Riemann solvers with non-ideal thermodynamics: exact, approximate, and machine learning solutions

by

Jeremy C. H. Wang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Mechanical & Mechatronics Engineering

Waterloo, Ontario, Canada, 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Karthik Duraisamy
Professor
Dept. of Aerospace Engineering
University of Michigan

Supervisor: Jean-Pierre Hickey
Assistant Professor
Dept. of Mechanical & Mechatronics Engineering
University of Waterloo

Internal Members: Zhao Pan
Assistant Professor
Dept. of Mechanical & Mechatronics Engineering
University of Waterloo

Xiao-Yu Wu
Assistant Professor
Dept. of Mechanical & Mechatronics Engineering
University of Waterloo

Internal-External Member: Sander Rhebergen
Associate Professor
Dept. of Applied Mathematics
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The following co-authored contributions are associated with this thesis:

1. Chapter 4 is published as Wang, J. C.-H. and Hickey, J.-P. (2020): Analytical solutions to shock and expansion waves for non-ideal equations of state, *Phys. Fluids* **32**, 086105, https://doi.org/10.1063/5.0015531, reproduced in accordance with the copyright policy of AIP Publishing

2. Chapter 5 was presented as Wang, J. C.-H. and Hickey, J.-P. (2020): Structurally complete approximate Riemann solvers for flows with non-ideal thermodynamics, *Annual Meeting of the American Physical Society Division of Fluid Dynamics*

3. Chapter 5 is published as Wang, J. C.-H. and Hickey, J.-P. (2022): A class of structurally complete approximate Riemann solvers for trans- and supercritical flows with large gradients, *J. Comput. Phys.* **468**, 111521, https://doi.org/10.1016/j.jcp.2022.111521, reproduced in accordance with the copyright policy of Elsevier

4. Chapter 6 is under review as Wang, J. C.-H. and Hickey, J.-P. (2022): FluxNet: a physics-informed learning-based Riemann solver for transcritical flows with non-ideal thermodynamics, *J. Comput. Phys.*

All sections of this thesis and the co-authored contributions above were written by Jeremy Wang, and reviewed and edited by Jean-Pierre Hickey. In all cases, Jeremy Wang performed the conceptualization, mathematical derivations, coding, generation of results, and analysis; Jean-Pierre Hickey supported the conceptualization and supervised the research in general.

Abstract

The Riemann problem is an important topic in the numerical simulation of compressible flows, aiding the design and verification of numerical codes. A limitation of many of the existing studies is the perfect gas assumption. Over the past century, flow technology has tended toward higher pressures and temperatures such that non-ideal state equations are required along with specific heats, enthalpy, and speed of sound dependent on the full thermodynamic state. The complexity of the resulting physics has compelled researchers to compromise on rigour in favour of computational efficiency when studying non-ideal shock and expansion waves. This thesis proposes exact, approximate, and machine learning approaches that balance accuracy and computational efficiency to varying degrees when solving the Riemann problem with non-ideal thermodynamics.

A longstanding challenge in the study of trans- and supercritical flows is that numerical simulations are often validated against prior numerical simulations or inappropriate ideal-gas shock tube test cases. The lack of suitable experimental data or adequate reference solutions means that existing studies face difficulties distinguishing numerical inaccuracies from the physics of the problem itself. To address these shortcomings, a novel derivation of exact solutions to shock and expansion waves with arbitrary equation of state is performed. The derivation leverages a domain mapping from space-time coordinates to characteristic wave coordinates. The solutions may be integrated into a suitable Riemann solution algorithm to produce exact reference solutions that do not require numerical integration.

The study of wave structures is also pertinent to the development of practical Riemann solvers for finite volume codes, which must be computationally simple yet entropy-stable. Using the earlier derivations, the idea of structurally complete approximate Riemann solvers (StARS) is proposed. StARS provides an efficient means for analytically restoring the isentropic expansion wave to pre-existing three-wave solvers with arbitrary thermodynamics. The StARS modification is applied to a Roe scheme and shown to have improved accuracy but comparable computational speed to the popular Harten-Hyman entropy fix. Four test cases are examined: a transcritical shock tube, a shock tube with periodic bounds that produce interfering waves, a two-dimensional Riemann problem, and a *gradient* Riemann problem—a variant on the traditional Riemann problem featuring an initial gradient of varying slope rather than an initial step function. Additionally, a scaling analysis shows that entropy violations are most prevalent and yield the greatest errors in trans- and supercritical flows with large gradients.

The final area of inquiry focuses on *FluxNets*, that is, learning-based Riemann solvers whose accuracy and efficiency fall in between those of exact and approximate solvers. Various approaches to the design and training of fully connected neural networks are assessed.

By comparing data-driven versus physics-informed loss functions, as well as neural networks of varying size, the results show that order-of-magnitude reductions in error compared to the Roe solver can be achieved with relatively compact architectures. Numerical validation on a transcritical shock tube test case and two-dimensional Riemann problem further reveal that a physics-informed approach is critical to ensuring smoothness, generalizability, and physical consistency of the resulting numerical solutions. Additionally, parallelization can be leveraged to accelerate inference such that the significant gains in accuracy are achieved at one quarter the runtime of exact solvers. The trade-off in accuracy versus efficiency may be justified in the case of non-ideal flows where even minor errors can result in spurious oscillations and destabilized solutions.

Keywords: computational fluid dynamics, numerical methods, transcritical thermodynamics, entropy, Riemann solver, machine learning, high-speed flow

Acknowledgements

Thank you Professor Hickey for your mentorship, your moral support during those bleak moments of research, your wisdom in navigating academic writing and publishing, and most of all, thank you for your trust, flexibility, and unwavering advocacy of my capabilities. You were the only researcher with whom I considered performing graduate research, and I am incredibly excited for the future graduate students who will find themselves privileged to work with you. I look forward to all our conversations in the years to come.

Thank you to the more than 90 teachers and professors, the countless colleagues, and those few close friends, who helped make me who I am over the years—from kindergarten through to university and beyond. Thank you especially to Chloé and Ralf for taking me on at DLR, and to UTAT for facilitating that wonderful summer experience. Were it not for that stint at DLR, Prof. Hickey and I may never have met in quite the way we did.

Thank you to the faculty, students, and administrators at the University of Waterloo. A particular debt of gratitude is owed to MPILAB, the broader department, Concept, Velocity, the Centre for Teaching Excellence, and Mitacs. You continue to foster what can only be described as the premier environment for graduate engineering research and entrepreneurship in Canada. Your tenacity, generosity, and support for student ambition are unparalleled.

Mom and Dad, we've had our share of tough times—thank you for your love and dedication. Betty, thank you for being there when I was growing up. Finally, Jen, thank you for your love, perseverance, and support through it all including proofreading this behemoth of a document. No journey or destination is worth it without good company.

Dedicated to the life of my love, Jen

Table of Contents

List of Figures

List of Tables

Abbreviations

CAL constraint-adapted loss 15, 50, 117

CFD computational fluid dynamics 8, 13, 20, 23, 26, 139, 145

CFL Courant-Friedrichs-Lewy 36, 40, 101, 111, 112, 164

EOS equation of state xix, 3, 5–7, 17, 20, 25–28, 54–59, 63, 64, 66, 67, 101, 107, 111, 113, 114, 116, 139, 141, 144

ML machine learning 3, 13–15, 18, 20, 41, 42, 44, 107, 110, 115, 139, 143

MSE mean-squared error 47, 48, 117, 118

PR Peng-Robinson 6, 27, 87, 91, 97, 101, 110, 116, 137

RK Redlich-Kwong 6, 27

RP Riemann problem 1, 3, 4, 6–10, 15, 20, 23, 30, 33, 35, 40, 42, 44, 45, 48, 50, 52, 54, 72, 86, 96, 97, 100–103, 107, 112–116, 128, 139, 140, 142

SRK Soave-Redlich-Kwong 6, 27

StARS structurally complete approximate Riemann solver 71, 141, 142, 144

List of Symbols

A_i	coefficient in the analytical solution to centred expansion waves
$\mathbf{A}$	Jacobian of the flux vector $\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}}$
a	parameter in the cubic equation of state $[\mathrm{Pa}\,(\mathrm{m}^3\,\mathrm{kmol}^{-1})^2\,\mathrm{K}^{0.5}]$
B_i	the coefficient of T^i in the polynomial fit of $c'(T)$
b	parameter in the cubic equation of state $[\mathrm{m}^3\,\mathrm{kmol}^{-1}]$
$C_{-,0,+}$	{left, entropy, right} characteristic curve
c	non-ideal speed of sound $[\mathrm{m\,s}^{-1}]$
c'	ideal speed of sound $[\mathrm{m\,s}^{-1}]$
c_v	non-ideal constant-volume specific heat capacity $[\mathrm{J\,kmol}^{-1}\,\mathrm{K}^{-1}]$
c'_v	ideal constant-volume specific heat capacity $[\mathrm{J\,kmol}^{-1}\,\mathrm{K}^{-1}]$
c_p	non-ideal constant-pressure specific heat capacity $[\mathrm{J\,kmol}^{-1}\,\mathrm{K}^{-1}]$
c'_p	ideal constant-pressure specific heat capacity $[\mathrm{J\,kmol}^{-1}\,\mathrm{K}^{-1}]$
E	total energy per unit volume $[\mathrm{J\,m}^{-3}]$
e	specific internal energy $[\mathrm{J\,kmol}^{-1}]$
$\mathbf{F}$	flux vector in the x direction

$\mathcal{F}$	neural network
$\mathbf{G}$	flux vector in the y direction
H	total enthalpy per unit mass $[\mathrm{J\,kg^{-1}}]$
h	non-ideal specific enthalpy $[\mathrm{J\,kmol^{-1}}]$
h'	ideal specific enthalpy $[\mathrm{J\,kmol^{-1}}]$
$J_{-,+}$	$\{$left, right$\}$ Riemann invariant $[\mathrm{m\,s^{-1}}]$
$\mathbf{K}$	(right) eigenvector of $\mathbf{A}$
$\mathcal{L}$	loss function
M	molar mass $[\mathrm{kg\,kmol^{-1}}]$
Ma	Mach number [unitless]
p	pressure [Pa]
p_c	critical pressure [Pa]
p_r	reduced pressure [unitless]
R	universal gas constant $= 8314.4621[\mathrm{J\,kmol^{-1}\,K^{-1}}]$
S	wavespeed $[\mathrm{m\,s^{-1}}]$
v	specific volume $[\mathrm{m^3\,kmol^{-1}}]$
T	temperature [K]
T_c	critical temperature [K]
T_r	reduced temperature [unitless]
t	time [s]
$\mathbf{U}$	vector of conserved variables
u	velocity in the x direction $[\mathrm{m\,s^{-1}}]$
w	velocity in the y direction $[\mathrm{m\,s^{-1}}]$

x	position [m], or input variable in machine learning
y	position [m], or output variable in machine learning

α	wave strength of $\mathbf{A}$ [$\mathrm{kg\,m^{-3}}$]
γ	ratio of specific heats [unitless]
δ	parameter in the cubic equation of state [$\mathrm{m^{3}\,kmol^{-1}}$]
ϵ	parameter in the cubic equation of state [$\mathrm{m^{6}\,kmol^{-2}}$]
ρ	density [$\mathrm{kg\,m^{-3}}$]
Θ	expression in the cubic equation of state [$\mathrm{Pa\,m^{6}\,kmol^{-2}}$]
θ	weights in a neural network
λ	eigenvalue of $\mathbf{A}$ [$\mathrm{m\,s^{-1}}$]
τ	number of temporally distinct floating point operations
ϕ	activation function in a neural network
ψ	regularization term added to a loss function
ω	acentric factor [unitless]

Chapter 1

Introduction

1.1 Literature Review

The Riemann problem (RP) is named after German mathematician Bernhard Riemann who first studied the initial value problem for a hyperbolic set of partial differential equations in 1860 [6]. Two constant flow states are initially separated by an imaginary membrane at $x = 0$. At time $t = 0$, the imaginary membrane disappears and the solution at any $t > 0$ must be found by evolving the one-dimensional time-dependent Euler equations for the given initial conditions. Four types of self-similar solutions to the RP are possible, composed of shock waves, expansion waves, and a contact discontinuity wave (Fig. 1.1).

There are multiple theoretical and practical uses of the RP. Exact solutions reveal the mathematical behaviour of generic systems of hyperbolic conservation equations. Intercell fluxes in the finite volume method may be computed using approximations to the RP [7]—

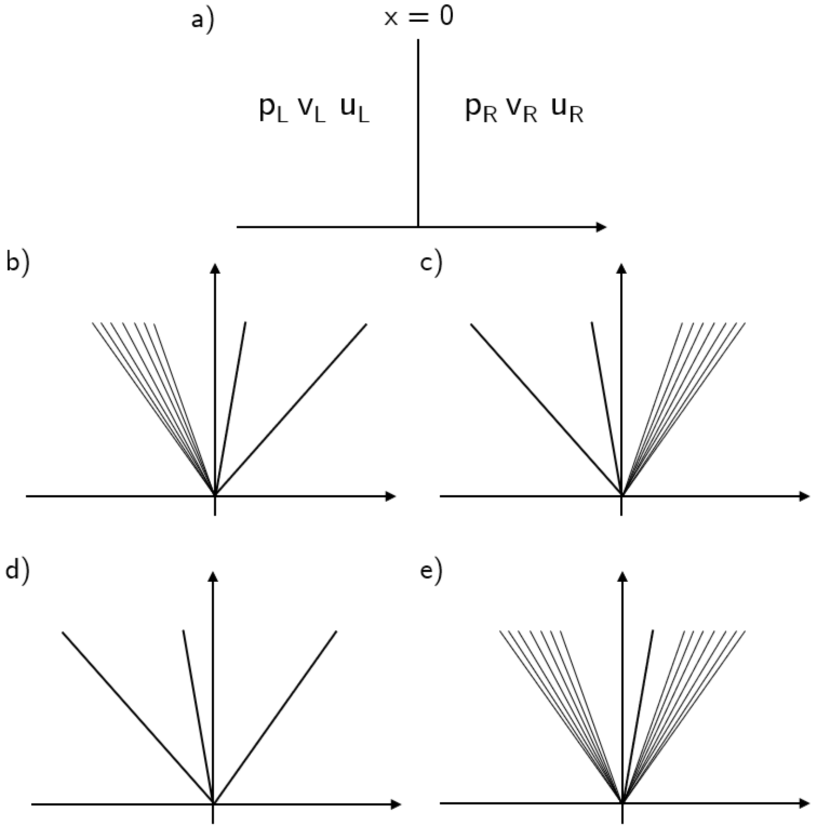


Figure 1.1: Overview of the Riemann problem in gas dynamics and the four possible configurations of self-similar solutions: a) initial conditions; b) rarefaction-contact-shock; c) shock-contact rarefaction; d) shock-contact-shock; e) rarefaction-contact-rarefaction. In practice, the wavespeeds may vary such that solutions are skewed more toward the left or right.

an idea that was originally proposed by Godunov [8]. Exact solutions may also be used to evaluate the accuracy of shock-capturing numerical methods, *e.g.* Sod shock tube test [9], Einfeldt's strong rarefaction test [10]. This is particularly valuable given the lack of experimental data to validate numerical simulations; few shock tube facilities operate at sufficiently high pressures and temperatures to mimic the conditions of modern engineering devices where non-ideal thermodynamics are present [11]. The RP also appears in the study of magnetohydrodynamics [12], astrophysical flows [13], and numerous practical problems such as SF_6 circuit breakers [14], dust explosions [15, 16, 17], and experimental shock ignition facilities [18, 19]. Studies of the RP thus come in two broad flavours: those which are concerned with accuracy and not computational speed (for the purposes of physical understanding and verifying numerical codes), versus those concerned with speed and adequate accuracy (for the purposes of computing fluxes to solve applied problems).

However, a longstanding shortcoming of existing studies on the RP is the tendency to assume that the gas is ideal and thermally or calorically perfect [6, 20, 8, 9, 21, 22, 23, 24]. For a number of modern applications, including jet and rocket engines, hypersonic aircraft, and supercritical diesel engines, it is necessary to use non-ideal EOS along with heat capacities, enthalpies, and speeds of sound that depend on the full thermodynamic state as defined by two independent state variables [25]. In this chapter, we explore the challenges with accuracy and practical computation when solving the RP especially with non-ideal thermodynamics. This discussion is divided into exact solutions, approximate solvers, and a new generation of machine learning (ML) approaches that strive to balance accuracy with fast runtime.

1.1.1 Exact Solutions

Finding the exact solution to an RP begins with deriving shock and expansion waves as two independent phenomena. Afterwards, an iterative scheme can be developed to solve the implicit system of equations that arises when these waves manifest simultaneously in the RP. Such solutions are said to be *exact* when it is possible to iterate to an arbitrary level of precision. Exact solutions offer physical insight as well as serve as a reference to test the accuracy of numerical schemes. It is worth distinguishing that while there exist closed-form solutions to ideal shocks and expansion waves as standalone problems, there are no exact closed-form solutions to the RP—not even for ideal gases—due to the highly nonlinear and coupled nature of the equations [24]. Many numerical studies of non-ideal flows do not even utilize exact solutions to verify the numerical method, instead they verify accuracy against prior numerical solutions [26, 27]. Nonetheless, the evaluation of a numerical code against either exact-iterative or numerical reference solutions to the RP is frequently termed a *shock tube test*.

For an ideal and calorically perfect gas, Riemann [6] is credited with deriving explicit analytical solutions to the centred expansion wave. Rankine [28] and Hugoniot [20] made major contributions to solving the stationary normal shock problem. Riemann applied the method of characteristics to the time-dependent Euler equations, while Rankine and Hugoniot performed direct algebraic manipulation of the steady Euler equations. In both cases, the solutions relied on the ideal and calorically perfect assumption, which in turn implies a polytropic gas. Polytropic gases possess immensely simplified thermodynamic behaviour because the product of pressure and volume raised to a constant exponent remains invariant with respect to temperature variations. This exponent depends on the type of thermodynamic process (*e.g.* for an isentropic flow, the exponent is the gas specific heat ratio), but it is guaranteed to be constant. Indeed, the ideal, calorically perfect, polytropic gas model lies at the foundation of much of classical compressible flow theory, including shocks, rarefactions, Prandtl-Meyer expansion fans, converging-diverging nozzles, Rayleigh and Fanno flows, and more [29].

Unfortunately, both the ideal gas law and the calorically perfect gas model are unrealistic under conditions such as transcritical or supercritical flow. Near the critical point, heat capacities exhibit highly nonlinear variations that induce nonlinear behaviour in enthalpy, specific heat ratio, and speed of sound [30]. Therefore, these thermodynamic quantities cannot be assumed constant, let alone functions of only one variable such as temperature. It is also necessary to use non-ideal EOS such as cubic or virial-type EOS to accurately model pressure-volume-temperature relationships [25]. The complicated form of the resulting equations renders impossible the direct integration of the Riemann invariants in any expansion waves, as well as prevents an explicit solution to shock ratios [31].

Shock tube tests for non-ideal thermodynamics have thus historically comprised highly resolved numerical solutions. These tests generally fall into one of two categories: high-order schemes [26, 27, 5, 32, 33, 11, 34, 35] or implicit solutions that take advantage of iterative solvers but use numerical integration in regions containing expansion waves [36, 37, 38, 39, 40, 41, 42, 43]. A handful of studies have also discussed the qualitative wave behaviour of non-ideal gases [44, 45, 46], advancing conceptual understanding though lacking quantitative answers.

Arina [26] was the first to perform numerical simulations on a RP under near-critical conditions. He extended the AUSM flux-splitting scheme to single-phase real gases where the van der Waals, Carnahan-Starling-De Santis, or Redlich-Kwong (RK) EOS could be used. Although novel, Arina's study was limited in that the numerical method was verified using the Sod shock tube test [9]—which involves an ideal gas. Arina's numerical simulations of the supercritical piston effect offered insight into relative agreement between real gas state equations, but were not verified against an exact solution in the trans- or supercritical regime where non-ideal effects are present.

Terashima *et al.* [27] subsequently developed a sixth-order differencing scheme with third-order total-variation-diminishing Runge-Kutta time integration, specifically intended for supercritical flows. They utilized the Soave-Redlich-Kwong (SRK) EOS, as well as a localized artificial diffusivity method to eliminate spurious oscillations near sharp discontinuities such as shocks. Their method successfully captured various qualitative features of a supercritical planar jet, such as flow instabilities, flapping motion, and jet entrainment, but the method was verified using Arina's [26] numerical results for a supercritical shock tube which were in turn verified on the Sod test as discussed above.

Recently, Ma *et al.* [5] devised an entropy-stable double-flux scheme with the Peng-Robinson (PR) EOS, verified against a transcritical nitrogen shock tube using the results from a high-resolution essentially non-oscillatory scheme as the reference solution. The method was then applied to simulate a variety of flow problems, including planar jets. Pantano *et al.* [32] applied a diffusive interface method with the van der Waals EOS to simulate supercritical flow, again verified using a supercritical RP whose reference solution was produced using a highly resolved quadrature method. Other contemporary works

[33, 34, 47] have followed a similar vein of simulation verified against simulation.

Finally, a number of iterative schemes have been devised that operate on the implicit equations of the non-ideal RP and numerically integrate the Riemann invariant when expansion waves are detected. These schemes are popular because the wave structure is modelled explicitly, thus yielding sharp interfaces between the waves. Closure of the implicit scheme is achieved through the contact discontintuity in the star-state region, across which pressure and velocity remain constant. Collela & Glaz [36] famously specified general algorithms to iteratively solve RP with arbitrary EOS by guessing, iterating, and converging on star-state velocity and pressure—this was a generalization of Godunov's original observation that was specific to ideal gases [8]. Saurel *et al.* [31] implemented the iterative scheme of Collela & Glaz [36] to produce reference solutions for testing different numerical schemes and nondimensionalized approximate Riemann solvers. Banks [39] also used a Collela & Glaz [36] iterative scheme for solving generic flow problems with complex state equations, while Kam [40] adopted a similar strategy for exactly solving detonation problems. In general, one can apply an iterative solver such as Newton or secant iteration to find these star-state conditions, and then calculate the remaining wavespeeds and flow properties in the RP. When rarefactions are present, the Riemann invariant is typically numerically integrated over density or another independent variable. The choice of density step size may require further iterations to ensure that star-state conditions are continuous with the rarefaction within acceptable tolerance [39, 40]. A major caveat with this class of iterative schemes is that stability is not guaranteed.

Ultimately, a noticeable issue is the absence of reference solutions that are free of spatial discretization and numerical integration errors. The current literature reveals a tenuous pattern of verifying numerical methods against other numerical methods or irrelevant ideal-gas test cases. Numerical inaccuracies in the reference solutions, such as artificial viscosity and truncation errors, can become indistinguishable from the physics of the problem itself. An exact non-ideal solution is lacking which would formally ground this area of inquiry. Such a solution might, for example, exploit a change of variable or specialized integration path to derive exact-iterative or partially closed-form solutions to the RP.

1.1.2 Approximate Riemann Solvers

Riemann solvers are numerical tools that can be used to compute the flux in the flux difference splitting approach. It is also possible to use flux vector splitting, which does not rely on Riemann solvers, but this is typically practised for steady flows such as in aerodynamics. By contrast, difference splitting is popular for unsteady and generic problems [24]. Many approximate Riemann solvers have been developed for practical use in computational fluid dynamics (CFD) [24], and are typically non-iterative to prioritize computational speed. Such solvers also tend to assume a perfect gas to relate the thermodynamic variables. From these flux estimates it is possible to construct efficient numerical schemes for simulating flow problems of interest to scientists and engineers [7]. With the increasing interest in systems operating at thermodynamic conditions that depart from the ideal assumptions—often characterized by highly non-linear thermodynamic coupling and computationally expensive evaluation of fluxes [48]—the accuracy and efficiency of Riemann solvers become increasingly relevant.

The general idea behind the construction of an approximate Riemann solver is that the propagation of information between cells can be modelled as an RP. If we know which region of the RP occupies the cell interface location at $x = 0$, then we can approximate the flux at the interface. This is usually achieved through a direct estimate of the flux in the numerical method [21, 49, 50, 23]. Alternatively, one can estimate the state in each region of the RP and then compute the flux [51] however this approach involves assuming a fixed, often false, wave structure of the exact solution [24].

Among the first approximate Riemann solvers are those of Osher [50], Roe [49], and Harten, Lax, & Van Leer [21]. Osher [50] approximated the flux by applying eigendecomposition then defining a Jacobian matrix splitting much like the flux vector splitting approach. By picking suitable integration paths, it is possible to integrate the Riemann invariant analytically. Roe [49] approximated the Riemann solution by assuming a constant Jacobian matrix of the flux vector with respect to conservative variables, requiring the constant Jacobian to satisfy certain properties, then solving the equations exactly. The HLL solver [21] approximated the Riemann solution as three constant states divided by the fastest left- and right-moving waves. The Rankine-Hugoniot equations relate the conditions across each wave. Using an appropriate wavespeed estimate, such as from Davis [52] or Einfeldt [22], it becomes practical to determine which state occupies the cell interface and to find the corresponding flux. Though suited to hyperbolic systems of two equations (*e.g.* the shallow water equations), these solvers often experience difficulties in resolving sharp features such as material interfaces, shear waves, contact surfaces, and strong shocks in the time-dependent Euler equations [24].

Indeed, for nearly four decades, there has been a continual pursuit to improve the representation of the wave structure within Riemann solvers. This is because improvements to the design of the Riemann solver can result in increased accuracy at minimal additional expense as compared to leveraging more involved discretization schemes. In some cases, it is impossible to resolve certain flow phenomena if the Riemann solver omits the necessary waves [53]. For instance, the HLLC solver [23] restored the missing contact discontinuity in the HLL flux [21], leading to improved resolution of material interfaces and sharp physical features in contexts such as supersonic and shallow water flows. Also, the HLLE [22] and HLLEM [10] solvers addressed issues with wavespeeds to ensure positively conservative results, particularly under vacuum conditions. To date, the highest fidelity approximate solvers consist of *three-wave* models [51, 54, 23, 55, 56], named as such because they account for shocks, contact discontinuities, and the heads of expansion waves.

A major limitation of most approximate solvers is that rarefactions in the solution to the RP are simplified to discontinuous jumps. The spatially varying nature of rarefactions is lost despite their presence in exact solutions as well as in the underlying physical problem. This occurs due to the consideration of piecewise constant states between the wave fronts, resulting from the linearization of the governing equations. At a fully subsonic or supersonic state (Fig. 1.2.a), the rarefaction does not enclose the cell interface and the omission of an exact expansion wave is benign; the intercell flux is determined by other regions in the solution. However, if a rarefaction is present in a transonic scenario (Fig. 1.2.b), and it is approximated as a discontinuous jump, one can prove that the resulting weak solution violates the entropy condition [57]. In the context of this thesis, we denote a *transonic* scenario when the head of the rarefaction is subsonic while the tail is supersonic. This often occurs in numerical simulations where the flow itself is transonic.

The issue of entropy violations is especially relevant to the study of trans- and supercritical flows, in which transonic conditions may arise more easily due to the immense energy of the fluids and the rapid variations in speed of sound around the critical point. Harten & Hyman [1], Osher [57], and Quirk [58] were among the first to explore entropy issues and fluxes due to transonic rarefactions in Riemann solvers. They showed that entropy violations are consistent with the mathematical definition of hyperbolic conservation laws but are thermodynamically inconsistent for the purposes of simulating real-world flows. Entropy-violating solutions frequently contain nonphysical phenomena such as expansion waves that suddenly decay into a shock front, also called *expansion shocks* or *rarefaction shocks*. Qu *et al.* [59] provide a contemporary review of Riemann solvers for high-speed flow problems, noting various issues and fixes therein.

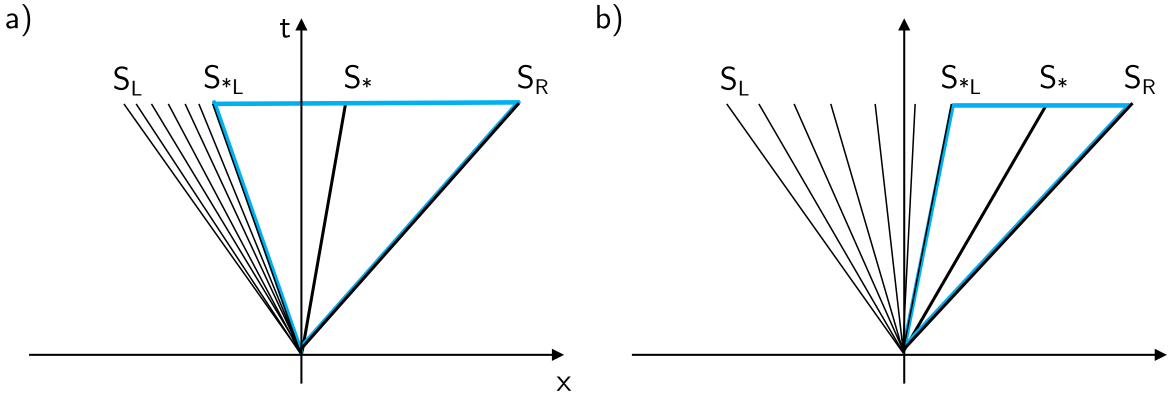


Figure 1.2: The general wave structure of the rarefaction-contact-shock solution (left pressure > pressure in the blue region > right pressure) as commonly drawn in textbooks and research papers (left), versus when the head and tail of the rarefaction sit on opposite sides of $x = 0$ (right). The blue lines enclose the so-called *star* or *star-state* region. S_L, S_{*L}, S_*, S_R are the speeds of the left expansion head, left expansion tail, contact discontinuity, and right normal shock, respectively.

Various entropy fixes have been developed over the years [1, 60], and they are generally modelled after Harten & Hyman's [1] approach of introducing a new intermediate state to approximate the lost rarefaction wave. The new intermediate state is often treated as a constant; alternatively, it can be linearly or polynomially interpolated between known states. This has the effect of introducing additional diffusivity in the flux terms to mitigate any expansion shocks. It has also been shown, in the case of perfect gases, that it is possible to calculate the flux analytically [61, 7, 24]. Even so, a simple and analytically correct means of restoring the expansion wave for arbitrary Riemann solvers—and especially under non-ideal thermodynamics—has not yet been demonstrated. Moreover, no studies to-date have investigated which flow conditions tend to cause transonic rarefactions. Such knowledge would be valuable for determining when entropy fixes ought to be used versus can be omitted for greater speed of computation.

Recent contributions in the area of entropy stability have instead primarily focused on extending fundamental entropy concepts to new applications. Few fundamental improvements have been made to the design of the Riemann solver itself. For example, studies have investigated entropy violations in boundary conditions [62], higher dimensions [63, 64], multicomponent flows [65], low-Mach number flows [66], or hybridized Riemann solvers that switch or average between different flux estimates [67, 68]. The ideal gas assumption is usually made, and any entropy fixes follow the classical implementation or with minor optimizations. Other works have also examined entropy stability in the context of discontinuous Galerkin schemes [69, 70, 71], magnetohydrodynamics [72], Lagrangian gas dynamics [73], relativistic hydrodynamics [74], and nonclassical dense gases where rarefaction shocks are physically admissible [75, 76, 77, 78, 79]. The case of a single-species gas with arbitrary state equation obeying the Euler equations has thus far been overlooked. Studying this particular problem would facilitate the analysis of nonphysicalities attributable only to the

Riemann solver.

1.1.3 Machine Learning for the Riemann Problem

Historically, the research and practice of CFD have been driven by domain expertise and heuristics—but with advances in high-performance computing, ML has emerged as a promising tool to address unresolved challenges [80].

One of the earliest and longstanding applications of ML to fluid mechanics has been the discovery and development of physical models. Kolmogorov, one of the early pioneers of both probability and turbulence theory, proposed turbulence closure as one of the key applications for statistical learning [80, 81]. Sirovich [82] developed the snapshot proper orthogonal decomposition to model the dynamics of coherent structures, a technique that would ultimately become the foundation for modern computer vision [83]. Jambunathan *et al.* [84] trained neural networks to predict convective heat transfer coefficients. Milano et al. [85] applied neural networks to reconstruct near-wall fields in turbulent flows. More recently, Ma *et al.* [86] leveraged ML to develop reduced order models for multiphase flow. Lusch *et al.* [87] investigated deep learning approaches to approximate nonlinear dynamics using linear embeddings. San *et al.* [88] and Pawar *et al.* [89] explored artificial neural networks for reduced order modelling in fluid dynamics. Milan *et al.* [48] used ML to accelerate thermodynamic calculations in real-fluid flows. Duraisamy et al. [90] offers a comprehensive review of applications of ML in turbulence modelling, with a particular focus on uncertainty quantification and predictive capability.

ML has also been successfully applied to flow control and optimization. Faller & Schreck

[91] reviewed a range of applications for ML to solve various problems in aeronautics with a particular emphasis on fault diagnostics and adaptive control systems. Lee *et al.* [92] and Mohan *et al.* [93] explored neural networks for turbulent flow control. Benard *et al.* utilized genetic algorithms for experimental mixing optimization [94]. Pierret & Van Den Braembussche [95] trained networks on databases of Navier-Stokes solutions to optimize turbine blade designs. A significant body of flow control research takes advantage of reinforcement learning spanning hydrological systems [96], laminar bluff-body flow [97], fish motion [98], gliders [99], and the kinematics of unmanned aerial systems [100].

Despite the explosive interest in ML, Brunton *et al.* [80] note that certain nuances unique to fluid mechanics have yet to be addressed. For example, few studies have examined how learning algorithms should contend with the vast physical scales, sensitivity to noise, presence of latent variables, sharp flow features, or transient states. For instance, Dissanayake & Phan-Thien [101], Gonzalez-Garcia *et al.* [102], Lagaris *et al.* [103] chose slow-evolving and smooth test problems to evaluate the feasibility of neural networks in solving ordinary and partial differential equations. Recent works are just beginning to push the boundary of neural networks for high-speed flows. Raissi *et al.* [104] proposed physics-informed neural networks for supervised learning tasks while abiding by physical laws specified as nonlinear partial differential equations. Mao *et al.* [105] examined physics-informed approaches for studying high-speed flow problems. Bezgin *et al.* [106] trained a convolutional neural network to estimate weights in a weighted essentially non-oscillatory scheme for nonclassical undercompressive shock problems. Whereas typical applications of ML are concerned mainly with predictive performance and generalizability, the use of ML in physics must continue to uphold the principle of interpretability. Together, these considerations have motivated the use of ML not to replace numerical schemes entirely, but rather, to address specific well-defined issues within fluid mechanics.

The RP may be one area where ML can enhance accuracy at minimal increase in computational expense and without compromising the explainability of the broader numerics. Current approximate solvers are designed to render a non-iterative means to estimate the star-state conditions, albeit at the expense of accuracy. The star-state conditions together with the initial conditions are then used to solve the remaining states of the RP, and in turn, the intercell flux. ML could provide a means to estimate star-state conditions at greater accuracy than traditional approximate solvers, but with less time complexity than exact solvers. Moreover, physics enforces that every set of initial conditions corresponds to only one set of star-state conditions.

