( \frac{k_B T}{mc^2} \right) \right] \left[ \frac{h}{c^2} - 4 \left( \frac{k_B T}{mc^2} \right) \right] \geq 1, \quad (12)$$



**Figure 1.** The effective adiabatic index  $\Gamma$  (top left), the reduced enthalpy  $\tilde{h}/c^2 := h/c^2 - 1$  (top right) as a function of temperature. Bottom panels show that Equation (10) approaches Equation (9) in both high- and low- $T$  limits, where the maximum relative errors  $|1 - \Gamma_{\text{TM}}/\Gamma_{\text{exact}}|$  and  $|1 - \tilde{h}_{\text{TM}}/\tilde{h}_{\text{exact}}|$ , are only 1.9 and 2.0 per cent, respectively.

implying that  $\Gamma$  must lie between  $4/3$  and  $5/3$  for any positive and finite value of temperature. Although the polytropic EoS is physically incorrect, we still reserve this feature in GAMER-SR for fast computation of a pure non-relativistic or relativistic gas.

The other two important quantities are the Mach number ( $\mathcal{M}$ ) and the sound speed ( $c_s$ ), given by

$$\mathcal{M} = \frac{\sqrt{U^i U_i}}{U_s}, \quad (13)$$

and

$$\frac{c_s}{c} = \sqrt{\frac{k_B T / mc^2}{3h/c^2} \left( \frac{5h/c^2 - 8k_B T / mc^2}{h/c^2 - k_B T / mc^2} \right)}, \quad (14)$$

for the TM EoS, where  $U_s = c_s / \sqrt{1 - (c_s/c)^2}$ . The sound speed approaches  $c/\sqrt{3}$  at ultra-relativistic temperature and will be used in the Riemann solver.

## 2.3. Conversion between primitive and conserved variables

In standard Riemann-type numerical schemes, conversion between conserved and primitive variables is a common procedure for data reconstructions and flux computations. For non-relativistic hydrodynamics, this conversion can be carried out in a straightforward and analytical manner. However, designing an accurate and efficient conversion algorithm for a relativistic problem in the presence of NR gases, which involves root-finding, is challenging. This is because catastrophic cancellations may arise in the non-relativistic gas.

Here we propose a new conversion scheme to solve this problem based on the TM EoS. The reduced

energy density (Equation 6) and the momentum density (Equation 5b) satisfy the relation

$$\begin{aligned} & \left( \frac{\tilde{E}}{Dc^2} \right)^2 + 2 \left( \frac{\tilde{E}}{Dc^2} \right) - \left( \frac{|\mathbf{M}|}{Dc} \right)^2 \\ &= \frac{\tilde{h}^2}{c^4} + \frac{2\tilde{h}}{c^2} - 2 \left( \frac{k_B T}{mc^2} \right) \left( \frac{\tilde{h}}{c^2} + 1 \right) + \frac{(k_B T / mc^2)^2 (\tilde{h} + c^2)^2}{(\tilde{h} + c^2)^2 + \left( \frac{|\mathbf{M}|c}{D} \right)^2} \\ &:= f(\tilde{h}), \end{aligned} \quad (15)$$

where  $f$  is positive definite,  $\tilde{h} := h - c^2$  is the reduced enthalpy, and the temperature  $k_B T / mc^2$  is related to  $\tilde{h}$  by inverting Equation (10):

$$\frac{k_B T}{mc^2} = \frac{2(\tilde{h}/c^2)^2 + 4(\tilde{h}/c^2)}{5(\tilde{h}/c^2) + 5 + \sqrt{9(\tilde{h}/c^2)^2 + 18(\tilde{h}/c^2) + 25}}. \quad (16)$$

The conserved variables  $\tilde{E}$ ,  $M^j$ , and  $D$  on the left-hand side are known quantities updated at every time step, from which one can solve for  $\tilde{h}$ .

We adopt  $\tilde{h} = h - c^2$  instead of  $h$  as the root because the latter is dominated by rest mass energy density in the low- $T$  limit and thus will suffer from catastrophic cancellation when numerically extracting temperature from trailing digits.

Equation (15) is suitable for the Newton-Raphson iteration method as it is a monotonically increasing function of  $\tilde{h}$ . That is, Equation (15) has no zero derivative of  $\tilde{h}$  that might otherwise lead to a divergence of the iterative procedure. The Newton-Raphson method requires an initial guess of  $\tilde{h}$  and the derivative of Equation (15) for iteration, both of which are presented in Appendix A. We adopt the convergence criterion  $|1 - \tilde{h}_i/\tilde{h}_{i+1}| \leq \epsilon_{\text{machine}}$ , where  $\tilde{h}_i$  is the approximate root at the  $i$ th iteration and  $\epsilon_{\text{machine}}$  is the machine precision ( $10^{-7}$  and  $10^{-16}$  for single and double precision, respectively).

After obtaining  $\tilde{h}$ , we substitute it into Equation (5b) to get four-velocity:

$$U^i = \frac{M^i c^2}{D(c^2 + \tilde{h})}. \quad (17)$$

Next, we compute the Lorentz factor and proper mass density from Equation (8) and Equation (5a) and then use Equation (16) to obtain temperature. Finally, the pressure is given by

$$p = \rho c^2 \left( \frac{k_B T}{mc^2} \right). \quad (18)$$

Justifying the superiority of our new conversion scheme using  $\tilde{E}$ , we estimate the relative error of

computing  $a - b$  by (Higham 2002)

$$\frac{|a| + |b|}{|a - b|} \epsilon_{\text{machine}}. \quad (19)$$

Thus, the error of the new conversion scheme can be estimated by substituting  $\left[ (\tilde{E}/Dc^2)^2 + 2(\tilde{E}/Dc^2) \right]$  and  $(|\mathbf{M}|/Dc)^2$  for  $a$  and  $b$ , respectively, in Equation (19). The error in terms of primitive variables reads

$$\begin{aligned} & \left[ \frac{\gamma^2 (\tilde{h} + 1)^2 (1 + \beta^2) + \frac{T^2}{\gamma^2} - 2(\tilde{h} + 1)T - 1}{(\tilde{h} + 1)^2 + \frac{T^2}{\gamma^2} - 2(\tilde{h} + 1)T - 1} \right] \epsilon_{\text{machine}} \\ & \approx (1 + \mathcal{M}^2) \epsilon_{\text{machine}}, \end{aligned} \quad (20)$$

where  $\beta = \sqrt{v_i}/c$ . The approximate equality in Equation (20) holds for all finite temperature.

According to Equation (13) and Equation (20), we can decide whether to adopt single or double precision before simulations. Taking ultra-relativistic pulsar wind ( $\mathcal{M}_{\text{wind}} \approx 10^6$ ) as an example, we must use double precision to suppress the error to  $10^{-4}$ . But for mild-relativistic AGN jets ( $\mathcal{M}_{\text{jet}} < 10$ ), single precision is sufficient to reduce the error to  $10^{-5}$ .

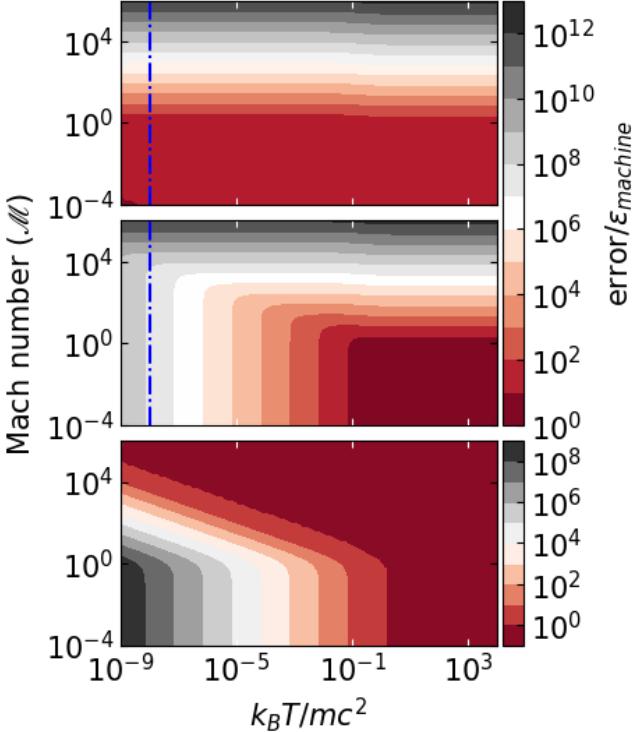
For the original scheme using the total energy density  $E$  instead of  $\tilde{E}$ , a similar error estimation can be performed by replacing  $\tilde{E}$  with  $E - Dc^2$  on the left-hand side of Equation (15), which gives

$$\left[ \frac{2\gamma^2 (\tilde{h} + 1)^2 + \frac{T^2}{\gamma^2} - 2(\tilde{h} + 1)T + (\tilde{h} + 1)^2 + 1}{(\tilde{h} + 2)\tilde{h} + \frac{T^2}{\gamma^2} - 2(\tilde{h} + 1)T} \right] \epsilon_{\text{machine}}. \quad (21)$$

Figure 2 shows the contour plots of Equation (20) for the new scheme (top panel) and Equation (21) for the original scheme (middle panel) as a function of  $\mathcal{M}$  and temperature. The bottom panel shows the ratio of Equation (21) to Equation (20). It demonstrates the advantage of using  $\tilde{E}$ . The top panel shows that using  $\tilde{E}$  in the conversion scheme is almost error-free when dealing with subsonic flows at any finite temperature, including the low- $T$  limit. In supersonic flows, the numerical errors proportional to  $\mathcal{M}^2$  are common and caused by finite digits of floating numbers. In comparison, the middle panel shows the error using  $E$ , which severely suffers from catastrophic cancellation in the low- $T$  limit even when  $\mathcal{M} \ll 1$ . See also Figure 16 in Appendix B.

On the other hand, conversion from primitive to conserved variables is also needed in the Riemann solver. This procedure involves straightforward substitution without the need of root-finding. We use

$$\frac{\tilde{h}}{c^2} = 2.5 \left( \frac{k_B T}{mc^2} \right) + \frac{2.25 (k_B T / mc^2)^2}{1 + \sqrt{2.25 (k_B T / mc^2)^2 + 1}}, \quad (22)$$



**Figure 2.** Numerical errors of the conversion from conserved to primitive variables as a function of  $\mathcal{M}$  and  $k_B T/mc^2$ . The top and middle panel show the errors of the new and original schemes estimated by Equation (20) and Equation (21), respectively. The bottom panel shows the ratio of Equation (21) to Equation (20). Figure 16 in Appendix B provides numerical evidences showing a remarkable consistency with the predicted values at  $k_B T/mc^2 = 10^{-8}$  (blue dashed-dotted line).

and

$$\frac{\tilde{E}}{Dc^2} = \frac{\left(\frac{|\mathbf{M}|}{Dc}\right)^2 + f(\tilde{h})}{1 + \sqrt{1 + \left(\frac{|\mathbf{M}|}{Dc}\right)^2 + f(\tilde{h})}}, \quad (23)$$

to compute  $\tilde{h}$  and  $\tilde{E}$ , where  $f(\tilde{h})$  can be computed from Equation (15) with known  $|\mathbf{M}|/Dc$ . Note that Equation (22) and Equation (23), following directly from Equation (10) and Equation (15) without any approximation, are written in a form without any subtraction to avoid catastrophic cancellation. In contrast, using Equation (5c) and Equation (6) to compute the reduced energy density  $\tilde{E}$  can suffer from catastrophic cancellation in the NR limit.

We close this section by providing a flowchart of the new conversion scheme in Figure 15 in Appendix B and by summarizing the equations actually solved by GAMER-SR. Other mathematically equivalent forms are unrecommended as they may suffer from catastrophic cancellation in the UR or NR limit.

- Evolution equations: Equation (4a, 4b, 7).

- Lorentz factor: Equation (8).
- Four-velocities: Equation (17).
- Temperature: Equation (16).
- Pressure: Equation (18).
- Reduced enthalpy: Equation (22).
- Reduced energy density: Equation (23).

### 3. NUMERICAL METHODS

#### 3.1. A GAMER Primer

Due to the flexibility and extensibility of GAMER (Schive et al. 2010; Schive et al. 2018), the SRHD module directly inherits the AMR structure and the MPI/OpenMP/GPU parallelization framework of hydrodynamics, and therefore we only provide a summary here. We define the base grid resolution as level-0 and the  $\ell$ th refinement as level- $\ell$ , where level- $\ell$  has a spatial resolution  $2^\ell$  times higher than that of the base level. Data in GAMER are always decomposed into patches, each of which consists of  $8^3$  cells, and the AMR implementation is realized by constructing a hierarchy of patches in an octree structure. According to user-defined refinement criteria, we can create or remove fine patches under the proper-nesting constraint.

In addition to the refinement criteria provided by the hydrodynamics module, we also implement two refinement criteria for SRHD: the gradient of the Lorentz factor and the magnitude of  $|\mathbf{M}|/D$ . The former aims to capture the thin and high- $\gamma$  shell in the Sedov-Taylor blast wave, while the latter ensures that the spine region in an over-pressured jet (cf. Figure 13) can be fully resolved. For all refinement criteria, the refinement thresholds on different levels can be set independently as run-time parameters.

We port the routines involving massive floating-point operations to GPUs such as the SRHD solvers and time-step calculations. On the other hand, we use CPUs to perform ghost-zone interpolation and patch refinement. As a result, we recommend using the refinement criteria only involving conserved variable for better performance because conserved variables are readily available from memory. By contrast, primitive variables can only be obtained by root-finding iteration, which is computationally expensive.