Magiera *et al.* [107] were the first to develop a neural network that predicts conditions in the star-state region. They trained networks of 5 to 7 layers with 20 to 70 nodes per hidden layer with exponential linear unit activation functions, adding a so-called *constraint-resolving* layer that ensures the star-state predictions satisfy the Rankine-Hugoniot shock jump conditions [20] within a specified tolerance. They also tested a more general constraint-adapted loss (CAL) method wherein the loss function penalizes deviation from the constraint but is not guaranteed to satisfy the loss exactly. The mean L1 errors show good agreement on the order of $O(10^{-2})$ and were able to resolve the shock front precisely, however a discussion of computational costs was not included. The study was also limited to perfect gases where the polytropic relation significantly simplifies the thermodynamic relationships.

This work was soon followed by Gyrya *et al.* [108] who used a two-layer network with 64 nodes per hidden layer and ReLU activation functions. No physical constraints were explicitly incorporated. Still, the network achieved root-mean squared errors of $O(10^{-2})$ for the star-state conditions. The authors also attempted to train their network to predict speeds of the left- and right-moving waves (which are highly nonlinear functions of the star-state conditions), yielding unacceptable errors of O(1). Compared to Magiera *et al.* [107], the network by Gyrya *et al.* was ostensibly simpler with similarly low star-state errors, but numerical results were prone to spurious discontinuities not reflected in the averaged error metrics associated with the learning curves. Most importantly, Gyrya *et al.* output 9 variables in the star-state region, of which many are redundant and can be calculated using standard equations in fluid mechanics. Thus, it is possible that their network added unwarranted computational overhead to produce these extra variables that may instead be computed more quickly and accurately by solving the appropriate fluids equations explicitly. In both the Magiera *et al.* [107] and Gyrya *et al.* [108] studies, there is limited analysis on network complexity and the implications for under- or overfitting. Additionally, it should be noted that the range of training and test data can affect the network size required to achieve a certain level of accuracy, and therefore the ideal network size is not universal.

Tangential studies have also considered other topics related to Riemann solver design. Fuks & Tchelepi [109] attempted to design a physics-informed neural network to predict fluxes given initial conditions in two-phase porous media, but noted challenges when training networks to minimize highly nonlinear loss surfaces with discontinuities. Dieselhorst *et al.* [110] trained neural networks to accelerate the primitive-to-conservative conversions in relativistic hydrodynamics, for which there is no analytical closed-form solution much like in non-ideal thermodynamics. Though these studies did not relate to flows with non-ideal thermodynamics, they suggest potential alternate network architectures and the possible benefits they may provide to the overall numerical scheme.

As such, there remains an opportunity to develop a learning-based Riemann solver that preserves the requisite wave structure of the solution, achieves low errors when predicting variables required for flux calculations, is computationally efficient compared to state-of-the-art approximate solvers, is physically consistent, and delivers these capabilities for flows with non-ideal thermodynamics. In transcritical flows especially, more sophisticated Riemann solvers are generally needed to preserve entropy stability—there may be opportunities to take advantage of parallel matrix algorithms to render efficient neural network computations relative to traditional methods.

1.1.4 Summary of Research Gaps

While analytical shock and rarefaction solutions are available for perfect gases, the lack thereof for non-ideal EOS has undermined the study of non-ideal flows in two ways. For one, numerical methods intended for non-ideal thermodynamics are regularly verified against shock tube tests where fluids behave as an ideal gas. Secondly, numerical results for trans- and supercritical problems are commonly verified against other numerical results despite sometimes circular validation. For non-ideal flows in particular, there is a need for exact solutions that do not involve numerical integration. This would permit more rigorous verification of codes designed for such flow conditions.

Within numerical codes, Riemann solvers have proven tremendously useful in the finite volume method. However, the occurrence of transonic rarefactions and entropy violations are discussed only tangentially to the derivation of Riemann solvers—such situations are considered an edge case in which so-called entropy fixes may be applied. Current entropy fixes are largely crude, introducing a heuristic amount of artificial diffusivity to cure non-physicalities in shock-capturing codes. These fixes are also highly simplistic, comprising constant-state averages of the head and tail of the rarefaction. Alternate methods for reconstructing the rarefaction should also be investigated, along with the scaling behaviour of errors that result when the rarefaction is ignored (*i.e.* modelled a discontinuous wave). Also in need of study are the flow conditions in which transonic rarefactions tend to occur.

If exact solutions prioritize accuracy over speed, and approximate Riemann solvers the opposite, then ML presents an exciting avenue to deliver a balance of both that has not been historically attainable. Early works have demonstrated the feasibility of training networks that can predict star-state conditions for ideal gases. However, for subcritical and ideal-gas flow problems, it may be argued that the computational demands of an ML approach outweigh any gains in accuracy. For transcritical or supercritical flows, where issues such as entropy violations and spurious oscillations warrant more complex Riemann solvers to ensure physically consistent solutions, a learning-based Riemann solver may be better received. Current literature has yet to explore this possibility. In addition, past studies have not rigorously studied the effect of network size and physical constraints on the accuracy and generalizability of predictions.

1.2 Research Objectives

The research gaps summarized at the end of §1.1 inform the objectives of this thesis:

1. Explore new approaches to solving shocks and rarefactions exactly in the case of non-ideal gases, leveraging analytical derivations wherever possible to uphold mathematical rigour.

2. Develop or modify practical Riemann solvers that offer improved entropy stability with minimal computational increase, especially for transonic or transcritical flows with large gradients where even small errors can destabilize solutions.

3. Develop a learning-based Riemann solver for flows with non-ideal thermodynamics, accounting for the role of network size and physical constraints on performance.

4. Compare and contrast exact, approximate, and machine learning approaches in terms of computational complexity and accuracy—thus enabling scientists and engineers to critically weigh each method's benefits and disadvantages when simulating modern compressible flow problems.

A list of key findings, research implications, and future research directions are provided in chapter 7, indicating that all objectives were satisfied.

1.3 Overview of Thesis Structure and Results

Chapters 2 and 3 provide the essential background theory on CFD and ML required to study the RP in the context of non-ideal gases. In the chapter on CFD theory, the full nonlinear Euler equations and the derivation of Godunov's first-order upwind scheme are discussed. Also examined are different EOS and thermodynamic departure functions needed to compute thermodynamic quantities such as heat capacity, speed of sound, and enthalpy for non-ideal thermodynamics. In the chapter on ML theory, such topics as supervised learning, fully connected neural networks (multi-layer perceptrons), loss functions, optimizers, regularization, and physics-informed approaches are covered. Together, these chapters explain the mathematical and physical concepts that are used in this thesis to investigate exact, approximate, and learning-based Riemann solvers.

Exact analytical solutions to non-ideal shocks and rarefactions are addressed in chapter 4. The stationary normal shock and the centred expansion wave cases are dealt with separately. It is shown that solving stationary normal shocks exactly requires the use of an iterative solver, however, such iterations converge quickly to a high tolerance within a few steps. Then, a novel domain mapping strategy is used to solve for the primitive variables inside a non-ideal rarefaction as explicit functions of space and time. The derivation reveals that the primitives vary according to a generic exponential function with constants determined by the boundary conditions at the head and tail of the wave. Most importantly, the derivation of the non-ideal expansion wave makes no assumptions on the EOS and is therefore valid for arbitrary thermodynamics and flow conditions. This promising result sets the stage for the subsequent chapter.

Chapter 5 explores the idea of structural completeness in the context of approximate Riemann solvers. Leveraging the prior analytical solutions for non-ideal rarefactions, a simple but effective analytical entropy fix is proposed that restores the expansion wave to linearized solvers such as the Roe solver. By rigorously analyzing the conditions under which rarefaction waves are transonic, it is also shown that entropy violations—and therefore entropy fixes—tend to be needed when flow problems possess large gradients and occur under trans- or supercritical conditions. Using various 1D and 2D numerical test cases, the new entropy fix achieves lower errors to the traditional Harten-Hyman fix while bearing similar computational complexity. The new fix is also analytical and thus requires no tuning of artificial diffusivity as with traditional fixes. The new entropy fix is shown to improve numerical stability, entropy satisfaction, and ensure positively conservative results on cases where the Roe solver would normally fail.

Having explored exact and approximate approaches, chapter 6 examines the feasibility and merits of a learning-based approach, termed *FluxNet*. An evaluation of different possible input and output variables reveals that the use of primitive variables yields simpler networks over attempts to incorporate conservatives or fluxes. The loss curves of data-driven and physics-informed FluxNets are analyzed, highlighting the importance of incorporating physical constraints to ensure smoothness, generalizability, and physical consistency. Numerical test cases show further reductions in error over the structurally complete approach of the prior chapter, albeit at some increase in computational expense. However, the bias and variance of errors over all train and test data is significantly smaller with FluxNets than with traditional Roe-type solvers. Ultimately, the FluxNet approach achieves greater accuracy than linearized solvers and less runtime complexity than exact techniques.

A review of the major findings, their implications, and potential future directions is given in chapter 7. While traditional approximate solvers may be suited for low-speed non-ideal gases, entropy-stable solvers are required for high-speed non-ideal flows with large gradients. In this regard, the structurally complete Riemann solver proposed in chapter 5, as enabled by the derivations of chapter 4, is arguably most appropriate due to its general validity. For flow problems where the stability of numerical schemes is especially sensitive to noise or errors, a FluxNet approach may be helpful to avoid phenomena such as pressure instabilities that naturally arise when simulating transcritical flows. For the purposes of scientific reproducibility, hyperlinks to external code repositories are included at the outset of chapters 4 through 6.

Chapter 2

Theory I: CFD

This chapter summarizes essential aspects of fluid mechanics, thermodynamics, and numerical schemes that are relevant to studying the RP with non-ideal gases.

2.1 Governing Equations

The flow is assumed to be inviscid, isentropic except at any shocks, and one-dimensional in the x-direction. Gravity is neglected. Therefore, the time-dependent Euler equations apply, which in differential and conservative form are:

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u)}{\partial x} = 0 \tag{2.1}$$

$$\frac{\partial (\rho u)}{\partial t} + \frac{\partial (p + \rho u^2)}{\partial x} = 0 \tag{2.2}$$

$$\frac{\partial}{\partial t}\left(\rho\left(\frac{e}{M}+\frac{u^2}{2}\right)\right)+\frac{\partial}{\partial x}\left(\rho u\left(\frac{e}{M}+\frac{u^2}{2}\right)\right)+\frac{\partial(pu)}{\partial x}=0 \tag{2.3}$$

where ρ is density, u is the velocity component in the x-direction, p is pressure, M is the molar mass of the fluid, e is specific internal energy on a molar basis, and t is time. The full multi-dimensional governing equations are provided in Appendix A. Additionally, the specific enthalpy on a molar basis is:

$$h = e + pv \tag{2.4}$$

where $v = M/\rho$ is the molar specific volume.

Sometimes it is more useful to work with total energy E per unit volume:

$$E=\rho\left(\frac{e}{M}+\frac{u^2}{2}\right)=\frac{e}{v}+\frac{\rho u^2}{2} \tag{2.5}$$

as well as total enthalpy H per unit mass:

$$H=\frac{E+p}{\rho}=\frac{h}{M}+\frac{u^2}{2} \tag{2.6}$$

which permit the energy equation (2.3) to be re-expressed:

$$\frac{\partial E}{\partial t}+\frac{\partial}{\partial x}\Big(u\,(E+p)\Big)=\frac{\partial E}{\partial t}+\frac{\partial}{\partial x}\Big(\rho u H\Big)=0 \tag{2.7}$$

With this information, it is possible to represent the Euler equations in matrix form:

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = 0 \tag{2.8}$$

where $\mathbf{U}$ is the vector of conserved variables and $\mathbf{F}$ is the flux vector:

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}; \qquad \mathbf{F} = \begin{bmatrix} \rho u \\ p + \rho u^2 \\ u(E+p) \end{bmatrix} \tag{2.9}$$

For steady flows, such as steady normal shocks that are aligned with the shock frame of reference, the time-dependent terms drop out. For centred expansion waves and most fluid problems, which are aligned with the lab frame of reference, the full Euler equations are applicable.

2.2 Equation of State

Herein, no requirements are imposed on the choice or nature of the EOS. However, for the purposes of computing numerical results, the general form of a cubic EOS [111] is employed:

$$p = \frac{RT}{v-b} - \frac{\Theta}{v^2 + \delta v + \epsilon} \tag{2.10}$$

where $R = 8314.4621\,\mathrm{J\,kmol^{-1}\,K^{-1}}$ is the universal gas constant, T is temperature, Θ is a function of temperature, and b, δ, ϵ are constants (Tab. 2.1). For an ideal gas, the parameters $b, \delta, \epsilon, \Theta$ are simply zero. This thesis focuses on single-species fluids, however Poling *et al.* [25] show how to generalize to mixtures via the appropriate mixture rules.

The use of pressure-explicit EOS is desirable for both analytical derivations and numerical studies because Maxwell relations enable all thermodynamic quantities to be expressed in terms of explicit partial derivatives and integrals of pressure. Cubic EOS in particular are widely used in CFD, where a balance needs to be struck between computational efficiency and thermodynamic accuracy. It is important to note that the choice of EOS for a given fluid problem is highly dependent on context, and as such, readers are referred to Poling *et al.*[25] for further discussion.

The partial derivatives and integrals of (2.10) are stated below. Constants of integration are left out for simplicity since these integrals are always evaluated over definite intervals in the derivations.

$$\left.\frac{\partial p}{\partial T}\right|_v = \frac{R}{v-b} - \frac{\frac{d\Theta}{dT}}{v^2+\delta v+\epsilon} \tag{2.11}$$

$$\left.\frac{\partial p}{\partial v}\right|_T = \frac{(2v+\delta)\Theta}{(v^2+\delta v+\epsilon)^2} - \frac{RT}{(v-b)^2} \tag{2.12}$$

$$\left.\frac{\partial^2 p}{\partial T^2}\right|_v = -\frac{\frac{d^2\Theta}{dT^2}}{v^2+\delta v+\epsilon} \tag{2.13}$$

$$\int \left.\frac{\partial p}{\partial T}\right|_v dv = R\ln(v-b) - \frac{2\frac{d\Theta}{dT}\operatorname{arctanh}(\frac{\delta+2v}{\sqrt{\delta^2-4\epsilon}})}{\sqrt{\delta^2-4\epsilon}} \tag{2.14}$$

$$\int v\left.\frac{\partial p}{\partial v}\right|_T dv = \frac{bRT}{v-b} - RT\ln(v-b) + \frac{2\Theta\operatorname{arctanh}(\frac{\delta+2v}{\sqrt{\delta^2-4\epsilon}})}{\sqrt{\delta^2-4\epsilon}} - \frac{\Theta v}{v^2+\delta v+\epsilon} \tag{2.15}$$

$$\int \frac{\partial^2 p}{\partial T^2}\bigg|_v dv = -\frac{2\frac{d^2\Theta}{dT^2}\operatorname{arctanh}(\frac{\delta+2v}{\sqrt{\delta^2-4\epsilon}})}{\sqrt{\delta^2-4\epsilon}} \tag{2.16}$$

Table 2.1: Abott parameters, constants, and derivatives of $\Theta(T)$ for common cubic EOS. ω is the acentric factor defined by Pitzer et al. [3][4]. It has a value $\omega = -\log_{10}(p_{ac}) - 1$ where $p_{ac} = \frac{p}{p_c}$ at $\frac{T}{T_c} = 0.7$. For CO_2, $\omega = 0.228$.

Name	Abbott Parameters	Constants	Relevant Derivatives
RK [112]	$b, \delta = b, \epsilon = 0,$ $\Theta = a(\frac{1}{T})^{\frac{1}{2}}$	$a = 0.4278\frac{R^2T_c^{\frac{5}{2}}}{p_c},$ $b = 0.0867\frac{RT_c}{p_c}$	$\frac{d\Theta}{dT} = -\frac{a}{2}(\frac{1}{T})^{\frac{3}{2}}$ $\frac{d^2\Theta}{dT^2} = \frac{3a}{4}(\frac{1}{T})^{\frac{5}{2}}$
SRK [113]	$b, \delta = b, \epsilon = 0,$ $\Theta = a(1 + (0.48 + 1.574\omega$ $-0.176\omega^2)(1 - (\frac{T}{T_c})^{\frac{1}{2}}))^2$	$a = 0.42747\frac{R^2T_c^2}{p_c},$ $b = 0.08664\frac{RT_c}{p_c}$	$\frac{d\Theta}{dT} = -\frac{\Omega a}{\sqrt{T_c}}(\frac{1-\Omega(\sqrt{\frac{T}{T_c}}-1)}{\sqrt{T}})$ $\frac{d^2\Theta}{dT^2} = \frac{\Omega(\Omega+1)a}{2\sqrt{T_c}T^{\frac{3}{2}}}$ $\Omega = 0.48 + 1.574\omega - 0.176\omega^2$
PR [2]	$b, \delta = 2b, \epsilon = -b^2,$ $\Theta = a(1 + (0.3746 + 1.5422\omega$ $-0.2699\omega^2)(1 - (\frac{T}{T_c})^{\frac{1}{2}}))^2$	$a = 0.45724\frac{R^2T_c^2}{p_c},$ $b = 0.07780\frac{RT_c}{p_c}$	$\frac{d\Theta}{dT} = -\frac{\Omega a}{\sqrt{T_c}}(\frac{1-\Omega(\sqrt{\frac{T}{T_c}}-1)}{\sqrt{T}})$ $\frac{d^2\Theta}{dT^2} = \frac{\Omega(\Omega+1)a}{2\sqrt{T_c}T^{\frac{3}{2}}}$ $\Omega = 0.37464 + 1.54226\omega - 0.2699\omega^2$

2.3 Heat Capacity, Enthalpy, and Speed of Sound

By relaxing the ideal gas assumption, a more general EOS may be considered. Thermodynamic functions of a single phase flow can be fully defined using two independent thermodynamic state variables, for example speed of sound $c(v, T)$ or enthalpy $h(v, T)$. Based on the selected EOS, the specific heat capacities, enthalpy, and speed of sound may be expressed as functions of v, T by applying thermodynamically-consistent departure functions to the corresponding ideal gas states.

The ideal gas specific heat capacities and enthalpies on a molar basis c'_p, c'_v, h' may be empirically modelled as polynomial functions of temperature alone. More specifically, c'_p, c'_v, h' are modelled as *thermally* perfect (dependent on T only), not *calorically* perfect (constant with respect to thermodynamic state). The subscripts p, v indicate constant pressure or constant volume, respectively, while the prime superscripts denote we are denoting an ideal gas. Thermally perfect gas specific heat capacities are commonly represented with polynomials of degree n [114]:

$$c'_p(T) = \sum_{i=0}^{n} B_i T^i \tag{2.17}$$

$$c'_v(T) = c'_p - R = \sum_{i=0}^{n} B_i T^i - R \tag{2.18}$$

where the B_i terms are constants fitted from experimental data. The thermally perfect gas specific enthalpy on a molar basis is then:

$$h'(T) = \int_0^T c'_p dT = \sum_{i=0}^{n} \frac{B_i}{i+1} T^{i+1} \tag{2.19}$$

where $h' = 0$ at $T = 0$. Now applying departure functions to Eqs. (2.17), (2.18), and (2.19), the non-ideal gas specific heat capacities are:

$$c_v(v, T) = c'_v + T \int_{\infty}^{v} \frac{\partial^2 p}{\partial T^2}\bigg|_v dv \tag{2.20}$$

$$c_p(v, T) = c_v - T\Big(\frac{\partial p}{\partial T}\bigg|_v\Big)^2\Big(\frac{\partial p}{\partial v}\bigg|_T\Big)^{-1} \tag{2.21}$$

and the non-ideal gas specific enthalpy h is:

$$h(p, T) = h'(T) + \int_0^p \left(v - T\frac{\partial v}{\partial T}\bigg|_p\right)\Bigg|_T dp \tag{2.22}$$

which can be converted to partial derivatives of pressure using the triple product rule and $dp = \frac{dp}{dv}\big|_T dv$:

$$h(p, T) = h'(T) + T\int_{\infty}^{v} \frac{\partial p}{\partial T}\bigg|_v dv + \int_{\infty}^{v} v\frac{\partial p}{\partial v}\bigg|_T dv \tag{2.23}$$

where the domain has been mapped from $[0, p]$ to $[\infty, v]$. Finally, the non-ideal gas speed of sound may be expressed:

$$c^2 = \frac{\partial p}{\partial \rho}\bigg|_s = -\frac{v^2}{M}\frac{\partial p}{\partial v}\bigg|_s \tag{2.24}$$

since $\rho = \frac{M}{v} \Rightarrow dv = -\frac{M}{v^2}d\rho$. For an isentropic process, $\frac{c_p}{c_v} = \frac{\partial v}{\partial p}\big|_T \frac{\partial p}{\partial v}\big|_s$ [115]. Thus:

$$c(v, T) = \sqrt{-\frac{v^2}{M}\frac{c_p}{c_v}\frac{\partial p}{\partial v}\bigg|_T} \tag{2.25}$$

Lastly, Eqs. (2.21) and (2.20) may be substituted into Eq. (2.25) to produce a monstrous equation that is left for the reader to expand at their extended leisure. It is easy to verify that in the case of an ideal gas, $\gamma = \frac{c_p}{c_v}$ and $\frac{\partial p}{\partial v}\big|_T = -\frac{RT}{v^2}$, yielding $c'(T) = \sqrt{\gamma RT/M}$ as expected. These equations for specific heat capacities, enthalpy, and speed of sound are analytical except for a polynomial fit of $c_p'(T)$ which for values of $n = 6$ have been shown to give errors of less than 0.5% compared to experimental values [116].

2.4 Characteristic Curves of the Riemann Problem

At the outset of Chapter 1, the RP was qualitatively described as an initial value problem whose solution comprises a combination of shock and expansion waves along with a single contact discontinuity wave. Mathematically, this wave structure arises from the method of characteristics. Textbooks [7, 24] normally derive the characteristic curves via eigendecomposition of the linearized time-dependent Euler equations. But when exact solutions are of interest, it is necessary to use the full nonlinear equations and thus avoid any eigenanalysis. It will be proven shortly that the characteristic curves obey the same compatibility equations as in the nonlinear case, giving rise to left-running (C_-), right-running (C_+), and entropy (C_0) wave characteristics.

In isentropic flow, $dp = c^2 d\rho$. Substituting into the continuity equation (2.1) and rearranging gives:

$$\frac{1}{\rho c}(\frac{\partial p}{\partial t} + u\frac{\partial p}{\partial x}) + c\frac{\partial u}{\partial x} = 0 \tag{2.26}$$

The momentum equation (2.2) may be reformulated by expanding the derivatives and using the continuity equation to cancel out terms, resulting in:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0 \tag{2.27}$$

Adding or subtracting Eq. (2.26) to/from Eq. (2.27) gives:

$$\left(\frac{\partial u}{\partial t} + (u \pm c)\frac{\partial u}{\partial x}\right) \pm \frac{1}{\rho c}\left(\frac{\partial p}{\partial t} + (u \pm c)\frac{\partial p}{\partial x}\right) = 0 \tag{2.28}$$

Now considering the total differential $du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial t}dt$, it is possible to choose characteristic curves $C_\pm : dx = (u \pm c)dt$ such that $\frac{du}{dt} = \frac{\partial u}{\partial t} + (u \pm c)\frac{\partial u}{\partial x}$, respectively. Similarly, $\frac{dp}{dt} = \frac{\partial p}{\partial t} + (p \pm c)\frac{\partial p}{\partial x}$. Substituting $\frac{du}{dt}$ and $\frac{dp}{dt}$ in (2.28) yields left-running characteristics $C_- : \frac{dx}{dt} = u - c$ along which the compatibility equation is $dp - \rho c du = 0$, as well as right-running curves $C_+ : \frac{dx}{dt} = u + c$ along which $dp + \rho c du = 0$. Finally, there are so-called entropy characteristics $C_0 : dx = udt$ where the compatibility equation is simply the isentropic speed of sound $dp - c^2 d\rho = 0$. This characteristic field is visualized in Fig. 2.1.

It is imperative to note that C_-, C_0, C_+ do not denote *specific* curves but rather *families* of all curves that obey the corresponding compatibility equation. There are infinitely many such curves, however the contact discontinuity is one particular C_0 while expansion wave heads and tails are particular cases of C_- or C_+ curves depending on if they appear on the left or right of the contact discontinuity. Shocks are regions where characteristic curves on either side coalesce.

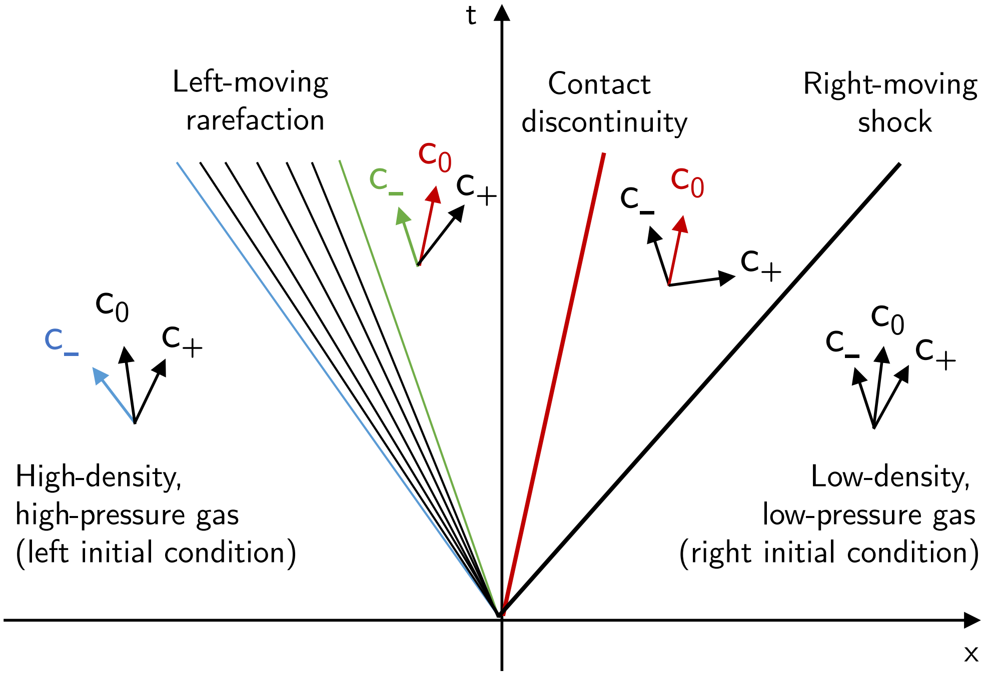

t
Left-moving rarefaction
Contact discontinuity
Right-moving shock
C₋
C₀
C₊
C₋
C₀
C₊
C₋
C₀
C₊
C₋
C₀
C₊
High-density, high-pressure gas (left initial condition)
Low-density, low-pressure gas (right initial condition)
x

Figure 2.1: Characteristic field of the Riemann problem in the case of a high-pressure, high-density left initial region and a low-pressure, low-density right initial region. It can be rigorously proven that C_- characteristics on either side of a left rarefaction are parallel to the nearest bound of the rarefaction (indicated by the matching blue and green slopes); C_0 characteristics in the central star-state region are aligned with the slope of the contact discontinuity (red slopes); and, all characteristics to the right of the shock and C_+ characteristics in the post-shock region coalesce into the shock front.

2.5 A Building Block for Numerical Methods

The RP is often termed a building block for numerical methods [24] because it is an intuitive model for how information propagates between cells. In the finite volume method, fluxes modelled via the RP ensure that solutions are physically consistent and correctly upwinded (Fig. 2.2). Godunov [8] was the first to recognize the potential of the RP and proposed a scheme that was first-order accurate in space and time. Unfortunately, Godunov's original scheme called for expensive exact-iterative solutions to the RP at each cell interface. It was not until van Leer's pioneering work [117] that an efficient second-order extension of Godunov's scheme led to a renewed interest and use of Riemann solvers. Since then, a number of Riemann solvers [49, 21, 23] have been developed and used in a wide range of numerical schemes including AUSM [118, 119] and finite-volume WENO [120]. Because a given fluids simulation may involve many thousands or millions of calls of a Riemann solver, such solvers must strike a careful balance between accuracy and computational simplicity.

2.6 First-Order Upwind Godunov Scheme

In this thesis, all numerical solutions are computed with a first-order upwind Godunov scheme—derived in detail for one dimension in the next paragraph. Extending the Godunov scheme to multiple dimensions is discussed in Appendix B. A low-order scheme was chosen for three reasons. Firstly, Godunov's theorem [8] states that monotone, linear numerical schemes are at most first-order accurate. That is, higher-order linear schemes do not guarantee monotonicity of solutions, which is a fundamental property of exact solutions

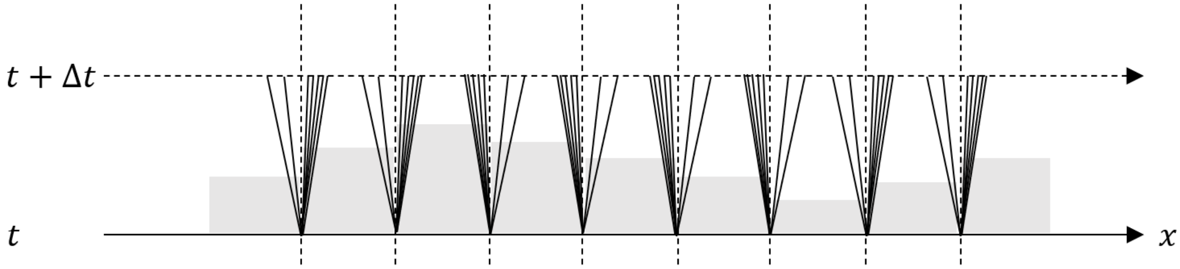


Figure 2.2: A visualization of the Riemann problem as a building block of the finite volume method. The conserved variables of each cell at time t are visualized by the grey rectangles. The cell interfaces are modelled as Riemann problems whereby different self-similar wave solutions emerge depending on the initial conditions flanking each interface. At time $t + \Delta t$, the waves have propagated such that the fluxes at each interface may be computed by solving, whether exactly or approximately, each Riemann problem.

of conservation laws. Secondly, Harten [21] proved that a scheme is monotone if and only if it is also total variation non-increasing, thus ensuring convergence. A first-order upwind Godunov scheme is therefore also convergent. Thirdly, and most importantly, a low-order scheme clearly demonstrates the impact of Riemann solver design upon the accuracy of results, whereas a higher-order scheme would require additional fixes (*e.g.* Total-Variation Diminishing) to preserve non-oscillatory behaviour [24]. In the higher-order scenario, it becomes challenging to decouple the error uniquely due to the Riemann solver. For these reasons, a low-order scheme is favoured. Time integration in this thesis is achieved via a strong stability-preserving third-order Runge-Kutta method [121]. Nevertheless, all Riemann solvers discussed in this thesis—whether exact, approximate, or learning-based—may be employed in higher-order schemes as required.

The derivation of the first-order upwind Godunov scheme follows. We are given a boundary value problem governed by (2.8) with initial conditions at $t = 0$:

$$\mathbf{U}(x, 0) = \mathbf{U}_{t=0}(x) \tag{2.29}$$

and boundary conditions at $x = 0$ and $x = L$:

$$\mathbf{U}(0, t) = \mathbf{U}_{x=0}(t); \qquad \mathbf{U}(L, t) = \mathbf{U}_{x=L}(t) \tag{2.30}$$

where the spatial domain is $x \in [0, L]$ and the time domain is $t \in [0, T]$ where $T \geq 0$ (not to be confused with temperature). It is assumed that there exists a unique entropy-satisfying solution to this problem. In order to permit discontinuous solutions, an integral form of the conservation laws (2.8) must be embraced. For a control volume $[x_1, x_2] \times [t_1, t_2]$ within the spatiotemporal domain, we have:

$$\int_{x_1}^{x_2} \mathbf{U}(x, t_2)dx - \int_{x_1}^{x_2} \mathbf{U}(x, t_1)dx + \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_2, t))dt - \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_1, t))dt = 0 \tag{2.31}$$

Now that the problem is defined, the numerical scheme may be developed in the following major steps: discretize the domain, apply suitable boundary conditions to the discretized domain, approximate the governing equations, and derive a flux function according to the solution of the RP at each cell boundary.

2.6.1 Discretization of the Domain

The spatial domain $[0, L]$ is discretized into M cells of width $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ where $i = 1, ..., M$ (not to be confused with molar mass). For simplicity, we uniformly distribute the cells so that $\Delta x = L/M$. For the i^{th} cell, the position of the cell centre x_i is:

$$x_i = (i - \frac{1}{2})\Delta x \tag{2.32}$$

and the cell boundaries $x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}$ are:

$$x_{i-\frac{1}{2}} = (i-1)\Delta x; \qquad x_{i+\frac{1}{2}} = i\Delta x \tag{2.33}$$

of which there are $M + 1$ cell boundaries in total.

The temporal discretization is performed in steps of Δt of varying size dependent on the maximum wavespeed in the spatial domain at each time t. The time step size must be limited to ensure that information from any one cell boundary does not propagate to any other cell boundary, or else the known cell conditions at the current time t cannot be used to solve for the next time step $t+\Delta t$. As discussed in §2.4, the fastest left-moving and right-moving wavespeeds originating from the j^{th} cell boundary, where $j = 0, ..., M$, are given by $S_{Lj} = u(x_j, t) - c(x_j, t)$ and $S_{Rj} = u(x_{j+1}, t) + c(x_{j+1}, t)$. Thus, a Courant-Friedrichs-Lewy (CFL) condition applies:

$$\Delta t = \text{CFL} \min_{j \in [0,M]} \left(\frac{\Delta x}{|S_{Lj}|}, \frac{\Delta x}{|S_{Rj}|} \right) \tag{2.34}$$

where CFL < 1. The resultant discretization scheme is shown in Fig. 2.3. Finally, the conserved variables in each cell are modelled as constant such that the overall solution is piecewise-constant at each time step. For notational purposes, we write $\mathbf{U}_i^n$ to denote the conditions in the i^{th} cell at the n^{th} time step.

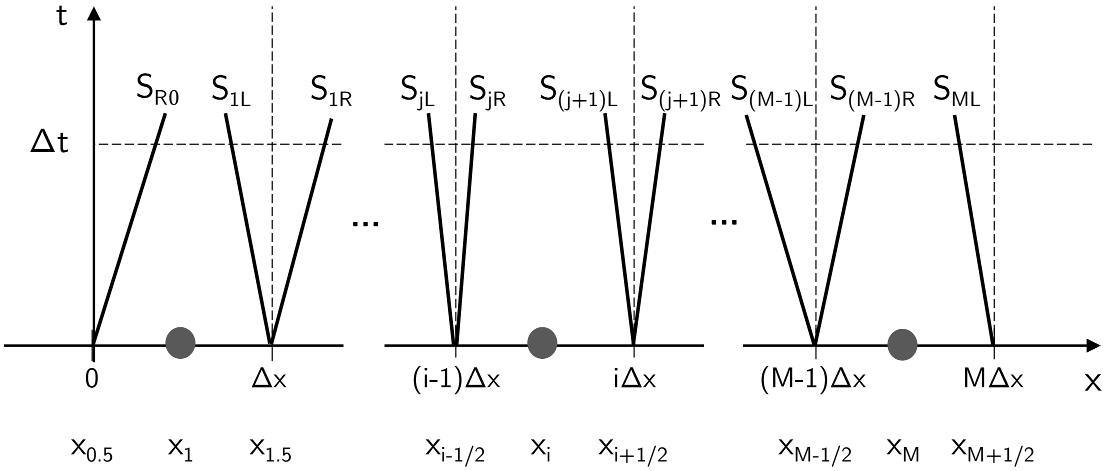


Figure 2.3: Discretization of the spatiotemporal domain. Dashed lines indicate the cell boundaries. Grey circles indicate the x position of the cell centres x_i. The time step size Δt changes from one time step to the next based on a Courant-Friedrichs-Lewy condition on the fastest wavespeeds $\min_{j\in[0,M]}(S_{Lj}, S_{Rj})$ in the spatial domain at time t.

2.6.2 Types of Boundary Conditions

The boundary conditions (2.30) are implemented via a so-called *ghost cell* approach due to its flexibility in accommodating various types of boundaries. A fictitious cell of width Δx is appended before $x = 0$, and another fictitious cell is appended after $x = L$. Their cell centres are $x_0 = -\frac{1}{2}\Delta x$ and $x_{M+1} = (M + \frac{1}{2})\Delta x$, respectively. When specifying the ghost cell conditions, there may be reflective, transmissive, or periodic boundaries.

A reflective boundary returns incoming signals with the same magnitude but opposite direction. The physical analogue to a reflective boundary is a stationary impermeable wall. This boundary condition is achieved by changing the sign of velocity but retaining the thermodynamic state ($i.e.$ p, ρ) of the flow. The corresponding ghost-cell states to the left and right of the domain are therefore:

$$\mathbf{U}_0^n = \begin{bmatrix} \rho_1^n \\ -\rho_1^n u_1^n \\ E_1^n \end{bmatrix}; \qquad \mathbf{U}_{M+1}^n = \begin{bmatrix} \rho_M^n \\ -\rho_M^n u_M^n \\ E_M^n \end{bmatrix} \tag{2.35}$$

Transmissive boundaries (sometimes called transparent, open-end, radiation, or far-field boundary conditions) permit incoming signals to propagate without any impedance. The use of transmissive boundaries is particularly useful in the simulation of small computational domains. Mathematically, a transmissive boundary is achieved by copying the conditions in the nearest cell. Ghost-cell states for transmissive boundaries are:

$$\mathbf{U}_0^n = \begin{bmatrix} \rho_1^n \\ \rho_1^n u_1^n \\ E_1^n \end{bmatrix}; \qquad \mathbf{U}_{M+1}^n = \begin{bmatrix} \rho_M^n \\ \rho_M^n u_M^n \\ E_M^n \end{bmatrix} \tag{2.36}$$

Periodic boundaries redirect incoming signals to the other end of the domain and while propagating in the same direction. This is useful when simulating periodic flows or wave interference while keeping the computational domain small. Periodic boundaries are thus similar to transmissive boundaries except the ghost-cell states are switched:

$$\mathbf{U}_0^n = \begin{bmatrix} \rho_M^n \\ \rho_M^n u_M^n \\ E_M^n \end{bmatrix}; \qquad \mathbf{U}_{M+1}^n = \begin{bmatrix} \rho_1^n \\ \rho_1^n u_1^n \\ E_1^n \end{bmatrix} \tag{2.37}$$

For further details on these boundary conditions, the reader may consult Toro [24].

2.6.3 Approximation of the Euler Equations

With the domain and boundary conditions discretized appropriately, we now turn our attention to the integral form of the conservation laws (2.31). For the i^{th} cell, let us consider the control volume defined by $x_1 = x_{i-\frac{1}{2}}$, $x_2 = x_{i+\frac{1}{2}}$, $t_1 = t_n$, and $t_2 = t_{n+1}$. Then:

$$\begin{aligned} &\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{U}(x, t_{n+1}) dx - \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{U}(x, t_n) dx \\ &+ \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i+\frac{1}{2}}, t)) dt - \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i-\frac{1}{2}}, t)) dt = 0 \end{aligned} \tag{2.38}$$

However, at the end of our discussion on discretization, we modelled the conditions in each cell as constant in space. That is, $\mathbf{U}(x, t_n) = \mathbf{U}_i^n$ for $x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ and $t = t_n$; and so forth. Thus, (2.38) simplifies to:

$$\mathbf{U}_i^{n+1} \Delta x - \mathbf{U}_i^n \Delta x + \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^n) \Delta t - \mathbf{F}(\mathbf{U}_{i-\frac{1}{2}}^n) \Delta t = 0 \tag{2.39}$$

where $\mathbf{U}^n_{i+\frac{1}{2}}$ are the conditions at the intercell boundary $x_{i+\frac{1}{2}}$ for $\Delta t > 0$, which corresponds to the solution at $x = 0$ of the RP with left and right initial conditions $(\mathbf{U}^n_i, \mathbf{U}^n_{i+1})$. Provided that the CFL condition is respected, the solution at $x = 0$ is constant for $\Delta t > 0$ since the slopes of the characteristic waves smoothly change sign passing through $x = 0$, becoming vertical exactly at $x = 0$ (Fig. 2.1). Although this observation is traditionally linked to the linearized, ideal-gas formulation of the Euler equations, a proof for the full non-linear equations and non-ideal thermodynamics is given in §2.4.

After some rearrangement, we obtain the Godunov first-order numerical scheme:

$$\mathbf{U}^{n+1}_i = \mathbf{U}^n_i + \frac{\Delta t}{\Delta x}\left(\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}\right) \tag{2.40}$$

where the intercell flux at $x_{i+\frac{1}{2}}$ is:

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}) \tag{2.41}$$

and the time step Δt is required to satisfy the CFL condition from earlier (2.34). To incorporate strong stability-preserving third-order Runge-Kutta time-stepping [121], the numerical scheme is modified to the following form:

$$\begin{aligned}
(\mathbf{U}^{n+1}_i)^{(1)} &= \mathbf{U}^n_i + \frac{\Delta t}{\Delta x}\left(\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}\right) \\
(\mathbf{U}^{n+1}_i)^{(2)} &= \frac{3}{4}\mathbf{U}^n_i + \frac{1}{4}(\mathbf{U}^{n+1}_i)^{(1)} + \frac{1}{4}\frac{\Delta t}{\Delta x}\left(\mathbf{F}^{(1)}_{i-\frac{1}{2}} - \mathbf{F}^{(1)}_{i+\frac{1}{2}}\right) \\
\mathbf{U}^{n+1}_i &= \frac{1}{3}\mathbf{U}^n_i + \frac{2}{3}(\mathbf{U}^{n+1}_i)^{(2)} + \frac{2}{3}\frac{\Delta t}{\Delta x}\left(\mathbf{F}^{(2)}_{i-\frac{1}{2}} - \mathbf{F}^{(2)}_{i+\frac{1}{2}}\right)
\end{aligned} \tag{2.42}$$

where $\mathbf{F}^{(k)}_{i+\frac{1}{2}}$ represents the flux at $x_{i+\frac{1}{2}}$ once the solution has advanced to the the k^{th} sub-step, *i.e.* $(\mathbf{U}_i^{n+1})^{(k)}$ for $k = 1, 2$.

2.6.4 Computing flux terms via a Riemann solver

The last step in designing any Godunov-type scheme is formulating an appropriate Riemann solver. A Riemann solver is any function that calculates, whether exactly or approximately, the intercell flux as described in (2.41). More precisely, a Riemann solver is a function $\mathcal{R}$ that computes the flux $\mathbf{F}$ at cell boundary $x_{i+\frac{1}{2}}$:

$$\mathbf{F}_{i+\frac{1}{2}} \approx \mathcal{R}(\mathbf{U}_i, \mathbf{U}_{i+1}) \tag{2.43}$$

where $(\mathbf{U}_i, \mathbf{U}_{i+1})$ are the cell states to the left and right of the boundary. The first-order upwind Godunov scheme is now completely specified. Indeed, the remainder of this thesis compares existing and novel approaches to designing exact, approximate, and ML representations of $\mathcal{R}$. As we can see, the Riemann solver significantly influences the accuracy and computational effort of a numeral scheme. $\mathcal{R}$ must be evaluated $M + 1$ times per time step in the case of 1D forward Euler time-stepping, or $3M + 3$ times per time step for the 1D strong stability-preserving third-order Runge Kutta method.

Chapter 3

Theory II: ML

This chapter summarizes essential ML topics needed to study the RP with non-ideal gases.

3.1 Types of Learning Algorithms

Mitchell's [122] definition of ML may be stated as the phenomenon by which a computer program improves its performance P on one or more tasks T from experience E. The program is said to *learn* from experience E, which contains *training data*. When the trained program is used to make predictions for practical purposes, this process is referred to as *inference*. Numerous algorithms have been developed over the years to realize learning behaviour in computer programs, and they may be generally classified as supervised, unsupervised, and semi-supervised learning [123].

In supervised learning, the task T is mapping inputs to outputs and the experience

E is a set of example input-output pairs. In unsupervised learning, the task T is finding patterns among input data (e.g. via clustering or principal component analysis) and the experience E is a set of unlabelled data. The approaches of reinforcement and semi-supervised learning strive to balance exploration and exploitation of knowledge, by incentivizing desirable performance (reinforcement) or providing partially unlabelled training data (semi-supervised).

For a Riemann solver, the task T is to map initial conditions to intercell fluxes by learning from training data E produced by exact-iterative computations—thus, a supervised learning approach is appropriate.

3.2 Supervised Learning

Goodfellow *et al.* [124] offer a precise definition of supervised learning that may be summarized as:

> Given a set of n training points $(\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2), ..., (\bar{x}_n, \bar{y}_n)$ where $\bar{x}_i$ is a feature tensor and $\bar{y}_i$ is the response tensor of an unknown function $\bar{y}_i = f(\bar{x}_i)$, find a model $\mathcal{F}$ that approximates the true function f.

In the context of Riemann solvers, the feature vector $\bar{x}_i$ should contain sufficient information from the initial conditions in order to uniquely determine the flux. This thesis employs a feature vector of the form $\bar{x}_i = [\rho_L, u_L, p_L, \rho_R, u_R, p_R] \in \mathbb{R}^6$, and the prediction or response vector is chosen to be $\bar{y}_i = [\rho_{*L}, u_{*L}, p_{*L}, u_*, \rho_{*R}, u_{*R}, p_{*R}] \in \mathbb{R}^7$. At first glance, it may seem odd to include u_*, u_{*L}, u_{*R} when all three are constant in the star-state—however this separation is necessary to enforce physical constraints as discussed later in §3.5. From these predicted variables, it is possible to determine the wavespeeds in the RP and estimate the intercell flux. §6.1 contains a comprehensive analysis of the possible choices of feature and predicted variables, and how they may be integrated with the broader numerical scheme.

The supervised learning task at hand can thus be made more specific: to find a regression model $\mathcal{F}(\bar{x}_i; \theta) \in \mathbb{R}^{6\times7}$ with learnable parameters or weights θ that approximates the true mapping from initial conditions to star-state conditions as stated above. The regression model's error, which affects the accuracy of the subsequent calculations of wavespeeds and flux, is the L1 norm:

$$\Delta \bar{y}_i = |\hat{y}_i - \bar{y}_i| \tag{3.1}$$

where $\hat{y}_i = \mathcal{F}(\bar{x}_i; \theta)$ is the neural network prediction. Only nitrogen gas is studied in this thesis' ML research (Chapter 6), otherwise additional mixture parameters should be represented in the feature vector.

3.3 Multi-Layer Perceptron

A feed-forward neural network architecture with non-linear activation functions and fully connected hidden layers is used in this thesis. This architecture, known as the multi-layer perceptron, was selected due to its popularity when performing non-linear regression tasks in fluid mechanics [125, 108, 107, 126, 127]. More complex networks such as convolutional and recurrent networks are found in image processing, speech analysis, and natural language processing [80] where the focus is instead on classification and ordered data. However for low-dimensional continuously real-valued inputs and outputs, the multi-layer perceptron is simple yet suitable.

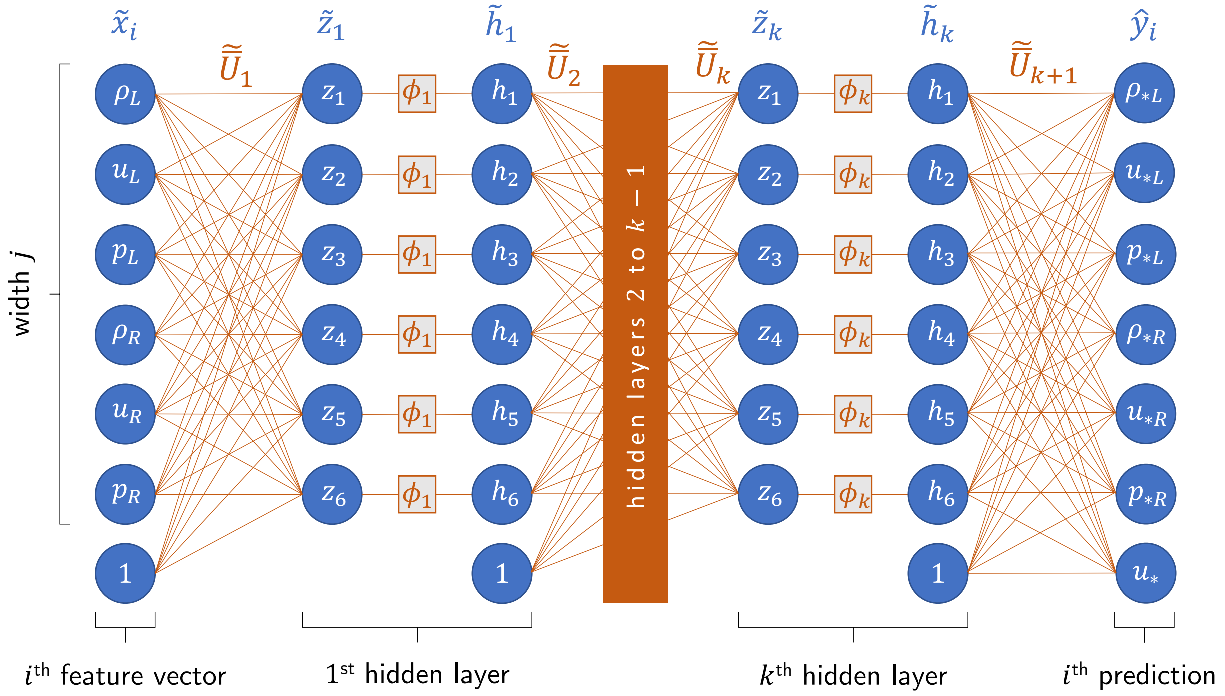


Figure 3.1: A generic multi-layer perceptron designed for the Riemann problem with depth $k+1$ and width j at each layer (which need not be equal to the dimension of $\bar{x}_i$ nor $\hat{y}_i$). Each hidden layer applies a non-linear activation function in an element-wise fashion. The matrices $\tilde{\tilde{U}}$ contain the weights of the network. The number of weights to be trained scales as $O(jjk)$.

The architecture of a generic multi-layer perceptron for the RP is shown in Fig. 3.1. There are $k+1$ layers (*i.e.* a depth of $k+1$) of which k are hidden layers, while each layer has a width of j which may vary from layer to layer. For simplicity, this thesis considers network designs where j is constant across all layers, similar to Magiera *et al.* [107] and

Gyrya *et al.* [108]. Each linear transform is of the form $\tilde{z} = \bar{\bar{U}}\bar{h} + \bar{b}$ where $\bar{\bar{U}} \in \mathbb{R}^{j \times j}$ and $\bar{z}, \bar{h}, \bar{b} \in \mathbb{R}^{j}$, which is equivalent to the more convenient form:

$$\tilde{z}_k = \tilde{\bar{\bar{U}}}_k \tilde{h}_k \tag{3.2}$$

where $\tilde{\bar{\bar{U}}} = [\bar{\bar{U}} \quad \bar{b}] \in \mathbb{R}^{j \times j+1}$ and $\tilde{h} = [\bar{h} \quad 1]^T \in \mathbb{R}^{j+1}$. For the first linear layer, $\tilde{h}_1 = \tilde{x}_1$. For the $(k+1)$th layer, $\tilde{z}_{k+1} = \hat{y}_i$. Within the k^{th} hidden layer, an element-wise non-linear activation function ϕ_k is applied:

$$\bar{h}_{kj} = \phi_k(\tilde{z}_{kj}) \tag{3.3}$$

Common activation functions are sigmoid, tanh, radial basis, and the rectified linear unit [124, 128]. To avoid potential saturation issues with vanishing gradients, a leaky rectified linear unit activation function with slope 10^{-2} is used in this thesis:

$$\phi_{LeakyReLU}(x) = \begin{cases} x & x \geq 0 \\ 0.01x & x < 0 \end{cases} \tag{3.4}$$

The term *artificial neural network*, or simply *neural network*, derives from the numerous variables and activation functions related in a highly interwoven fashion much like in animal brains. However, this similarity is more art than science—artificial neural networks ought to be treated as mathematical constructs only loosely inspired by biology.

It is also worth noting that the training dataset $(\bar{x}_i, \bar{y}_i)$ typically needs to be scaled in order to prevent saturation of activation functions, particularly when there are different orders of magnitude or units in the feature and prediction vectors. For this thesis, mean normalization is used:

$$x' = \frac{x - \mathrm{mean}(x)}{\max(x) - \min(x)} \tag{3.5}$$

where $x \in \mathbb{R}$ may be any of the feature or predicted variables, and x' is the corresponding normalized value. Alternative normalization schemes include Z-score and min-max normalization, however we know *a priori* that the data is not normally distributed, while min-max normalization rescales inputs to $[0, 1]$ which does not take advantage of the leaky rectified linear unit's nonzero outputs for negative inputs.

3.4 Loss Function and Optimizer

In addition to specifying the network design, it is necessary to determine an appropriate loss function and optimizer. The loss function is designed to take on smaller values when the weights θ yield better performance, *i.e.* lower values of some error metric of interest. Training involves executing the search algorithm, also called the optimizer, to find weights that minimize the loss function for some set of training data. For every data point $(\bar{x}_i, \bar{y}_i)$ and choice of weights θ, there is a loss $\mathcal{L}$ given by some loss function $\mathcal{L}(\bar{x}_i, \bar{y}_i; \theta)$ that represents a penalty related to poor performance. Often, the loss is taken as a mean average across the training data:

$$\mathcal{L}(\mathcal{D}; \theta) = \frac{1}{n} \sum_{(\bar{x}, \bar{y}) \in \mathcal{D}} \mathcal{L}(\bar{x}, \bar{y}; \theta) \tag{3.6}$$

where $\mathcal{D} = \{(\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2), ..., (\bar{x}_n, \bar{y}_n)\}$ is the dataset used for training. Common loss functions for regression include the L_1 and L_2 norms, Huber, as well as mean-squared error (MSE) [124, 128]. Due to the suitability of MSE for generic regression problems, it is utilized in this thesis:

$$\mathcal{L}_{MSE}(\mathcal{D};\theta) = \frac{1}{n}\sum_{(\bar{x}_i,\bar{y}_i)\in\mathcal{D}} ||\bar{y}_i - \mathcal{F}(\bar{x}_i;\theta)||^2 \tag{3.7}$$

An optimization algorithm may then be applied that searches for weights θ such that the loss function is minimized over the dataset $\mathcal{D}$. Common optimizers include stochastic gradient descent, root-mean square propagation [129], and Adam [130], among others. This thesis employs the Adam [130] optimizer, named after the algorithm's use of **ad**aptive **m**oments. By varying the learning rate via first and second gradient moments with exponential decay, Adam facilitates fast convergence even in high-dimensional non-linear search spaces such as those expected in the RP with non-ideal thermodynamics. The specific variant of the Adam algorithm used in this study is provided in Fig. 3.2.

3.5 Regularization and a Physics-Informed Approach

The universal approximation property [131, 132, 133] guarantees the existence of $\mathcal{F}$, however finding weights θ that yield a desired level of accuracy is generally an NP-hard problem. Thus, the tuning of hyperparameters (*i.e.* parameters that control the learning process), selection of appropriate train and test data, and use of regularization techniques (*e.g.* ridge regression, lasso, data augmentation, bagging, dropout, batch normalization) are additional considerations for mitigating overfitting and reducing generalization error. A balance must be struck between accuracy on the training data versus accuracy on unseen

```
Require: Learning rate $\gamma$
Require: Moment decay rates $\beta_1, \beta_2 < 1$
Require: Vector of initial weights $\bar{\theta}_0$
Require: Weight decay $\lambda$
Require: Small constant $\epsilon$ to improve numerical stability
Require: Loss function $\mathcal{L}$
  Initialize first and second moment vectors $\bar{m}_0, \bar{v}_0 \leftarrow 0$
  for t = 1 to max. epochs do
      $\bar{g}_t \leftarrow \nabla_{\bar{\theta}} \mathcal{L}(\mathcal{D}; \bar{\theta}_{t-1})$
      $\bar{g}_t \leftarrow \bar{g}_t + \lambda \bar{\theta}_{t-1}$
      $\bar{m}_t \leftarrow \beta_1 \bar{m}_{t-1} + (1 - \beta_1) \bar{g}_t$
      $\bar{v}_t \leftarrow \beta_2 \bar{v}_{t-1} + (1 - \beta_2) \bar{g}_t \cdot \bar{g}_t$
      $\hat{m}_t \leftarrow \frac{1}{(1-\beta_1^t)} \bar{m}_t$
      $\hat{v}_t \leftarrow \frac{1}{(1-\beta_2^t)} \bar{v}_t$
      $\bar{\theta}_t \leftarrow \bar{\theta}_{t-1} - \frac{\gamma \hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$        ▷ Note: operations are applied element-wise in this step
  end for
  return $\bar{\theta}_t$
```

Figure 3.2: Adam optimization algorithm with weight decay. Variables with an overhead bar or circumflex (hat) indicate vectors of length equal to the number of weights in the network. The notation used above is restricted to the algorithm definition, and should not be confused with other variables in fluid mechanics (*e.g.* v for molar specific volume).

data that may be encountered during inference.

Since the present research involves a regression task governed by known physical laws, traditional data-driven regularization techniques are eschewed in favour of a *physics-informed* approach. A physics-informed approach [105, 104] involves directly integrating physical laws into the neural network or loss function such that physical violations are penalized during training. Unlike data-driven regularization, physical constraints are free of subjective heuristical criteria for smoothness and generality of predictions. Weights that yield low losses on training data are still penalized if they violate physical laws, just as how traditional regularization penalizes weights that might yield higher losses on unseen test data. As with all regularization techniques, the physics-informed approach enables smaller training datasets to be used for accurate predictions across a wide feature space.

Physical constraints are incorporated through a CAL approach [107]. Alternatively, the network complexity itself may be increased, but this would reduce runtime efficiency. A generic CAL loss function is of the form:

$$\mathcal{L}_{CAL}(\mathcal{D};\theta) = \frac{1}{n}\sum_{(\bar{x}_i,\bar{y}_i)\in\mathcal{D}} ||\bar{y}_i - \mathcal{F}(\bar{x}_i;\theta)||^2 + \kappa|\psi(\bar{x}_i;\theta)| \tag{3.8}$$

where κ is a user-defined hyperparameter and $\psi(\bar{x}_i;\theta) \in \mathbb{R}$ is the residual calculated from a physics-informed constraint function such that $\psi = 0$ when the constraint is satisfied.

A natural set of physical constraints for the RP is furnished via the Rankine-Hugoniot jump conditions applied to the Euler equations (Eq. 2.8). Suppose there is a moving jump at $x_s(t)$, moving at speed $u_s(t)$. The situation is considered on a control volume of spatial and temporal lengths $\Delta x = x_2 - x_1$ and $\Delta t = t_2 - t_1$ such that $x_1 < x_s < x_2$. Then, a

Taylor expansion of Δx with respect to Δt yields:

$$\Delta x = u_s(t_1)\Delta t + O(\Delta t^2) \tag{3.9}$$

Similarly, applying a Taylor expansion to $\mathbf{U}(x,t)$ in the regions left and right of the jump results in:

$$\mathbf{U}(x_1 < x < x_s, t) = \mathbf{U}(x_1, t) + O(\Delta t) \tag{3.10}$$

$$\mathbf{U}(x_s < x < x_2, t) = \mathbf{U}(x_2, t) + O(\Delta t) \tag{3.11}$$

Substituting Eqs. (3.10) and (3.11) into the integral conservation laws (Eq. 2.38) gives:

$$\Delta x \mathbf{U}(x_1, t_1) = \Delta x \mathbf{U}(x_2, t_1) + \Delta t \mathbf{F}(\mathbf{U}(x_1, t_1)) - \Delta t \mathbf{F}(\mathbf{U}(x_2, t_1)) + O(\Delta t^2) \tag{3.12}$$

and substituting in Eq. (3.9) then dividing by Δt produces:

$$u_s(t_1)\big(\mathbf{U}(x_2, t_1) - \mathbf{U}(x_1, t_1)\big) = \mathbf{F}(\mathbf{U}(x_2, t_1)) - \mathbf{F}(\mathbf{U}(x_1, t_1)) + O(\Delta t) \tag{3.13}$$

Now taking the limit $\Delta t \to 0$ and generalizing t_1 to be any t, we have the Rankine-Hugoniot condition:

$$u_s = \frac{\mathbf{F}(\mathbf{U}(x_2, t)) - \mathbf{F}(\mathbf{U}(x_1, t))}{\mathbf{U}(x_2, t) - \mathbf{U}(x_1, t)} \tag{3.14}$$

and it is easily observed that the faster that the flux changes with respect to the conserved variables, the faster the jump speed—as one might intuitively expect.

The Rankine-Hugoniot constraints are conveniently prescribed at the contact discontinuity in the RP, since velocity and pressure remain constant across such waves. In terms of the variables that $\mathcal{F}$ is designed and trained to predict, it may be formally stated that $p_{*L} = p_{*R}$ and $u_{*L} = u_{*R} = u_*$ where the subscripts $(*L, *R)$ denote the conditions immediately to the left and right of the contact discontinuity. Expanding and rearranging Eq. (3.14) yields individual constraint functions for mass, momentum, and energy:

$$\psi_{mass} = u_s(\rho_{*R} - \rho_{*L}) - \rho_{*R}u_{*R} + \rho_{*L}u_{*L} \tag{3.15}$$

$$\psi_{mom} = u_s(\rho_{*R}u_{*R} - \rho_{*L}u_{*L}) - (\rho_{*R}u_{*R}^2 + p_{*R}) + (\rho_{*L}u_{*L}^2 + p_{*L}) \tag{3.16}$$

$$\psi_{energy} = u_s(E_{*R} - E_{*L}) - u_{*R}(E_{*R} + p_{*R}) + u_{*L}(E_{*L} + p_{*L}) \tag{3.17}$$

and it is straightforward to verify through direct substitution that $\psi_{mass}, \psi_{mom}, \psi_{energy} = 0$ across a contact discontinuity. Should $\mathcal{F}(\bar{x}_i; \theta)$ incorrectly predict one or more star-state conditions of interest, a nonzero physics-informed penalty is applied as follows:

$$\psi(\bar{x}_i; \theta) = \kappa_{mass}\psi_{mass}(\bar{x}_i; \theta) + \kappa_{mom}\psi_{mom}(\bar{x}_i; \theta) + \kappa_{energy}\psi_{energy}(\bar{x}_i; \theta) \tag{3.18}$$

where the $\kappa_{mass}, \kappa_{mom}, \kappa_{energy}$ terms are additional user-defined hyperparameters that weigh the relative importance of each constraint in the training process.

Furthermore, the total energy on a per unit volume basis E is neither a feature nor prediction of $\mathcal{F}$, so it is necessary to restore the original magnitudes of the variables to calculate the total energy terms. The resultant values can subsequently be normalized by some factor suitable for the training data in use. Also, the energy constraint is the most functionally complex of the three constraints, since energy calculations require the evaluation of the energy-enthalpy relationship (Eq. 2.4) and non-ideal enthalpy (Eq. 2.23), which in turn depend on several thermodynamic departure functions. This issue does not arise in existing literature, which is based upon the perfect gas assumption. But for non-ideal thermodynamics, preliminary training attempts in §6.1 revealed that omitting the energy constraint improved convergence and accuracy. The energy constraint is therefore not used in this thesis.

Chapter 4

Analytical Solutions with Non-Ideal Thermodynamics

A generalization is performed to the theory of inviscid shocks and rarefactions, yielding well-behaved exact implicit shock solutions as well as the first known exact closed-form solutions for isentropic expansion waves with non-ideal EOS. Generalized shock ratios based on the Rankine-Hugoniot equations are formulated, while a novel domain mapping is used to analytically integrate the Riemann invariant for isentropic rarefactions. It is also shown that the essential mathematical structure of the isentropic expansion wave is *constitutively* invariant—that is, does not depend on EOS. Although the present derivations are performed for stationary shocks and centred expansion waves, Anderson [29] and Collela & Glaz [36] offer guidance on constructing solutions to incident waves and the full RP.

This chapter is structured as follows: section 4.1.1 defines the stationary normal shock problem formally, and provides generic shock functions that are applicable to arbitrary EOS. In section 4.1.2, numerical results are computed and discussed for various shock ratios, using both ideal and non-ideal EOS.

This chapter is published as Wang, J. C.-H. and Hickey, J.-P. (2020): Analytical solutions to shock and expansion waves for non-ideal equations of state, *Phys. Fluids* **32**, 086105, https://doi.org/10.1063/5.0015531.

The codes that contributed to the results of this chapter are available at https://git.uwaterloo.ca/jc9wang/analytical_shocks_and_rarefactions.

4.1 Normal Shocks

4.1.1 Problem Setup and Derivation: Stationary Normal Shock

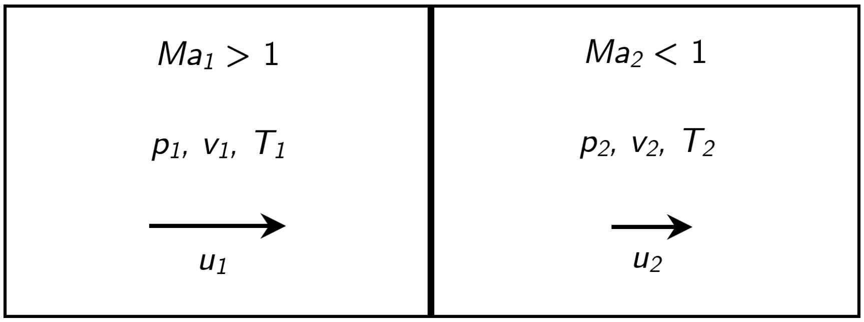


Figure 4.1: The stationary normal shock problem, with the shock fixed at $x = 0$.

The stationary normal shock (Fig. 4.1) features a centred flow discontinuity with known pre-shock conditions, indicated by the subscript 1. To be determined are the post-shock conditions, indicated by the subscript 2. For ideal gases, exact shock ratios can be derived as explicit functions of pre-shock conditions [29].

For non-ideal gases, an implicit solution is unavoidable due to the mathematical form of the EOS and thermodynamic properties dependent on two independent state variables. However, it is still possible to use the implicit equations to find the pre- and post-shock conditions such that the final result is piecewise-explicit in space and time due to inviscid shocks manifesting as step discontinuities that travel at finite wavespeeds (zero in the case of a stationary shock). Traditionally, such exact implicit solutions have been solved by iterating on the default form of Rankine-Hugoniot relations [40, 42], where a total of six variables (four pressure, two enthalpy) contain highly involved expressions comprising the non-ideal EOS or its partial derivatives and integrals.

It is possible to condense the governing equations into generalized Rankine-Hugoniot conditions comprising continuity-momentum and continuity-energy shock functions where the number of terms containing non-ideal thermodynamics is reduced. This yields better-behaved Jacobian and Hessian expressions that are conducive to finding implicit solutions. These generic shock functions are:

$$CM(v_2, T_2) = p_2 - p_1 - \frac{Mu_1^2}{v_1^2}(v_1 - v_2) \tag{4.1}$$

$$CE(v_2, T_2) = \frac{h_2}{M} - \frac{h_1}{M} - \frac{u_1^2}{2}\left(1 - \frac{v_2^2}{v_1^2}\right) \tag{4.2}$$

which contain only four variables (pressure and enthalpy of pre- and post-shock conditions) where the non-ideal thermodynamics appear. Since p is defined by the EOS $p = p(v, T)$, and the real-gas enthalpy is of the form $h = h(v, T)$, the two unknowns T_2, v_2 are found by solving for $CM = 0$ and $CE = 0$.

The set of nonlinear equations resulting from the substitution of cubic EOS and departure functions into Eqs. (4.1) and (4.2) may be readily solved using iterative methods. For this study, the Trust-Region-Dogleg nonlinear solver was used, which handles both convex and non-convex systems adequately [134].

4.1.2 Shocks Results for Non-Ideal Equations of State

Normal shock ratios for carbon dioxide gas were computed using ideal and non-ideal EOS (Figs. 4.2 and 4.3). For all cases, a constant pre-shock velocity is maintained in order to isolate the role of the EOS on the shock ratios. Each curve represents the shock ratios for a particular EOS as pre-shock conditions are varied under either constant pressure or temperature. Different pre-shock Mach numbers are achieved for the same pre-shock conditions due to the choice of EOS which affects the speed of sound. Properties of carbon-dioxide were taken from Poling *et al.* [25] and Yaws [135]. The subscript r indicates reduced values (normalized by their critical values). The iterative solutions converged to 10^{-7} error within 3 to 15 steps.

The ideal gas solutions overpredict temperature and Mach ratios but underpredict pressure and specific volume ratios relative to the non-ideal EOS. At first glance, this may appear counter-intuitive since the ideal gas law ignores intermolecular forces and thus should overpredict pressure and volume. In actuality, this result is due to this very phenomenon. Neglecting intermolecular forces resulted in pressure differences of $\mathcal{O}$(1-

6MPa), corresponding to an approximately 20% increase in the pre-shock pressures of $\mathcal{O}$(5-24MPa), but a <10% increase in the high post-shock pressures of $\mathcal{O}$(60-70MPa). This causes a lower p_2/p_1. Similar logic holds for the specific volume ratio. Also, Ma_1 was lower in the ideal gas curves due to the higher temperatures predicted by the ideal gas law. Since Ma_1 and p_2/p_1 are traditionally used as the independent variables in solving stationary and moving shocks respectively [29, 31], these graphs reveal that the choice of EOS can yield significant discrepancies in the apparent shock ratios as a function of these variables.

The tendency for ideal and non-ideal ratios to agree within a 6-10% error in low-pressure subcritical as well as high-temperature supercritical conditions in Fig. 4.2 is consistent with the recent finding [136] that the ideal gas law is valid within 10% error for $T_r > 2$ and $p_r < 6$, where intermolecular forces are negligible. It is also worth discussing why the constant-volume results (Fig. 4.2) are functions in Ma_1 and vary monotonically, while the constant-temperature results (Fig. 4.3) do not possess these properties. This peculiarity is due to the fact that cubic EOS are named as such because they are cubic in volume, but are functions of T and $T^{0.5}$. Isotherms in the p-v plane change curvature near the critical point, resulting in the shock state response curving back on itself in Fig. 4.3. By contrast, any linear combination of linear and square root functions is always concave or always convex, leading to the monotonic results of Fig. 4.2 even near the critical point.