For enhancing software portability and reusability, GAMER not only supports both CPU-only and GPU modes but also allows the same physics modules to be shared by both CPU and GPU computations. Specifically, in the CPU-only mode, we compute different grid patches in the same MPI process in

parallel with OpenMP. In the GPU mode, we replace these OpenMP parallel clauses with CUDA thread blocks and then use threads within the same thread block to update all cells within the same grid patch. This scheme maximizes the reuse of physics routines, avoids redundant code development and maintenance, and significantly lowers the barrier of code extension, especially for developers not acquainted with GPU programming. We have utilized this CPU/GPU integration infrastructure in the SRHD implementation.

GAMER-SR supports the MUSCL-Hancock (Toro 2011) and VL (Falle 1991; van Leer 2006) schemes for numerical integrations and both piece-wise linear method (PLM; van Leer 1979) and piecewise parabolic method (PPM; Woodward & Colella 1984) for data reconstruction. For the Riemann solver, it supports both relativistic HLLC and HLLE solvers (Mignone & Bodo 2005; Mignone & Bodo 2006), which have been adapted not only to be compatible with the TM EoS by using the corresponding sound speed, Equation (14), but also to evolve the reduced energy density (i.e. replacing  $E$  with  $\tilde{E} + Dc^2$ ).

### 3.2. Flexible Time-step

GAMER-SR provides two Courant-Friedrichs-Lowy (CFL) conditions for time-step determination. The first one is based on the local signal propagation speed,  $S_{\max}$ , which gives maximum allowed time-steps in a wide dynamical range. Thus, it can significantly improve performance when the maximum  $v/c$  is not close to unity. The other is based on the speed of light, where we simply replace  $S_{\max}$  by  $c$ . It gives the most conservative estimation of time-steps and is more time-consuming when the flow speed is far less than  $c$ , although it is simple to implement and requires less computation.

To calculate  $S_{\max}$ , we first define  $\hat{\mathbf{u}}_s$  to be a spatial unit vector in the direction of sound propagation, we then apply the Lorentz boost with velocity  $-\boldsymbol{\beta}$  to the four-velocity of sound speed  $(\gamma_s, U_s \hat{\mathbf{u}}_s)$  from local rest frame to laboratory frame. We finally obtain the four-velocity of signal that travels in laboratory frame as follows:

$$(\gamma\gamma_s + \gamma U_s (\boldsymbol{\beta} \cdot \hat{\mathbf{u}}_s), U_s \hat{\mathbf{u}}_s + (\gamma - 1) U_s (\boldsymbol{\beta} \cdot \hat{\mathbf{u}}_s) \boldsymbol{\beta} + \boldsymbol{\beta} \gamma \gamma_s), \quad (24)$$

where  $\gamma$  and  $\gamma_s$  are the Lorentz factor of flow and of sound speed.  $U_s$  is the four-velocity of sound speed defined by  $c_s / \sqrt{1 - c_s^2}$ . Since the direction of the fastest signal propagation is in general parallel to flow velocity, we assume that both sound and flow propagate in the same direction (i.e.  $\hat{\mathbf{u}}_s = \hat{\boldsymbol{\beta}}$ ). The spatial components of Equation (24) then reduce to

$$(\beta \gamma_s + \gamma U_s) \hat{\boldsymbol{\beta}}. \quad (25)$$

Motivated by Equation (25), we simply choose  $|U^i| \gamma_s + \gamma U_s$  as the bound of each spatial component and sum over  $|U^i| \gamma_s + \gamma U_s$  for each spatial component to obtain

$$U_{\max} = \gamma_s (|U_x| + |U_z| + |U_{\bar{z}}|) + 3\gamma U_s, \quad (26)$$

where  $U_{x/y/z}$  is the  $x/y/z$ -component of the four-velocities of flow.

Note that Equation (26) is essentially the addition of flow speed and sound speed in special relativity theory. Converting Equation (26) back to three-velocity

$$S_{\max} = U_{\max} / \sqrt{1 + (U_{\max}/c)^2}, \quad (27)$$

and substituting Equation (27) into the CFL condition, we finally obtain the flexible time-step based on the local signal propagation speed for SRHD:

$$\Delta t = C_{\text{CFL}} \left( \frac{\Delta h}{S_{\max}} \right), \quad (28)$$

where  $\Delta h$  is the cell spacing and  $C_{\text{CFL}}$  the safety factor with a typical value of  $\sim 0.5$  for MUSCL-Hancock and VL schemes.

Note that Equation (28) can be reduced to its non-relativistic counterpart,

$$\Delta t = C_{\text{CFL}} \left( \frac{\Delta h}{|v_x| + |v_y| + |v_z| + 3c_s} \right), \quad (29)$$

when  $\gamma \sim 1$  and to  $\Delta t = C_{\text{CFL}} \Delta h / c$  when  $\gamma \gg 1$ .

### 3.3. Handling unphysical results

Unphysical results, for example, negative pressure, negative density and superluminal motion, can stem from the failure of the following criterion:

$$\left( \frac{\tilde{E}}{Dc^2} \right)^2 + 2 \left( \frac{\tilde{E}}{Dc^2} \right) - \left( \frac{|\mathbf{M}|}{Dc} \right)^2 > \epsilon_{\text{machine}}, \quad (30)$$

where the left-hand side involves the numerically updated quantities and  $\epsilon_{\text{machine}}$  is the machine epsilon – typically,  $2 \times 10^{-16}$  for double precision and  $1 \times 10^{-7}$  for single precision. The failure may take place in one of the following four steps:

#### (1) SRHD solver

SRHD solver is responsible for updating the conserved variables by a given time-step. If unphysical result occurs in a cell, we redo data reconstruction by reducing the original minmod coefficient by a factor of 0.75. If the failure still occurs, we further reduce the minmod coefficient repeatedly until Equation (30) passes or the reduced minmod coefficient vanishes. Note that interpolating with a vanished minmod coefficient is

essentially equivalent to the piece-wise constant spatial reconstruction.

### (2) Grid refinement

Unphysical results may occur during grid refinement when performing interpolations on parent patches. The remedy here is the same as that in the SRHD solver. We repeat the interpolation process with a reduced minmod coefficient on the conserved variables until Equation (30) passes or the minmod coefficient vanishes. A vanished minmod coefficient is essentially equivalent to directly copying data from the parent patch without interpolation.

### (3) Ghost-zone interpolation

To preserve conservation, where the volume-weighted average of child patch data are equal to its parent patch data, we normally fill the ghost zones of the patches on level  $\ell + 1$  by interpolating the conserved variables on level  $\ell$  when the ghost zones lie on level  $\ell$ . However, if unphysical results occur, we interpolate primitive variables instead. Interpolating primitive variables is more robust than interpolating conserved variables since Equation (30) is always satisfied. After interpolation, we fill the ghost zones with the conserved variables derived from the interpolated primitive variables. Note that this procedure still preserves conservation because ghost zones do not affect conservation.

### (4) Flux correction operation

For a leaf coarse patch adjacent to a coarse-fine interface, the flux difference between the coarse and fine patches on the interface will be used to correct the coarse-patch conserved variables adjacent to this interface. If unphysical results are found after this flux correction, we simply ignore the correction on the failed cells. Skipping the correction will break the strict conservation but it only occurs rarely.

## 4. TEST PROBLEMS

To understand how evolving the total energy density may deteriorate simulation results and to demonstrate how much the new scheme improves, we compare the results from evolving  $E$  by the flux  $M$  (original scheme) with that from evolving  $\tilde{E}$  by the flux  $(\tilde{E} + p)U_x/\gamma$  (new scheme). Since catastrophic cancellation is likely to occur in UR and NR limits, we will conduct several test problems in these two limits. All simulations throughout this paper adopt the HLLC Riemann solver and PLM data reconstruction unless otherwise specified.

### 4.1. Convergence test for sinusoidal waves

We perturb proper mass density in the high- and low- $T$  limits to compare the accuracy of both schemes over a wide dynamical range. We construct the initial

conditions as follows. All cases share homogeneous and static background with proper mass density  $\rho_0 = 1$  on uniform grids, whereas the ambient temperatures are set to  $k_B T/mc^2 = 10^{10}$  and  $10^{-10}$  for the high- and low- $T$  limits, respectively. We then sinusoidally perturb the background with a tiny amplitude,  $\delta\rho/\rho_0 = 10^{-6}$ .

To monitor how errors in the numerical solution decrease as a function of increasing spatial resolution in the three-dimensional space, we adopt a propagating wave along the diagonal direction of the simulation cubic box with the periodic boundary condition. Thus, the analytical solution is  $\rho(\mathbf{x}, t) = \rho_0 + \delta\rho \sin[(x + y + z)/\sqrt{3} - c_s t]$ , where  $c_s$  is the sound speed given by Equation (14).

We define the L1-norm error as

$$L1(Q) = \frac{1}{N} \sum_{i=1}^N \left| 1 - \frac{Q_{\text{numerical}}(\mathbf{x}_i)}{Q_{\text{analytical}}(\mathbf{x}_i)} \right|, \quad (31)$$

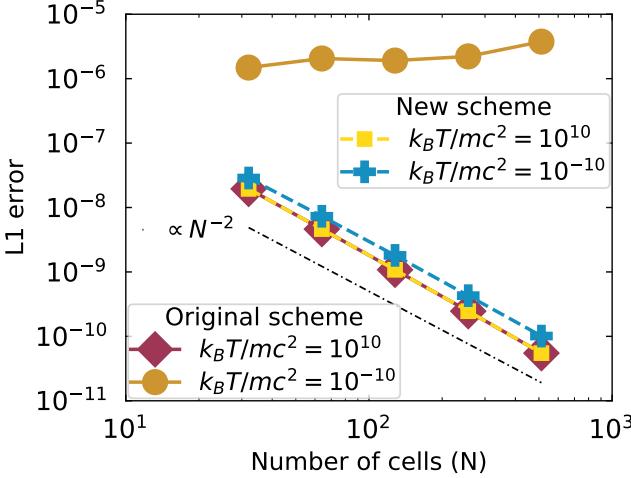
where  $Q_{\text{numerical}}(\mathbf{x}_i)$  is the numerical solution of  $i$ -th cell at  $\mathbf{x}_i$  and  $Q_{\text{analytical}}(\mathbf{x}_i)$  is the corresponding analytical solution. We then calculate the L1 error of the proper mass density along the wave propagating direction. As shown in Figure 3, the L1 errors of the new scheme in both the high- $T$  limit (■) and low- $T$  limit (+) decrease as  $N^{-2}$ , consistent with the second-order accuracy of the MUSCL-Hancock scheme with PLM data reconstruction. However, the error of the original scheme in the low- $T$  limit (●) is much larger and roughly equal to a constant of  $2 \times 10^{-6}$ . This is expected because the error arising from the original scheme can be estimated from Equation (21) in the NR limit:

$$\frac{4}{3(k_B T/mc^2)} \epsilon_{\text{machine}}, \quad (32)$$

where  $k_B T/mc^2 = 10^{-10}$  and  $\epsilon_{\text{machine}} \sim 10^{-16}$  for double precision.

We thus conclude that for the original scheme in the NR limit, the cancellation between  $(E/Dc^2)^2$  and  $[(|\mathbf{M}|/Dc)^2 + 1]$  leads to an error of  $\sim 10^{-6}$  when computing primitive variables, roughly consistent with the L1 error (●). For the opposite high- $T$  limit (♦), the discretization error, however, completely overwhelms the error ( $\sim 4\epsilon_{\text{machine}} \sim 4 \times 10^{-16}$ ) estimated from Equation (21) in the high- $T$  limit, thus dominating the L1 error. The error arising from the cancellation in the new scheme,  $(\tilde{E}/Dc^2)^2 + 2(\tilde{E}/Dc^2) - (|\mathbf{M}|/Dc)^2$ , on the left side of Equation (15), is close to  $\epsilon_{\text{machine}}$  in both the high- and low- $T$  limits when  $\mathcal{M} < 1$  (see Section 2.3 and Appendix B for details).

### 4.2. 1-D relativistic Riemann problems



**Figure 3.** Convergence test for sinusoidal waves. The L1 errors of the new scheme in both the high- $T$  limit (■) and low- $T$  limit (+) decrease as  $N^{-2}$ , consistent with the second-order accuracy of the MUSCL-Hancock scheme with PLM data reconstruction. However, the L1 error of the original scheme in the low- $T$  limit (○) is much larger and roughly equal to a constant of  $2 \times 10^{-6}$ . This is expected because the error in the original scheme can be estimated from Equation (32). For the original scheme in the opposite high- $T$  limit (◆), the discretization error, however, completely overwhelms the error ( $\sim 4 \times 10^{-16}$ ) estimated from Equation (21) in the high- $T$  limit when  $\gamma \sim 1$ , thus dominating the L1 error.

The 1-D Riemann problem (Sod 1978) has played an important role by providing exact nonlinear solutions against which (relativistic) hydrodynamic codes can be tested. Riemann problem is an initial-value problem with a piece-wise constant initial data that has a single discontinuity in the domain of interest. In this subsection, we directly compare the new and original schemes by simulating two relativistic Riemann problems. We then demonstrate that the new scheme handles both the UR and NR limits very well. By contrast, the original scheme severely suffers from numerical errors in the NR limit. Both schemes share the same numerical setup, e.g., MUSCL-Hancock integration, PLM data reconstruction, hybrid van-Leer, generalized minmod slope limiter, and uniform grids with the outflow boundary condition. In addition, we have numerically derived the exact solution of a nontrivial relativistic Riemann problem with the TM EoS (see Appendix C for details) in order to verify the numerical results.