A sensitivity analysis of the continuity-momentum and continuity-energy functions to v_2, T_2 further revealed that although variations are $\mathcal{O}$(10^7 kJ kg^{-1}) and $\mathcal{O}$(10^4 Pa) respectively, the functions are locally convex. In particular, Fig. 4.4 contains sensitivity plots generated at transcritical conditions where thermodynamic derivatives are typically the largest in magnitude. Still, the variations are smooth and convex. Thus, this form of the implicit solution lends itself well to iterative solution methods, contrary to traditional wisdom on the computational intractability of non-ideal shocks [29]. Compared to existing methods [38, 41, 42], the present solution generalizes to all pressure-explicit non-ideal EOS while reducing the number of terms containing non-ideal thermodynamics and yielding convex surfaces in the solution space.

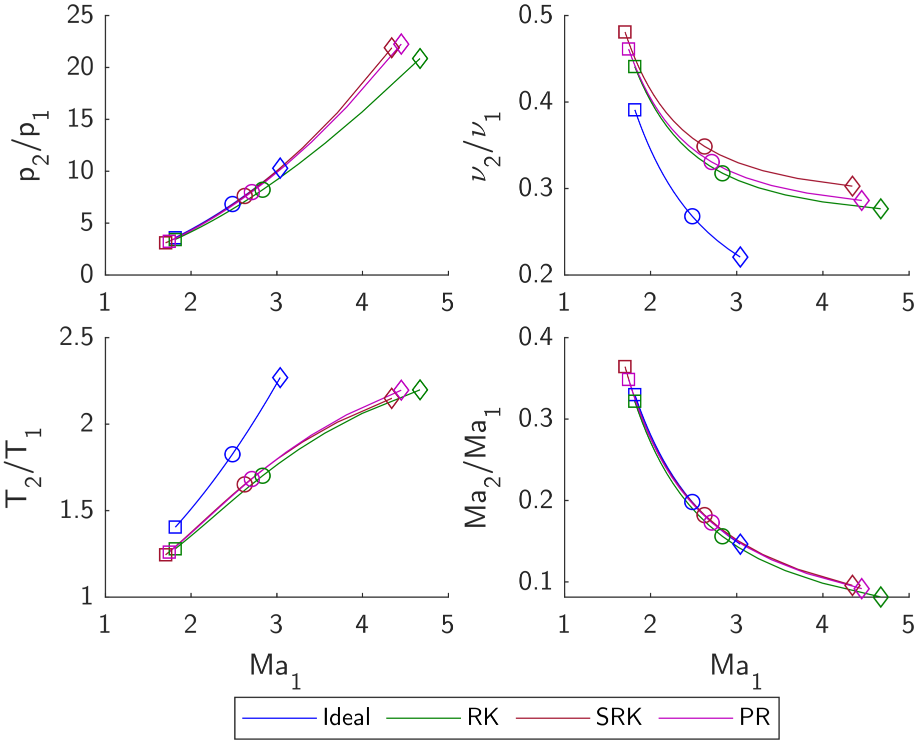


Figure 4.2: Shock ratios for carbon dioxide gas with constant v_1 = 0.293 m^{3}mol^{-1}, u_1 = 750 ms^{-1}. Diamonds compare results for a particular subcritical case ($p_{1r,PR}$ = 0.68, T_{1r} = 0.95); circles, transcritical ($p_{1r,PR}$ = 1.00, T_{1r} = 1.15); squares, supercritical ($p_{1r,PR}$ = 3.22, T_{1r} = 2.63).

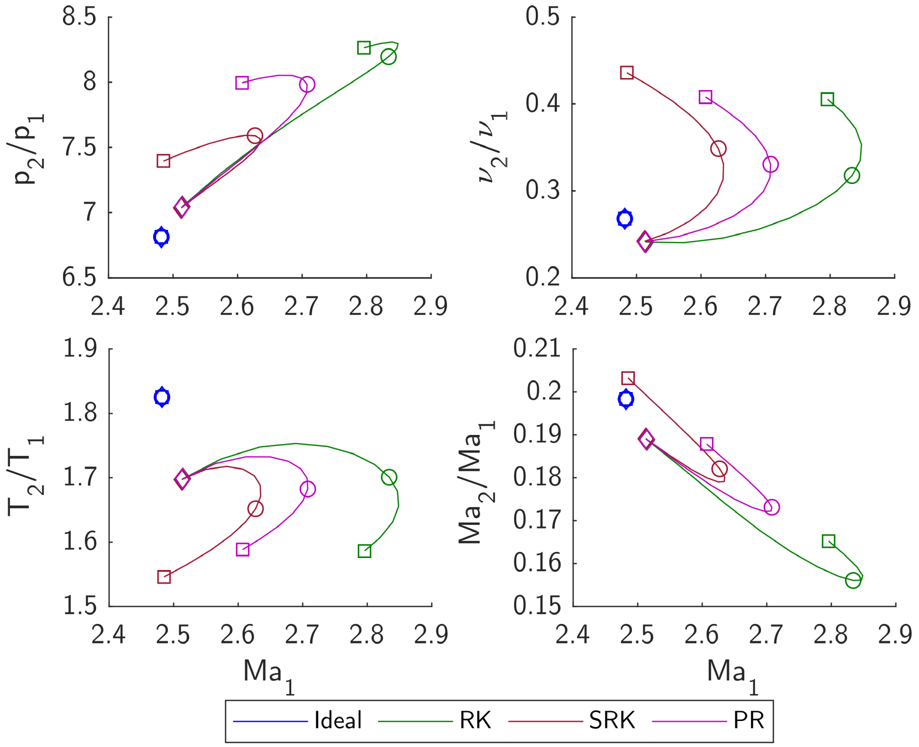


Figure 4.3: Shock ratios for carbon dioxide gas with constant T_1 = 375K, T_{1r} = 1.23, u_1 = 750 ms^{-1}. Diamonds compare results for a particular subcritical case ($p_{1r,PR}$ = 0.02); circles, transcritical ($p_{1r,PR}$ = 1.28); squares, supercritical ($p_{1r,PR}$ = 1.91).

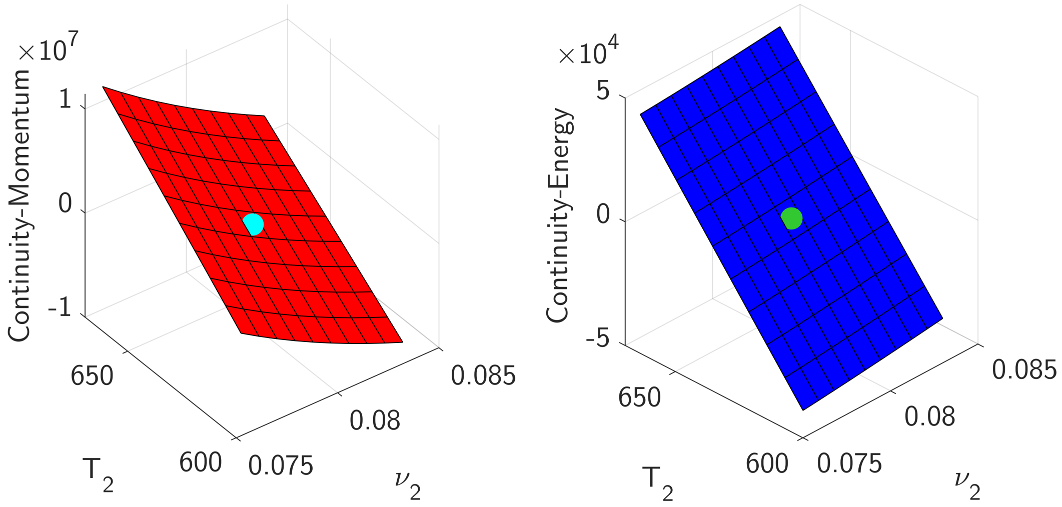


Figure 4.4: The Continuity-Momentum function [kJ kg^{-1}] and Continuity-Energy function [Pa] versus T_2 [K] and v_2 [m^{3}mol^{-1}], for a transcritical pre-shock condition ($p_{1r,PR} = 1.28$, $T_{1r} = 1.23$). The blue dot is the solution that zeros both functions.

4.2 Expansion Waves

4.2.1 Problem Setup and Derivation: Centred Expansion Wave

The centred expansion wave features a widening expansion region (Fig. 4.5) that propagates left when $p_1 > p_3$ or propagates right when $p_1 < p_3$. For the present analysis, the former case is considered; the derivations remain similar for right-propagating waves. The boundary conditions in regions ① and ③ are known and uniform, as is typically the case in a shock tube or finite difference splitting method. To be determined are the properties of the rarefaction wave region in between these bounds.

Applying the method of characteristics produces the widely known compatibility equa-

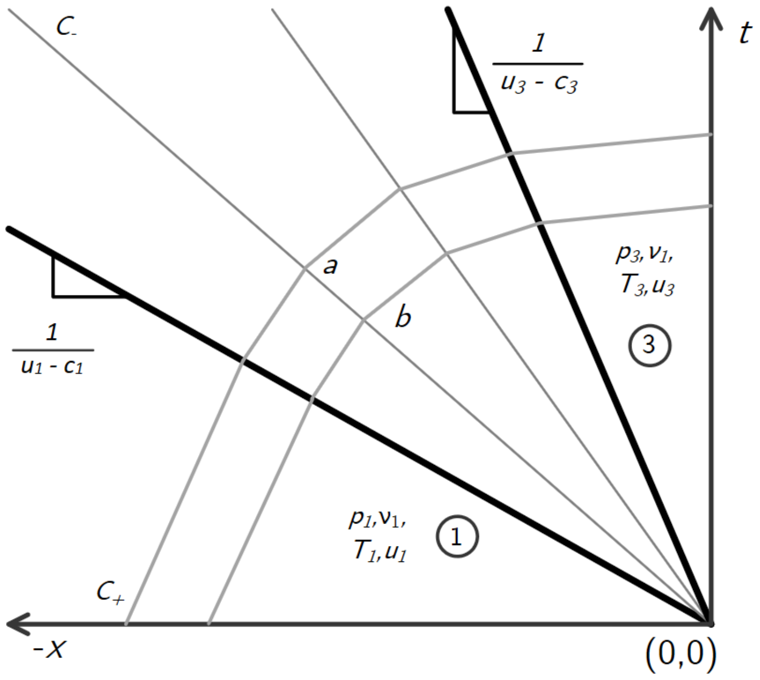


Figure 4.5: The centred expansion wave where $p_1 > p_3$. The points a, b lie along an arbitrary C_- characteristic inside the wave.

tions along $C_\pm$ curves:

$$du \pm \frac{dp}{\rho c} = 0 \iff J_\pm = u \pm \int \frac{dp}{\rho c} = constant \tag{4.3}$$

where $J_\pm$ are the corresponding Riemann invariants. Authors such as Anderson [29] have shown that for arbitrary points a, b along any C_- line:

$$J_{\pm_a} = J_{\pm_b} \tag{4.4}$$

which implies that J_+ is constant *throughout* the wave. Also, by adding or subtracting J_+

and J_- at a point:

$$u = \frac{1}{2}(J_+ + J_-) \quad , \quad \int \frac{dp}{\rho c} = \frac{1}{2}(J_+ - J_-) \tag{4.5}$$

For an ideal gas, it is possible to use polytropic relations and the ideal gas speed of sound to directly compute the integral. However, in a real gas, the ratio of specific heat capacities is not constant. The inability to analytically evaluate this integral for a real gas has traditionally prevented the derivation of exact closed-form solutions [46]. However, it will be seen that $u = \frac{1}{2}(J_+ + J_-)$ and a novel domain mapping approach are sufficient to complete the derivation. This strategy is inspired by the Riemann solver of Osher [50], which attempts to identify generalized Riemann invariants and admissible integration paths in phase space.

Eqs. (4.4) and (4.5) prove that C_- characteristics are lines of constant velocity u. According to Eq. (4.3), this implies $dp = 0$, which in an isentropic flow ($dp = c^2 d\rho$) further implies $d\rho = 0$. Because two state variables p, ρ are constant, there must be constant T and $c(v, T)$. Thus, the C_- characteristics represent straight, constant-property lines passing through the origin in the xt plane. Interestingly, this result is independent of the chosen EOS and therefore holds for any state equation: the essential mathematical structure of the rarefaction is an artefact of mass continuity, momentum conservation, and constant entropy. Although EOS properties may be used to prove it, such as in traditional ideal gas derivations [29], this is not necessary. The role of the EOS is only to influence the speed at which the head and tail propagate by way of c_1 and c_3.

To ultimately obtain v, p, u as functions of x and t, Eq. (4.3) will be reformulated, then v, p, u will be solved in the (C_-, t) domain and mapped to (x, t). This is possible because every point (x,t) in the rarefaction region can be equivalently specified by a choice of (C_-,t). Finally, T is found via the EOS.

Substituting $dp = c^2 d\rho$ and $d\rho = -\frac{M}{v^2}dv$ into Eq. (4.3) gives the following for $C_\pm$ lines:

$$du \mp \frac{c}{v}dv = 0 \tag{4.6}$$

Now viewing the problem in the (C_-,t) domain, it is possible to partially integrate Eq. (4.6) along each characteristic to produce J_- in terms of $\frac{x}{t}$, *i.e.* the slope of each C_-. Since speed of sound is constant along each C_-:

$$J_- = u + c\int_{C_-}\frac{dv}{v} = u + c\ln v \tag{4.7}$$

Substituting $u = \frac{x}{t} + c$ corresponding to C_-, replacing J_- with J_+ via $u = \frac{1}{2}(J_+ + J_-)$ from Eq. (4.5), and rearranging yields $\ln v = \frac{1}{c}(\frac{x}{t} - J_+ + c)$. But, the choice of C_- and hence $\frac{x}{t}$ is arbitrary. Thus, v is of the form:

$$v(x,t) = \exp\left(A_1\frac{x}{t} - A_2\right) \tag{4.8}$$

where A_1, A_2 are constants. We have successfully mapped $v(C_-,t) \to v(x,t)$.

To find p and u, the problem is considered in (C_-,t) and one can partially integrate Eq. (4.3):

$$J_- = u - \frac{1}{\rho c}\int_{C_-} dp = u - \frac{pv}{Mc} \tag{4.9}$$

which implies that the product $pv = f(\frac{x}{t})$. Knowing v:

$$p(x,t) = \exp\Big(- A_3 \frac{x}{t} - A_4 \Big) \tag{4.10}$$

Also, substituting Eq. (4.8) into Eq. (4.7) and utilizing $u = \frac{1}{2}(J_+ + J_-)$ again produces u:

$$u(x,t) = A_5 + A_6 \frac{x}{t} \tag{4.11}$$

where A_3, A_4, A_5, A_6 are constants.

The constants A_i can be determined by applying the head and tail boundary conditions at x_1, x_3 where $x_1 = (u_1 - c_1)t$ and $x_3 = (u_3 - c_3)t$. Their full expressions are provided below:

$$A_1 = \frac{\ln(\frac{v_1}{v_3})}{u_1 - c_1 - u_3 + c_3} \tag{4.12}$$

$$A_2 = \frac{u_1 - c_1 + u_3 - c_3}{2(u_1 - c_1 - u_3 + c_3)} \ln(\frac{v_1}{v_3}) - \frac{1}{2} \ln v_1 v_3 \tag{4.13}$$

$$A_3 = \frac{-\ln(\frac{p_1}{p_3})}{u_3 - c_3 - u_1 + c_1} \tag{4.14}$$

$$A_4 = \frac{u_1 - c_1 + u_3 - c_3}{2(u_1 - c_1 - u_3 + c_3)} \ln(\frac{p_1}{p_3}) - \frac{1}{2} \ln p_1 p_3 \tag{4.15}$$

$$A_5 = \frac{1}{2}(u_1 + u_3) - \frac{(u_1 - u_3)(u_1 - c_1 + u_3 - c_3)}{2(u_1 - c_1 - u_3 + c_3)} \tag{4.16}$$

$$A_6 = \frac{u_1 - u_3}{u_1 - c_1 - u_3 + c_3} \tag{4.17}$$

Clearly, these constants act to stretch and shift the general shape of the solutions, which have an exponential form in the case of p, v or linear form in the case of u. Governing equations therefore enforce solutions to p, v, u, while solving the EOS $p = p(v, T)$ permits computation of T from p, v.

This generalized interpretation is consistent with ideal gas solutions such as those by Anderson [29]. In ideal gases, the integral in Eq. (4.5) is evaluated analytically and it is unnecessary to perform the domain mapping and partial integration that gives p, v as exponential functions and A_i as constants. Instead, flow speed is classically found to be: $u(x, t) = \frac{2}{\gamma+1}\left(c_1 + \frac{x}{t}\right)$ where p, v, T are powers of linear functions of $u(x, t)$. The non-(x, t) coefficients and terms in the ideal gas solutions are simply constants composed of γ and u, c at either the head or tail (head and tail conditions are equivalent due to polytropic relations). Thus, in both the ideal and general case, an analytical solution is found by relating these coefficients and terms to the boundary conditions.

4.2.2 Expansion Wave Results for Non-Ideal Equations of State

The speed of sound in cubic EOS, found via Eq. (2.25), is used to calculate $x_1 = (u_1 - c_1)t$ and $x_3 = (u_3 - c_3)t$. p, v, u is solved by applying the head and tail boundary conditions to Eqs. (4.8), (4.10), and (4.11) at a given time t. A nonlinear solver such as the Trust-Region-Dogleg method [134] may be used to find T via the EOS. Results for a subcritical region ① are provided in Fig. 4.6 and the curves are identical, as expected.

Expansion waves in transcritical and supercritical carbon dioxide gas were computed using ideal and non-ideal EOS (Figs. 4.7 and 4.8). u_1 was set to zero as is the case in shock tube experiments. v and T were independently specified at the left boundary, and p was found with the EOS. On the right boundary, T was specified, ideal gas polytropic relations were used to specify v, and p was again found with the EOS. This was done only to ensure the boundary v, T would be identical across ideal and non-ideal results for the purpose of comparison, however in real problems this may not be the case.

At transcritical head conditions, the ideal EOS overpredicts pressure, and the differences in head and tail speed of sound cause overestimates of 0.5 MPa to 2 MPa or $\sim 20\%$ more than the corresponding non-ideal values at the same point along x. At supercritical values, the ideal non-ideal results converge to nearly identical curves except with pressure and specific volume differences of 15% near the centre of the rarefaction wave. This once again supports the notion [136] that ideal EOS may be acceptable under certain high-temperature and low-pressure supercritical conditions.

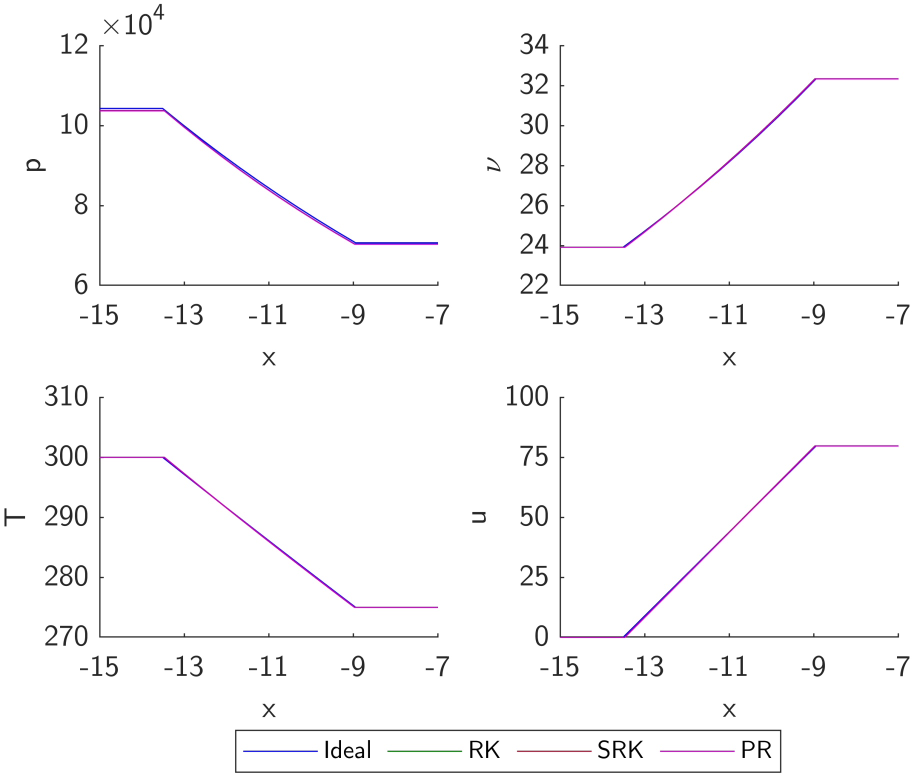


Figure 4.6: Pressure [Pa], specific volume [m^{3}mol^{-1}], temperature [K], and flow speed [m s^{-1}], versus x [m] at time t=0.05s for a subcritical head ($p_{1r,PR}$=0.014, T_{1r}=0.99).

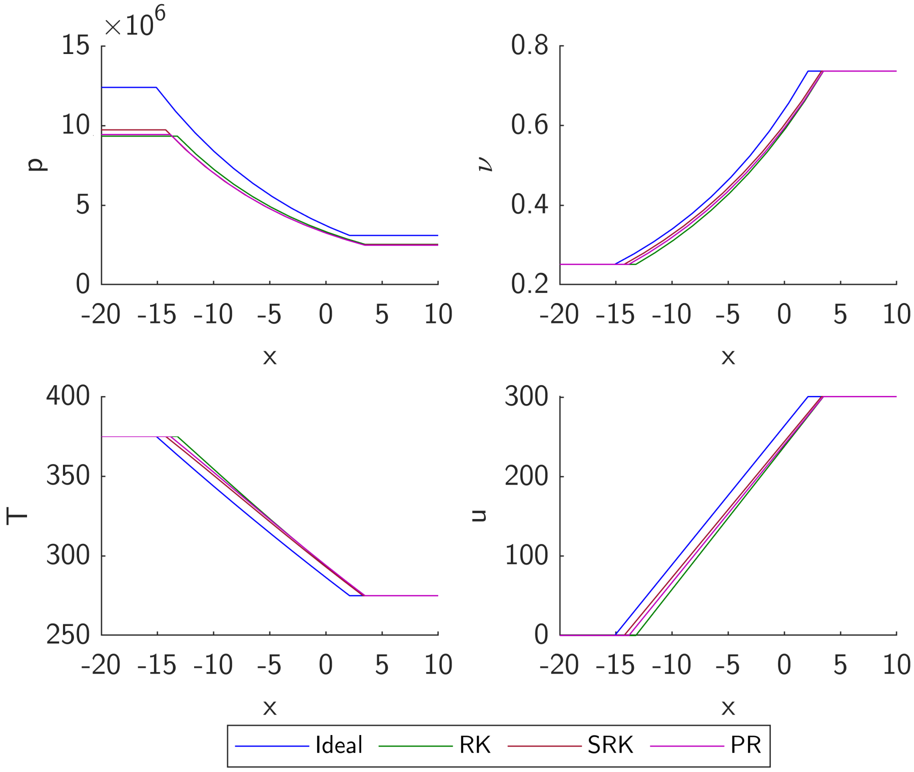


Figure 4.7: Pressure [Pa], specific volume [m^{3}mol^{-1}], temperature [K], and flow speed [m s^{-1}], versus x [m] at time t=0.05 s for a transcritical head ($p_{1r,PR}$=1.28, T_{1r}=1.23)

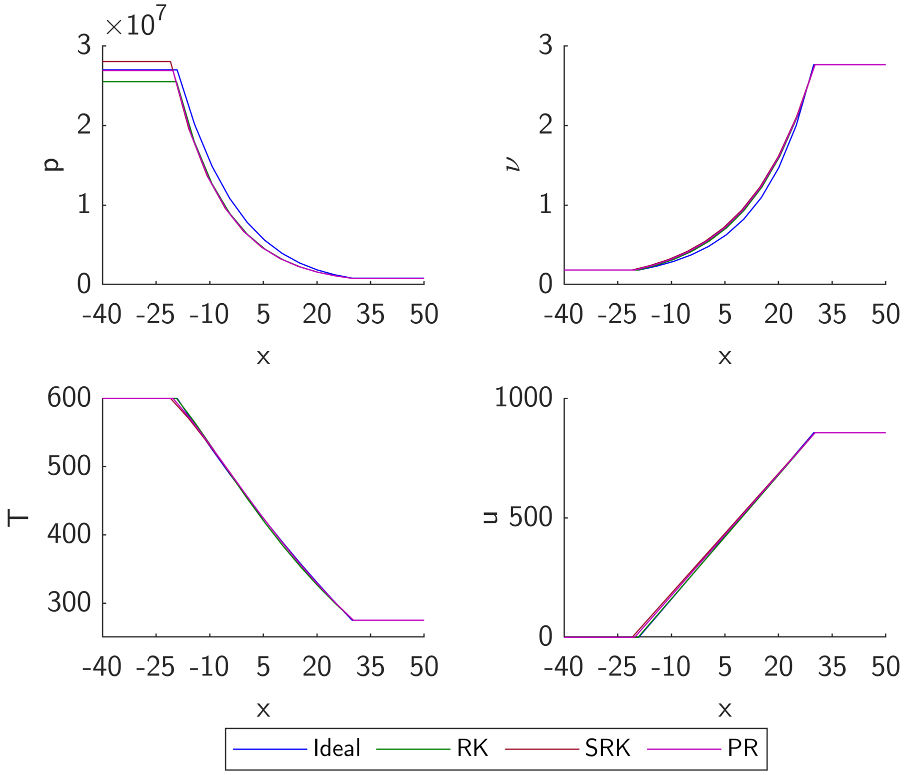


Figure 4.8: Pressure [Pa], specific volume [m^{3}mol^{-1}], temperature [K], and flow speed [m s^{-1}], versus x [m] at time t=0.05 s for a supercritical head ($p_{1r,PR}$=3.65, T_{1r}=1.97)

Chapter 5

Structurally Complete Approximate Riemann Solvers (StARS)

The idea of structurally complete approximate Riemann solver (StARS) is defined and explored, which use recent derivations by Wang & Hickey [137] to analytically restore the expansion wave in pre-existing three-wave solvers. By *structurally complete* and *approximate*, it is meant that StARS provides explicit non-iterative means to compute:

1. wave speeds associated with the method of characteristics, *i.e.* normal shocks, contact discontinuities, rarefaction heads, and rarefaction tails; and

2. primitive and conservative variables as well as fluxes in each region between these waves—in particular, expansion waves are analytically reconstructed and not approximated as a constant or interpolated state.

The result is a class of efficient entropy-stable approximate solvers that offer improved accuracy, the benefits of which are especially apparent under transonic flux conditions. Most importantly, this property of structural completeness is valid for both ideal and non-ideal thermodynamics. In this chapter, the Roe solver [49], whose entropy properties have been widely studied, is compared to a structurally complete version of the Roe solver (dubbed Roe-StARS) across compressible flow test cases where entropy violations arise. Also performed is a comprehensive scaling analysis of flow conditions that give rise to transonic rarefactions, yielding a clear conceptual understanding of the thermodynamic and flow conditions in which such rarefactions occur. It is shown that transonic fluxes are particularly prevalent in trans- and supercritical flows with large thermophysical gradients.

The chapter is organized as follows: section 5.1 describes the general approach in restoring the expansion wave to an arbitrary three-wave solver, so that transonic fluxes are correctly accounted for. Section 5.2 conducts a scaling analysis of the flow conditions under which transonic fluxes occur and the errors that can arise if they are omitted. Finally, section 5.3 provides numerical results for a transcritical shock tube, shock tube with periodic bounds resulting in interfering shocks and rarefactions, a so-called gradient RP, and a two-dimensional RP.

This chapter is published as Wang, J. C.-H. and Hickey, J.-P. (2022): A class of structurally complete approximate Riemann solvers for trans- and supercritical flows with large gradients, *J. Comput. Phys.* **468**, 111521, https://doi.org/10.1016/j.jcp.2022.111521

The codes that contributed to the results of this chapter are available at https://git.

uwaterloo.ca/jc9wang/structurally-complete-riemann.

5.1 Restoring the rarefaction wave

5.1.1 Detecting the presence of a rarefaction at the cell interface

Fig. 1.2 depicts the rarefaction-contact-shock solution configuration that is often shown in textbooks and papers, although rarefactions and/or shocks can occur on the left, right, or both flanks of the star-state region (outlined in blue). The speed of various characteristic waves are denoted as S followed by the appropriate subscript. Pressure and velocity are uniform throughout the star-state region, and there is always a contact discontinuity wave located within the star-state region. Of interest are the conditions under which the rarefaction head and tail sit on opposite ends of the cell interface at $x = 0$, denoted as a *transonic* state as defined in the introduction. It is assumed that the properties of the left and right star-state regions are available through the choice of a pre-existing three-wave Riemann solver.

Classical derivations of the expansion wave structure assume an ideal and calorically perfect gas [29], however Wang & Hickey [137] recently proved that this wave structure is common to both ideal and non-ideal thermodynamics. Specifically, left-running characteristics in left rarefactions and right-running characteristics in right rarefactions are always straight lines through the origin, irrespective of the choice of state equation. Riemann invariants along the characteristics that pass through the head and tail of a simple wave are constant throughout the simple wave. These universal properties justify the validity of the expansion wave diagram whether cubic, virial, or other non-ideal state equations are used in conjunction with thermodynamic properties that depend on two state variables.

Left and right rarefactions require different treatments. Left rarefactions occur when $p_L > p_*$ as drawn in Fig. 1.2, whereas right rarefactions occur when $p_R > p_*$ (representing a symmetry about $x = 0$ in 1.2). The trivial case of no rarefactions occurs only when p_L and p_R are lower than p_*, in which case the discussion is moot. For simplicity, the remaining derivations assume a left rarefaction, however the derivation for right rarefactions is in Appendix D.

For left rarefactions, the characteristics along $dx/dt = u - c$ follow straight lines, $x/t = u - c$, that pass through the origin. Thus, the position of the head and tail may be determined as:

$$x_L = S_L t = (u_L - c_L)t \tag{5.1}$$

$$x_{*L} = S_{*L} t = (u_* - c_{*L})t \tag{5.2}$$

where L indicates the initial properties of the left cell and $*L$ indicates the properties of the left star-state region. Therefore, the cell interface at $x = 0$ is enclosed by a left rarefaction when $p_L > p_*$, $S_L < 0$, and $S_{*L} > 0$. Similarly, the cell interface is inside a right rarefaction when $p_R > p_*$, $S_R > 0$, and $S_{*R} < 0$. Evidently, the expansion head is subsonic, the expansion tail is supersonic, and such rarefactions are therefore transonic.

5.1.2 Computing the flux of a transonic rarefaction

Wang & Hickey [137] also proved that inside expansion waves, the primitive variables v, p, u obey generalized analytical expressions of the form in (4.8), (4.10), (4.11) which include constants A_i determined by substituting the known conditions at the head and tail of the expansion wave. Since the boundary conditions are indeed known from the pre-existing three-wave solution, and since $x = 0$ defines the cell interface, these analytical expressions can be recast for the purposes of computing a transonic flux at $t > 0$:

$$v(0,t) = \exp\left(\frac{1}{2}\ln v_L v_{*L} - \frac{u_L - c_L + u_* - c_{*L}}{2(u_L - c_L - u_* + c_{*L})}\ln\frac{v_L}{v_{*L}}\right) \tag{5.3}$$

$$p(0,t) = \exp\left(\frac{1}{2}\ln p_L p_* - \frac{u_L - c_L + u_* - c_{*L}}{2(u_L - c_L - u_* + c_{*L})}\ln\frac{p_L}{p_*}\right) \tag{5.4}$$

$$u(0,t) = \frac{1}{2}(u_L + u_*) - \frac{(u_L - u_*)(u_L - c_L + u_* - c_{*L})}{2(u_L - c_L - u_* + c_{*L})} \tag{5.5}$$

As a sanity check, the properties at the cell interface should remain constant for all $t > 0$ since properties along straight-line characteristics are constant and $x = 0$ defines a vertical straight-line characteristic within the transonic rarefaction. As expected, the expressions do not depend on time. Similar expressions for right transonic rarefactions are provided in Appendix D. Thus, by calculating the expansion tail properties through an existing three-wave solver, it is possible to analytically solve for the primitive variables at the cell interface when a transonic rarefaction is present. Finally, the flux vector may be computed in the usual manner as:

$$\mathbf{F}_{Roe-StARS} = \begin{bmatrix} \rho u \\ p + \rho u^2 \\ u(E+p) \end{bmatrix}_{(x,t)=(0,t)} \tag{5.6}$$

substituting in the values of ρ, u, p at the interface. Compared to the original three-wave solver, the StARS version consists of two steps: a check that determines whether the wavespeeds imply a transonic flux, and if so, the computation of (5.3) through (5.5) to determine the flux. This is of comparable computational expense to the Harten and Hyman [1] entropy fix. A caveat to this analytical restoration is that while the structure of the expansion wave is exact given the head and tail conditions, the tail conditions are still approximate—therefore, like other entropy fixes, entropy-violating discontinuities such as expansion shocks are not fully removed, they are only mitigated. Additionally, in the event that a transonic flux is not present, then the original Riemann solver solution is used.

5.2 Scaling Analysis

Here we perform a scaling analysis on the errors when transonic rarefactions are omitted from the Riemann solver. For demonstrative purposes, the Roe solver [49] and the Roe-StARS version are analyzed with nitrogen gas as the working fluid. The approach may be trivially extended to other three-wave solvers and media of interest.

5.2.1 Flow conditions causing transonic rarefactions in the Riemann solver

Earlier we showed that in order for a rarefaction to occur at the cell interface, the expansion head must be subsonic and the expansion tail must be supersonic. Head properties are based on the initial conditions of the nearby cell, so it is necessary that the magnitude of the Mach number obeys:

$$|Ma| = \frac{|u|}{c} < 1 \tag{5.7}$$

for transonic rarefactions. Additionally, for a supersonic tail, it is necessary that the magnitude of the nearest star-state region's Mach number:

$$|Ma_{*K}| = \frac{|u_{*K}|}{c_{*K}} > 1 \tag{5.8}$$

where $K = \{L, R\}$ for left or right rarefactions, respectively. In this context, the nearest star-state region refers to the subset of the star-state region (highlighted in blue in Fig. 1.2) bounded by S_* and S_{*K}. Ma_{*K} can be evaluated with the given solver's estimates for the star-state region and the speed of sound equation defined in (2.25).

In the case of Roe, the Roe averages provide an approximation to the conditions in the star-state region. A version of the Roe solver [49] modified for real gases is provided in Appendix E. By substituting the applicable Roe averages, the supersonic inequality above becomes:

$$\widetilde{Ma} = \frac{1}{\tilde{c}} \left| \frac{\sqrt{\rho_L} u_L + \sqrt{\rho_R} u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \right| = \frac{1}{\tilde{c}} \left| \frac{u_L + \sqrt{\frac{\rho_R}{\rho_L}} u_R}{1 + \sqrt{\frac{\rho_R}{\rho_L}}} \right| > 1 \tag{5.9}$$

where $\tilde{c} = c(\tilde{v}, \tilde{T})$ is the Roe-averaged non-ideal gas speed of sound. It is possible to compute $\widetilde{Ma}$ directly from the known left and right states, however to fully grasp the nature of these variations, it is insightful to recast the equation as a first-order approximation of flow variables, derivatives, and discretization parameters. To first-order accuracy, (5.9) may be expanded as:

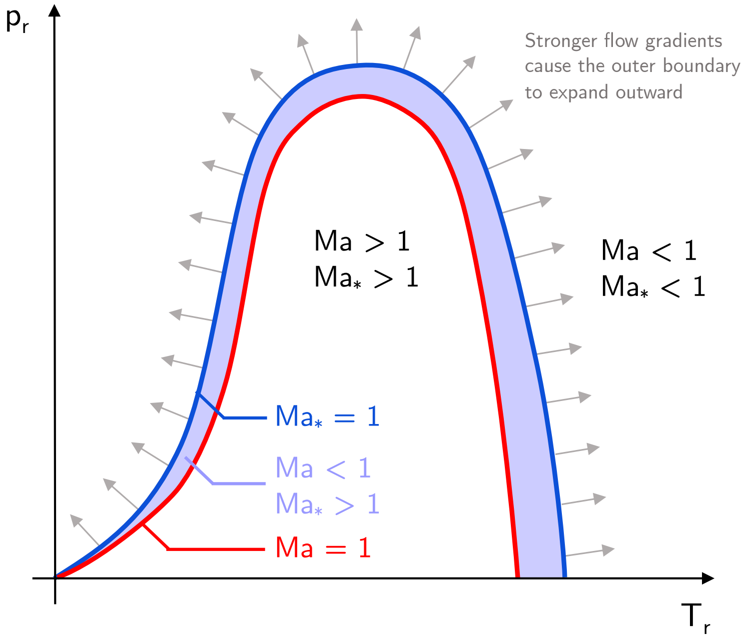
$$\widetilde{Ma} = \frac{1}{\tilde{c}} \left| \frac{u + \left(\sqrt{1 + \frac{1}{\rho}\frac{d\rho}{dx}\Delta x}\right)(u + \frac{du}{dx}\Delta x)}{1 + \sqrt{1 + \frac{1}{\rho}\frac{d\rho}{dx}\Delta x}} \right| > 1 \tag{5.10}$$

where $u_L = u$, $u_R \approx u + \frac{du}{dx}\Delta x$, $\rho_L = \rho$, and $\rho_R \approx \rho + \frac{d\rho}{dx}\Delta x$, $\rho \neq 0$. The expansion of $\tilde{c}$ in terms of primitives is avoided for tractability and, as we shall see shortly, does not hinder the analysis. Rearranging (5.10) and assuming $u \neq 0$:

$$\widetilde{Ma} = \frac{|u|}{\tilde{c}} \left| \frac{1 + (1 + \frac{1}{u}\frac{du}{dx}\Delta x)\sqrt{1 + \frac{1}{\rho}\frac{d\rho}{dx}\Delta x}}{1 + \sqrt{1 + \frac{1}{\rho}\frac{d\rho}{dx}\Delta x}} \right| > 1 \tag{5.11}$$

Thus, $\widetilde{Ma}$ varies primarily as a function of the Roe-averaged speed of sound, velocity, density, and the spatial derivatives of velocity and density which, in turn, depend on other flow variables linked through the governing equations.

We can now paint a clear conceptual picture of the thermodynamic region, $\mathcal{D}$, in the reduced temperature/pressure plot $T_r \times p_r$, within which transonic fluxes occur (Fig. 5.1) for a given set of hydrodynamic flow conditions. The ratio $|u|/\tilde{c}$ from (5.11) may be interpreted as a close approximation of the Mach number magnitude $|Ma| = |u|/c$ especially as $\Delta x \to 0$ in fine grids. This means that the sonic curve $|\widetilde{Ma}| = 1$ may be approximated as the sonic curve $|Ma| = 1$ multiplied by a positive factor equal to the large fraction in (5.11). In the remainder of this paper, this factor is referred to as the *stretching* factor due its role in determining the area between $|Ma| = 1$ and $|\widetilde{Ma}| = 1$ that defines $\mathcal{D}$.



p_r
Stronger flow gradients cause the outer boundary to expand outward
Ma > 1
Ma* > 1
Ma < 1
Ma* < 1
Ma* = 1
Ma < 1
Ma* > 1
Ma = 1
T_r

Figure 5.1: Conceptual diagram of the thermodynamic state space $\mathcal{D}$ where transonic rarefactions occur, shown in the shaded area. p_r, T_r are the reduced pressure and temperature. In between the sonic curves lies $\mathcal{D}$. As flow gradients increase in magnitude, the blue curve corresponding to $Ma_* = 1$ stretches further, causing $\mathcal{D}$ to grow and cover more of the thermodynamic state space.

When flow gradients are weak, the stretching factor is small and $\mathcal{D}$ therefore occupies a relatively small area in the thermodynamic state space. As the magnitudes of flow gradients increase, the stretching factor increases in magnitude, causing $\widetilde{Ma} = 1$ to widen in all directions and therefore resulting in a larger domain $\mathcal{D}$ where transonic fluxes—and hence entropy fixes—should be considered. These differences are exemplified in a) and b) of Fig. 5.2, which correspond to the following small- and large-gradient conditions:

$$\begin{bmatrix} dp/dx \\ dT/dx \\ du/dx \end{bmatrix}_{small} = \begin{bmatrix} -1 \text{ MPa m}^{-1} \\ -1\times 10^{2} \text{ K m}^{-1} \\ 1\times 10^{3} s^{-1} \end{bmatrix}; \quad \begin{bmatrix} dp/dx \\ dT/dx \\ du/dx \end{bmatrix}_{large} = \begin{bmatrix} -1\times 10^{1} \text{ MPa m}^{-1} \\ -1\times 10^{3} \text{ K m}^{-1} \\ 1\times 10^{4} s^{-1} \end{bmatrix} \tag{5.12}$$

with $\Delta x = 5\times 10^{-3}$ m and $u = 300$ m s^{-1}.

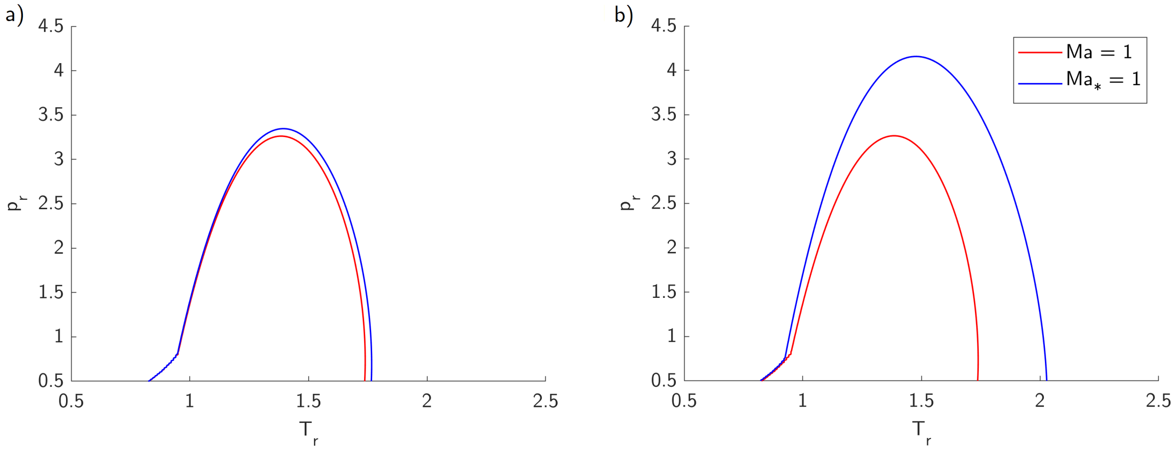


Figure 5.2: Sonic Mach number $Ma = 1$ and sonic star-state region Mach number $Ma_* \approx \widetilde{Ma} = 1$ plotted with respect to thermodynamic state, computed for nitrogen gas under different conditions: a) small-gradient case $dp/dx = -1$ MPa m^{-1}, $dT/dx = -100$ K m^{-1}, $du/dx = 1000s^{-1}$, $u = 300ms^{-1}$, $\Delta x = 5\times 10^{-3}m$; b) large-gradient case $dp/dx = -10$ MPa m^{-1}, $dT/dx = -1{,}000$ K m^{-1}, $du/dx = 10{,}000$ s^{-1}, $u = 300$ m s^{-1}, $\Delta x = 5\times 10^{-3}$ m. The figures demonstrate the effect of flow gradients on the size of the region in which transonic fluxes can occur.

There is an important nuance to these trends: at small density gradients where the velocity gradient is positive and density gradient is negative, the stretching factor can grow as gradients continue to decrease. This can result in seemingly peculiar behaviour, such as nonphysical shock magnitudes that increase by a small amount as gradients become smaller (see results in 5.3.2).

It is also worth highlighting a few observations on the nature of the sonic curve $|Ma| = 1$. Since Ma is a function only of flow speed and thermodynamic state, we should not expect the $|Ma| = 1$ curve to change under different spatial flow gradients. Indeed, $|Ma| = 1$ is identical between Figs. 5.2 a) and b). The shape of the sonic curves is also heavily influenced by the speed of sound c (Fig. 5.3). The left leg of the sonic curve closely follows the Widom pseudo-boiling line, where c is known to attain a local minimum [30]. The right leg of the sonic curve is determined by how quickly the fluid's speed of sound increases from the local minimum as pressure and temperature continue to increase.

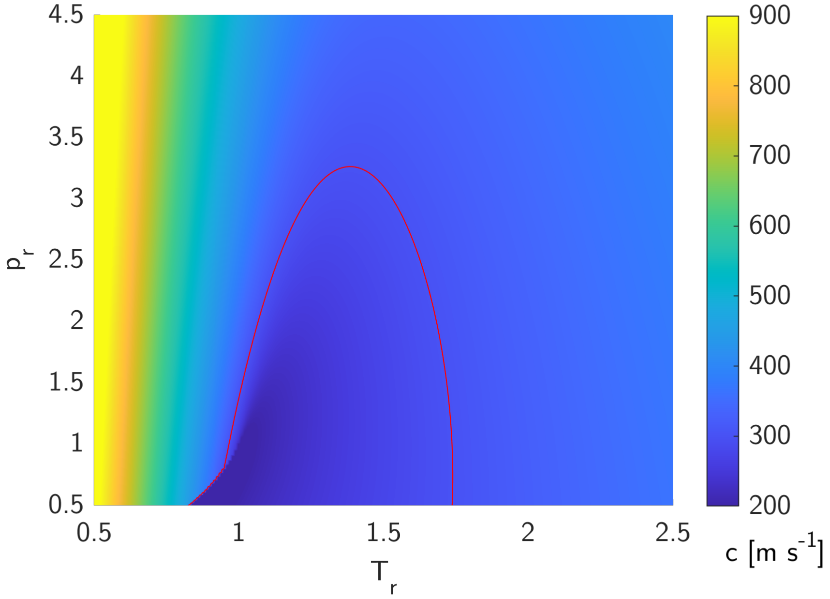


Figure 5.3: Speed of sound as a function of thermodynamic state. The sonic curve $|Ma| = 1$ is plotted along the red curve for $u = 300$ ms^{-1}. Values below 200 or above 900 ms^{-1} were truncated to the nearest colour to more clearly highlight visualize near the sonic curve.

5.2.2 Errors due to omitting the rarefaction

When approximate Riemann solvers linearize the time-dependent Euler equations, the rarefaction collapses into a single jump discontinuity that lies somewhere within the region of the rarefaction in the exact solution. The cell interface is then situated in either the nearest initial condition or the nearest star-state region. Whether it is the initial condition or the star-state region that determines the flux depends on both the flow conditions and where the linearized jump discontinuity falls within the rarefaction. For simplicity, we will consider the error scaling with respect to the star-state region although the behaviour would be similar with respect to the nearest initial condition. Thus, the magnitude of the error in the primitive variables for omitting a left transonic rarefaction is:

$$|\Delta v| = \Big|v(0,t) - v_{*L}\Big| = \Big|\exp\left(\frac{1}{2}\ln v_L v_{*L} - \frac{u_L - c_L + u_* - c_{*L}}{2(u_L - c_L - u_* + c_{*L})}\ln\frac{v_L}{v_{*L}}\right) - v_{*L}\Big| \tag{5.13}$$

$$|\Delta p| = \Big|p(0,t) - p_{*L}\Big| = \Big|\exp\left(\frac{1}{2}\ln p_L p_* - \frac{u_L - c_L + u_* - c_{*L}}{2(u_L - c_L - u_* + c_{*L})}\ln\frac{p_L}{p_*}\right) - p_{*L}\Big| \tag{5.14}$$

$$|\Delta u| = \Big|u(0,t) - u_*\Big| = \Big|\frac{1}{2}(u_L + u_*) - \frac{(u_L - u_*)(u_L - c_L + u_* - c_{*L})}{2(u_L - c_L - u_* + c_{*L})} - u_*\Big| \tag{5.15}$$

Similar expressions exist for right transonic rarefactions. These expressions are highly nonlinear functions and therefore it is less straightforward to perform a first-order analysis as with the sonic curves of the previous subsection. Instead, it is easier to imagine how this truncation error would vary depending on the proximity to the $|\widetilde{Ma}| = 1$ curve within the $\mathcal{D}$ domain. The closer that the conditions are to $|\widetilde{Ma}| = 1$, the closer the tail of the rarefaction is to the cell interface $x = 0$ and therefore the smaller the truncation error.

Conversely, the errors tend to increase towards $|Ma| = 1$.

In addition to the errors in the primitives, it is useful to compute the flux error:

$$|\Delta\mathbf{F}| = \left|\mathbf{F}_{Roe} - \mathbf{F}_{Roe-StARS}\right| \tag{5.16}$$

which can also be plotted and analyzed in terms of its three vector components. The flux components are functions of the primitives and therefore analysis by inspection is less straightforward. It is sufficient to note that the flux components comprise products and sums of the primitives, and therefore the errors in the primitives are magnified in the flux vector. Even relatively small errors in the primitives can yield order-of-magnitude greater errors in the fluxes. This observation is particularly noteworthy given that Riemann solvers ultimately output a flux, and not a set of primitives, for use in the finite volume method. The primitives provide additional insight into the underlying variables that affect the flux estimate.

Figs. 5.4 and 5.5 show the primitive and flux errors for two sets of flow conditions corresponding to the small-gradient and large-gradient cases from Fig. 5.2. As predicted by the error analysis above, errors under small-gradient conditions are relatively small, while in the large-gradient case they are larger in magnitude and occupy a greater portion of the thermodynamic state space. It is prudent to emphasize that these plots show errors on a *per-cell* basis computed at a single point in time, and only in transonic regions of the flow. Over the course of millions of Riemann solver evaluations in a flow simulation, errors have the potential to accumulate and propagate throughout the spatio-temporal computational domain. The question of producing entropy-stable solutions with proper

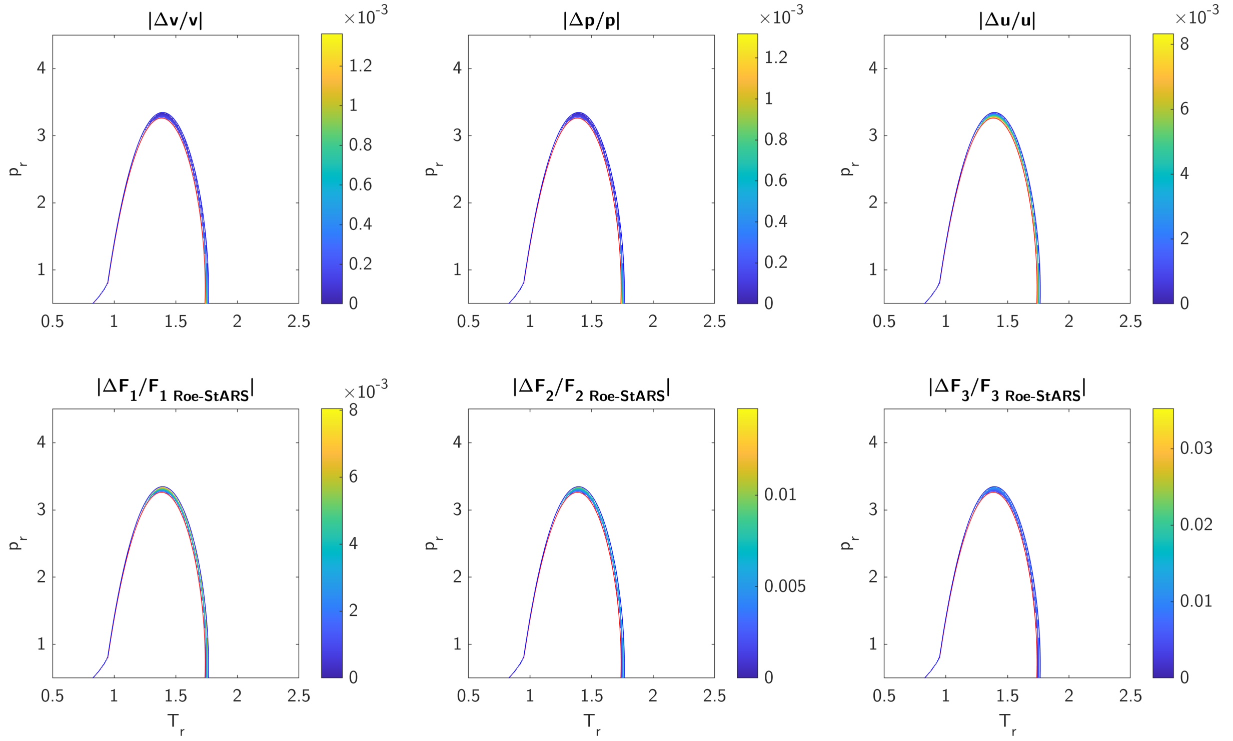


Figure 5.4: Relative errors in the primitive and flux vector components between Roe and Roe-StARS, computed for N_2 gas under the small-gradient conditions of Fig. 5.2.a). As predicted by the theory, there is a relatively narrow region where transonic fluxes occur, and the errors are relatively low in magnitude compared to Fig. 5.5. The sonic curves $Ma = 1$ and $Ma_* = 1$ are indicated with the red and blue curves. F_1, F_2, F_3 are the mass flux, momentum flux, and energy flux, respectively.

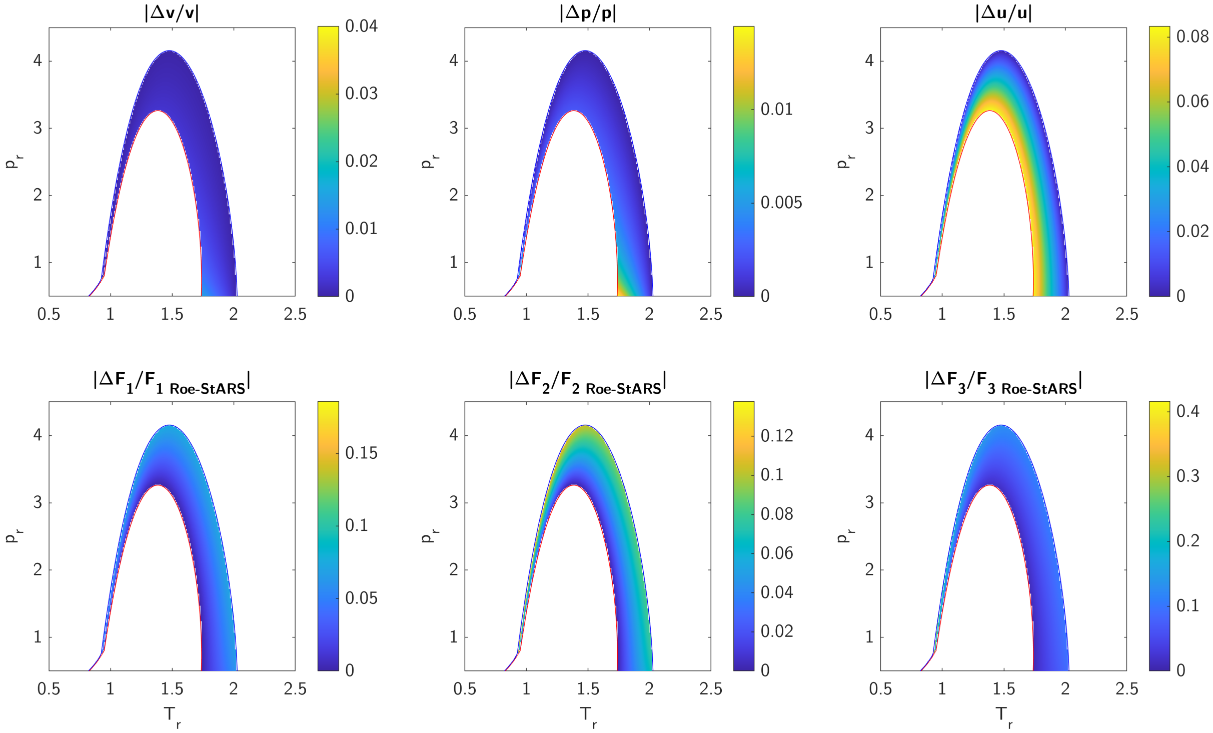


Figure 5.5: Relative errors in the primitive and flux vector components between Roe and Roe-StARS, computed for nitrogen gas under the same large-gradient conditions as Fig. 5.2.b). As predicted by the theory, there is a relatively large region where transonic fluxes occur, and the errors are relatively large in magnitude compared to Fig. 5.4. The sonic curves $Ma = 1$ and $Ma_* = 1$ are indicated with the red and blue curves. F_1, F_2, F_3 are the mass flux, momentum flux, and energy flux, respectively.

transonic flux modelling is therefore paramount for trans- and supercritical flows with large flow gradients.

5.3 Numerical Results

Numerical results are compared for four test cases involving shocks and rarefactions with nitrogen gas: 1) a transcritical shock tube, 2) a shock tube with periodic boundaries and interfering waves, and 3) a novel interpretation of the RP as the limiting case of general flow gradients; 4) a two-dimensional RP. The results of the third test case are analyzed together with the scaling analysis. The ensuing discussion focuses on the adverse effects of entropy violations and the merits of the simple and accurate entropy fix through a structurally complete solution to the RP. All results are computed using the first-order upwind scheme described in §2.6 in order to isolate the benefits of the improved Riemann solver.

5.3.1 Transcritical shock tube

Here we examine the results of solving a transcritical shock tube. Results are compared between the classical Roe solver [49], the Roe solver with Harten-Hyman entropy fix as originally proposed in [1], and the Roe-StARS solver derived in this study. Exact solutions are computed using a Collela & Glaz-type iterative solver [36] leveraging Wang & Hickey's [138] shock and expansion wave equations for non-ideal gases, and the moving normal shock equations of Appendix C. A comparison between ideal and non-ideal results is also provided. At time $t = 0$, the flow field is initialized to the following conditions:

$$\begin{bmatrix} \rho_L \\ u_L \\ p_L \end{bmatrix} = \begin{bmatrix} 180\ \mathrm{kg\,m^{-3}} \\ 150\ \mathrm{m\,s^{-1}} \\ 11\ \mathrm{MPa} \end{bmatrix}; \qquad \begin{bmatrix} \rho_R \\ u_R \\ p_R \end{bmatrix} = \begin{bmatrix} 7.4\ \mathrm{kg\,m^{-3}} \\ 50\ \mathrm{m\,s^{-1}} \\ 0.2\ \mathrm{MPa} \end{bmatrix} \tag{5.17}$$

The computational domain consists of 256 cells that are uniformly distributed over $x \in [-1, 1]$. A CFL number of 0.5, based on the acoustic time scale, is used. Transparent (*i.e.* transmissive, non-reflecting) boundary conditions are implemented via ghost cells, as described in Toro [24]. The solution at time $t = 0.0009$ s is shown in Figs. 5.7 (calorically perfect gas) and 5.8 (non-ideal gas modelled with the PR state equation [2]). A grid convergence study was performed, showing that the selected number of cells falls within the region of asymptotic convergence (Fig. 5.6). Plotted are the L_2 and L_∞ norms of the density error across the domain, normalized by the exact solutions. Linear convergence is observed, consistent with other studies involving sharp flow features [5, 139].

In Fig. 5.7, the Roe solver admits an obvious expansion shock at $x = 0$. Approaching $x = 0$ from either side, the slope of every flow variable tends toward zero followed by a sudden and nonphysical discontinuous jump. This coincides with the flow passing through a transonic state, as seen in the Ma subplot of Fig. 5.7. To quantify the nonphysical shock mitigation achieved by the Roe-StARS versus Roe-Harten, it is helpful to examine the percentage reduction in error (*i.e.* at $x = 0$). The L1 error in a variable y may be computed:

$$\Delta y = \frac{\hat{y} - y_{exact}}{y_{exact}} \tag{5.18}$$

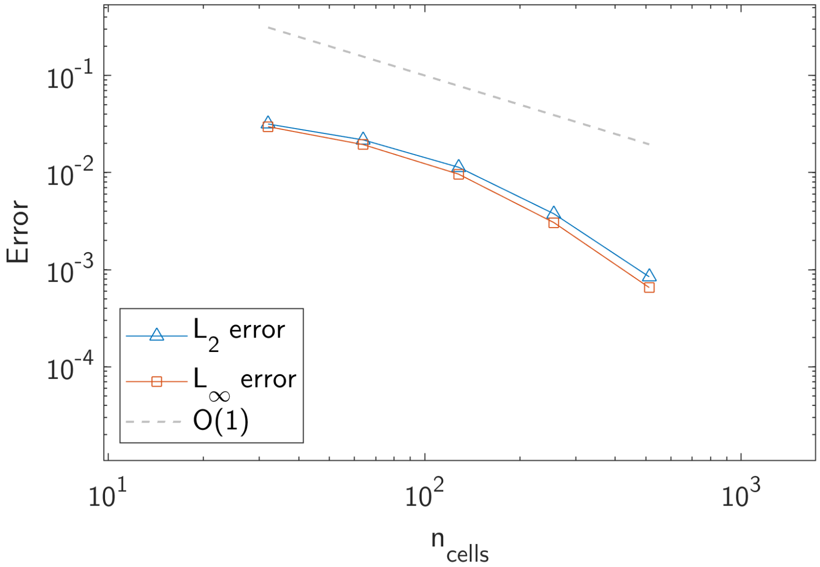


Figure 5.6: Grid convergence study for the case simulated in Fig. 5.7, with n_{cells} ranging from 32 to 512. Linear convergence is observed.

where $\hat{y}$ is the solution calculated via the Riemann solver of interest (*i.e.* Roe, Roe-Harten, Roe-StARS) and y_{exact} is the exact solution. The smaller the error, the smaller the magnitude of the nonphysical feature. The errors, and percent reductions in error offered by the entropy-fixed solvers, are summarized in Tabs. 5.1 and 5.2.

Table 5.1: Absolute L1 errors at $x = 0$ for primitive variables ρ, u, p_r achieved by the Roe, Roe-Harten, and Roe-StARS solvers in the ideal transcritical shock tube of Fig. 5.7. Percentage reduction in the Roe error is shown in parentheses.

Error	Roe	Roe-Harten	Roe-StARS
$\Delta\rho$	26.0	4.19 (-83.9%)	2.62 (-89.9%)
Δu	-53.6	-9.26 (-82.7%)	-5.80 (-89.2%)
Δp_r	0.584	0.104 (-82.1%)	0.0705 (-87.9%)

The Roe-Harten solver reduces the magnitude of the jump by 82% to 84% for all the

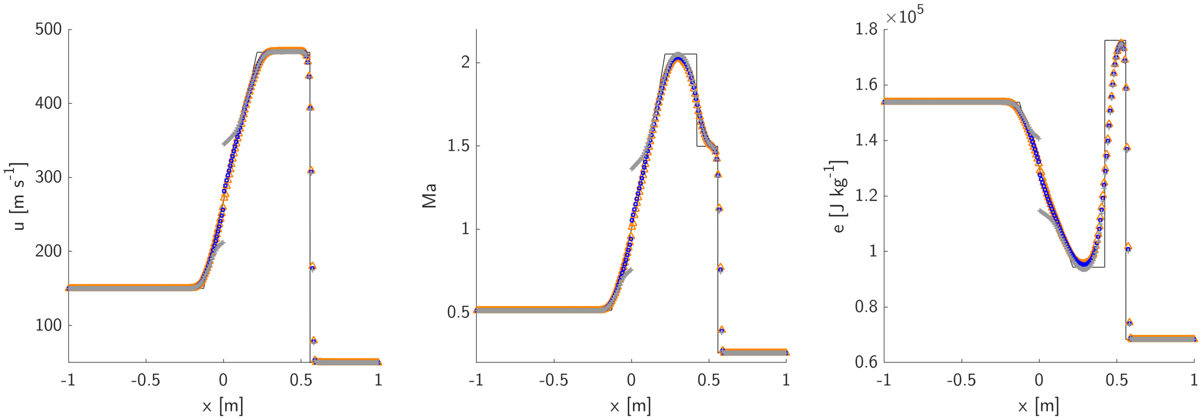


Figure 5.7: Numerical results for a transcritical nitrogen shock tube solved with perfect gas thermodynamics. Results for the Roe, Roe solver with Harten-Hyman entropy fix [1] (Roe-Harten), and Roe-StARS solver are shown.

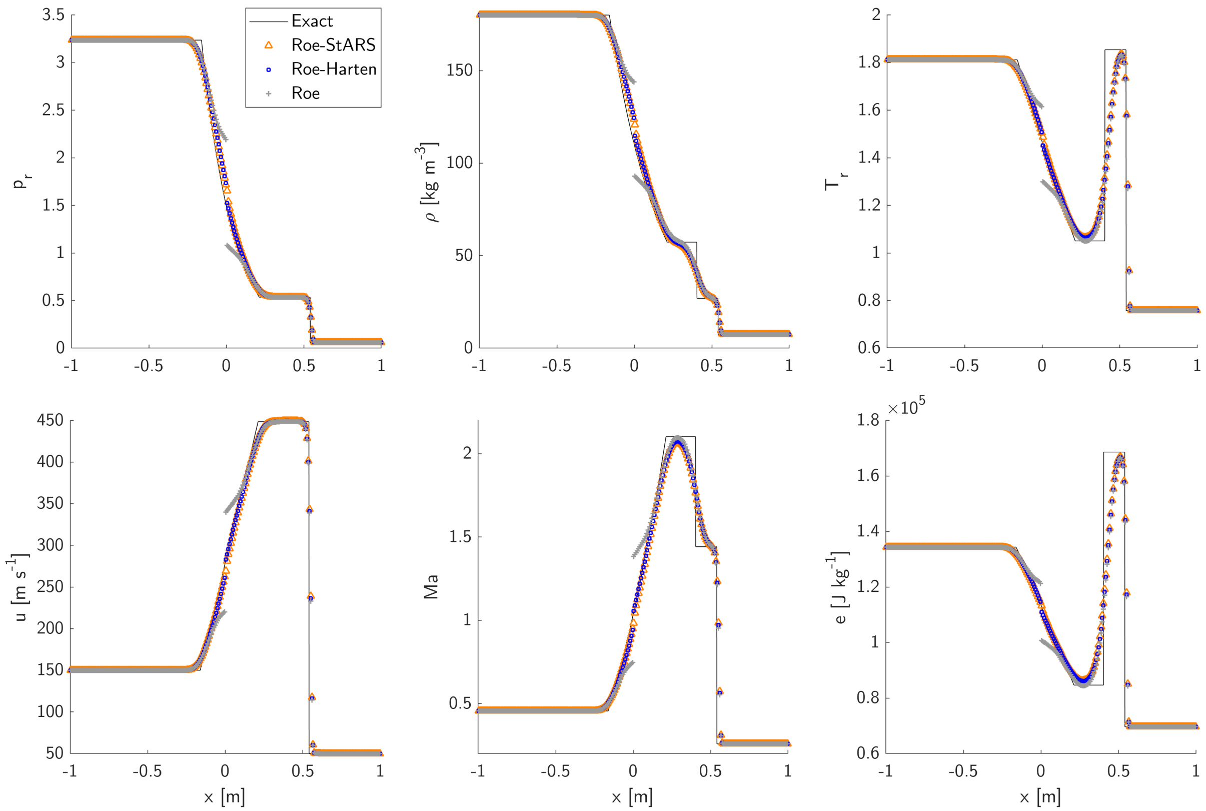


Figure 5.8: Numerical results for a transcritical nitrogen shock tube solved with non-ideal thermodynamics and the Peng-Robinson equation of state [2]. Results for the Roe, Roe solver with Harten-Hyman entropy fix [1] (Roe-Harten), and Roe-StARS solver are shown.

Table 5.2: Absolute L1 errors at $x = 0$ for primitive variables ρ, u, p_r achieved by the Roe, Roe-Harten, and Roe-StARS solvers in the non-ideal transcritical shock tube of Fig. 5.8. Percentage reduction in the Roe error is shown in parentheses.

Error	Roe	Roe-Harten	Roe-StARS
$\Delta\rho$	33.0	13.5 (-59.2%)	12.9 (-61.1%)
Δu	-56.2	-15.2 (-73.0%)	-13.8 (-75.4%)
Δp_r	0.680	0.222 (-67.4%)	0.208 (-69.4%)

Table 5.3: Floating point operations required to solve the flux of a single cell interface using the Roe, Roe-Harten, or Roe-StARS solvers.

	No Entropy Violation		With Entropy Violation	
Solver	Ideal Gas	PR Gas	Ideal Gas	PR Gas
Roe	123	157	123	157
Roe-Harten	229 (+86%)	331 (+111%)	239 (+94%)	341 (+177%)
Roe-StARS	229 (+86%)	331 (+111%)	285 (+132%)	387 (+215%)

primitive variables. However, an inconvenience with Harten-Hyman-type entropy fixes is the need to tune parameters that control the amount of artificial dissipation introduced during an entropy violation [58]. Thus, it is possible that other parameter selections could result in greater or lesser smoothness for a given flow problem, which impacts the generalizability of these fixes. On the other hand, the Roe-StARS solver achieves a 88% to 90% reduction in the expansion shock magnitude and requires no tuning. Away from the expansion shock, the pressure and velocity are nearly identical between Roe, Roe-Harten, and Roe-StARS. The entropy-fixed solvers yield within the star-state region a slight reduction in density and Mach number, as well as a slight increase in temperature and specific internal energy.

Fig. 5.8 yields similar conclusions regarding the relative improvement of the Roe-StARS versus Roe-Harten solvers. Both solvers noticeably reduce and smoothen the expansion shock. The Harten-Hyman fix lessens the expansion shock error by 59% to 73% while the Roe-StARS solver achieves a 61% to 75% reduction. It is also worth noting that, when compared against the perfect gas solution, the use of a non-ideal state equation yields a different temperature profile, seen in Fig. 5.8 as compared to Fig. 5.7. Also, no spurious oscillations are observed for this test despite their prevalence in transcritical flows with non-ideal thermodynamics [140, 5]. Any differences between the solutions in Fig. 5.8 can thus be attributed to the Riemann solvers and not additional errors that may arise in the simulation of high-speed transcritical flows.

Finally, a comparison of the computational requirements of each Riemann solver is provided in Tab. 5.3. The number of floating point operations per cell interface are shown, determined by a line-by-line review of the flux functions in the present study's MATLAB implementations of the Roe, Roe-Harten, and Roe-StARS solvers (a link to the code is provided in the conclusion). Although Roe-StARS involves more floating point operations than the baseline Roe solver, it is of identical complexity to the Harten-Hyman entropy fix when no entropy violations are present (*i.e.* no fix is applied to the flux). That is, both traditional entropy fixes and Roe-StARS perform similar checks on the wavespeeds and therefore the computational requirements are identical when entropy issues do not occur. In the presence of an entropy violation, Roe-StARS is substantially more demanding than either the Roe or Roe-Harten solvers, however for most flow problems, a very small fraction of the total cells contain entropy-violating solutions. For instance, in the current transcritical shock tube problem, only the two cells adjacent $x = 0$ admit entropy problems, or approximately 0.8% of the computational domain. Roe-StARS' improvements in accuracy and generalizability versus the additional computational expense may therefore advantageous in cases where transonic fluxes are present and entropy stability is desired.