#### 4.2.1. Ultra-relativistic limit

We simulate a head-on collision of two identical gases with  $\gamma = 10^6$  and  $k_B T/mc^2 = 10^5$  with uniform 512 grids. The computational domain is in the interval  $[0, 1]$ . The initial discontinuity is located at  $x = 0.5$ . The first row of

Table 1 presents the initial right and left states. Figure 4 shows the results at  $t = 1.0$ . The left panels show the entire simulation domain, while the right panels show the zoom-in image of the post-shock region, which has been violently heated up to ultra-relativistically hot temperature ( $k_B T/mc^2 \sim 10^{11}$ ) by the extremely high- $\gamma$  gases flowing inwards from both sides. As can be seen, the new scheme (○) fully agrees with the original scheme (★) on the large-scale profile but also on the small-scale errors, meaning that the new scheme does not sacrifice the numerical accuracy in the UR limit. In addition, we notice that the non-negligible and spurious waves occur in the post-shock region, which are not due to root-finding iterations but to spatial discretization errors as the spurious waves can be reduced by increasing spatial resolution.

#### 4.2.2. Mixed limits

To demonstrate that the new scheme can handle a large dynamical range covering both extremely hot and extremely cold gases, we simulate a nontrivial Riemann problem where the temperature straddles between the high- and low- $T$  limits. This initial condition evolves into a cold left-traveling rarefaction wave separated by a contact discontinuity to match an extremely hot downstream of an ultra-relativistic shock traveling toward the right. Also, we have numerically derived the exact solution of this particular Riemann problem with the TM EoS (see Appendix C). The second row in Table 1 shows the initial left and right states. The simulation adopts a computational domain  $[0, 100]$  with 102,400 cells. Since the speed of the right-traveling shock is 276 times faster than that of the left-traveling rarefaction wave, we put the initial discontinuity at  $x = 5 \times 10^{-2}$  to provide an ample space for the right-traveling shock.

Figure 5 shows the results at  $t = 80$ , where there are three points to be emphasized. First, we find not only that the shock front at  $x = 26$  is well resolved by 3–4 cells but also that the new scheme (★) agrees well with the exact solution of the TM EoS (—), as shown in all insets. Second, the L1 error, defined by Equation (31), of the density profile from the original scheme (○) is 23 per cent within the region between the head of rarefaction wave ( $x = 2.67 \times 10^{-2}$ , the third number from top in the leftmost column of Table 2 in Appendix C) and initial discontinuity ( $x = 5 \times 10^{-2}$ ), consistent with the 20 per cent error estimated by Equation (32) with  $k_B T/mc^2 = 8 \times 10^{-6}$ . Similar conclusions can be drawn for other physical quantities. However, in the region  $5 \times 10^{-2} < x < 0.3$  swept by the right-traveling contact discontinuity, errors of the original scheme are much larger than the

**Table 1.** The left and right initial states of the Riemann problems in Section 4.2. We denote the left/right states by the subscript L/R.

	$p_L$	$\rho_L$	$U_L$	$p_R$	$\rho_R$	$U_R$	Floating-point format
Ultra-relativistic limit	1.0	$10^{-5}$	$10^6$	1.0	$10^{-5}$	$-10^6$	Double precision
Mixed limits	$10^{-4}$	$10^2$	$10^{-3}$	$10^{-10}$	$10^{-12}$	$-10^2$	Single precision

estimate, which requires further investigation. Third, the solutions of the TM EoS (—) match well with both  $\Gamma = 5/3$  (----) in the NR region ( $x < 0.21$ ) and  $\Gamma = 4/3$  (-----) in the UR region ( $x > 0.27$ ). It demonstrates the capability of capturing the transition from  $\Gamma = 5/3$  (for  $k_B T/mc^2 \rightarrow 0$ ) to  $\Gamma = 4/3$  (for  $k_B T/mc^2 \rightarrow \infty$ ) for the new scheme. The exact solutions of this test are shown in Table 2 in Appendix C.

#### 4.3. Multi-dimensional grid effects for high- $\mathcal{M}$ flows

To investigate the detrimental impact of grid effects on the evolution of ultra-relativistic and high Mach number hydrodynamic problems, we separately simulate two identical three-dimensional mono-direction flow with different flow directions. One flow is along the diagonal direction of the simulation box and the other is parallel to the grid direction. Both simulations share the same numerical set-up as follows. Flows are initially represented by cylinders extending to the boundaries of a periodic cubic box with a width  $L$ . The cylinder diameter is  $D = 0.028L$ . The proper mass density ratio of the flow and the ambient is  $\rho_{\text{flow}}/\rho_{\text{amb}} = 10^{-5}$ . The temperatures of the flows and the ambient are  $k_B T_{\text{flow}}/mc^2 = 1.0$  and  $10^{-5}$ , respectively. The four-velocity ( $\gamma\beta$ ) profile inside the flow source is  $10^6(1 + \cos(2\pi r/D))$ , where  $r$  is the distance from the flow axis inside the source. Other physical quantities are uniformly distributed inside the source.

The AMR base level is covered by  $64^3$  cells in all cases. We adopt the gradient of the proper mass density and the magnitude of  $|\mathbf{M}|/D$  as the two inclusive refinement criteria. We refine a patch if the gradient of a cell satisfies

$$\frac{\Delta h_\ell}{Q} \left( \left| \frac{\partial Q}{\partial x} \right| + \left| \frac{\partial Q}{\partial y} \right| + \left| \frac{\partial Q}{\partial z} \right| \right) > C_Q, \quad (33)$$

where  $Q = \rho$ ,  $C_Q = 0.3$ , and  $\Delta h_\ell$  is the cell size at refinement level  $\ell$ . This criterion aims to capture the finger structure due to instabilities at the interface between the flow and the ambient gases. Also, a patch will be refined when any cell satisfies  $|\mathbf{M}|/D > 10^4$  so that the high-speed region is refined to the finest level.

Figure 6 shows the simulation results at  $t = 0.4L/c$ . In Figures 6(a) and 6(b), we adopt four AMR levels to ensure that the flow diameter can be resolved by 28 cells. The extremely high Mach number ( $\mathcal{M} \sim 10^6$ ) flow

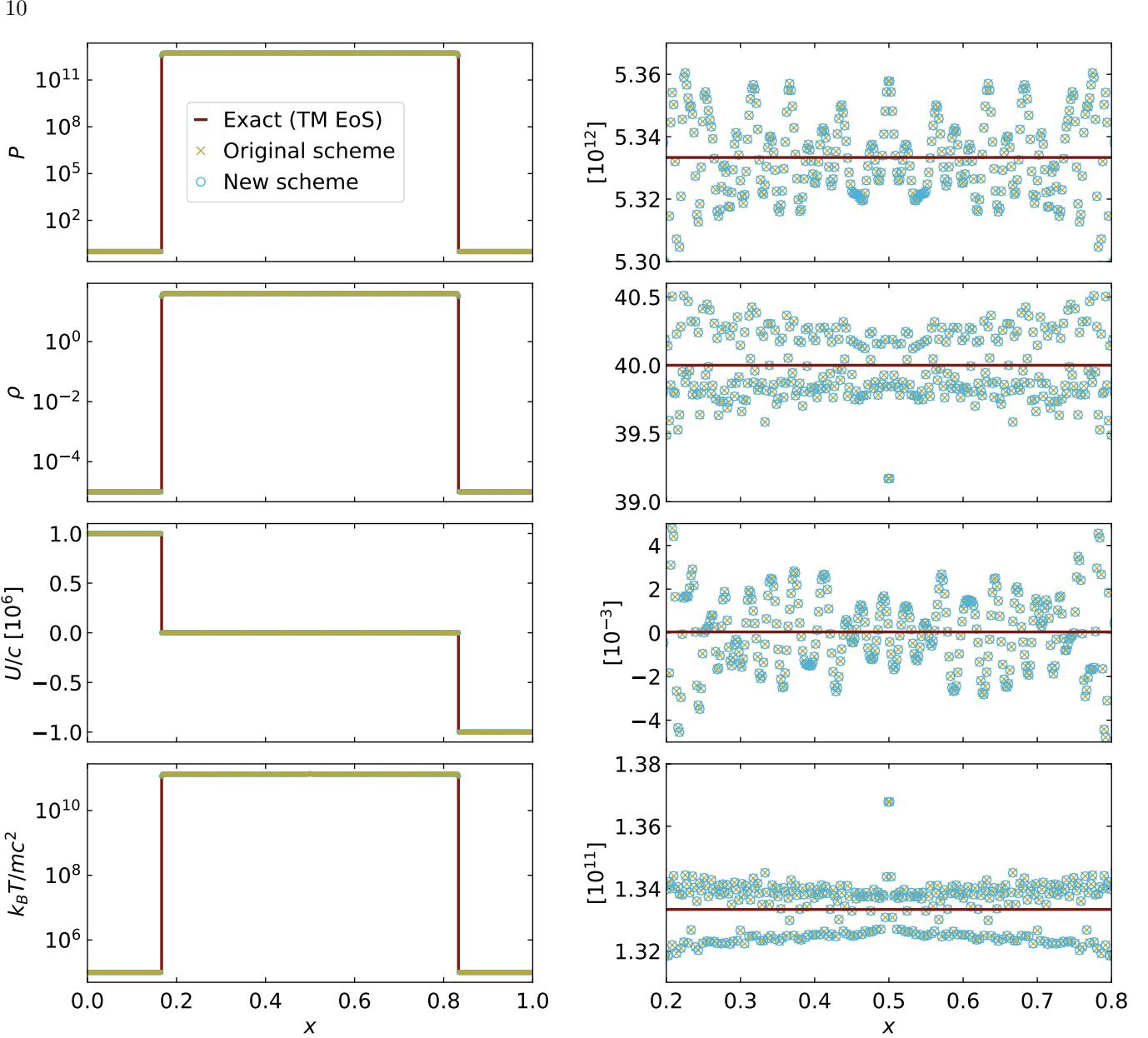
leaves any instability short of time to develop, and one expects a smooth flow-ambient interface. However, the interface of the oblique flow turns out to be subject to severe dissipation. The fuzzy-looking cross-sections in the transverse slices of the oblique flow (right column in Figure 6(b)) suggest that the dissipation is caused by numerical instabilities when high Mach number flow travels obliquely across Cartesian grids. This numerical problem is not limited to relativistic high Mach number flows but also occurs in non-relativistic high Mach number flows.

To examine this issue further, we increase the spatial resolution by a factor of 2 and decrease the time-step by a factor of 0.3 from the standard Courant condition. The results (Figures 6(c) and 6(d)) indicate that increasing spatial and temporal resolution can neither significantly ameliorate the dissipation nor help the oblique flow converge to the horizontal flow. This artificial grid effect can adversely influence the study of high-speed jets, especially for hydrodynamical instabilities near the jet boundaries.

An example of this boundary instability is the finger-like pattern observed immediately outside the parallel flow (right column in Figure 6(a)), which we believe to arise from a genuine instability seeded by discretization noise. The finger-like pattern has a higher temperature than the ambient, and in fact consists of two-dimensional flat sheets along the flow. This is demonstrated in Figure 7 with transverse slices cut through ‘B’ and ‘C’. The patterns are identical to that cut through ‘A’ in Figure 6(c). These 2-D sheet pattern persists even after adding 1 per cent level of white noise into the background density, illustrating that the coherence of sheets along the flow direction is genuinely generated by the high-speed flow boundaries. This finger pattern is similar, but not identical, to the curvature-driven fingers of a knotted jets reported recently (Gourgouliatos & Komissarov 2017). Our flow has a smooth and parallel boundary without any curvature to drive the fingers.

## 5. PERFORMANCE SCALING

We measure both strong and weak scalings of GAMER-SR with AMR and hybrid MPI/OpenMP/GPU parallelization. The simulations were conducted on the Piz-Daint supercomputer that provides a 12-core Intel

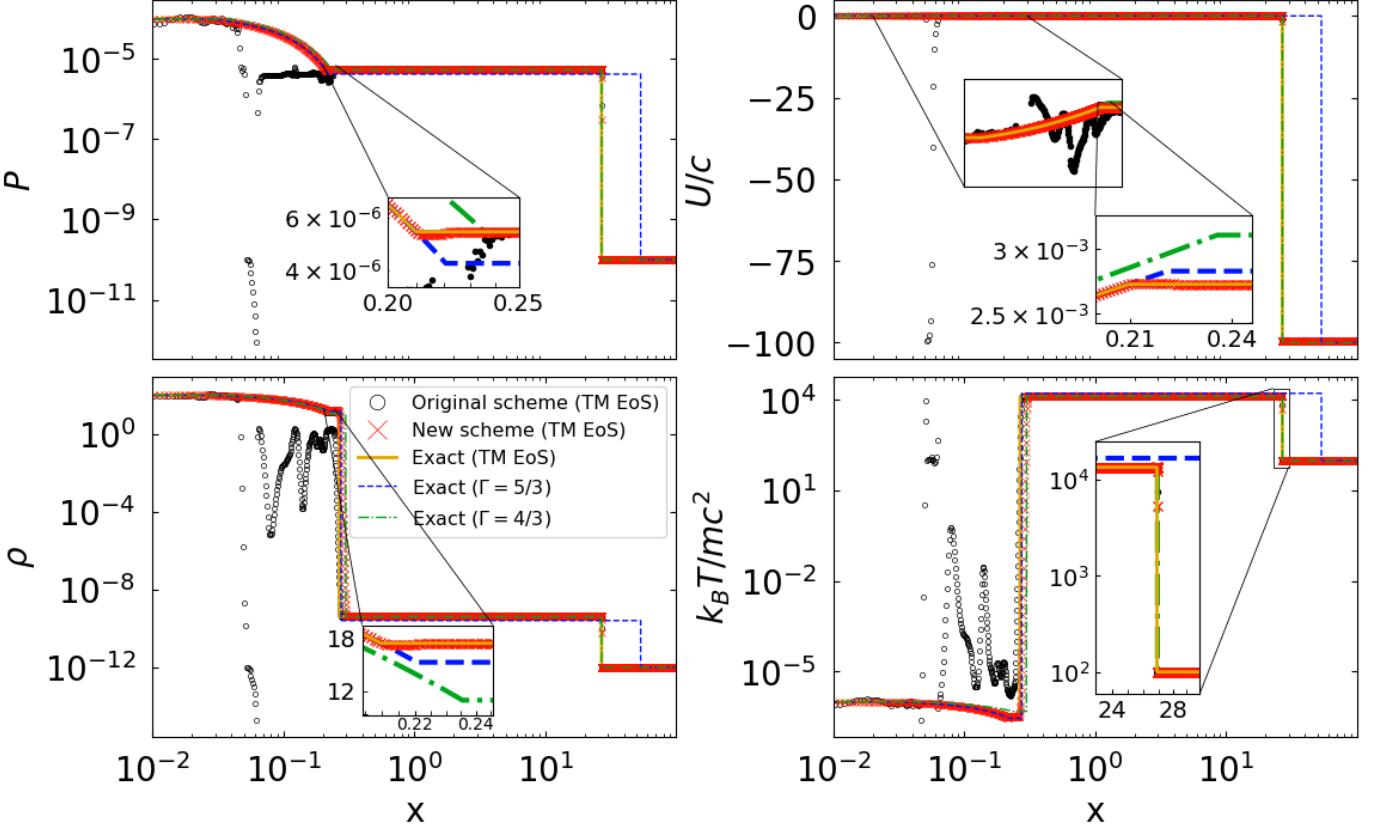


**Figure 4.** Riemann problem in the UR limit with a head-on collision of two identical gases with  $\beta\gamma = 10^6$  and  $k_B T / mc^2 = 10^5$  at  $t = 1.0$ . The left column shows the entire simulation domain, while the right column shows the zoom-in image of the post-shock region. From top to bottom: pressure, proper mass density, four-velocity, and temperature. Note that we plot the Mach number in the zoom-in image (right column) of four-velocity so as to readily compare the amplitude of velocity oscillation with sound speed. The new scheme ( $\circ$ ) fully agrees with the original scheme ( $\star$ ) not only on the large-scale profiles but also on the small-scale errors, meaning that the new scheme ( $\circ$ ) does not sacrifice (or improve) the numerical accuracy in the UR limit for that in the NR limit.

Xeon E5-2690 CPU and a Tesla P100 GPU on each computing node. Strong and weak scalings are defined as how the simulation wall time varies with the number of computing nodes for a fixed total problem size and for a fixed problem size per node, respectively. We launch one MPI process with 12 OpenMP threads per node and enable GPU acceleration with single precision.

We divide this section into two parts. First, we measure the strong scaling of a relativistic jet

simulation. The simulation set-up, such as initial condition, boundary condition, and grid refinement, follows those described in Section 6.2. Second, we present the weak scaling for periodic and spherical multi-blast waves test (see Figure 8). Note that the overall performance (i.e. total cell updates per second) in both tests have excluded ghost zones in the cell number counts.



**Figure 5.** Riemann problem in the mixed UR and NR limits at  $t = 80$ . The second row in Table 1 shows the initial condition. Clock-wise from top-left: pressure, 4-velocity, temperature, and proper mass density. We find not only that the shock front at  $x = 26$  is well resolved by 3–4 cells but also that the new scheme ( $\star$ ) agrees well with the exact solution of the TM EoS (—), as shown in all insets. The L1 error of the density profile from the original scheme (○) is 23 per cent within the region between the head of rarefaction wave and initial discontinuity (i.e.  $2.67 \times 10^{-2} < x < 5 \times 10^{-2}$ ), consistent with the 20 per cent error estimated by Equation (32) with  $k_B T / mc^2 = 8 \times 10^{-6}$ . However, in the region  $5 \times 10^{-2} < x < 0.3$  swept by the right-traveling contact discontinuity, errors of the original scheme are much larger than the estimate, which requires further investigation. The TM profiles match well with both the  $\Gamma = 5/3$  profiles (----) in the NR region ( $x < 0.21$ ) and the  $\Gamma = 4/3$  profiles (-----) in the UR region ( $x > 0.27$ ).

#### (1) Strong scaling:

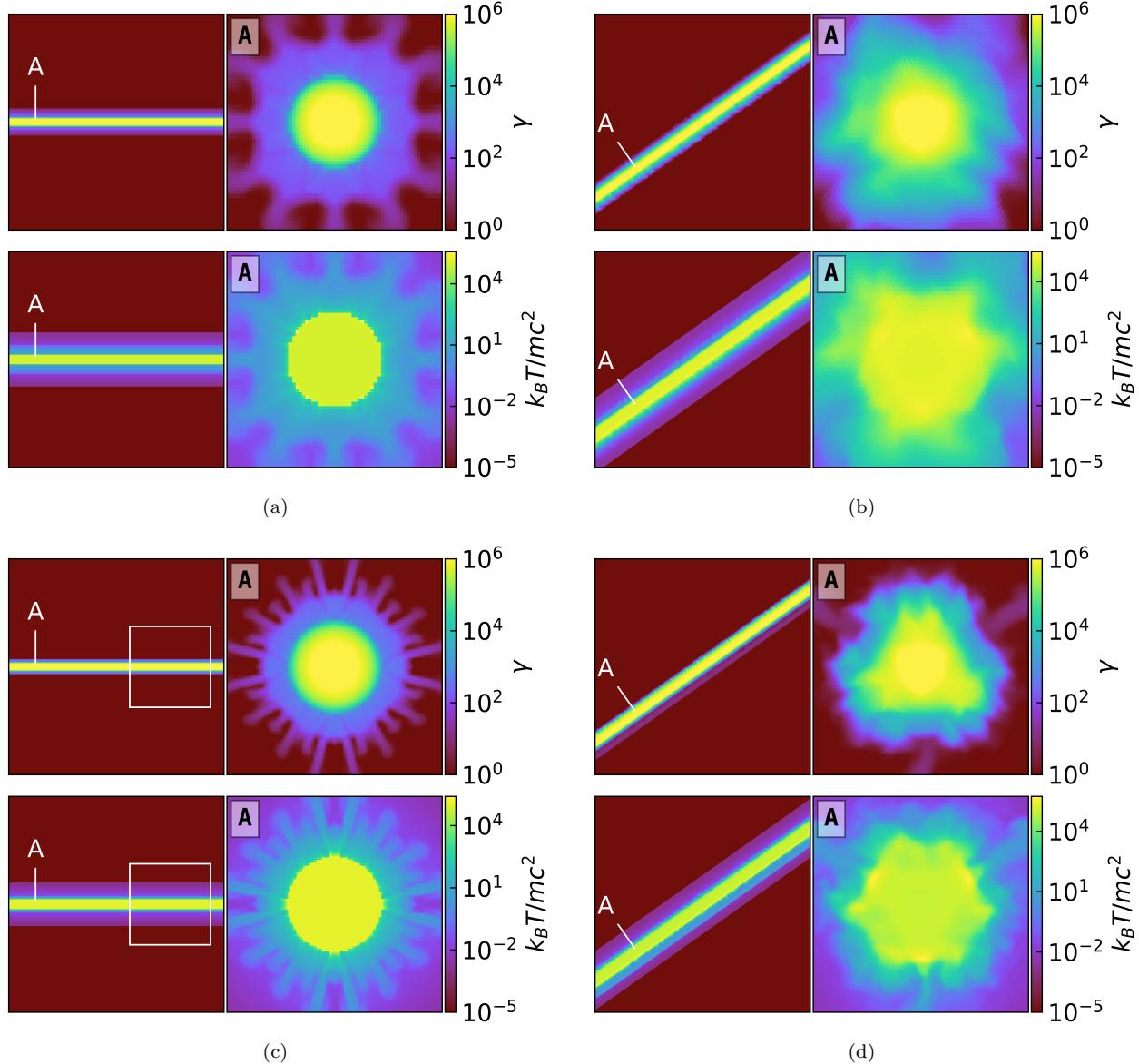
Figure 9 shows the strong scaling results. The parallel efficiency for strong scaling is defined by  $[T(N_{\text{ref}})/T(N_{\text{node}})] / (N_{\text{node}}/N_{\text{ref}})$ , where  $T(N_{\text{node}})$  is the simulation wall time using  $N_{\text{node}}$  nodes.  $N_{\text{ref}}$  is the number of nodes for reference and is fixed to 16 in our test. The overall performance reaches  $5 \times 10^{10}$  cell updates per second with 2048 GPU nodes, corresponding to a parallel efficiency of  $\sim 45$  per cent. The deviation from the ideal scaling is mainly due to MPI communication as (1) the load-imbalance in the EoS iteration is found to be only subdominant, and (2) the MPI communication time accounts for 45 per cent of the wall time with 2048 nodes. Thus, the parallel efficiency would increase from 0.45 to  $0.45/(1 - 45\%) \sim 0.82$  if excluding MPI communication.

#### (2) Weak scaling:

The periodic computational domain is composed of

identical cubic subdomain, each of which has a volume of  $L^3$  and has an explosion source at its own centre with a radius of  $r_{\text{src}} = 0.4L$  and an ultra-relativistic temperature of  $k_B T_{\text{src}}/mc^2 = 10^5$ . The uniform ambient gas has a non-relativistic temperature of  $k_B T_{\text{amb}}/mc^2 = 10^{-5}$  and a density of  $\rho_{\text{src}} = \rho_{\text{amb}} = 1.0$ . Each subdomain is composed of a  $64^3$  base-level grid with three refinement levels, where we refine patches based on the gradient of the reduced energy density. All blast waves evolve from  $t = 0$  to  $t = 0.5L/c$ . We measure the overall performance and parallel efficiency using 1–2048 nodes, where each node computes one subdomain. Figure 8 shows a temperature slice ( $z = 1.5L$ ) through the centre of four blast waves at  $t = 0.5L/c$ , with the grid patches overlaid.

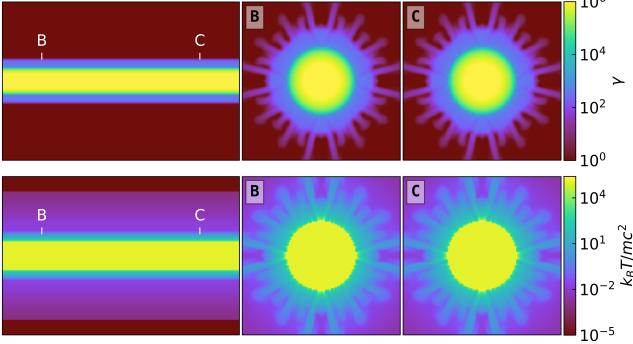
Figure 10 shows the weak scaling results. The parallel efficiency for weak scaling is defined by  $T(1)/T(N_{\text{node}})$ , where  $T(N_{\text{node}})$  is defined the same as the strong scaling.



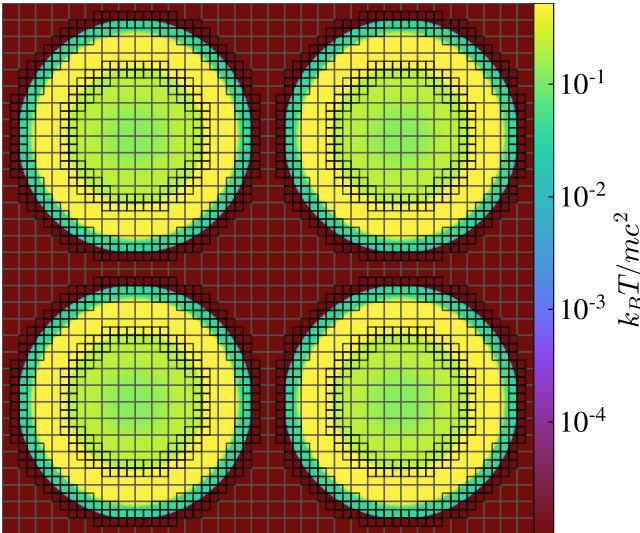
**Figure 6.** Ultra-relativistic flows propagating along different spatial directions with respect to the grids. In all subfigures (a)–(d), the left and right columns are the longitudinal and transverse slices, respectively. Longitudinal slices are taken through the flow source while the transverse slices are taken through the label ‘A’. The flow diameter is resolved by 28 cells in (a) and (b) and by 56 cells in (c) and (d). The flow has an extremely high Mach number ( $\mathcal{M} \sim 10^6$ ) leaving any instability short of time to develop, and one expects a smooth flow-ambient interface. However, the fuzzy-looking cross-sections in the transverse slices of the oblique flow (right columns in (b) and (d)) suggest that the high-speed flow induces false instability when the flow travels obliquely across Cartesian grids. Increasing the spatial and temporal resolution does not help the numerical solution of an oblique flow to converge to that of a parallel flow.