5.3.2 Periodic shock tube: interfering shock and expansion waves

Historically, the investigation of entropy violations has been characterized by the study of nonphysical features at a fixed time after the initial conditions. For example, in the last subsection, a transcritical shock tube was studied at $t = 0.9$ ms. With limited opportunity for waves to propagate throughout the domain, there is no interference between rarefactions, shocks, or contact discontinuities. By contrast, real-world flow simulations call for thousands of time steps that would inevitably obscure and dissipate entropy errors so as to render them indistinguishable from physically consistent flow features. This subsection aims to highlight one example of this occurrence.

A flow field is initialized with the same conditions as in the previous subsection, using the same uniform mesh, and the same time advancement to solve a transcritical shock tube problem. However, the boundary conditions are now periodic: the fluxes at one end of the spatial domain propagate to the other end. This is implemented by setting the 0^{th} ghost cell conditions equal to the $(n_{cells})^{th}$ cell conditions, and the $(n_{cells} + 1)^{th}$ ghost cell conditions equal to the 1^{st} cell conditions, then proceeding with flux calculations as usual. Additionally, the time marching is allowed to proceed until $t = 7.5$ ms, or more than 8 times the final simulation time of the earlier transcritical shock tube. Perfect gas thermodynamics are used. The results at time $t = \{0.9, 2.0, 4.0, 7.5\}$ ms are shown for the primitive variables ρ, u, p_r in Fig. 5.9. Both Roe-StARS and Roe-Harten would remove nonphysical shocks, thus the results of Roe-StARS compared with Roe.

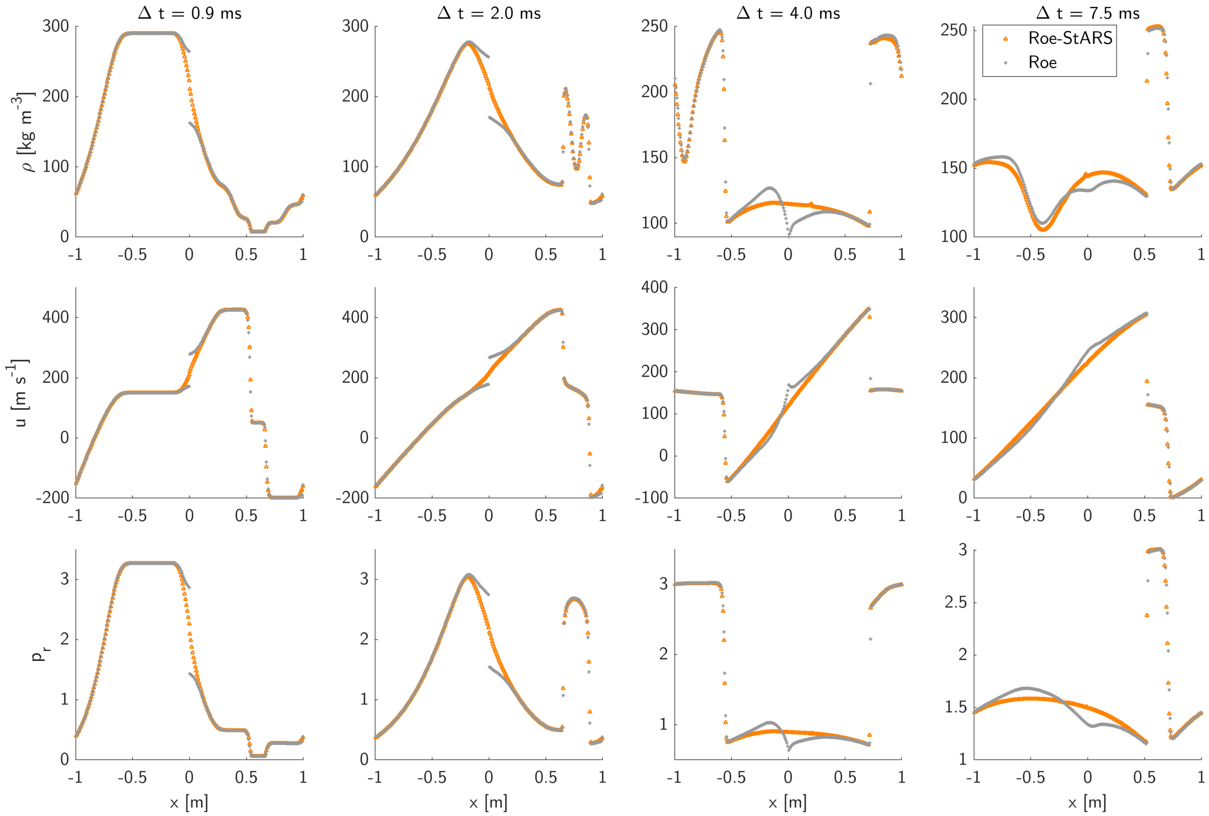


Figure 5.9: Numerical results at various times for a transcritical nitrogen shock tube with periodic boundary conditions. The solution initially comprises distinct waves that interfere as time proceeds, causing the errors from the expansion shock to propagate throughout the domain.

At time $t = 0.9$ ms, the solution has just started to evolve and the wave regions are distinct. Two rarefactions are generated: a central expansion wave that encloses the erroneous shock at $x = 0$, and an expansion wave whose head is located at approximately $x = -0.55$ m and whose tail wraps around to the right side just left of $x = 1$ m (bear in mind the periodic boundaries). The first rarefaction shall be referred to as the *central* rarefaction; the second rarefaction, the *edge* rarefaction. A second, significantly smaller expansion shock is observed at the rarefaction tail near $x = 1$ m, however its effect on the solution at later time snapshots is marginal. Two shocks are also generated: a right-moving shock at approximately $x = 0.5$ m and a left-moving shock at approximately $x = 0.7$ m. Two contact discontinuities are observed, however these dissipate rapidly in the subsequent time snapshots. Overall, the solution at $t = 0.9$ ms resembles the usual rarefaction-contact-shock structure except with periodic bounds. Outside the vicinity of the expansion shocks, the solution is largely similar between Roe and Roe-StARS.

By $t = 2.0$ ms, the heads of the central and edge rarefactions meet at approximately $x = -0.18$ m. The region where Roe and Roe-StARS differ has grown, now spanning $x \in [-0.17, 0.20]$ m as the central expansion wave spreads out in space thereby propagating errors from the original expansion shock at $x = 0$ m. The two shocks have also passed each other, the left-moving shock at $x = 0.64$ m and the right-moving shock at $x = 0.90$ m. In the interstitial post-shock region $x \in [0.64, 0.90]$ m, the density and pressure of the Roe-StARS solution is minutely lower than that of Roe. The contact discontinuities have interfered with the strong shocks and are no longer uniquely identifiable. The differences between Roe-StARS and Roe continue to be concentrated around the expansion shock locations.

The snapshot at $t = 4.0$ ms offers the first glimpse of the solution after the expansion waves have started to interfere, revealing significant discrepancies between Roe and Roe-StARS. By now, the right-moving shock has reappeared on the left at $x = -0.55$ m and overlaps with the edge and central rarefactions. The slower-moving left shock has been pushed back to $x = 0.73$ m as it has started to interfere with the right-moving tail of the central rarefaction. Whereas the Roe-StARS solution is smooth throughout the region $x \in [-0.55, 0.73]$ m, the Roe solution contains a kink as well as overshooting and undershooting to the left and right of $x = 0$.

The last snapshot shown at $t = 7.5$ ms shows the right shock having made it all the way around to meet the left shock again near $x = 0.6$ m. All waves have interfered with one another in at least one location, and unlike at $t = 4.0$ ms, the errors due to the central expansion shock have dissipated. In particular, no obvious kinks are visible and the solution is relatively smooth save for the two known shocks. Unless one were carefully tracking each wave as we have in this analysis, it would be easy to assume that no entropy violation has occurred in the Roe results. The Roe-StARS solver, however, nearly completely removes the nonphysicalities and their propagation throughout the domain.

5.3.3 Gradient Riemann problem

The RP is typically initialized with a discontinuous step function at $x = 0$. However, computationally, the act of discretizing the spatial domain implies that step functions are mathematically equivalent to regions of sufficiently large slope. This begs the question: if numerical schemes treat discontinuous jumps as regions of large flow gradients, then how do entropy violations change as the initial condition jump is replaced by shallower gradients?

Here we study versions of the RP where the centre of the initial conditions is not a step function but a gradient region of finite slope. These problems are dubbed *gradient* Riemann problems. We demonstrate that entropy violations persist and change in magnitude with respect to the initial gradients, and that these changes—though sometimes non-monotonic—are consistent with the earlier scaling analysis.

The initial conditions are as follows. Suppose ΔL is the thickness of the gradient region centred at $x = 0$. Then for $x < -\frac{\Delta L}{2}$, the density and pressure are 180kg/m^3 and 11MPa. For $x > \frac{\Delta L}{2}$, the initial conditions are 7.4 kg/m^3 and 0.2 MPa. The flow speed at $t = 0$ is set to 150 m/s everywhere. However, for the gradient region $|x| \leq \frac{\Delta L}{2}$, the initial conditions are:

$$\begin{bmatrix} \rho \\ u \\ p \end{bmatrix} = \begin{bmatrix} -\frac{\rho_R - \rho_L}{\Delta L} x + \frac{\rho_R + \rho_L}{2} \\ -\frac{u_R - u_L}{\Delta L} x + \frac{u_R + u_L}{2} \\ -\frac{p_R - p_L}{\Delta L} x + \frac{p_R + p_L}{2} \end{bmatrix} \tag{5.19}$$

where L and R signify the conditions at $x < -\frac{\Delta L}{2}$ and $x > \frac{\Delta L}{2}$, respectively. This is a linear region that connects the left and right primitive variables. The spatial domain is made up of 128 points uniformly distributed between $x \in [-0.5, 0.5]$ m, yielding the same level of discretization as the previous subsection. Non-ideal thermodynamics with the PR state equation [2] are used and the time step is selected based on a constant CFL number of 0.5. The solution is advanced to $t = 0.5$ ms, and transparent boundary conditions are used although the time advancement is so short that waves would not interfere even if periodic boundary conditions were applied instead. Results are compared for Roe and Roe-StARS for values of $\Delta L = \{0.78, 2.34, 3.91, 5.46\}$ cm (Fig. 5.10). The magnitude of the expansion shock at $x = 0$ is tabulated in Tab. 5.4.

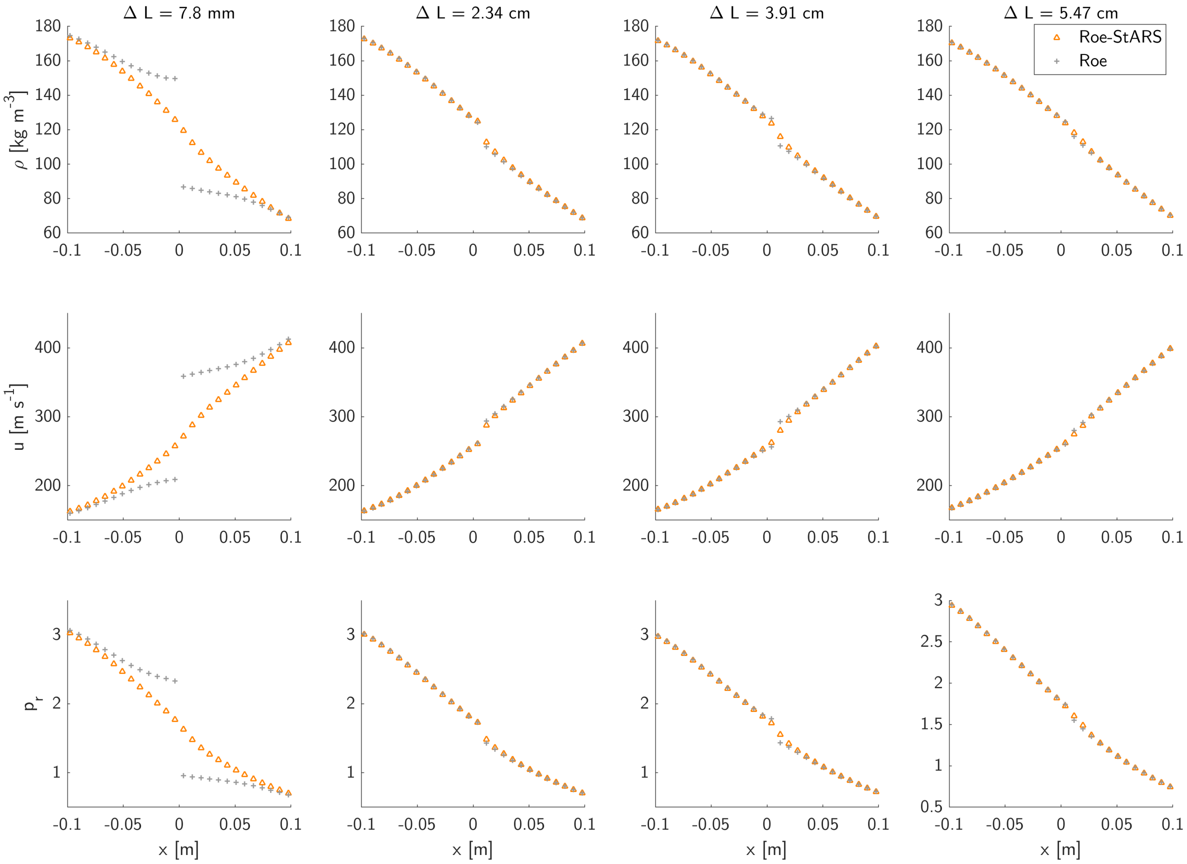


Figure 5.10: Pressure, flow speed, and reduced pressure plots for four gradient Riemann problems in which left and right initial conditions are identical but the width of the initial gradient region varies from $\Delta L = 7.8$ mm to 5.47 cm. The magnitude of the expansion shock decreases at first, then increases slightly as gradients become smaller.

The leftmost plots in Fig. 5.10 are effectively up-close snapshots of the typical Riemann problem with a discontinuous jump. At $\Delta L = 7.8$ mm, the gradient is indiscernible from a step function that has been discretized. The limits in the x-axis are confined to ±0.1 m to

Table 5.4: Magnitude of the nonphysical shock at $x = 0$ for primitive variables ρ, u, p_r achieved by the Roe versus Roe-StARS solvers for the gradient Riemann problem of Fig. 5.10. Note the non-monotonic behaviour of the Roe solver as the gradient becomes smaller.

	Roe				Roe-StARS			
ΔL	7.8mm	2.34cm	3.91cm	5.47cm	7.8mm	2.34cm	3.91cm	5.47cm
$\Delta\rho$	-62.8	-14.1	-15.9	-8.6	-6.4	-12.2	-7.7	-5.5
Δu	150.0	32.5	36.6	19.6	14.3	27.1	17.4	12.4
Δp_r	-1.37	-0.30	-0.35	-0.19	-0.14	-0.25	-0.17	-0.12

offer a clearer picture of the entropy violation. As expected, an expansion shock appears at $x = 0$ for the Roe solver, but this is almost completely mitigated with Roe-StARS.

What seems anomalous at first glance is that as ΔL grows (*i.e.* the initial gradients become shallower), the magnitude of the Roe solver's expansion shock *rises* slightly for $\Delta L = 3.91$ cm, then falls again for $\Delta L = 5.47$ cm. Similar behaviour is observed for the Roe-StARS solver's expansion shock, which rises slightly for $\Delta L = 2.34$ cm and falls again for $\Delta L = 3.91$ cm. The magnitude of the nonphysical shock at $x = 0$ continues to decrease monotonically thereafter for both solvers. It is also worth noting that the Roe-StARS entropy violation is nevertheless always smaller than the Roe entropy violation.

To explain this non-monotonic behaviour, we return to (5.11) which describes the conditions needed for supersonic rarefaction tails. Recall that subsonic rarefaction heads and rarefaction tails result in transonic conditions that give rise to entropy errors. Let us substitute $z = \sqrt{1 + \frac{1}{\rho}\frac{d\rho}{dx}\Delta x}$ and $f = 1 + \frac{1}{u}\frac{du}{dx}\Delta x$, where $f > 0$ at $x = 0$ in this test case. Then (5.11) simplifies to:

$$\widetilde{Ma} = \frac{|u|}{\tilde{c}} \left| \frac{1 + fz}{1 + z} \right| \tag{5.20}$$

As negative density gradients decrease in magnitude, z increases, which in turn causes the entire stretching factor to increase in magnitude. Of course, $\frac{du}{dx}$ and thus f is also decreasing, and these two competing effects determine whether the overall expression grows or shrinks. In the presence of suitable flow conditions, the stretching factor can grow and thus expand the region D in the thermodynamic state space where transonic fluxes arise. It is therefore imperative in the analysis of entropy violations to consider not only the *magnitude* of these violations when they occur, but also how *often* they occur in the spatio-temporal domain. In this example, the relative errors for a given cell and time step ((5.13) to (5.15)) decrease as gradients become smaller—yet the thermodynamic state space under which entropy violations occur is growing. The result is non-monotonic expansion shock magnitude as gradients change. All approximate Riemann solvers are thus capable of producing this non-monotonic expansion shock behaviour.

5.3.4 Two-dimensional Riemann problem

Although the RP is typically studied in 1D, it is easy to extend the problem to higher dimensions in which more complex flow structures and anomalies can occur. Here we consider a test case comprising the RP in 2D. Details on extending the first-order Godunov scheme and Roe solver to higher dimensions are available in Appendices A, B, and E. In higher dimensions, it is possible to construct more elaborate schemes and meshes that can improve the accuracy and stability of results. However, uniform Cartesian grids are used here for simplicity and consistency with the 1D cases studied earlier. It is the ability of the Riemann solver, and less so the scheme, that is of central interest.

It should be noted that there is extensive literature on the nature of solutions to the

2D RP. Glimm *et al.* [141] performed early studies exploring the nature of the 2D RP, examining the interaction of various wave structures and proving essential properties and corollaries for two-dimensional elementary waves. This was soon followed by Zhang & Zheng [142] who proposed criteria on the initial conditions required to achieve certain wave arrangements. Major contributions have also been made by Schulz-Rinne [143], Schulz-Rinne *et al.* [144], Lax & Liu [145], and others [146, 147, 148, 149, 150], who discovered further wave combinations, structures, flow instabilities, and spurious flow features. Like the 1D RP, the 2D RP serves as a test case for verifying numerical codes, in addition to offering critical insights on wave interactions when interpreting complex real-world phenomena such as Mach reflection and diffraction. What renders the multi-dimensional RP challenging to solve is that waves propagate not only in orthogonal directions, but also at oblique angles especially at the origin where information from all four quadrants interfere. For the present purpose, it suffices to consider just one of the potential configurations of a 2D RP comprising two rarefaction and two contact discontinuity waves under transcritical conditions.

Fig. 5.11 shows the initial conditions and expected wave evolution. u is the velocity in the x-direction, while w is the velocity in the y-direction. Transparent boundary conditions are set at $x = -1m, 1m$ and $y = -1m, 1m$. A CFL number of 0.45 is used. The computational domain consists of 256×256 cells. The numerical scheme is advanced until time $t = 0.0006s$. The medium is transcritical nitrogen gas, modelled with the PR EOS and full non-ideal thermodynamics. The results for the Roe and Roe-StARS solvers are provided in the contours of Figs. 5.12 and 5.13. Unlike the 1D RP, an analytical expression for expansion waves is lacking due to the interference of various waves near the origin.

Therefore, the ensuing analysis is based on the concepts discussed earlier and knowledge of the 1D RP.

For both the Roe-StARS and Roe solvers, slices of the solution along $x < 0$ and $y < 0$ are similar to the earlier transcritical shock tube as one should expect. Shock waves are observed at $y \sim 0.4m, x < 0$ and $x \sim 0.4m, y < 0$, while transonic rarefactions whose heads and tails move in opposite directions are seen enclosing the negative x and y axes. The regions near the positive x and y axes comprise slip lines across which pressure and velocity are equal but density varies—these correspond to 2D contact discontinuities. In both solutions, a high pressure region is observed just to the upper right of the origin. This local maximum occurs because the pressurized gas in the bottom left interferes with the dense gas in the top right, resulting in a localized peak. However, this is where the similarities end.

The Roe solution exhibits nonphysical expansion shocks along the negative x and y axes. As is typical for a rarefaction wave, there are a number of successive contour curves approaching each nonphysical shock from within the bottom left quadrant. These contours—which are akin to characteristic curves—change drastically across the expansion shock and become more sparse on the other side. That is, the characteristic lines effectively terminate as they intersect with the nonphysical shock. The nonphysical shocks also extend past the origin before connecting with the wavefront where the two slip lines meet. Due to the severity of the expansion shock, information from the bottom left quadrant does not propagate effectively into the remaining domain, limiting the peak values of the compressed region that forms between the bottom left and top right regions. Most critically, $t = 0.0006s$ is the maximum simulation time that the Roe solver can manage. Spurious flow instabilities manifest around $x = 0.3m, y = 0.4m$ that cause negative pressure, temperature, and specific internal energy. The simulation would crash for any additional time steps, and indeed, the results presently shown are physically incorrect. It is worth noting that for the Euler equations in any number of spatial dimensions, Einfeldt *et al.* [10] proved that Godunov's scheme is positively conservative, however no Godunov-type scheme based on linearized Riemann solvers can be guaranteed as such. In particular, they proved that for the Roe scheme, certain choices of initial data will result in a nonphysical vacuum, producing instabilities despite the existence of solutions. This appears to be the cause of the present anomaly, where pressure, temperature, and internal energy are negative and density is extremely low in the vicinity of the spurious oscillations.

Fortunately, the Roe-StARS solution admits physically consistent results and is capable of advancing past $t = 0.0006s$. The rarefaction shocks are greatly mitigated, with the characteristics maintaining a fan-like appearance as expected in uninterrupted rarefaction waves. This permits information from the bottom left quadrant to reach other areas of the domain, such that the peak density is $\sim 270\mathrm{kg\,m^{-3}}$, or about 50% higher than in the case of the Roe solution. In place of the spurious flow instabilities of the Roe solution, there are minor kinks in the contour lines. This is visible in the top-right-most contour of the specific internal energy plot, near the corner where the contour line changes direction. Additionally, there appear to be small vortical structures near $x = 0.2m, y = 0.2m$ owing to the slip lines and complex pressure and density interactions between the four quadrants. Similar types of structures manifest in ideal-gas test cases examined by others [144, 145].

Overall, the 2D RP test case further demonstrates how entropy violations can lead to flow inaccuracies far away from the space-time coordinates of the original violation, as well as produce spurious oscillations. Roe-StARS therefore offers a heuristics-free approach to mitigating these challenges in cases where flow interactions are complex and exact reference solutions are generally unavailable.

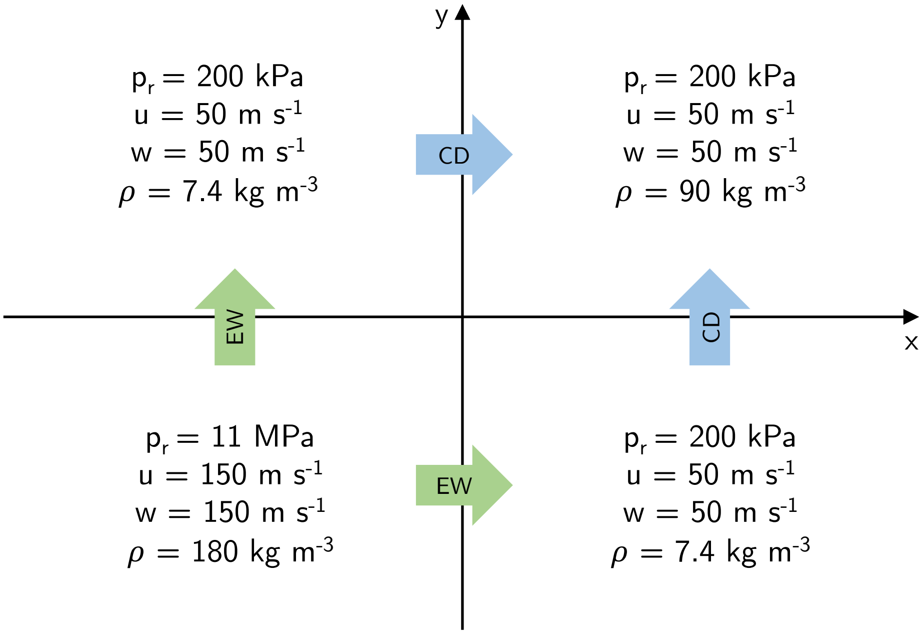


Figure 5.11: Initial conditions for the 2D Riemann problem test case. The gas in the bottom left quadrant is set to the highest pressure and density. The top left and bottom right quadrants are set to the lowest density. Density in the top right is higher than the densities of the top left and bottom right. Pressure and velocities are the same in the top left, top right, and bottom right quadrants. Based on these initial conditions, the expected waves are shown consisting of two contact discontinuities (*i.e.* slip lines in 2D) and two expansion waves.

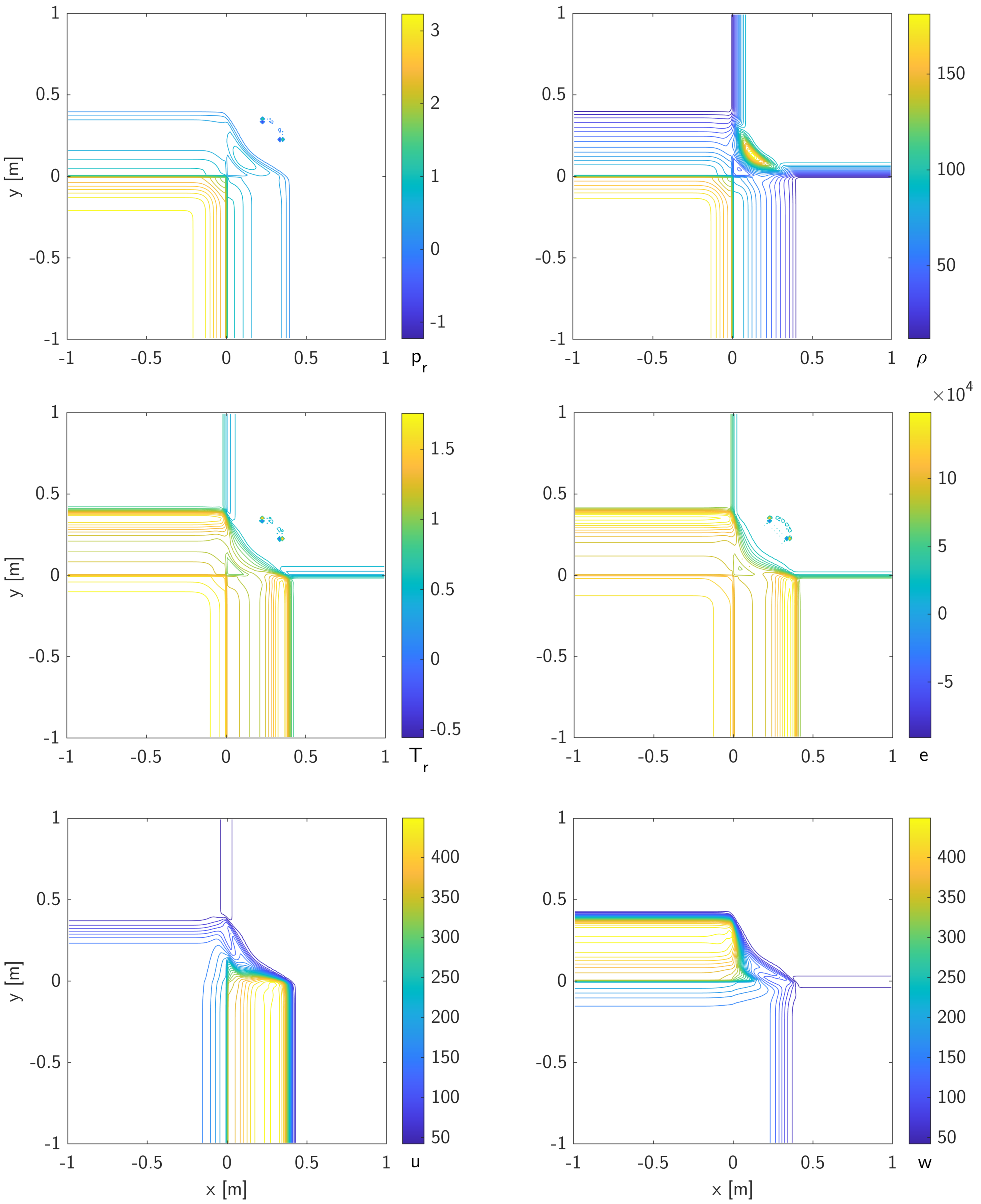


Figure 5.12: Solution to the 2D Riemann problem test case at $t = 0.0006s$, featuring two rarefaction waves and two slip lines, as computed with a Roe solver. Note the nonphysical expansion shocks along the negative x and y axes, the negative primitive variables, and the spurious oscillations in the top-right quadrant. e here is presented in units of $\mathrm{J\,kg^{-1}}$.

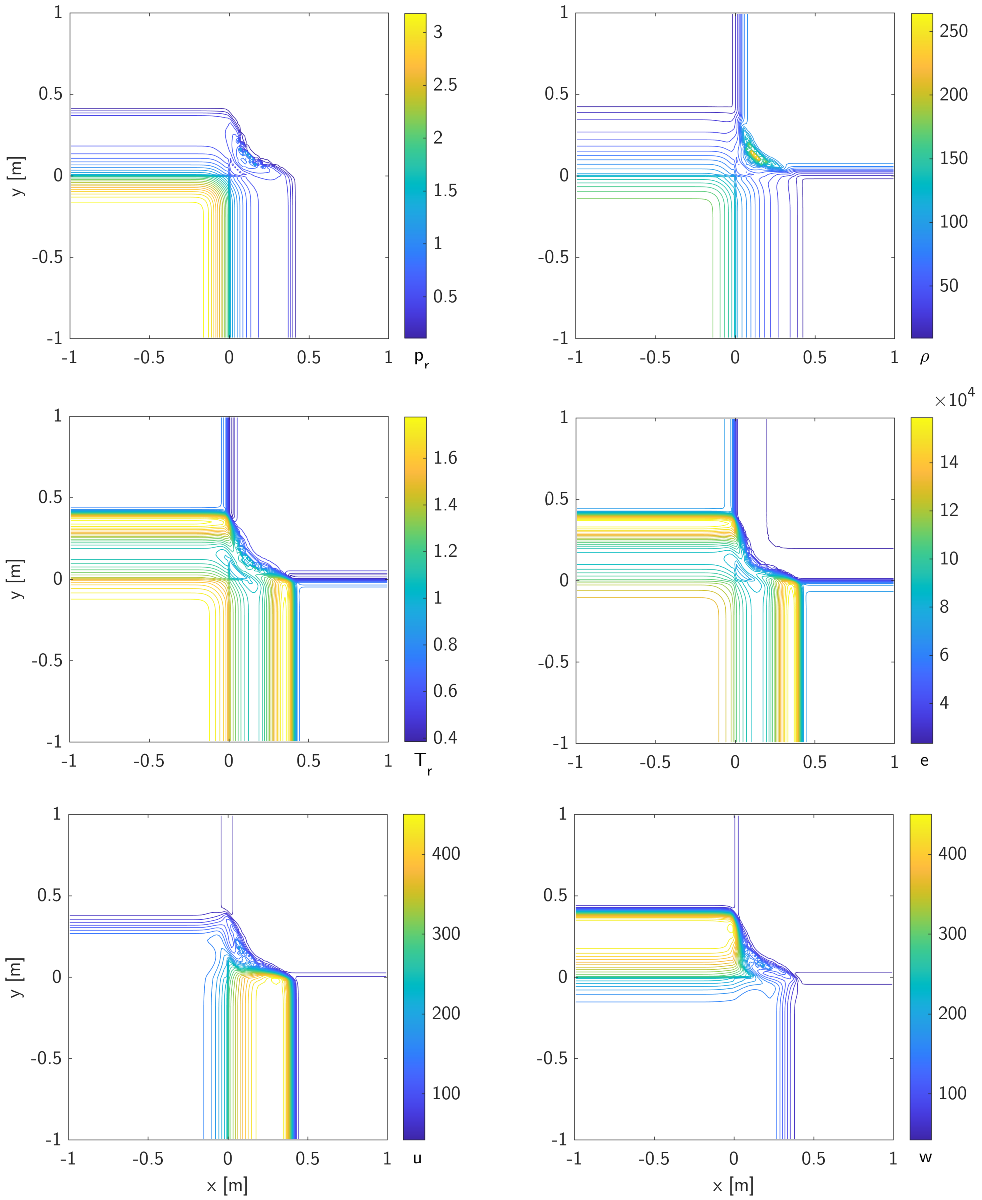


Figure 5.13: Solution to the 2D Riemann problem test case at $t = 0.0006s$, featuring two rarefaction waves and two slip lines, as computed with the Roe-StARS solver. Unlike the Roe solution, the Roe-StARS solutions significantly mitigate any nonphysicalities or spurious oscillations. e here is presented in units of $\mathrm{J\,kg^{-1}}$.

Chapter 6

FluxNet: a Physics-Informed Learning-Based Riemann Solver

Different multi-layer perceptron networks are trained to predict star-state variables in the RP, which in turn can be used to compute the intercell flux for Godunov schemes. For the present purpose, neural networks used for flux computations are termed *FluxNets*. Prior efforts to leverage ML for the RP have tended to assume perfect gas thermodynamics and have used physics-informed loss functions and arbitrary network dimensions without much justification. Given the growing interest in non-ideal flow problems, where flux and thermophysical computations can jointly account for approximately 80% of total runtime [48], there may be performance gains in adopting a learning-based Riemann solver for arbitrary EOS. In this chapter, it is shown that a compact physics-informed FluxNet can be trained whose errors are an order-of-magnitude less than with approximate solvers, and whose time complexity is $\sim$ 25% that of exact solvers when matrix computations are parallelized. Virtually all FluxNet designs examined in this chapter predict star-state conditions with 0.1 to 0.3% error during training, however the physics-informed approach ensures smooth generalizable predictions. Additionally, it is shown that a learning-based approach produces smaller, less biased distributions of error in the star-state variables compared to the Roe family of solvers.

This chapter is organized in the following manner. Section 6.1 analyzes preliminary designs of different FluxNet approaches, and comments on the feasibility that these FluxNets could offer competitive performance (*i.e.* balance of accuracy and efficiency) relative to traditional Riemann solvers. In section 6.2, the data preparation process is discussed, including the generation and post-processing of exact data used for training and testing. A total of six FluxNet designs detailed in section 6.3 are trained and tested. Their learning curves are studied in subsection §6.4.1, with the final two candidates tested on transcritical 1D and 2D test cases in subsections 6.4.2 and 6.4.3. A final comparison of accuracy and time complexity between various Riemann solvers is offered in subsection 6.4.4.

Chapter 6 is under review as Wang, J. C.-H. and Hickey, J.-P. (2022): FluxNet: a physics-informed learning-based Riemann solver for transcritical flows with non-ideal thermodynamics, *J. Comput. Phys.*

The codes (including hyperparameters) that contributed to the results of this chapter are available at `https://git.uwaterloo.ca/jc9wang/fluxnet`.