The parallel efficiency is measured to be 90 per cent with 64 nodes and 78 per cent with 2048 nodes, achieving a peak overall performance of  $1.3 \times 10^{11}$  cell updates per second with 2048 nodes. The drop of parallel efficiency is mainly due to MPI communication, the time fraction of which increases from 10.3 ( $N_{\text{node}} = 64$ ) to 18.8 per cent ( $N_{\text{node}} = 2048$ ).

We remark that the strong and weak scaling tests demonstrate 55 and 80 per cent parallel efficiencies, respectively, with 1024 nodes on the **Piz-Daint** supercomputer. Moreover, the peak performance on a single Tesla P100 GPU achieves  $7 \times 10^7$  cell updates per second, which is about one-third of the peak performance of non-relativistic hydrodynamics (Schive et al. 2018).



**Figure 7.** Close-up view of the rectangular region in Figure 6(c) with longitudinal (first column from the left) and transverse (other columns) slices passing through ‘B’ and ‘C’, respectively. Compared with Figure 6(c), it clearly shows that the finger pattern consists of 2-D flat sheets along the flow.

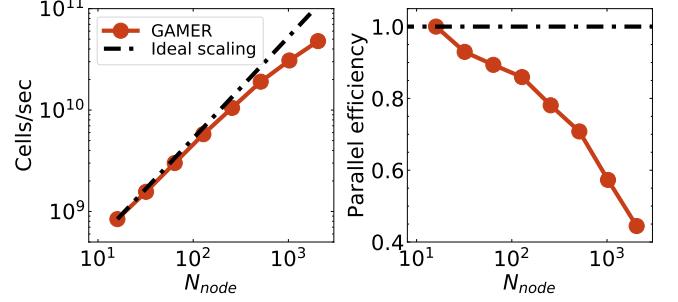


**Figure 8.** Temperature slice through the centre of blast waves at  $t = 0.5L/c$ , with the grid patches overlaid in the case of 8 nodes in the weak scaling test.

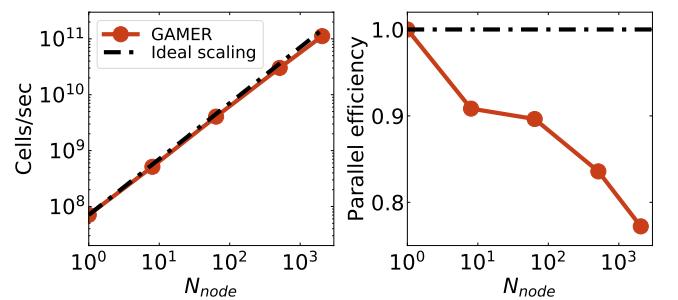
## 6. ASTROPHYSICAL APPLICATIONS

### 6.1. Triaxial blast wave

This triaxial relativistic blast wave problem models a hypothetical astrophysical mega-explosion driven by an ultra-relativistically hot plasma source absent of particular symmetry. It is an atypical test for which we verify the code’s ability to deal with strong 3D shocks. The simulation evolves a blast wave from a triaxial source in a homogeneous medium. The triaxial source has aspect ratios  $1 : 1.5 : 2$  with a semi-major axis  $0.01L$  aligned with the diagonal direction, where  $L$  is the width of a cubic computational box. The source is filled with a uniform ultra-relativistic ( $k_B T_{\text{src}}/mc^2 = 10^6$ ) plasma and the ambient is filled with a uniform



**Figure 9.** Strong scaling of the relativistic jet simulation with five AMR levels and 16–2048 GPU nodes (left panel: cell updates per second; right panel: parallel efficiency). The maximum total number of cells, excluding ghost cells, is  $7.9 \times 10^9$ . The deviation from the ideal scaling is mainly due to MPI communication, the time fraction of which increases by a factor of 10 when increasing the number of nodes from 64 to 2048.



**Figure 10.** Weak scaling of the multi-blast waves simulation with three AMR levels and 1–2048 GPU nodes (left panel: cell updates per second; right panel: parallel efficiency). The maximum total number of cells, excluding ghost cells, is  $6.9 \times 10^7$  per node. The parallel efficiency drops from 0.90 to 0.78 when  $N_{\text{node}}$  increases from 64 to 2048 mainly because the MPI communication time fraction increases from 10.3 to 18.8 per cent.

non-relativistic ( $k_B T_{\text{amb}}/mc^2 = 10^{-9}$ ) HII gas. The density is homogeneous throughout the entire domain with  $\rho_{\text{src}} = \rho_{\text{amb}} = 1.0$ . After the system quickly relaxes, the hot plasma rapidly expands driving a forward shock traveling almost at the speed of light.

The AMR base level is covered by  $32^3$  cells with the periodic boundary condition. The highest refinement level is 9 so that the initial source can be adequately resolved by approximately 82 cells along the minor axis. To refine both the initial source and the thin shell of the blast wave shock, we adopt the gradient of the reduced energy density as the refinement criterion, with  $Q = \tilde{E}$  and  $C_Q = 1.0$  in Equation (33).

For comparison, we also simulate a spherical blast wave to understand how the initially triaxial shape affects the evolution of the ultra-relativistic blast wave. Both the spherical and triaxial cases have the same

simulation set-up and the same source volume, i.e.  $r/L = \sqrt[3]{0.01 \times 0.0075 \times 0.005}$ , where  $r$  is the radius of the spherical source.

Figure 11 shows the results. We observe that the interior hot plasma pushes out a contact discontinuity immediately inward of the shock and that the thickness of the shell between the contact discontinuity and the shock diminishes in time. In early time, the triaxial profiles (green) at  $t = 0.05L/c$  deviate from the spherical counterparts (red), especially in the pressure and proper mass density, although the shock positions almost coincide. However, at a later time, the profiles at  $t = 0.3L/c$  show no significant difference between the triaxial (yellow) and spherical (blue) blast waves, indicating the initial shape of the source does not have a great impact on the asymptotic evolution of ultra-relativistic blast waves.

To further investigate how the triaxial blast wave evolves into a spherical one, we extract the radii  $R_L(t)$  and  $R_S(t)$  of the triaxial blast wave along the semi-major ( $r_L$ ) and semi-minor axes ( $r_S$ ) of the initial source from simulation data. As shown in Figure 12, we find that the dimensionless quantity  $\left(\ln\left(\frac{R_L - R_S}{\sqrt{R_L R_S}}\right) / \left(\frac{r_L - r_S}{\sqrt{r_L r_S}}\right)\right)^2$  is approximately equal to  $0.66\left(\sqrt{R_L R_S} / \sqrt{r_L r_S} - 1\right)$ . This dependence suggests that the triaxiality is damped out with the blast wave propagation by the relation:

$$\frac{R_L(t) - R_S(t)}{\sqrt{R_L(t)R_S(t)}} = \frac{\left(\frac{r_L - r_S}{\sqrt{r_L r_S}}\right) \exp\left[-0.81\left(\frac{\sqrt{R_L(t)R_S(t)}}{\sqrt{r_L r_S}} - 1\right)^{1/2}\right]}{(34)}$$

We remark that this test problem also demonstrates that GAMER-SR can successfully handle ultra-relativistic gases embedded in a cold HII region, which can be difficult for conventional SRHD codes.

### 6.2. Limb-brightened jet

Most active galactic nuclei (AGN) jets in VLBI observations appear ridge-brightened, while limb-brightened jets are rare and have been reported only in a few nearby radio galaxies, such as Mrk 501 (Giroletti et al. 2004), M87 (Asada & Nakamura 2012; Kim et al. 2018), Cygnus A (Boccardi et al. 2015), and 3C84 (Nagai et al. 2014; Giovannini et al. 2018). Motivated by these observations, we simulate a three-dimensional SRHD jet using GAMER-SR to study its acceleration and collimation in the hope to shed light on the limb-brightened jets.

We adopt the gradient of the reduced energy density and the magnitude of  $|\mathbf{M}|/D$  as the two inclusive refinement criteria. A patch is refined if any cell satisfies

either Equation (33) with  $Q = \tilde{E}$  and  $C_Q = 0.1$  or  $|\mathbf{M}|/D > 20$ . The first criterion aims to capture the strong terminal shock and the cocoon, while the second one ensures that the central ‘spine’ region of the jet can be properly resolved.

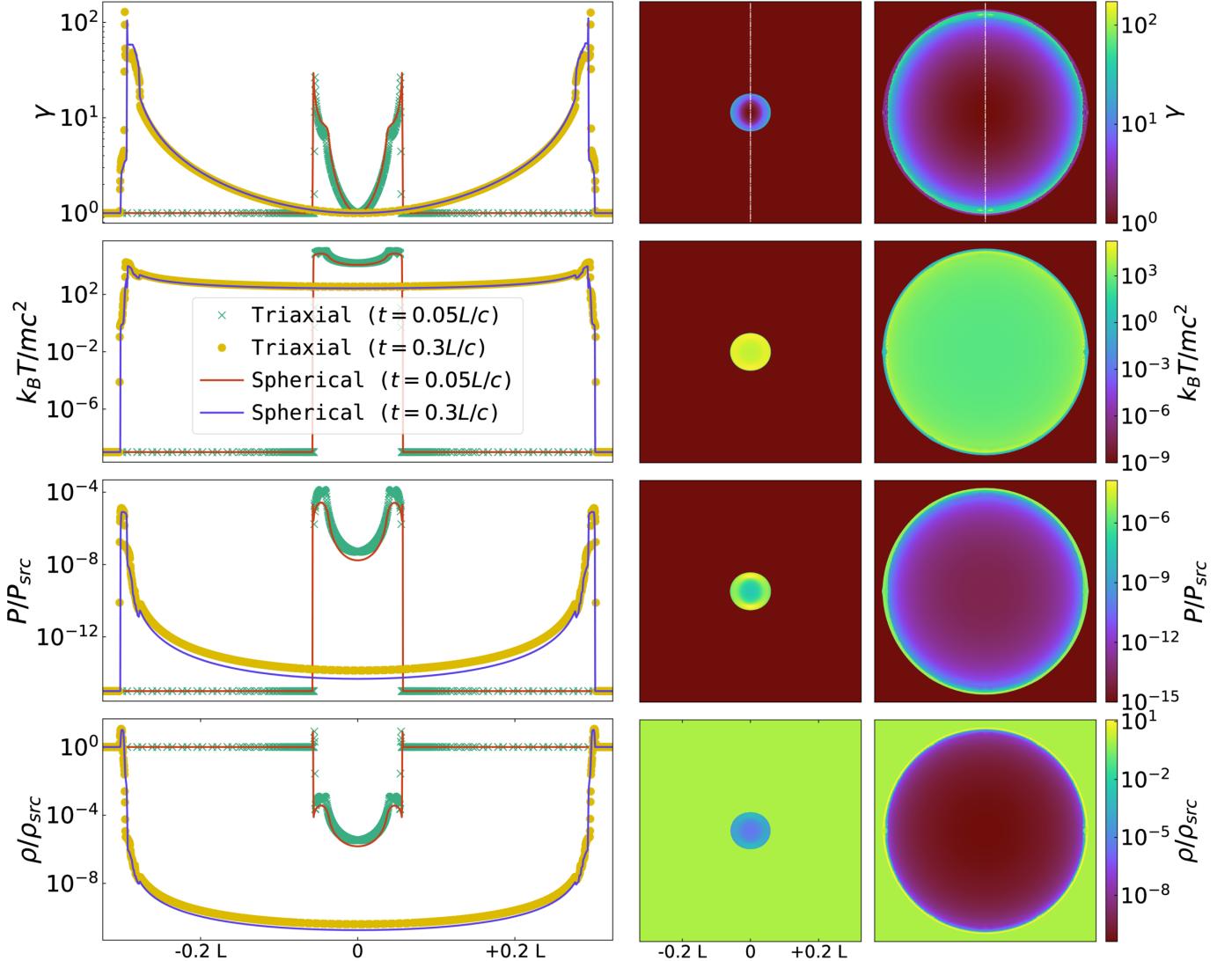
The jet is continuously ejected from a cylindrical source with four-velocity  $\beta\gamma = 10.0$  ( $\gamma \sim 10.05$ ). The proper mass density ratio between the jet source and the ambient gases is set to 1.0. The temperature ( $k_B T/mc^2$ ) of the source and the ambient gases are set to 0.5 and  $10^{-5}$ , respectively. The outflow is thus an extremely under-pressured jet. Both the diameter and the length of the cylindrical source are well resolved by 32 cells.