6.1 Preliminary Design and Feasibility Assessment

Prior to the systematic training and evaluation of any FluxNets, it is wise to thoroughly consider the alternative choices for feature and prediction vectors, their potential implications for network performance, and how the resulting network must be integrated with the broader numerical scheme. As part of this analysis, preliminary, informal training attempts were conducted in Python 3.9 and PyTorch 1.10.2, using the datasets described shortly in §6.2. These initial training sessions helped to determine which candidate designs were suitable for continued study. The dimensions used for preliminary training were 6 hidden layers with 64 nodes per hidden layer, benchmarked on the multi-layer perceptron designs in existing literature [108, 107].

It is crucial to note that if calculations are parallelized, then the time complexity of multiplying two square matrices is $O(\ln(n))$ where n is the height or width of each matrix [151]. Therefore, a multi-layer perceptron with n nodes per layer and h hidden layers would have a time complexity τ that scales as:

$$\tau = h(\ln(n) + n) \tag{6.1}$$

in which the addition of n accounts for the element-wise activation functions per layer. So, a 64×6 FluxNet scales as $O(6(\ln(64) + 64)) \approx 409$. A further $O(10^2)$ operations are required to compute the intercell flux from the star-state conditions predicted by each FluxNet. Thus, a 64×6 FluxNet possesses a time complexity of roughly 509, as compared to the 331 to 387 floating point operations when computing the Roe-StARS flux in the case of a PR gas (Tab. 5.3). In traditional Riemann solvers, many of these calculations are nonlinear and scalar, and therefore must be performed serially. Although an ML approach does indeed require more total computations, the runtime consequences can be mitigated through parallelized linear algebra realized on GPU-equipped servers. A mild sensitivity analysis of network performance to network size is discussed later in the systematic training and evaluation phase beginning in §6.3. If the 64×6 initial network specification produced inadequate performance during the preliminary design phase, then hyperparameters and dimensions were varied until the network complexity exceeded or would exceed that of traditional solvers by an order of magnitude. In this case, the design was abandoned.

A total of four candidates for feature and prediction vectors were examined (Tab. 6.1). Although finer variations are possible, these four represent distinct philosophies in the development of a learning-based Riemann solver: 1) a conservative variable-only approach, 2) a primitive-to-conservative solver, 3) a purely primitive variable solver that specifies all independent thermodynamic state variables, and 4) a primitive variable solver with only the minimum required predictions. The analysis herein concludes that it is indeed feasible to develop a learning-based Riemann solver whose accuracy and computational complexity fall in between current approximate and exact approaches, with the fourth design eventually selected for the development of detailed FluxNet candidate designs in §6.3. The first three candidates were discarded for theoretical reasons and initial training observations that will be discussed below.

Design 1 represents an end-to-end Riemann solver in the sense that no further mathematical steps are required to compute each step of the Godunov scheme as shown in Eq. (2.40). In Design 1, it is possible to go directly from the known conservative variables at

Table 6.1: Potential choices of feature and prediction vectors for designing a FluxNet (*i.e.* a learning-based Riemann solver). The subscripts (L, R) refer to the left and right initial conditions, while $(*L, *R)$ refer to the left and right star-state conditions. $\mathbf{F}$ with no subscript is shorthand for the intercell flux, that is, $\mathbf{F}(\mathbf{U}_{i+\frac{1}{2}})$ assuming that (L, R) correspond to the $(i, i+1)$ cells as per notation in Chapter 2.

Design No.	Features (Input)	Predictions (Output)
1	$\mathbf{U}_L, \mathbf{U}_R$	$\mathbf{F}, u_*$
2	$p_L, u_L, \rho_L, p_R, u_R, \rho_R$	$\mathbf{F}, u_*$
3	$p_L, u_L, \rho_L, T_L, p_R, u_R, \rho_R, T_R$	$\rho_{*L}, u_{*L}, p_{*L}, T_{*L}, u_*, \rho_{*R}, u_{*R}, p_{*R}, T_{*R}$
4	$p_L, u_L, \rho_L, p_R, u_R, \rho_R$	$\rho_{*L}, u_{*L}, p_{*L}, u_*, \rho_{*R}, u_{*R}, p_{*R}$

one time step, to the desired conservative variables at the next time step, using a direct mapping from $\mathbf{U}$ to $\mathbf{F}$. This design lends itself to a rather simple numerical scheme by using entirely conservative variables and eliminating the usual computational overhead of alternating between conservative variables (useful for fluid dynamics) and primitive variables (useful for thermodynamics), which become highly computationally expensive under non-ideal thermodynamic conditions. The Rankine-Hugoniot conditions, in conservative form (Eq. 3.18), may be used to define a physics-informed loss function.

Despite its seeming simplicity and elegance, Design 1 possesses two critical flaws. Firstly, recall that the CFL condition (Eq. 2.34) requires the speed of sound to be computed at each time step. However, the speed of sound is a thermodynamic quantity that, in general, depends on the complete thermodynamic state defined by two independent primitive variables, *e.g.* v, T when using cubic EOS. Although Design 1 would simplify flux computations, the overall numerical scheme still requires the overhead of conservative-to-primitive conversion. Secondly, the use of conservative variables forces the FluxNet to discover the mapping from conservatives to primitives and vice-versa. This is because the characteristic waves that influence the behaviour of the RP are functions of primitive variables, and there are multiple combinations of primitive variables that correspond to the same momentum and energy components of $\mathbf{U}$. Design 1 requires that the FluxNet learn the complete primitive to conservative mapping in order to preserve a one-to-one correspondence. Design 1 would therefore increase network complexity while simultaneously still requiring primitives to be calculated separately for CFL purposes.

Design 2 attempts to overcome the weaknesses of Design 1 by mapping primitive variables to the intercell flux $\mathbf{F}$, eliminating the need for the FluxNet to uncover the highly nonlinear conservative-to-primitive relationship. In practice, attempts to train a multilayer perceptron to predict $\mathbf{F}$ were unfruitful. Despite varying hyperparameters and testing other optimizers, it was observed that network widths and depths more than triple the initial dimensions could not reduce training or test errors below 30%. This poor performance may be attributed to the sheer complexity of the underlying mathematical relationships. The flux terms are nonlinear expressions of primitive variables that depend on which state of the RP encloses the cell interface. The correct state is, in turn, dependent on the accurate estimation of characteristic waves that separate the states, which further depend on the complex mathematical form of the non-ideal speed of sound (Eq. 2.25). Indeed, if a FluxNet could be designed to predict fluxes accurately, it is likely that inference would be far slower than running exact-iterative solvers, let alone approximate solvers, even when computations are parallelized. Fuks & Tchelepi [109] have also noted difficulties in predicting flux variables from initial conditions, however that was in the context of two-phase porous media.

In Design 3, a primitive-to-primitive mapping is proposed that also includes T. Temperature together with density may be used to calculate the wavespeeds of the RP, determine which region of the RP occupies the intercell boundary, and compute the flux. It should be noted that estimating all three thermodynamic state variables does not necessarily overconstrain the prediction space, provided that the network learns the underlying EOS. The Rankine-Hugoniot constraints in primitive form (Eqs. 3.15 to 3.17) may be employed to define a physics-informed loss function. Once again, practical attempts to train such a network revealed difficulties in predicting p, ρ, T that were consistent with the EOS. Despite varying hyperparameters, network width and depth, and optimizers, the best-performing FluxNets predicted only two of three thermodynamic state variables within $< 30\%$ error during any training session. The percent error in the third remaining state variable was observed oscillating near $O(10^2)$. An example of a preliminary training run on Design 3 is shown in Fig. 6.1. An additional loss term based on the EOS was tested, thereby incorporating an explicit constraint to assist learning of the EOS, but yielded similar results. As with Design 2, Design 3 experienced difficulty producing physically consistent predictions in an efficient manner.

Design 4 was selected for further exploration. By removing T, initial training attempts with a 64×6 FluxNet revealed that all predicted star-state variables fell well below 5% error after a mere 200 epochs, with potential for improvement as the loss function was still decreasing at the point where the training process was stopped prematurely for the purposes of preliminary evaluation. It is worth acknowledging that a version of Design 4 was trained to predict ρ, T rather than p, ρ, but like with Design 3, either ρ or T tended to diverge. Compared to Design 3, the only difference in the broader numerical scheme is that Design 4 requires a post-inference computation of T_{*L} and T_{*R} from the corresponding

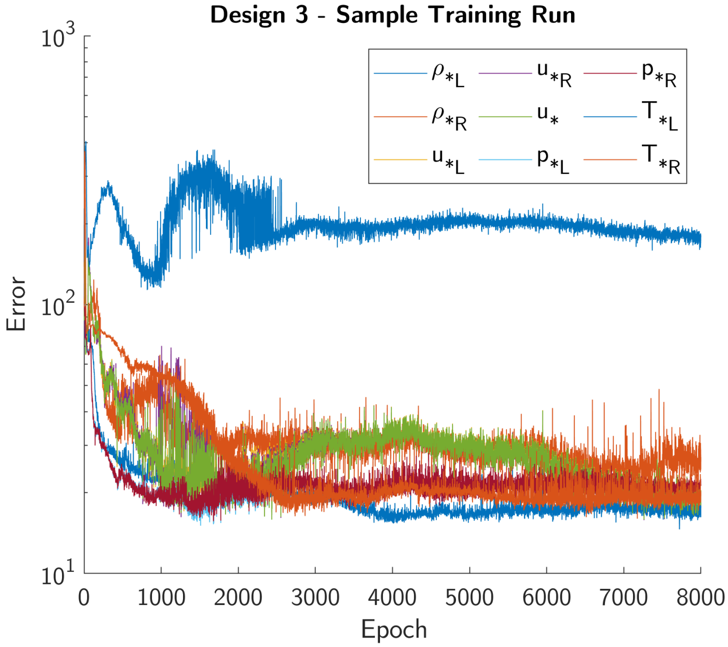


Figure 6.1: Learning curve from a sample training run on Design 3, showing mean L1 test errors that stall around 12% in all predicted variables except T_{*L} which oscillates at $> 100\%$ error. Accordingly, Design 3 was not advanced further.

p, ρ using the EOS. It is also interesting that although the physics of the RP depends on the choice of EOS, the absence of T and any EOS-based physical constraint produced improved convergence, stability, and accuracy. Moreover, this performance was achieved only with the mass and momentum constraints. In fact, the inclusion of the energy constraint hampered the convergence and stability of the training process, likely due to the extremely nonlinear relationships underpinning the energy terms.

It should be emphasized that Chapter 5 revealed how the accurate prediction of star-state conditions and wavespeeds in turn leads to more accurate numerical results. This is true not only of transcritical and transonic conditions, but compressible flows in general:

recall from §1.1 that the HLLC solver improved resolution of material interfaces and sharp flow features by restoring the contact discontinuity wave to the star-state estimates. Hence, the remaining ML research focuses primarily on the network development process as it relates to accurate and efficient prediction of the star state, culminating in a perfunctory demonstration of FluxNets for solving the earlier transcritical shock tube of §5.3.1.

6.2 Data Preparation

Datasets for training and testing were generated using a MATLAB code that iteratively computes exact solutions to the non-ideal RP using a Collela & Glaz [36] scheme modified to use the analytical derivations of Chapter 4 and the moving normal shock equations of Appendix C. Each sample involved 2 to 6 iterations to solve the RP within $< 0.1\%$ error in all primitive star-state variables. Intercell fluxes were also computed for the preliminary training attempts on Design 2 of §6.1. Many samples failed to converge since the Newton-Raphson iterations therein guarantee neither stability nor convergence, especially in the case of highly nonlinear non-ideal thermodynamics.

The thermodynamic state space around the critical point for nitrogen gas was sampled a total of 64,000 times using uniform probability distributions across the 6 feature variables representing initial conditions. Velocities were selected so that sub-, trans-, and supersonic conditions would be represented. The limits of sampling are shown in Tab. 6.2. It is important to note that the range of training data in this study was selected to encapsulate the highly nonlinear behaviour of thermodynamic variables about the pseudo-boiling curve [136]. Network performance, and thus network size and complexity, will generally vary depending on the desired range of features and predictions that the network is expected to map accurately.

Of the 64,000 points for which exact solutions were attempted, only 21,237 points or approximately one-third converged. Liquid points were then removed (i.e. points above the boiling curve but below the critical temperature). Repeated points were also deleted to ensure a uniform distribution of samples. The resulting dataset comprised $2^{13} = 8192$ points, which was doubled to 16,384 points by swapping left and right initial conditions via symmetry of the RP. Finally, the data was mean-normalized as described in Eq. (3.5) to avoid gradient saturation during the eventual training process. Training and test splits were taken with ten-fold cross-validation on a randomly sampled subset of $16,000$ training and $1,600$ test points. Although this data preparation approach was used for nitrogen gas assuming PR EOS, it is expected to hold for other gases since the corresponding states principle involves shifting the critical locus but maintaining the mathematical form of the thermodynamic relationships [25].

Table 6.2: Ranges of initial conditions represented in the training and test datasets. The ranges are representative of transcritical flow problems [5].

Feature	Range
p_L, p_R	100 kPa to 13.5 MPa
u_L, u_R	-200 to 200 ms^{-1}
ρ_L, ρ_R	1 to 200 $\mathrm{kg\,m}^{-3}$

6.3 Network Architectures and Losses

Six variations in multi-layer perceptron architecture and loss function were trained and evaluated (Tab. 6.3), all based on Design 4 from §6.1. These specific architectures and losses were chosen to assess the impact of network complexity and loss function on performance. Network depth was varied in order to evaluate the sensitivity of final losses and errors to the network complexity—a more exhaustive investigation of network size is left to future studies, as the current research is framed around feasibility and practicality, not optimality. Each FluxNet was trained via both an MSE loss (Eq. 3.7) and a CAL version of MSE (Eq. 3.8), in order to compare traditional and physics-informed approaches. The CAL version included a Rankine-Hugoniot physical constraint of the form Eq. (3.18) with mean-normalized variables, $\kappa_{mass}, \kappa_{mom} = 1$, and $\kappa_{energy} = 0$ thus omitting the energy constraint. For non-ideal thermodynamics, the energy constraint was determined to have a deleterious effect on training as stated in §3.5 and §6.1.

All FluxNets were trained in double-precision Python 3.9 and PyTorch 1.10.2 using an Adam optimizer (Fig. 3.2). The optimizer was configured with learning rate 2.5×10^{-5}, first and second moment decay rates 0.9 and 0.999, and weight decay rate 10^{-9}. For the physics-informed loss functions, $\kappa = 10^{-6}$ was set in Eq. (3.8). These hyperparameters were empirically found to yield adequate stability and convergence. A maximum of 8000 training epochs were conducted per FluxNet, with the final weights selected based on the lowest training loss obtained in the last 4000 epochs. In order to control for varying weight initializations and choices of cross-validation splits, the same random seed was used for all FluxNets.

Table 6.3: List of FluxNet architectures and loss functions evaluated in this thesis. The time complexity refers to the number of temporally distinct floating point operations assuming parallel inference, including an additional $O(10^2)$ operations required to compute the intercell flux from the star-state conditions predicted by each FluxNet. This overhead can vary slightly depending on the number of $T = T(p, \rho)$ iterations required.

Name	Nodes	Hidden Layers	Loss	Time Complexity
64×5 MSE	64	5	MSE	441
64×5 MSE-RH	64	5	MSE + ϕ	441
64×6 MSE	64	6	MSE	509
64×6 MSE-RH	64	6	MSE + ϕ	509
64×7 MSE	64	7	MSE	577
64×7 MSE-RH	64	7	MSE + ϕ	577

6.4 Results

Learning curves and numerical results are presented for the learning-based Riemann solvers considered above. The learning curves show the progression of errors and losses during the training process of each neural network, revealing that the training process for all networks seem similarly promising in terms of mean L1 errors. The FluxNet approach also yields a significantly tighter distribution of L1 errors as compared to the Roe family of solvers. Moreover, the numerical results indicate the importance of a physics-informed approach while demonstrating improved numerical stability and accuracy relative to Roe-type solvers. The MSE and MSE-RH versions of the 64×5 FluxNet were used to solve a transcritical shock tube test case, showing that the 64×5 MSE-RH FluxNet yielded smoother, more generalizable solutions. A two-dimensional Riemann problem was also solved in which the Roe solution admits various nonphysicalities and errors that were mitigated by the 64×5 MSE-RH FluxNet. Overall, the 64×5 MSE-RH neural network achieved an order of magnitude improvement in accuracy with a 23% increase in time complexity relative to Roe-StARS, assuming parallelized computations.

6.4.1 Learning curves

The learning curves for each FluxNet in Tab. 6.3 are shown in Figs. 6.2 and 6.3. Plotted are the progression of test errors, train losses, and test losses across all epochs. The vertical axes are logarithmic to help easily distinguish the curves from each other—careful interpretation is therefore warranted since the seemingly large spikes are in fact relatively small in magnitude. The final errors and losses are compiled in Tab. 6.4. Mean L1 errors across the test dataset are reported. The mean L1 errors are also tabulated for the Roe solver and the Collela & Glaz [36] iterative solver, using the same dataset on which the FluxNets were tested.

Table 6.4: Mean L1 test errors (%) and losses (unitless) for six different learning-based Riemann solvers. For reference, typical errors for approximate and exact-iterative solvers are shown.

Solver	ρ_{*L}	u_{*L}	p_{*L}	u_*	ρ_{*R}	u_{*R}	p_{*R}	$\mathcal{L}_{train}[10^{-6}]$	$\mathcal{L}_{test}[10^{-6}]$
64 × 5 MSE	0.28	0.09	0.11	0.10	0.17	0.10	0.09	3.14	1.44
64 × 5 MSE-RH	0.29	0.08	0.17	0.07	0.22	0.08	0.13	3.21	1.40
64 × 6 MSE	0.26	0.09	0.15	0.12	0.19	0.09	0.16	3.04	1.45
64 × 6 MSE-RH	0.25	0.13	0.17	0.14	0.22	0.14	0.17	2.94	1.43
64 × 7 MSE	0.28	0.14	0.13	0.13	0.27	0.13	0.16	2.92	1.51
64 × 7 MSE-RH	0.94	1.08	0.96	1.05	0.84	1.13	0.77	5.82	3.7
Roe	4.20	1.04	1.74	1.04	0.27	1.04	1.74	N/A	N/A
Exact	0.01	0.01	0.01	0.01	0.01	0.01	0.01	N/A	N/A

All FluxNets except 64 × 7 MSE-RH achieved final test errors of $O(10^{-1})\%$. The networks tended to elicit greater errors in density (around 0.2 to 0.3%) and smaller errors in velocity and pressure (around 0.1 to 0.2%). The lower errors on velocity and pressure

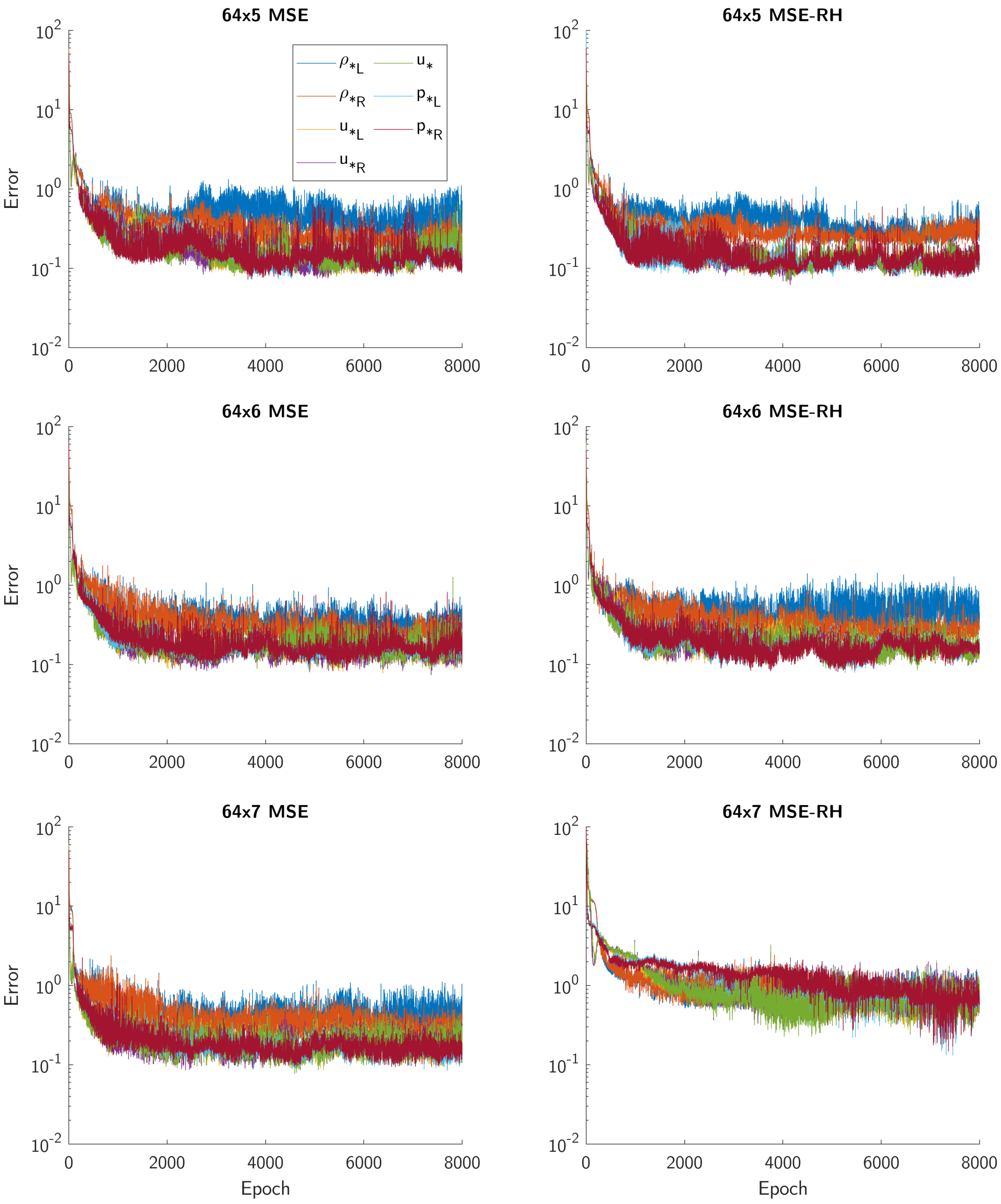


Figure 6.2: Learning curves showing mean L1 test errors for all multi-layer perceptrons in Tab. 6.3. Network depth does not significantly affect the convergence rates nor the final errors and losses. All networks except 64×7 MSE-RH converged to errors $< 1\%$.

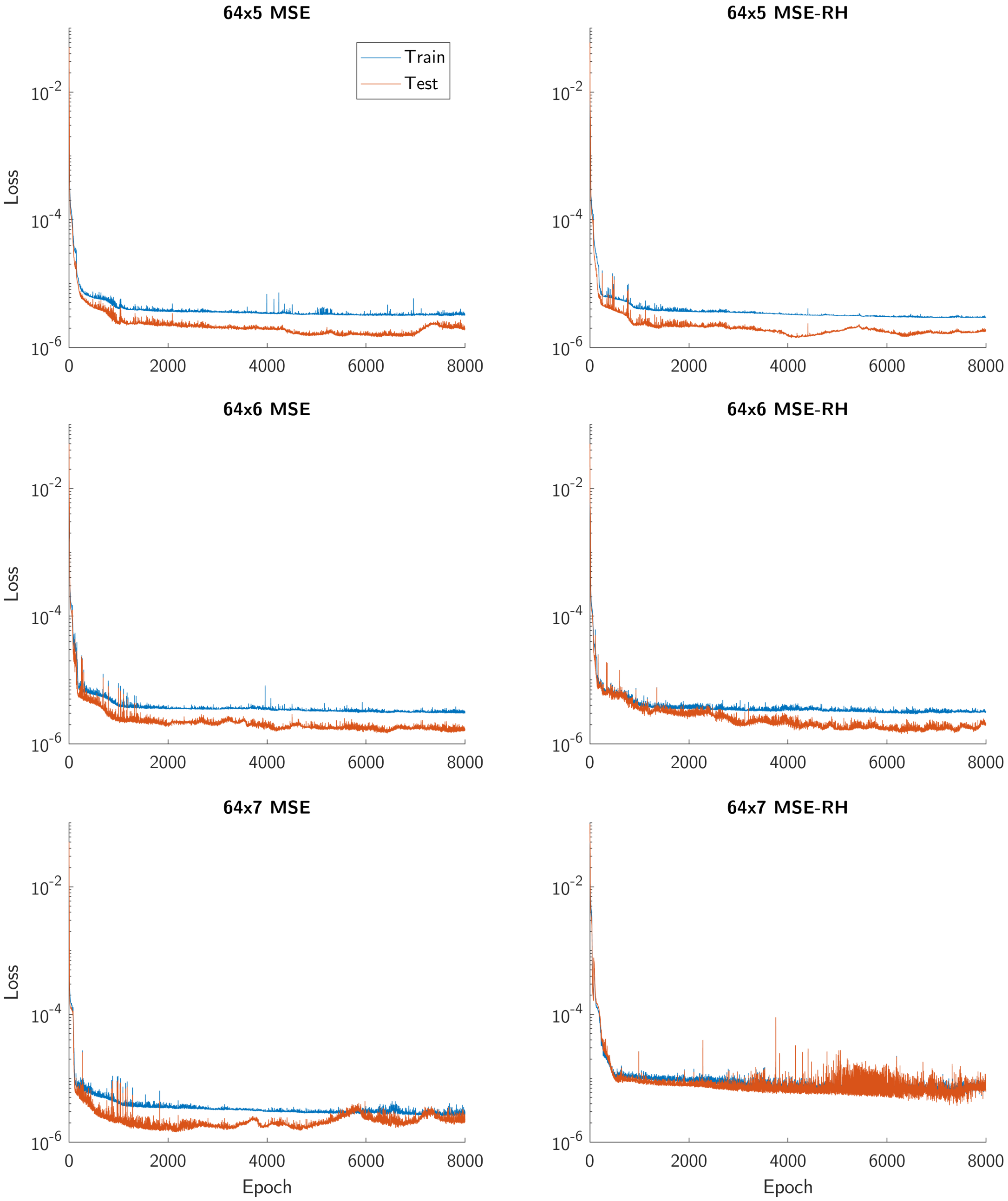


Figure 6.3: Learning curves showing train and test losses for all multi-layer perceptrons in Tab. 6.3. All 64 × 5 and 64 × 6 networks converged to similar losses, while the 64 × 7 MSE-RH network exhibited poor convergence. The lower test losses are likely caused by bias in the test sample, however final train and test errors were similar.

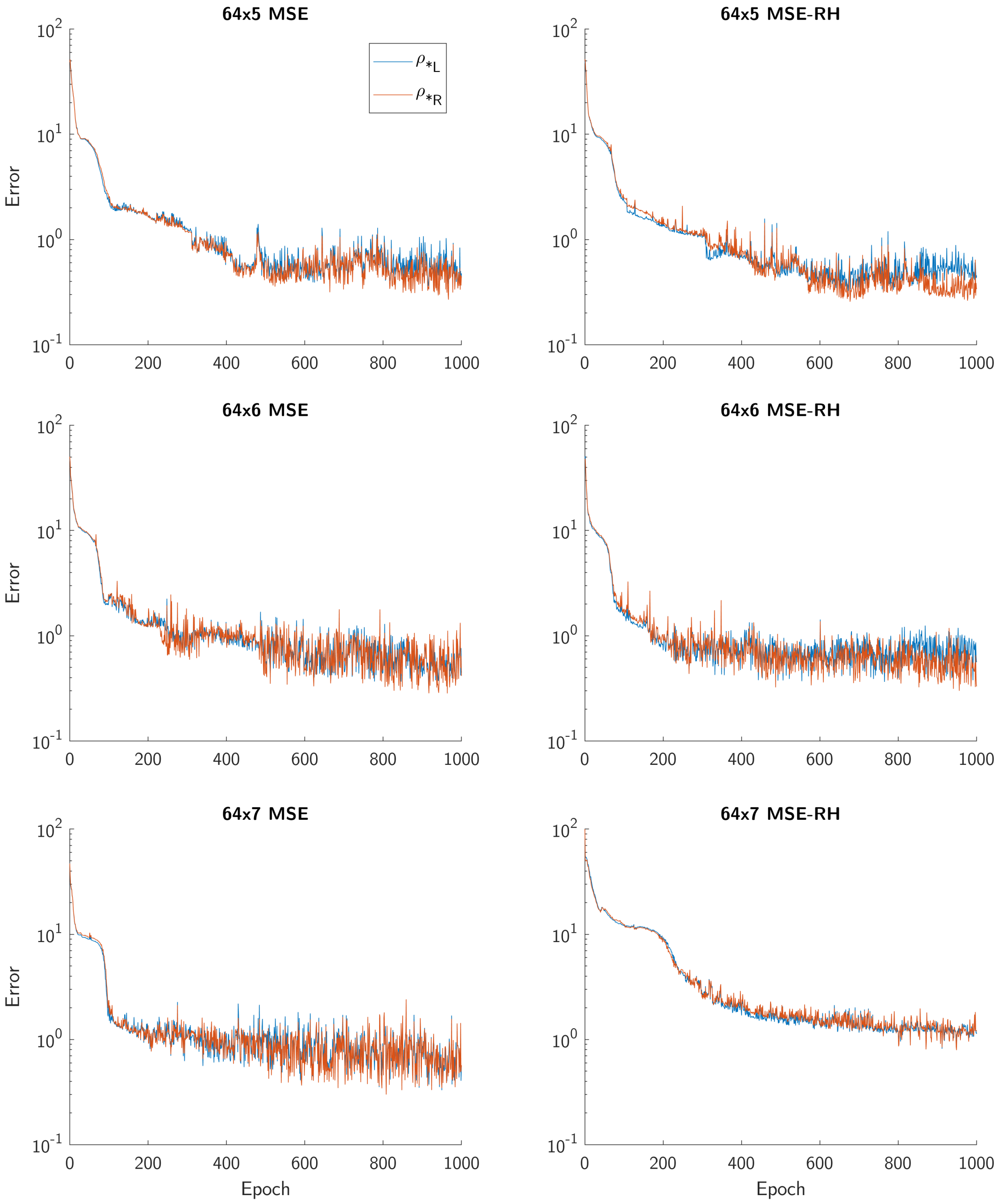


Figure 6.4: Learning curves truncated at 1000 epochs, showing test errors only for ρ_{*L}, ρ_{*R} across all multi-layer perceptrons trained from Tab. 6.3. Here it is visible that the MSE networks exhibit less stable convergence compared to the physics-informed MSE-RH networks.

are likely due to the physics of the contact discontinuity wave, across which pressure and velocity should be constant. Such lower errors are observed for both the MSE and MSE-RH FluxNets, which suggests that the networks were able to learn this pattern at least partially from the data alone. By contrast, density can change across a contact discontinuity (*i.e.* since the contact discontinuity represents the moving material interface between the initial left and right states). Learning of the pressure-density relationship may have been hampered by the absence of the Rankine-Hugoniot energy constraint—which, again, was omitted due the highly nonlinear physics causing unstable training behaviour.

The 64×7 MSE-RH network bears anomalous final test errors, losses, and training curves. Final train and test losses are nearly double that of the other FluxNets, and unlike all other FluxNets, the loss curves start to oscillate noticeably around 4000 epochs. It is worth noting that attempts to vary the learning rate still resulted in the same convergence behaviour and noisy loss curves. Taken together, these observations imply that the difference is related to both the number of weights and the Rankine-Hugoniot constraint in 64×7 MSE-RH. Neural networks become more prone to overfitting as the number of weights and number of training epochs increase. It is plausible that in 64×7 MSE-RH, the oscillations arise as the Rankine-Hugoniot constraint competes against overfitting tendencies due to greater number of weights. This would also explain why the same oscillations and poor performance are not observed in the 64×7 MSE network. The 64×7 MSE network is free to overfit, and indeed, it has the lowest train loss and greatest test loss of all networks examined. The overfitting explanation is further supported by the noisy numerical results in the next subsection, §6.4.2. This hypothetical limit on the number of weights to prevent overfitting is, however, also a function of the training data. Were a different range of the thermodynamic state space sampled in the training data, then this threshold would likely change as well. Different sampling techniques, such as a gradient-weighted approach that increases point density in regions of steep thermodynamic changes, could also shift the threshold at which overfitting becomes problematic.

In comparing the networks among each other, it is clear that the 64×5 MSE and MSE-RH networks are favourable for further evaluation, due to their relative simplicity yet similar test errors to the more complex architectures. While the error plots in Fig. 6.2 show all curves spanning all training epochs and predicted variables, finer insights may be gleaned by considering only a subset of the curves. Fig. 6.4 shows the density learning curves for the first 1000 epochs. The error curves for the MSE-RH networks appear to oscillate less and at smaller amplitudes than the equivalent MSE networks, potentially due to the less naive, physics-informed search process. Except for the 64×7 MSE-RH network, all networks also approached $O(1)\%$ error around 200 epochs with both MSE and MSE-RH networks converging at roughly the same rate. Despite the apparent noise in mean L1 errors, we can be satisfied that all networks converged to an acceptable degree given the decreasing, relatively noise-free trends in the loss curves, excepting 64×7 MSE-RH.

Neural networks thus provide an order of magnitude reduction in the mean L1 test errors compared to the Roe family of solvers, which ranged from 1 to 4% as shown in Tab. 6.4. Though this improvement may seem marginal, even minor inaccuracies in transcritical numerical simulations can destabilize solutions and lead to spurious results. In addition, only the mean L1 test errors have been discussed. It is compelling to analyze the *distribution* of L1 errors across the entire dataset.

Fig. 6.5 contains box plots of the L1 errors across the entire dataset including train and test samples. Data points are identified as outliers if they are greater than $q_e + w(q_3 - q_1)$ or less than $q_1 - w(q_3 - q_1)$ where w corresponds to 3σ, q_1 is the first quartile, and q_3 is the third quartile. The 64×5 MSE-RH FluxNet is evidently more precise in that its error distributions are significantly tighter across all star-state variables. Additionally, the mean errors of the 64×5 MSE-RH FluxNet are more closely aligned with zero, whereas the mean error of the Roe family of solvers overpredicts density and pressure by $\sim 2\%$, and underpredicts velocity by $\sim 1\%$. The error distribution may also be analyzed through a probability density plot, such as Fig. 6.6 which plots the error probability density function in ρ_{*L} by normalizing a histogram of errors against the total number of data points. Compared to 64×5 MSE-RH's dense clustering around zero, the mode of the Roe distribution is at zero although the errors are more evenly spread out with a smaller peak at 2.5%. Interestingly, there is a local maximum in error probability for the 64×5 MSE-RH network around 4%. Further analysis reveals that this is due to points in the dataset where feature and response variables are identical save for $O(10^{-2})$ or smaller errors arising due to the limited precision and stability of the exact-iterative solver when initial conditions are similar.