Figure 13 shows the simulation results. It demonstrates that the jet flow is entirely confined by a turbulent cocoon at all time. Two points are worth noting from these longitudinal slices. First, the Lorentz factor (first row) rises from 10 to 26, and meanwhile the temperature (second row) drops from 0.5 to 0.01 along the jet. Second, the relativistic Bernoulli number minus  $c^2$  (fifth row), defined as  $hy - c^2$ , remains nearly constant within the spine region. According to the de Laval nozzle effect, these suggest that thermal energy is converted to kinetic energy by the expansion of a supersonic flow. Surprisingly, the gases are still accelerated in the region between the label ‘C’ and the confinement point close to ‘D’. These images seem to suggest acceleration during flow convergence, which in fact does not contradict the de Laval nozzle effect. The gases still expand away from the jet axis after passing ‘C’, which can be confirmed by examining the transverse slice of the radial component of the Mach number (the last row),

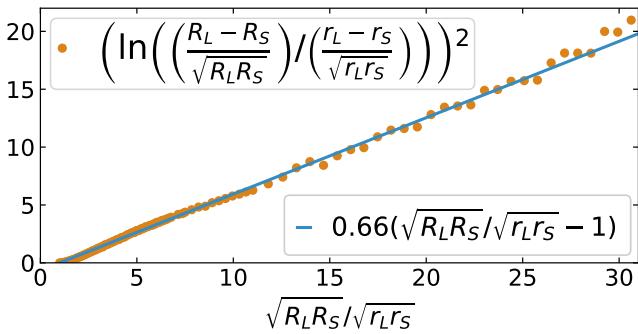
$$\mathcal{M}_{\text{radial}} = (\hat{\mathbf{r}} \cdot \mathbf{U}) / (c_s \gamma_s), \quad (35)$$

where  $\hat{\mathbf{r}}$  is the cylindrical unit radial vector,  $\mathbf{U}$  is the four-velocity of flow, and  $\gamma_s = c_s / \sqrt{1 - c_s^2}$ . The definition of the radial Mach number given by Equation (35) is Lorentz invariant when the transforming direction is along the jet. Obviously, gases expand not only between the jet source and ‘C’, but also inside the entire central spine region. Thus, the flow convergence in between ‘C’ and ‘D’ is a false impression.

Associated with this expanding jet flow is the limb-brightened phenomenon. Confined by the cocoon, the radial flow of the cooler jet imparts onto the cocoon with a boundary shock, as signified by the edge  $\mathcal{M}_{\text{radial}} \gg 1$ . Hot gases in the post-shock region then diffuse into the cocoon transverse to the jet through some instabilities composed of high-density and low-temperature fingers. This finger pattern is similar to that reported in Section 4.3.



**Figure 11.** Triaxial blast wave test. The middle ( $t = 0.05L/c$ ) and right ( $t = 0.3L/c$ ) columns show the slices passing through the triaxial source at the mid-plane of its intermediate axis (i.e. the horizontal and vertical axes are along the major and minor axes, respectively). The left column shows the profiles along the minor-axis (i.e. the white dotted-dashed line in the  $\gamma$  map).

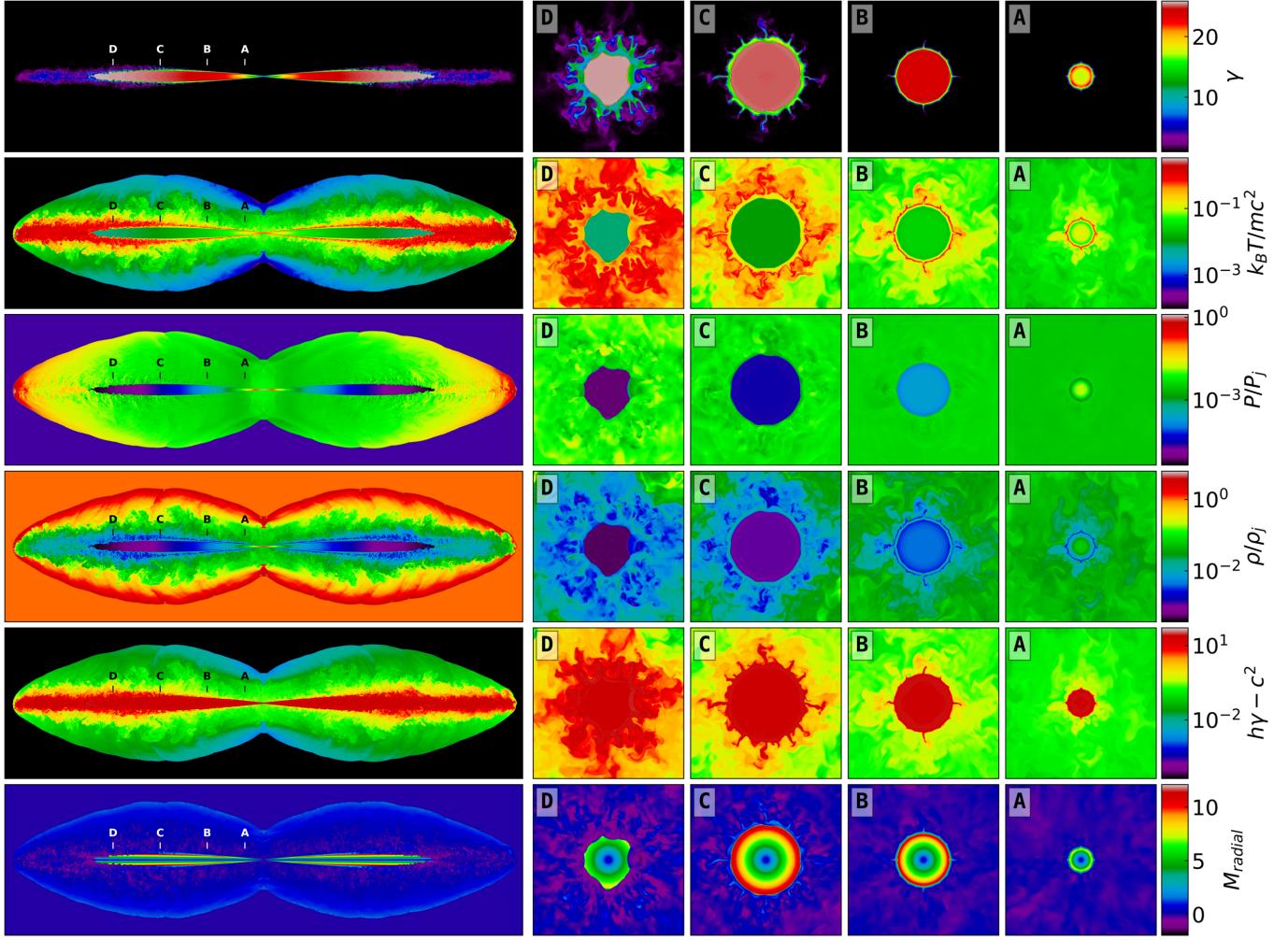


**Figure 12.** Damping of the triaxiality in the triaxial blast wave test, where  $R_L$  and  $R_S$  are the radii of the triaxial blast wave along the semi-major ( $r_L$ ) and semi-minor axes ( $r_S$ ) of the initial source.

Certainly the boundary shock can generate particle acceleration and produce extra synchrotron brightness at the jet edge, thus yielding limb brightening. Since the boundary shock is relatively weak, the extra synchrotron brightness cannot be immense. This may explain why limb brightening is mostly observed in nearby AGN jets.

## 7. CONCLUSIONS

In this paper, we have presented a novel special relativistic hydrodynamics code, GAMER-SR, which incorporates a new, well-tailored conversion scheme (cf. Figure 15) between primitive and conserved variables, together with the Taub-Mathews equation of state (TM EoS) covering both the ultra-relativistic and non-relativistic limits. The new scheme adopts the four-velocity  $U^j$ , the reduced energy density  $\tilde{E}$ , and the



**Figure 13.** From top to bottom: Lorentz factor, temperature, pressure (normalized by the jet source pressure  $p_j$ ), proper mass density (normalized by the jet source density  $\rho_j$ ), relativistic Bernoulli number minus  $c^2$  (i.e.  $hy - c^2$ ), and radial component of Mach number defined by Equation (35) at  $t = 0.73L/c$ . Left column: longitudinal slices passing through the jet source. Right four columns: transverse slices passing through the locations labelled by A, B, C, and D.

reduced enthalpy  $\tilde{h}$  to effectively avoid the catastrophic cancellation in subsonic flows at all finite temperature, including the particularly challenging low-temperature regime, with errors decreasing as  $\mathcal{M}^2 \epsilon_{\text{machine}}$  when  $\mathcal{M} \gg 1$ .

We have numerically derived the exact solution of a Riemann problem covering both extreme cold and ultra-relativistically hot gases with the TM EoS. Simulation results using our new scheme are in very good agreement with the exact solution in both the ultra-relativistic and non-relativistic regimes. (cf. Figure 5). In comparison, the catastrophic cancellation arising from the original (unoptimized) scheme can be much more severe than the truncation error in the non-relativistic limit, especially in the region swept by a traveling contact discontinuity.

The new scheme has been integrated into the code GAMER to facilitate the hybrid OpenMP/MPI/GPU

parallelization and adaptive mesh refinement. Thanks to that, the performance of the root-finding iterations in the TM EoS can be significantly improved by GPU. The parallel efficiency using 2048 computing nodes is measured to be  $\sim 45$  per cent for strong scaling (cf. Figure 9) and  $\sim 75$  per cent for weak scaling (cf. Figure 10) on the **Piz-Daint** supercomputer.

GAMER-SR has been demonstrated to be able to handle ultra-relativistic flow with a Lorentz factor as high as  $10^6$ . However, we also find that the Cartesian grids can lead to artificial dissipation when the direction of a high Mach number flow is not aligned with grids. This problem cannot be mitigated by increasing spatial and temporal resolution.

We have examined two astrophysical problems with coexisting relativistically hot and cold gases to demonstrate the power of GAMER-SR. The first problem

deals with a relativistic blast wave with a triaxial source. Not only do we find that the code is able to capture the ultra-relativistic strong shock very well, but we also discover a simple rule governing how the triaxiality of the blast wave diminishes as a function of the blast wave radius.

The second problem addresses the flow acceleration and limb-brightening of a relativistic AGN jet. We find that the jet, from its head to its source, is always enclosed inside a turbulent cocoon. The jet is accelerated all the way up to the first confinement point, where an internal shock appears. We attribute such flow acceleration to the relativistic Bernoulli's law. In addition, the synchrotron limb-brightening is found to be caused by the jet transverse boundary shock, outside which the post-shock cosmic-ray particles are mixed with the turbulent cocoon and give out extra synchrotron emission.

GAMER-SR does not implement explicit grid derefinement criteria, dual-energy formalism, and magnetic fields (e.g., Stone et al. 2020). The magnetic fields play a critical role in many high energy astrophysical problems, such as the evolution

of accretion discs (e.g., Blandford & Payne 1982), and the interaction between cosmic rays and thermal gases in the Fermi bubbles (e.g., Yang & Ruszkowski 2017). We leave them to future work.

## ACKNOWLEDGEMENT

P.T. thanks Kuo-Chuan Pan for helping conduct the parallel scaling tests at CSCS (*Piz-Daint*) under Grant No. 661. H.S. acknowledges funding support from the Jade Mountain Young Scholar Award No. NTU-109V0201, sponsored by the Ministry of Education, Taiwan. This research is partially supported by the Ministry of Science and Technology (MOST) of Taiwan under Grants No. MOST 107-2119-M-002-036-MY3 and No. MOST 108-2112-M-002-023-MY3, and the NTU Core Consortium project under Grants No. NTU-CC-108L893401 and No. NTU-CC-108L893402.

## DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

## REFERENCES

- Asada K., Nakamura M., 2012, *The Astrophysical Journal*, 745, L28
- Blandford R. D., Payne D. G., 1982, *Monthly Notices of the Royal Astronomical Society*, 199, 883
- Blandford R., Meier D., Readhead A., 2019, *Annual Review of Astronomy and Astrophysics*, 57, 467
- Boccardi B., Krichbaum T. P., Bach U., Mertens F., Ros E., Alef W., Zensus J. A., 2015, *A and A*, 585, A33
- Chiueh T., 1989, *Phys. Rev. Lett.*, 63, 113
- Chiueh T., Li Z.-Y., Begelman M. C., 1991, *ApJ*, 377, 462
- Chiueh T., Li Z.-Y., Begelman M. C., 1998, *The Astrophysical Journal*, 505, 835
- Falle S. A. E. G., 1991, *Monthly Notices of the Royal Astronomical Society*, 250, 581
- Fong W., et al., 2019, *ApJL*, 883, L1
- Fryxell B., et al., 2000, *The Astrophysical Journal Supplement Series*, 131, 273
- Ghirlanda G., et al., 2019, *Science*, 363, 968
- Giovannini G., et al., 2018, *Nature Astronomy*, 2, 472
- Giroletti M., et al., 2004, *The Astrophysical Journal*, 600, 127
- Gourgouliatos K. N., Komissarov S. S., 2017, *Nature Astronomy*, 2, 167
- Higham N. J., 2002, *Accuracy and Stability of Numerical Algorithms*, 2nd edn. Society for Industrial and Applied Mathematics, USA
- Jüttner F., 1911, *Annalen der Physik*, 339, 856
- Kennel C. F., Coroniti F. V., 1984a, *ApJ*, 283, 694
- Kennel C. F., Coroniti F. V., 1984b, *ApJ*, 283, 710
- Kim J.-Y., et al., 2018, *A&A*, 616
- Li Z.-Y., Chiueh T., Begelman M. C., 1992, *ApJ*, 394, 459
- Lora-Clavijo F. D., Cruz-Osorio A., Guzmán F. S., 2015, *The Astrophysical Journal Supplement Series*, 218, 24
- Martí J. M., Müller E., 1994, *Journal of Fluid Mechanics*, 258, 317
- Mathews W. G., 1971, *ApJ*, 165, 147
- Mignone A., Bodo G., 2005, *Monthly Notices of the Royal Astronomical Society*, 364, 126
- Mignone A., Bodo G., 2006, *Monthly Notices of the Royal Astronomical Society*, 368, 1040
- Mignone A., McKinney J. C., 2007, *Monthly Notices of the Royal Astronomical Society*, 378, 1118
- Mignone A., Plewa T., Bodo G., 2005, *The Astrophysical Journal Supplement Series*, 160
- Mooley K. P., et al., 2018a, *Nature*, 554, 207
- Mooley K. P., et al., 2018b, *Nature*, 561, 355
- Nagai H., et al., 2014, *The Astrophysical Journal*, 785, 53

- Noble S. C., Gammie C. F., McKinney J. C., Zanna L. D., 2006, *The Astrophysical Journal*, 641, 626
- Núñez-de la Rosa J., Munz C.-D., 2016, *MNRAS*, 460, 535
- Rezzolla L., Zanotti O., 2018, Relativistic hydrodynamics. Oxford University Press
- Rezzolla L., Zanotti O., Pons J. A., 2001, *Journal of Fluid Mechanics*, 449, 395
- Ryu D., Chattopadhyay I., Choi E., 2006, *The Astrophysical Journal Supplement Series*, 166, 410
- Schive H.-Y., Tsai Y.-C., Chiueh T., 2010, *The Astrophysical Journal Supplement Series*, 186, 457
- Schive H.-Y., ZuHone J. A., Goldbaum N. J., Turk M. J., Gaspari M., Cheng C.-Y., 2018, *Monthly Notices of the Royal Astronomical Society*, 481, 4815
- Sod G. A., 1978, *Journal of Computational Physics*, 27, 1
- Stone J. M., Tomida K., White C. J., Felker K. G., 2020, *The Astrophysical Journal Supplement Series*, 249, 4
- Synge J. L., 1957, North-Holland Pub. Co.; Interscience Publishers
- Taub A. H., 1948, *Physical Review*, 74, 328
- Toro E. F., 2011, Riemann solvers and numerical methods for fluid dynamics : a practical introduction. Springer, Berlin
- Turk M. J., Smith B. D., Oishi J. S., Skory S., Skillman S. W., Abel T., Norman M. L., 2011, *The Astrophysical Journal Supplement Series*, 192, 9
- Woodward P., Colella P., 1984, *Journal of Computational Physics*, 54, 115
- Woosley S. E., 1993, *ApJ*, 405, 273
- Yang H.-Y. K., Ruszkowski M., 2017, *The Astrophysical Journal*, 850, 2
- van Leer B., 1979, *Journal of Computational Physics*, 32, 101
- van Leer B., 2006, pp 192–206

## APPENDIX

Note that the speed of light, the particle mass, and the Boltzmann constant are set to unity in Appendix for simplicity.

### A. INITIAL GUESS FOR NEWTON-RAPHSON ITERATION

We use the Newton-Raphson iteration to find the root  $\tilde{h}$  of Equation (15). The iteration requires the derivative of Equation (15) with respect to  $\tilde{h}$ :

$$\frac{df}{d\tilde{h}} = 2\tilde{h} + 2 - 2T - 2(\tilde{h} + 1)\frac{dT}{d\tilde{h}} + \frac{2T(\tilde{h} + 1)^4 \frac{dT}{d\tilde{h}} + 2T(\tilde{h} + 1)\left(\frac{|\mathbf{M}|}{D}\right)^2 \left[(\tilde{h} + 1)\frac{dT}{d\tilde{h}} + T\right]}{\left[(\tilde{h} + 1)^2 + \left(\frac{|\mathbf{M}|}{D}\right)^2\right]^2}, \quad (\text{A1})$$

where

$$\frac{dT}{d\tilde{h}} = \frac{4\tilde{h} + 4}{5\tilde{h} + 5 + \sqrt{9\tilde{h}^2 + 18\tilde{h} + 25}} - \frac{(2\tilde{h}^2 + 4\tilde{h})\left(\frac{9\tilde{h} + 9}{\sqrt{9\tilde{h}^2 + 18\tilde{h} + 25}} + 5\right)}{\left(5\tilde{h} + 5 + \sqrt{9\tilde{h}^2 + 18\tilde{h} + 25}\right)^2}, \quad (\text{A2})$$

follows from Equation (16).

The root-finding iteration also requires an initial guess of  $\tilde{h}_{\text{guess}}$ , for which we suggest the following procedure. In the low- $T$  limit, we Taylor expand Equation (15) in powers of  $\tilde{h}$  and keep the first- and second-order terms:

$$\left(\frac{\tilde{E}}{D}\right)^2 + 2\left(\frac{\tilde{E}}{D}\right) - \left(\frac{|\mathbf{M}|}{D}\right)^2 = \frac{6}{5}\tilde{h} + \left\{ \frac{43}{125} + \frac{4}{25\left[1 + \left(\frac{|\mathbf{M}|}{D}\right)^2\right]}\right\}\tilde{h}^2. \quad (\text{A3})$$

Solving Equation (A3) for the unknown  $\tilde{h}$  gives the positive solution:

$$\tilde{h}_{\text{guess}} = \frac{\sqrt{125\left[1 + \left(\frac{|\mathbf{M}|}{D}\right)^2\right]\left[43\left(\frac{|\mathbf{M}|}{D}\right)^2 + 63\right]\left[\left(\frac{\tilde{E}}{D}\right)^2 + 2\left(\frac{\tilde{E}}{D}\right) - \left(\frac{|\mathbf{M}|}{D}\right)^2\right]}}{43\left(\frac{|\mathbf{M}|}{D}\right)^2 + 63} \left\{ 75\left[1 + \left(\frac{|\mathbf{M}|}{D}\right)^2\right] + \sqrt{125\left[1 + \left(\frac{|\mathbf{M}|}{D}\right)^2\right]\left[43\left(\frac{|\mathbf{M}|}{D}\right)^2 + 63\right]\left[\left(\frac{\tilde{E}}{D}\right)^2 + 2\left(\frac{\tilde{E}}{D}\right) - \left(\frac{|\mathbf{M}|}{D}\right)^2\right] + 75^2\left[1 + \left(\frac{|\mathbf{M}|}{D}\right)^2\right]^2} \right\}. \quad (\text{A4})$$

In the opposite high- $T$  limit, Equation (15) can be reduced to

$$\left(\frac{\tilde{E}}{D}\right)^2 + 2\left(\frac{\tilde{E}}{D}\right) - \left(\frac{|\mathbf{M}|}{D}\right)^2 = \frac{9}{16}\tilde{h}^2, \quad (\text{A5})$$

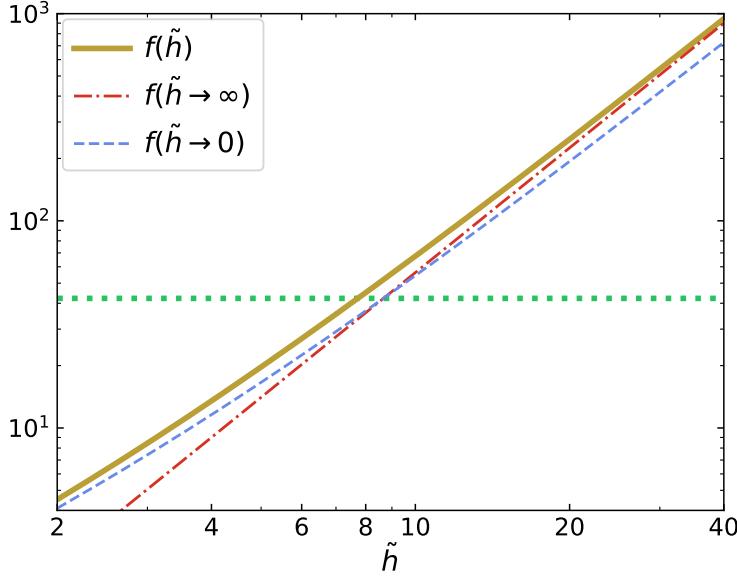
which leads to

$$\tilde{h}_{\text{guess}} = \frac{4}{3}\sqrt{\left(\frac{\tilde{E}}{D}\right)^2 + 2\left(\frac{\tilde{E}}{D}\right) - \left(\frac{|\mathbf{M}|}{D}\right)^2}. \quad (\text{A6})$$

Equation (A4) and Equation (A6) provide two initial guesses for ‘cold’ and ‘hot’ gases, respectively. The threshold to distinguish between ‘cold’ gases and ‘hot’ gases is given by

$$\left\{ \frac{1800\left[1 + \left(\frac{|\mathbf{M}|}{D}\right)^2\right]}{437\left(\frac{|\mathbf{M}|}{D}\right)^2 + 117} \right\}^2, \quad (\text{A7})$$

which is obtained by equating Equations (A4) and (A6) (see Figure 14). If  $(\tilde{E}/D)^2 + 2(\tilde{E}/D) - (|\mathbf{M}|/D)^2$  is greater than Equation (A7), we choose Equation (A6) as an initial guess for the Newton-Raphson iteration (hot gases); otherwise, we choose Equation (A4) (cold gases).



**Figure 14.**  $f(\tilde{h}; |\mathbf{M}|/D = 1)$  (—) and its asymptotes when  $\tilde{h} \rightarrow \infty$  (----) and  $\tilde{h} \rightarrow 0$  (----). The horizontal line (----) is given by Equation A7, which passes through the intersection of two asymptotes and provides a threshold to distinguish between ‘cold’ and ‘hot’ gases for the initial guess of  $\tilde{h}$  in the Newton-Raphson iteration.

## B. NUMERICAL ERROR ANALYSIS FOR ROOT-FINDING

Figure 15 provides a detailed flowchart of the conversion between primitive and conserved variables. Figure 16 demonstrates that the numerical errors of root-finding arising from the new and original conversion schemes are consistent with the predicted values given by Equation (20) and Equation (21). We measure this conversion error by first converting the input primitive variables  $(\rho_0, U_0^i, p_0)$  into conserved variables  $(D_1, M_1^i, \tilde{E}_1)$ . Next, we convert  $(D_1, M_1^i, \tilde{E}_1)$  back to  $(\rho_2, U_2^i, p_2)$  and then measure the relative error between  $\tilde{h}(p_0/\rho_0) := \tilde{h}_0$  and  $\tilde{h}(p_2/\rho_2) := \tilde{h}_2$ . Since the catastrophic cancellation is more prominent in the low- $T$  limit, we measure the error  $|1 - \tilde{h}_0/\tilde{h}_2|$  as a function of Mach number from  $10^{-4}$  to  $10^6$  with a fixed non-relativistic temperature  $k_B T/mc^2 = 10^{-8}$  (i.e. the blue dashed-dotted line in Figure 2). To verify the accuracy in three-dimensional space, we choose the direction of four-velocity to be parallel to the line  $x = y = z$  (i.e.  $U_0^i = |\mathbf{U}_0|/\sqrt{3}$  for all  $i$ ). Double precision is adopted to handle the large dynamic range. Figure 16 confirms that the numerical errors of the new and original schemes are mainly caused by round-off errors in the calculation of  $(\tilde{E}/D)^2 + 2(\tilde{E}/D) - (|\mathbf{M}|/D)^2$  and  $(E/D)^2 - (|\mathbf{M}|/D)^2 - 1$ , respectively.

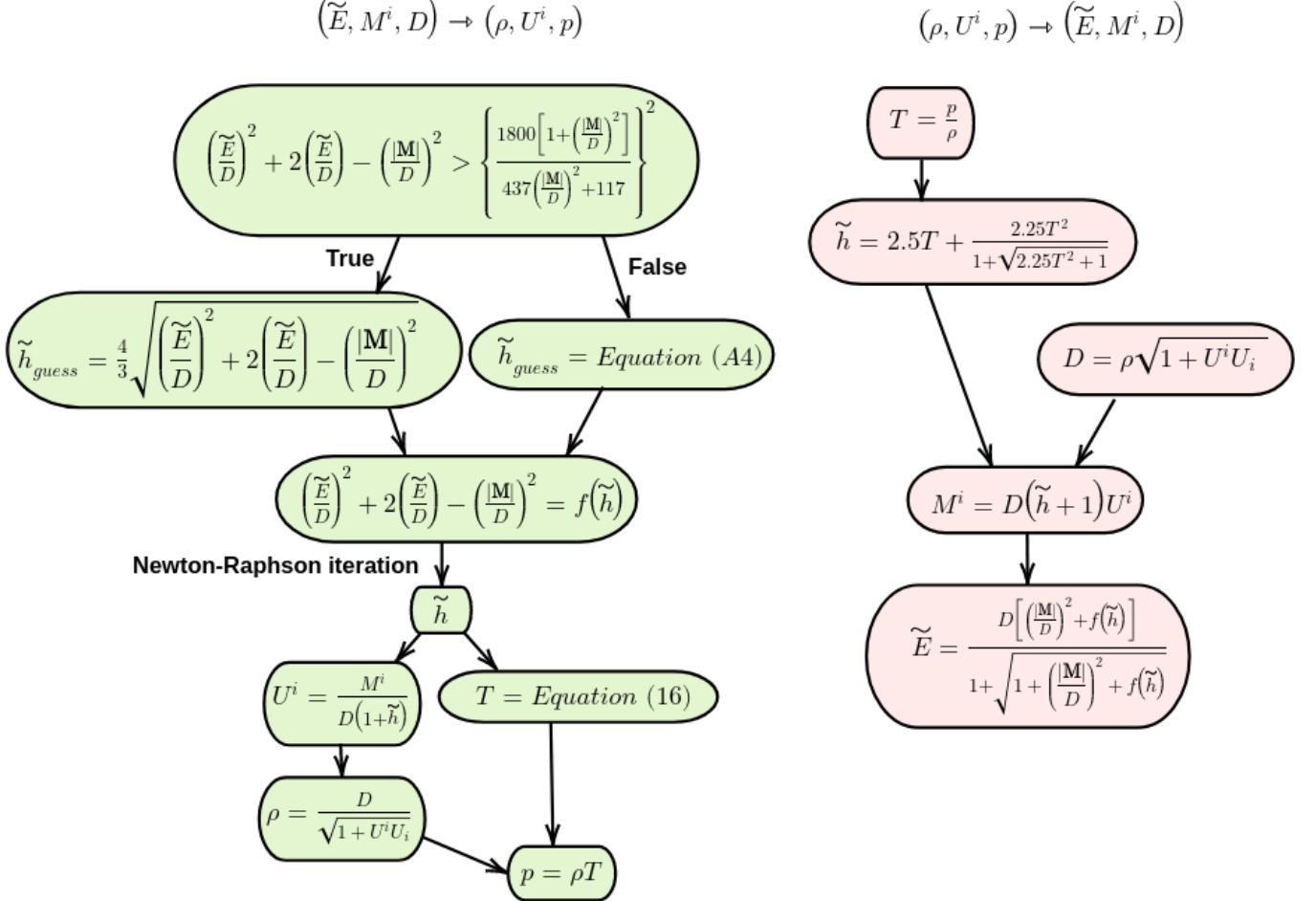
## C. EXACT SOLUTIONS OF RELATIVISTIC RIEMANN PROBLEMS WITH THE TM EQUATION OF STATE

To derive the exact solutions of relativistic Riemann problems with the TM EoS, we have generalized the previous framework of a constant polytropic EoS (Martí & Müller 1994; Rezzolla et al. 2001) to the TM EoS. More precisely, this approach can be applied to any EoS once we know the relationship between enthalpy and temperature. Here we only summarize the important equations and highlight salient differences from the polytropic EoS. We use the subscripts  $L/C_L/C_R/R$  to refer to the left/left-contact/right-contact/right regions and define the relative four-velocity of  $U_1$  with respect to  $U_2$  as  $\text{Relative}(U_1, U_2)$ . Note that we have replaced three-velocity with four-velocity again to avoid catastrophic cancellation in the ultra-relativistic limit.

The exact solution of a relativistic Riemann problem with the TM EoS can be obtained through the following three steps:

1. For a given initial condition, we can determine the wave pattern by comparing the relative velocity between the two unperturbed initial states with the three limiting values. These values mark the transition from one wave pattern to another and can be directly computed from the initial condition. See Rezzolla et al. (2001) for details.
2. We determine the unknown pressure  $p^*$  between the left and right waves by numerically solving

$$U_{LR} = \text{Relative}(U_{LC_L}(p^*), U_{RC_R}(p^*)), \quad (\text{C8})$$



**Figure 15.** Flowchart of converting conserved variables to primitive variables (left) and the opposite (right).

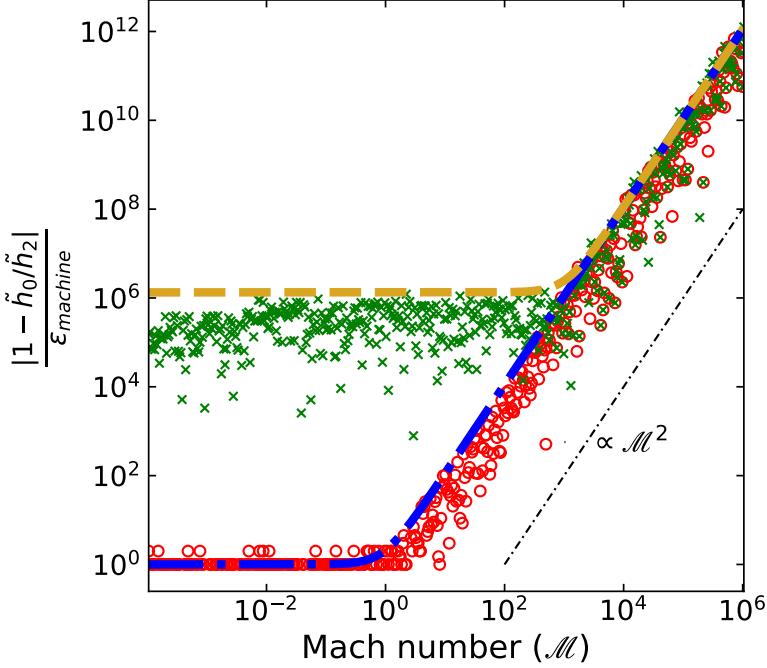
where  $U_{LR} = \text{Relative}(U_L, U_R)$ ,  $U_{LC_L}(p^*) = \text{Relative}(U_L, U_{C_L}(p^*))$ , and  $U_{RC_R}(p^*) = \text{Relative}(U_R, U_{C_R}(p^*))$ . Note that the four-velocity in the left-/right-contact region,  $U_{C_{L/R}}$ , is different for each of the three possible wave patterns. For example, if the left wave is rarefaction and the right wave is shock, then  $U_{C_L} = U_{\mathcal{R}}(p^*)$  and  $U_{C_R} = U_{\mathcal{S}}(p^*)$  where  $U_{\mathcal{R}}(p^*)$  and  $U_{\mathcal{S}}(p^*)$  are defined as follows.

- (a)  $U_{\mathcal{R}}(p^*)$  represents the relation between pressure and flow four-velocity behind the *rarefaction* wave:  
Given the pressure behind the rarefaction wave (i.e.  $p^*$ ) during the Newton-Raphson iteration for solving Equation (C8), we can determine  $U_{\mathcal{R}}(p^*)$  by numerically solving the system of equations (10), (14), (18), (C9a), and (C9b):

$$\frac{dU}{d\rho} = \pm \frac{c_s \gamma}{\rho}, \quad (C9a)$$

$$\frac{p}{\rho^{5/3}}(h - T) = \text{constant}. \quad (C9b)$$

Hereafter, the upper/lower sign applies to the right/left wave. The ordinary differential equation (C9a), known as the Riemann invariant (Rezzolla & Zanotti 2018), relates the dynamical ( $U$ ) and thermal ( $c_s$ ) quantities. Equation (C9b), derived from Equation (10) and the second law of thermodynamics, results from the fact that entropy is constant through the rarefaction wave. The ‘constant’ in Equation (C9b) is a function of entropy and can be determined by the thermal quantities in the region unperturbed by the rarefaction wave. In the case of the constant polytropic EoS, Equation (C9b) reduces to a familiar form:  $p/\rho^\Gamma = \text{const.}$



**Figure 16.** Numerical errors of the conversion between primitive and conserved variables. It shows the relative error  $|1 - \tilde{h}_0/\tilde{h}_2|$  as a function of Mach number with a given non-relativistic temperature ( $k_B T/mc^2 = 10^{-8}$ ) for the new scheme (○) and the original scheme (✖). These errors are mainly caused by the cancellation in  $(\tilde{E}/D)^2 + 2(\tilde{E}/D) - (|\mathbf{M}|/D)^2$  and  $(E/D)^2 - (|\mathbf{M}|/D)^2 - 1$ , which are inevitably introduced during the root-finding iteration and consistent with the predicted values given by Equation (20) (—) and Equation (21) (----).

(b)  $U_{\mathcal{S}}(p^*)$  represents the relation between pressure and flow four-velocity behind the *shock* wave:

Let ‘up/down’ denote the upstream/downstream state of the shock wave. Under the condition that  $p_{\text{down}}$  ( $p_{\text{down}} = p^*$  in this case) is given during the Newton-Raphson iteration for solving Equation (C8), we can compute  $h_{\text{down}}$  by numerically solving the jump condition of the enthalpy:

$$h_{\text{up}}^2 - h_{\text{down}}^2 = \left( \frac{h_{\text{down}}}{\rho_{\text{down}}} + \frac{h_{\text{up}}}{\rho_{\text{up}}} \right) (p_{\text{up}} - p_{\text{down}}). \quad (\text{C10})$$

Equation (C10) is known as the Taub adiabat (Taub 1948), where  $\rho_{\text{down}}$  can be eliminated using Equations (16) and (18). After determining  $h_{\text{down}}$  by a root-finding routine, the mass flux across the shock can be calculated by

$$J = \left( \frac{p_{\text{down}} - p_{\text{up}}}{h_{\text{up}}/\rho_{\text{up}} - h_{\text{down}}/\rho_{\text{down}}} \right)^{0.5}. \quad (\text{C11})$$

The four-velocities of shock and post-shock then follow from

$$U_{\text{shock}} = \pm \left( \frac{J}{\rho_{\text{up}}} \right) \sqrt{1 + U_{\text{up}}^2} \pm U_{\text{up}} \sqrt{1 + \left( \frac{J}{\rho_{\text{up}}} \right)^2}, \quad (\text{C12})$$

and

$$U_{\mathcal{S}}(p^*) = \mp \left( \frac{J}{\rho_{\text{down}}} \right) \sqrt{1 + U_{\text{shock}}^2} + U_{\text{shock}} \sqrt{1 + \left( \frac{J}{\rho_{\text{down}}} \right)^2}, \quad (\text{C13})$$

respectively. Equations (C12) and (C13) are essentially the Lorentz boost that takes four-velocity from the shock rest frame to the lab frame. Note that the mass flux  $J$  is an invariant under the Lorentz boost in the flow direction.

**Table 2.** Exact solution of a relativistic Riemann problem with the TM EoS at  $t = 80.0$ . Columns from left to right give  $x$ -coordinate, proper mass density, four-velocity, and pressure. The initial condition is given in the last row of Table 1, with the initial discontinuity at  $x = 5 \times 10^{-2}$ . The blue solid line in Figure 5 plots the solution. The exact solution is calculated with double precision and shown in 16 digits to reach machine accuracy. The data are available in the supplement.

$x$	$\rho$	$U_x$	$p$
0.000000000000000e+00	1.000000000000000e+02	+1.000000000000000e-03	1.000000000000000e-04
2.5743971630613077e-02	1.000000000000000e+02	+1.000000000000000e-03	1.000000000000000e-04
2.6720534130613077e-02	9.999999999999986e+01	+1.000000000000002e-03	1.000000000000000e-04
2.8673659130613080e-02	9.8588362795909134e+01	+1.0183105709873300e-03	9.7658360819209613e-05
3.1603346630613080e-02	9.6495914226915929e+01	+1.0457764274335814e-03	9.4228343648087098e-05
3.6486159130613087e-02	9.3074657510006034e+01	+1.0915528549894106e-03	8.8726314406083176e-05
4.2345534130613087e-02	8.9077088828567909e+01	+1.1464845688462534e-03	8.2466343818876140e-05
5.2111159130613087e-02	8.2671707042031258e+01	+1.2380374291888151e-03	7.2821829638494934e-05
6.4806471630613094e-02	7.4813960019366874e+01	+1.3570561601099081e-03	6.1655409508574801e-05
8.4337721630613108e-02	6.3723430244968533e+01	+1.5401619444173906e-03	4.7188055213551995e-05
1.1168147163061304e-01	5.0121316453652021e+01	+1.7965101817141935e-03	3.1625521037347636e-05
1.5172053413061287e-01	3.3922515604881056e+01	+2.1718776864619303e-03	1.6499866085321606e-05
2.0836115913061276e-01	1.7591444669719621e+01	+2.7028867092515813e-03	5.5229310921865207e-06
2.0933772163061276e-01	1.7369735883307754e+01	+2.7120420544404751e-03	5.4074080094554571e-06
2.0972078270771960e-01	1.7283280852025452e+01	+2.7156332803617649e-03	5.3626249948767070e-06
2.6627329885804002e-01	1.7283280852025452e+01	+2.7156332803617649e-03	5.3626249948767070e-06
2.6724986145813656e-01	4.0108528993879889e-10	+2.7156332816129858e-03	5.3626249948767070e-06
2.6909288248391281e+01	4.0108528993879889e-10	+2.7156332816129858e-03	5.3626249948767070e-06
2.6910264810891281e+01	9.999999999999998e-13	-1.000000000000000e+02	1.000000000000000e-10
1.000000000000000e+02	9.999999999999998e-13	-1.000000000000000e+02	1.000000000000000e-10

3. Once  $p^*$  is known,  $\rho_{\text{down}}$  follows from Equations (16) and (18), which in turn allows for computing  $U_{\text{down}}$  and  $U_{\text{shock}}$  through Equation (C12) and Equation (C13). On the other hand,  $\rho$  behind the rarefaction wave follows from solving the system of equations (10), (18), and (C9b). Finally, given the self-similar and isentropic character of the rarefaction wave,  $\rho$  and  $U$  within the rarefaction fan can be computed by solving the system of equations (C9a) and  $U(\xi) = \text{Relative}(\xi/\sqrt{1-\xi^2}, \pm c_s/\sqrt{1-c_s^2})$ , where  $\xi = x/t$ .

On the basis of above procedure, we show in Table 2 the exact solution of the relativistic Riemann problem given in the last row of Table 1 with the TM EoS at  $t = 80.0$ . The source code is available at (<https://github.com/zengbs/ExactSolutionRelativisticRiemannProblem>).