6.4.2 Transcritical shock tube

The 64×5 MSE and 64×5 MSE-RH FluxNets were used to solve the transcritical shock tube test case from §5.3.1, leveraging the first-order Godunov scheme described in 2. Unfortunately, the 64×5 MSE network caused significant numerical instabilities that produced

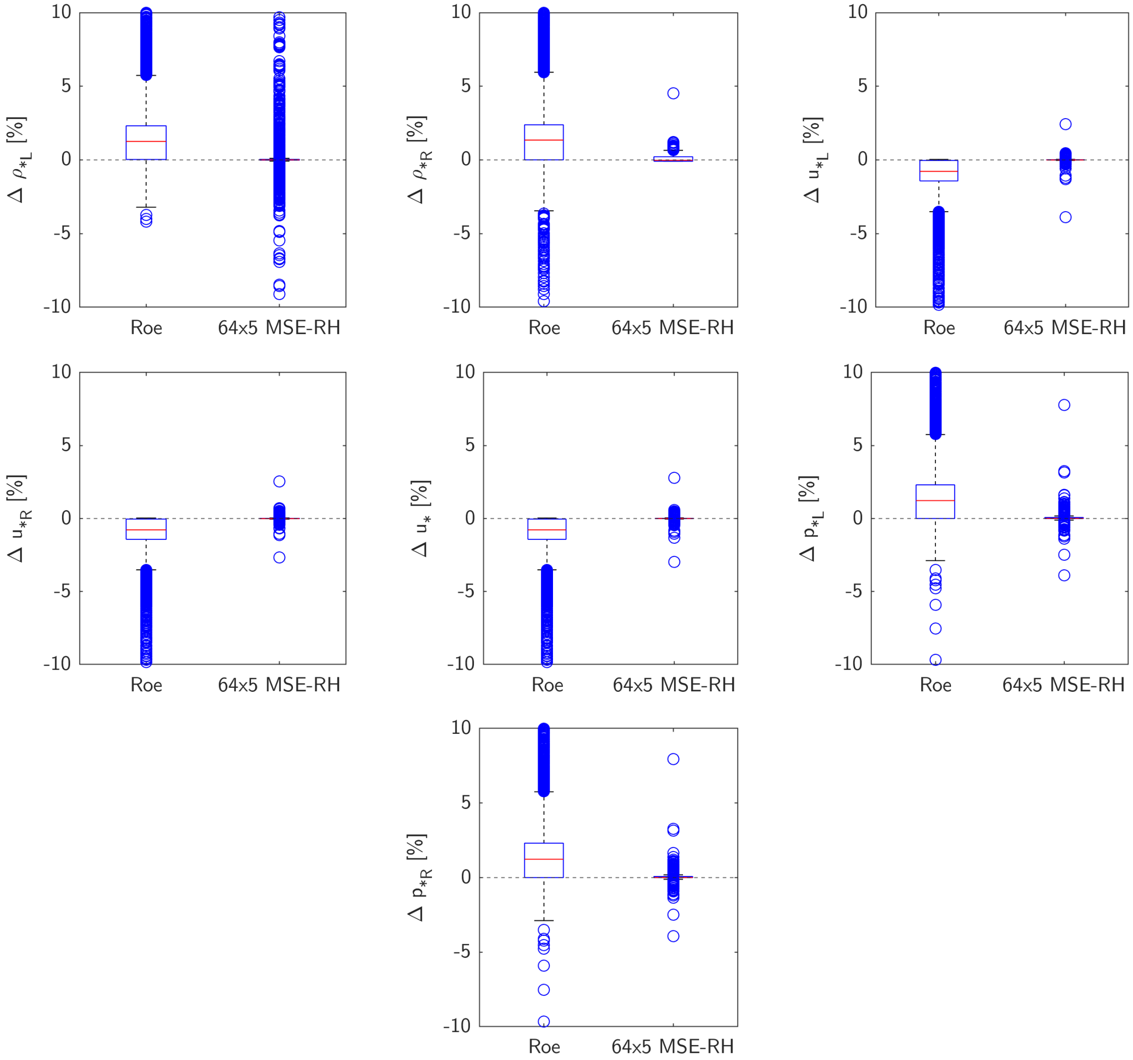


Figure 6.5: Box plots showing the distribution of errors in all star-state errors estimated by 64 × 5 MSE-RH FluxNet versus the Roe family of solvers. The red lines indicate the mean, the boxes enclose data from the 25^{th} to the 75^{th} percentile, and the circles indicate outliers. Vertical axes are rescaled to make boxes more visible—many outliers, especially for the Roe results, sit beyond the vertical limits. Errors were calculated across the entire dataset including train and test samples.

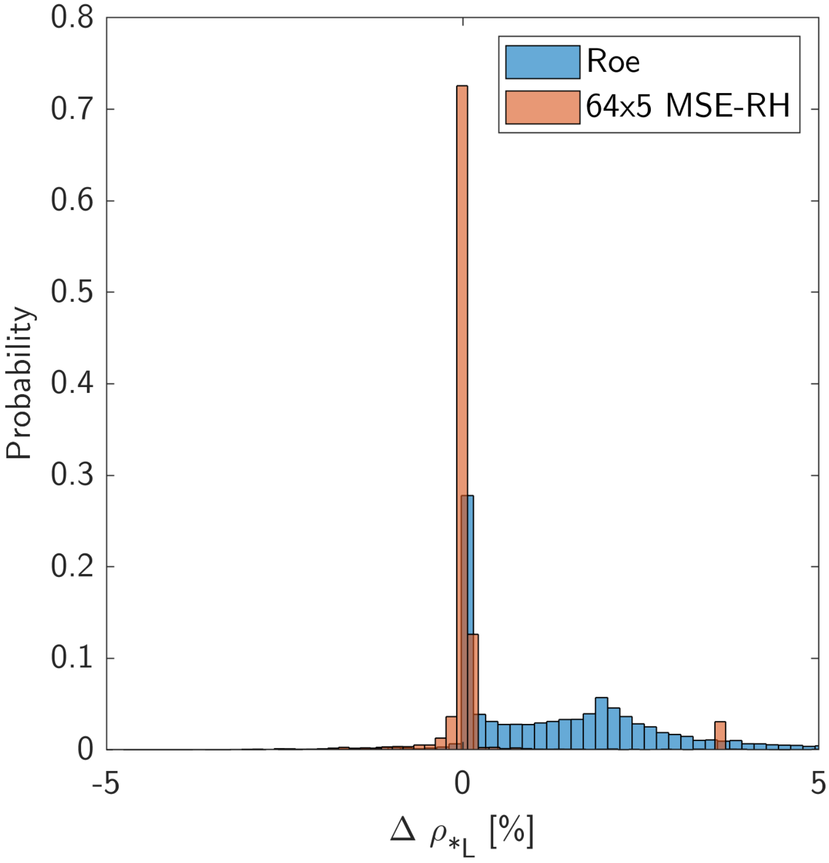


Figure 6.6: Probability of errors in ρ_{*L} for the 64 × 5 MSE-RH FluxNet versus the Roe family of solvers. The horizontal axis is rescaled to improve resolution near the mean—many outliers sit beyond the horizontal limits. Errors were calculated across the entire dataset including train and test samples.

negative temperatures and internal energies. To determine whether this would always be the case, the 64×5 MSE network was tested on a slightly modified version of the original problem initialized to a lower left pressure, density, and velocity as well as slightly higher right pressure, density, and velocity. This change was designed to reduce the gradients during the initial time steps before the self-similar wave structure of the RP has stabilized. On the modified problem, 64×5 MSE did not produce negative thermodynamic variables (Fig. 6.7). Notwithstanding, the results are unacceptable. There are visible errors throughout the rarefaction wave region, as well as significant noise near the tail of the rarefaction where the wave meets the left star state. The region to the right of the shock is also unusually diffusive considering that information should not propagate upstream of a normal shock even in a first-order scheme. This would only be possible if the wavespeeds calculated from the star-state predictions by the FluxNet resulted in fluxes taken from an incorrect state of the RP. This may also explain why oscillations are most severe at interface between the rarefaction tail and left star state: the flux solver needs to accurately calculate wavespeeds in order to ensure entropy correctness and proper upwinding. Considering that the 64×5 MSE network had lower train losses and higher test losses than the 64×5 MSE-RH network during training (Tab. 6.4), this adds to the pool of evidence that the non-physics-informed network overfitted the data.

Meanwhile, 64×5 MSE-RH generated smooth and physically consistent solutions with the original transcritical shock tube as posed in §5.3.1. To facilitate comparison with the exact and Roe-StARS solutions computed earlier, Fig. 6.8 plots solutions to the original problem rather than the modified problem used for 64×5 MSE. Looking at $x = 0$ where Roe normally suffers from an expansion shock nonphysicality, it is immediately clear that

Table 6.5: Absolute L1 errors at $x = 0$ for primitive variables ρ, u, p_r achieved by the Roe, Roe-StARS, and 64×5 MSE-RH solvers in the non-ideal transcritical shock tube of Fig. 6.8. Errors for Roe-Harten are repeated here from Tab. 5.2 for reference. Percentage reduction in the Roe error is shown in parentheses.

Error	Roe	Roe-Harten	Roe-StARS	64×5 MSE-RH
$\Delta\rho$	33.0	13.5 (-59.2%)	12.9 (-61.1%)	11.8 (-64.2%)
Δu	-56.2	-15.2 (-73.0%)	-13.8 (-75.4%)	-11.8 (-79.1%)
Δp_r	0.680	0.222 (-67.4%)	0.208 (-69.4%)	0.171 (-74.9%)

64×5 MSE-RH not only alleviates this error but is more smooth and accurate in this region than even the Roe-StARS solver (Tab. 6.5). Minor differences between the Roe-StARS and 64×5 MSE-RH solution are also observed throughout the solution. For instance, 64×5 MSE-RH predicts a normal shock position slightly to the left of that predicted by Roe-StARS, closer to the exact solution.

6.4.3 Two-dimensional Riemann problem

In order to further evaluate the implications of a FluxNet approach, a 2D Riemann problem test case was solved with the 64×5 MSE-RH FluxNet and the Roe solver [49]. Similar to the 1D Riemann problem, the 2D problem serves as a test case for verifying numerical codes. Additionally, in higher dimensions, more complex wave interactions such as Mach diffraction and reflection can occur. More elaborate schemes and meshes may also be deployed to improve accuracy and stability because waves propagate not only in orthogonal directions but at angles oblique to the unit basis vectors [24, 7]. However, for the present purposes, it suffices to consider a uniform Cartesian grid and just one of the potential configurations of a 2D Riemann problem—it is the the Riemann solver, and less so the

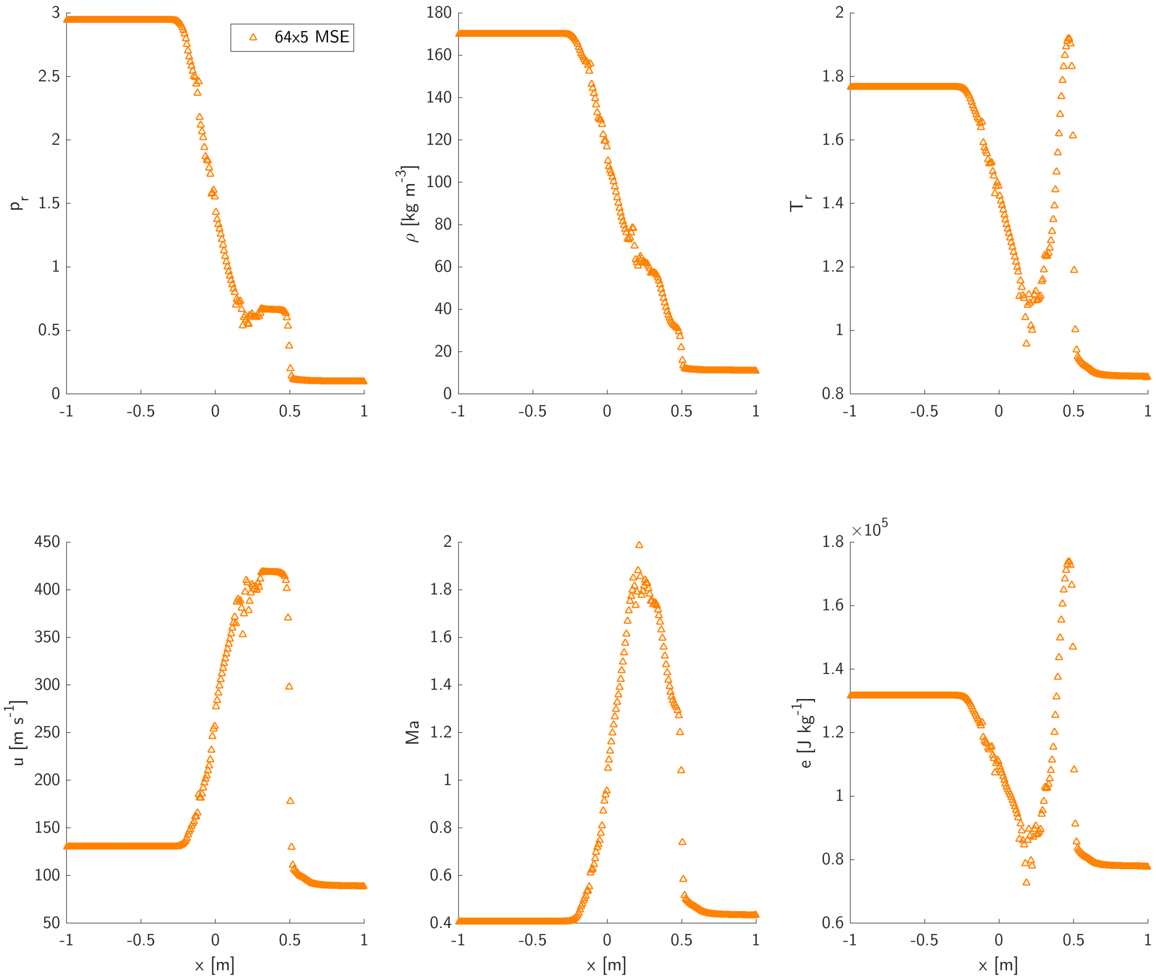


Figure 6.7: Numerical results for a transcritical nitrogen shock tube solved with the 64×5 MSE FluxNet, showing oscillatory and spurious solutions. Non-ideal thermodynamics and the Peng-Robinson equation of state [2] are applied.

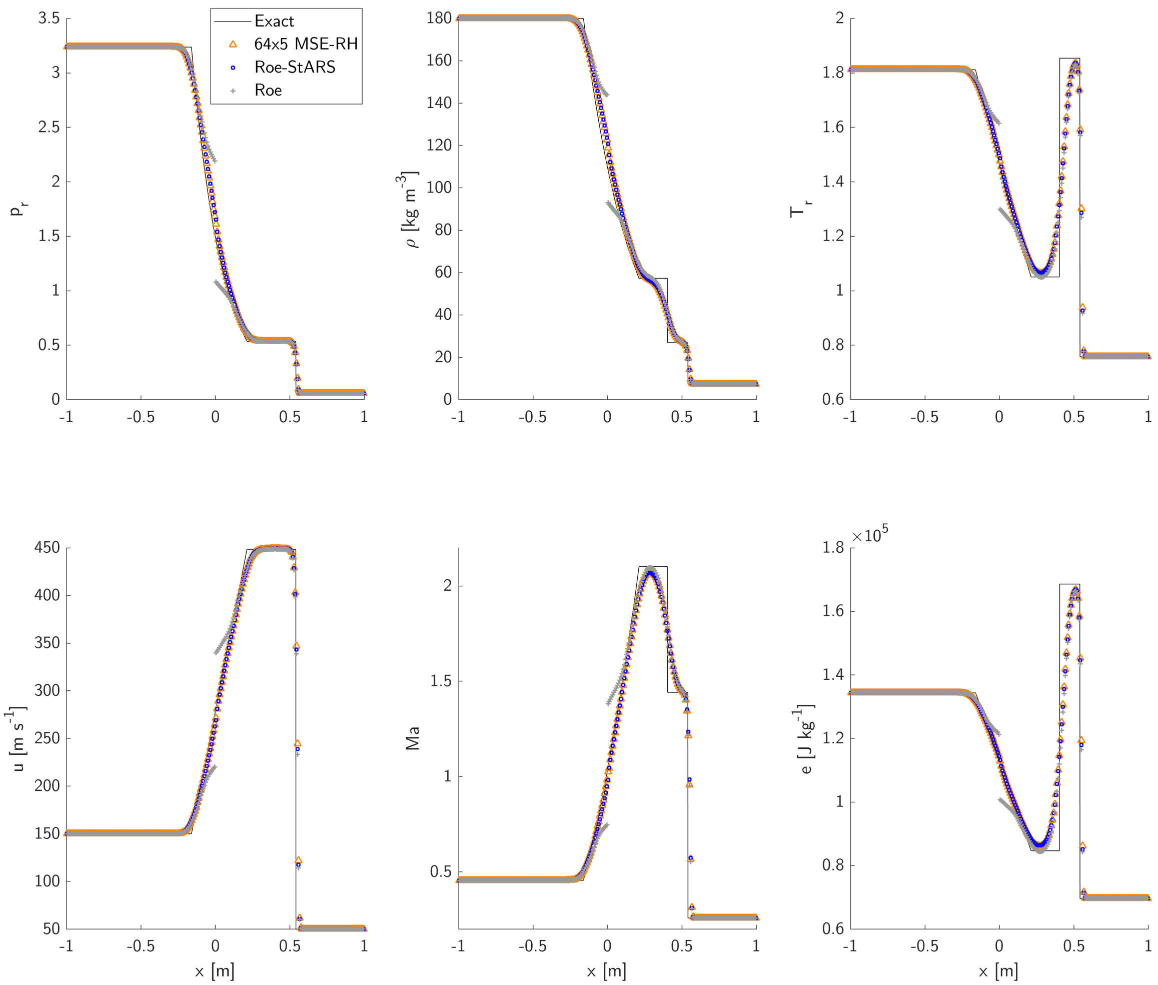


Figure 6.8: Numerical results for a transcritical nitrogen shock tube solved with non-ideal thermodynamics and the Peng-Robinson equation of state [2]. Results for the Roe, Roe-StARS, and 64×5 MSE-RH FluxNet solver are shown.

scheme, that is under consideration.

Fig. 6.9 provides the initial conditions and expected wave evolution for a transcritical 2D Riemann problem inspired by the case studied earlier in §5.3.4. Here, u is the velocity in the x-direction and w is the velocity in the y-direction. Transparent boundary conditions were set at $x = -1m, 1m$ and $y = -1m, 1m$. A CFL number of 0.45 was used. The computational domain consisted of 256 × 256 cells. The numerical scheme was advanced until $t = 0.0009s$. The medium comprised transcritical nitrogen gas modelled with the Peng-Robinson state equation [2] and full non-ideal thermodynamics. Results for the Roe and 64 × 5 MSE-RH solvers are shown in Figs. 6.10 and 6.11. Analytical solutions are unavailable for the 2D Riemann problem, therefore the subsequent discussion is based on concepts developed for the 1D Riemann problem [24, 152].

The most obvious difference between the Roe and FluxNet results is that nonphysical shocks are greatly diminished in the FluxNet solution. In the Roe solution, expansion shock fronts manifest along the negative x and y axes, even extending slightly past the origin. These shocks remain stationary with respect to time, thus preventing information from propagating across these spurious boundaries. By contrast, the FluxNet solution shows a finer, more even distribution of characteristic waves that is typical of isentropic expansion fans. The nonphysical shock fronts also do not extend as far past the origin. The mitigation of expansion shocks further changes the nature of the peak that forms just to the upper right of the origin. In the FluxNet solution, the pressure, density, temperature, and internal energy are all higher in this region. This is in contrast to the Roe solution, where the artificial shocks produce relatively low post-shock conditions. Interestingly, the maximum internal energy is higher in the Roe solution, occurring near $x = 0.4m, y < 0$

and $x < 0, y = 0.4m$. This is likely due to the highly nonlinear thermodynamics whereby small variations in pressure, volume, or temperature can result in noticeable differences in more complex thermodynamic quantities such as internal energy.

Finally, it is worth noting that the FluxNet solution contains small kinks in u along the $y = 0$ expansion shock and in w along the $x = 0$ expansion shock. These kinks do not appear in the Roe solution. These errors may be attributed to the simplistic Cartesian formulation of the numerical scheme. Expansion shocks that run orthogonal to the flow direction (*e.g.* $x = 0$ expansion shock in w, or $y = 0$ expansion shock in u) are not always effectively mitigated due to the directional independence of the unit basis vectors. Such behaviour can likely be corrected with advanced meshing or flux weighting techniques [24]. The essential insights from this 2D test case are that the FluxNet approach successfully mitigates nonphysicalities, and even small errors in thermodynamic state can result in significant errors when dealing with full non-ideal thermodynamics.

6.4.4 Comparison of accuracy and time complexity

Sufficient data is now available to perform a critical comparison of the different types of Riemann solvers examined throughout this thesis. Entropy violation errors are summarized in Tab. 6.5, time complexity is shown in Tab. 6.6, mean errors are contained in Tab. 6.4, and error distributions are plotted in Fig. 6.5.

As far as computational efficiency is concerned, it is manifest that any entropy-stable Riemann solver involves $O(10^2)$ greater time complexity than the unmodified Roe solver. Whereas Roe-Harten and Roe-StARS exhibit minor differences in time complexity between

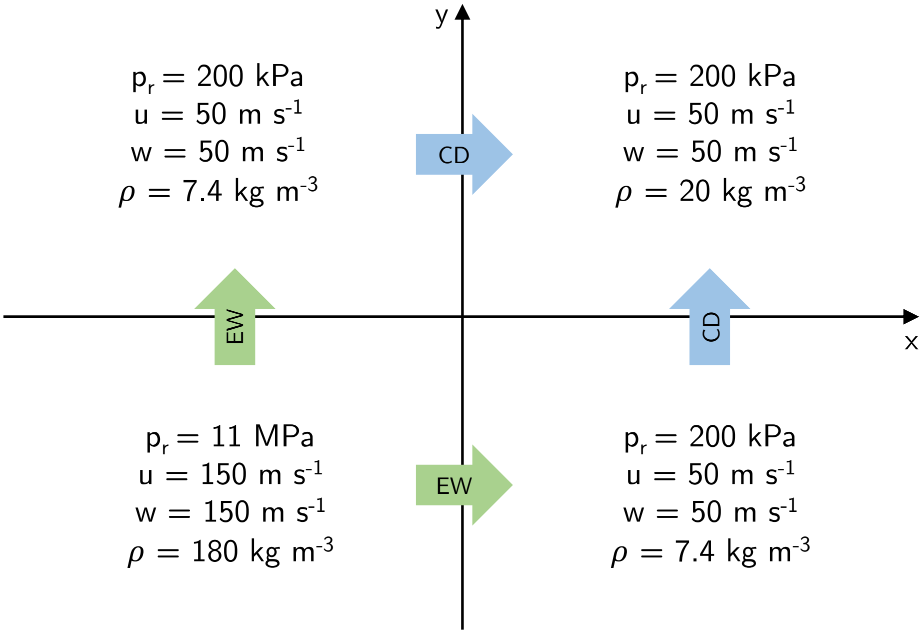


Figure 6.9: Initial conditions for the 2D Riemann problem test case considerd in this study, based loosely on the case considered in §5.3.4. The expected wave pattern is shown consisting of two contact discontinuities (*i.e.* slip lines in 2D) and two expansion waves.

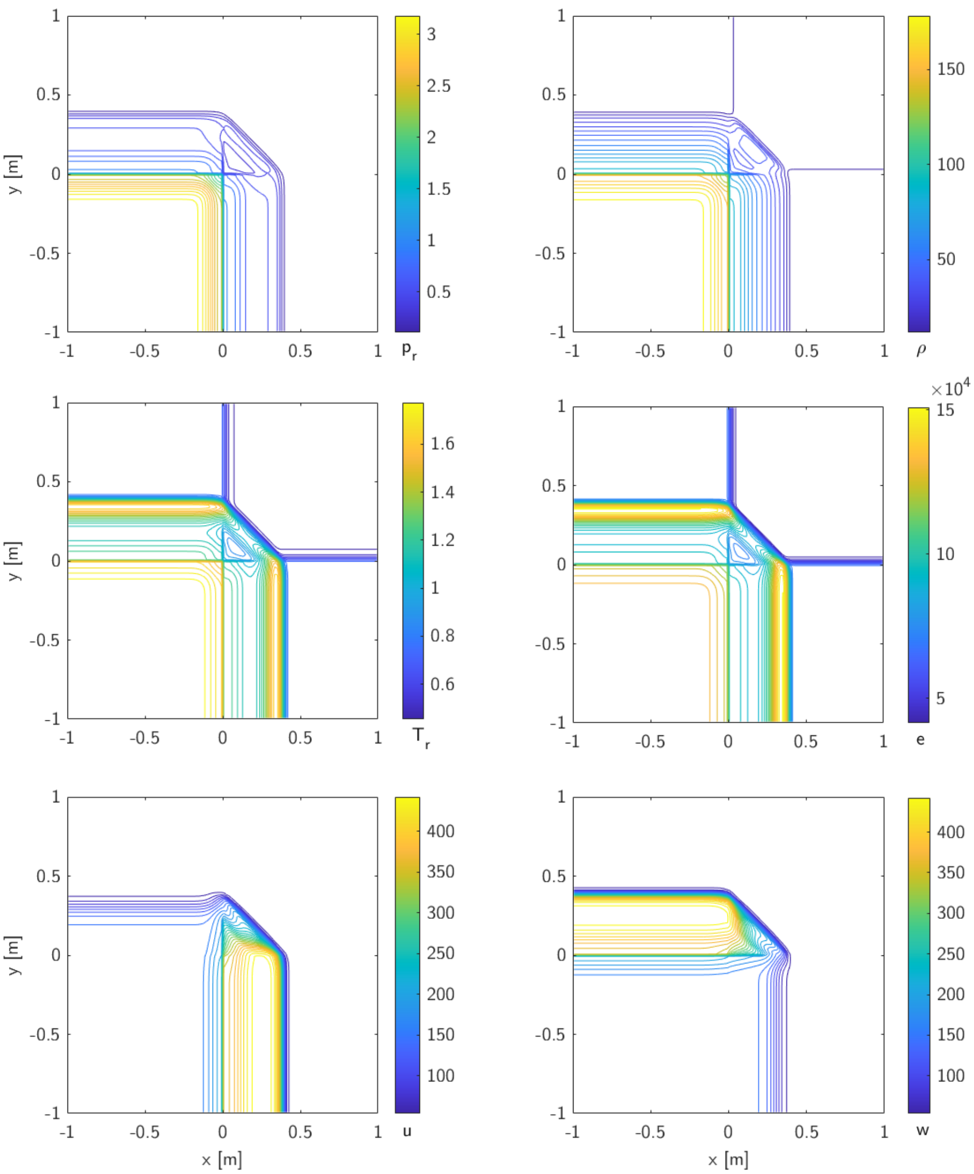


Figure 6.10: Solution to the 2D Riemann problem test case at $t = 0.0009s$, as computed with a Roe solver. Two expansion waves and two slip lines are visible, along with non-physical expansion shocks along the negative x and y axes. e here is presented in units of $\mathrm{J\,kg^{-1}}$.

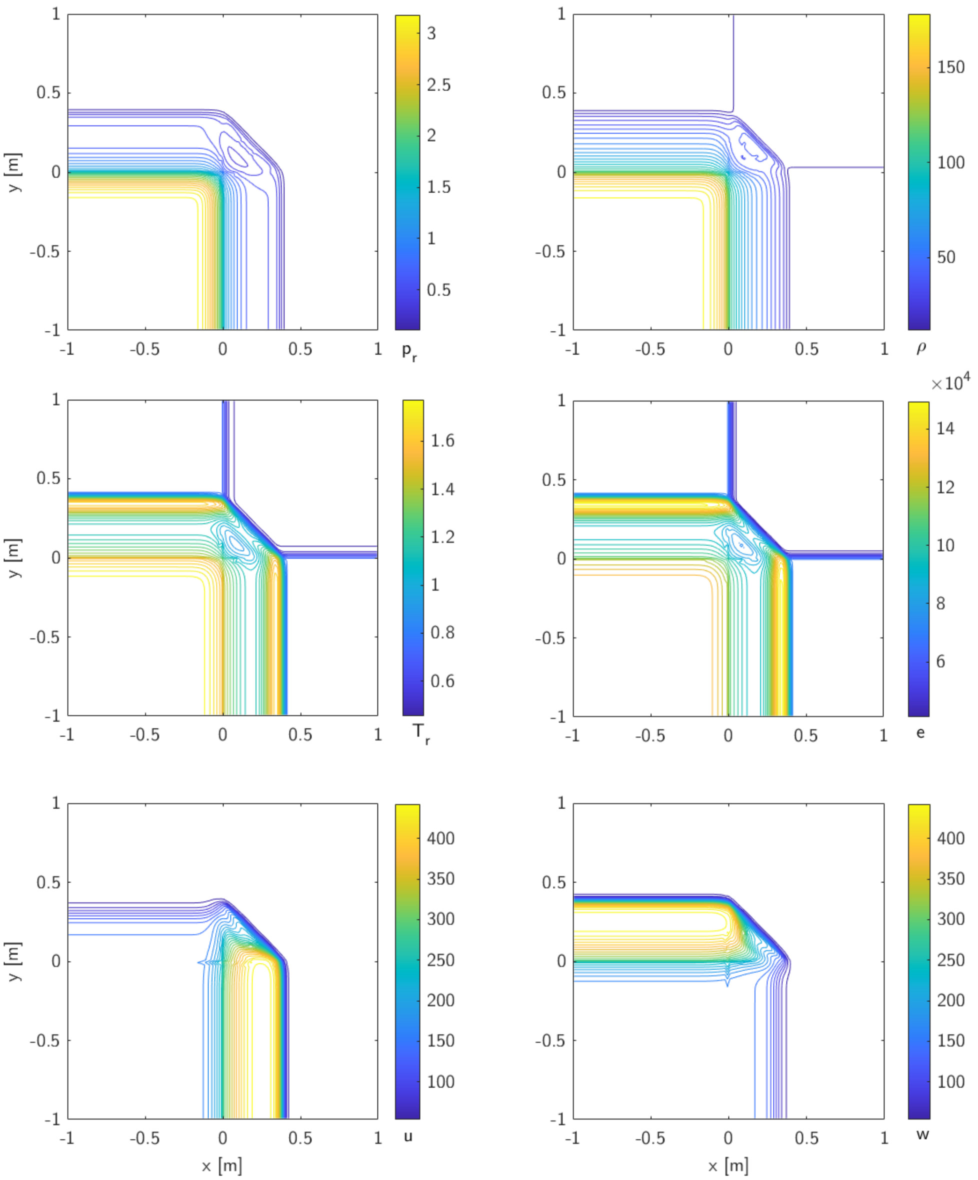


Figure 6.11: Solution to the 2D Riemann problem test case at $t = 0.0009s$, as computed with the 64×5 MSE-RH FluxNet. Two expansion waves and two slip lines are visible, and unlike the Roe solution, the FluxNet solutions significantly mitigate any nonphysicalities. e here is presented in units of $\mathrm{J\,kg^{-1}}$.

Table 6.6: Time complexity required to solve the flux of a single cell interface using the Roe, Roe-StARS, 64 × 7 MSE-RH, or exact-iterative solvers. Time complexity for Roe-Harten is repeated here from Tab. 5.3 for reference.

	No Entropy Violation		With Entropy Violation	
Solver	Ideal Gas	PR Gas	Ideal Gas	PR Gas
Roe	123	157	123	157
Roe-Harten	229 (+86%)	331 (+111%)	239 (+94%)	341 (+117%)
Roe-StARS	229 (+86%)	331 (+111%)	285 (+132%)	387 (+146%)
64 × 5 MSE-RH	441 (+258%)	441 (+181%)	441 (+258%)	441 (+181%)
Exact	524 (+326%)	1799 (+1046%)	524 (+326%)	1799 (+1046%)

ideal and non-ideal gases, as well as whether an entropy violation is present, the FluxNet approach is equally demanding under all circumstances. It is necessary to highlight as well that methods such as Roe-Harten require tuning the appropriate amount of artificial diffusivity to balance entropy stability with sharpness, whereas Roe-StARS needs no tuning and is thus more generic in its design and application. FluxNets, on the other hand, are indirectly the product of heuristic decisions such as the training data and hyperparameters that influence the learning process.

In terms of errors and accuracy, the Roe solver is likely inappropriate for many transcritical flows where entropy violations must be avoided. Roe-Harten, Roe-StARS, and FluxNets offer varying levels of entropy stability, mean error, and error distribution that are each beneficial for different flow problems. When transcritical flow problems span relatively small regions of the thermodynamic state space, the Roe-Harten or Roe-StARS solvers are most appropriate due their efficiency yet effective mitigation of entropy errors. Roe-StARS in particular avoids re-tuning and unnecessary diffusivity. However, when flow problems span relatively large thermodynamic variations and numerous time steps, a learning-based approach may be more sensible. The error distribution plots suggest that over a wide thermodynamic state space, traditional Roe-type solvers exhibit greater biases and a wider range of errors in star-state conditions compared to the physics-informed FluxNet. Such errors would naturally accumulate as simulation time progresses, as discussed in the earlier numerical results of §5.3.

Chapter 7

Conclusion

Earlier in §1.2, we stated that the research objectives were to explore novel exact, approximate, and learning-based approaches to solve the RP for non-ideal gases, as well as compare the performance between them. Now with the objectives satisfied, the major research contributions and the key findings therein may be summarized. The ideas presented in this thesis endeavour to improve numerical simulations of high-speed transcritical and supercritical flows, where the accuracy, efficiency, and thermodynamic consistency of Riemann solvers is essential to achieving physically consistent results.

In Ch. 1, the history and modern literature on the RP—particularly with non-ideal thermodynamics—was critically reviewed for knowledge gaps. Chs. 2 and 3 established the prerequisite CFD and ML theory related to this research. The first contribution centred on the derivation of novel approaches to solving normal shocks and centre expansion waves subject to arbitrary EOS. The second contribution leveraged the newly derived analytical equations for an expansion wave to create so-called *structurally complete* approximate Riemann solvers from pre-existing three-wave solvers. Also analyzed were the occurrence and behaviour of entropy violations in 1D and 2D problems. Finally, the third contribution demonstrated the feasibility of training compact neural networks to solve the RP in a physics-informed manner.

7.1 Key Findings and Implications

Amidst the broad research contributions of this thesis, a number of critical conclusions may be established in relation to high-speed flow simulations with non-ideal thermodynamics.

The continuity-momentum and continuity-energy formulation of the stationary normal shock problem is conducive to fast convergent and stable iterations when seeking exact solutions. Contrary to what some studies seem to assume, the surfaces with respect to post-shock v and T are smooth and well-behaved. By initializing to the ideal gas solutions, it is relatively quick to iteratively compute exact solutions for normal shocks with non-ideal thermodynamics.

There exists a closed-form analytical solution to expansion waves with non-ideal thermodynamics, which may be derived by leveraging a novel domain mapping from space and time coordinates to characteristic curves. As a result of these derivations, the mathematical shape of expansion waves may be expressed as an explicit function of space and time. Also, this functional form arises from the governing equations only—thermodynamics acts to stretch and shift the wave in space. This implies that the derivations may be used to design or improve future Riemann solvers regardless of the thermodynamic conditions, be it low or high pressure flow, ideal or non-ideal EOS, or whether one EOS is used or another.

A structurally complete approximate Riemann solver, or StARS, approach may be used to restore the expansion wave analytically to pre-existing linearized solvers, thus rendering them entropy-stable. An unfortunate limitiation of linearized Riemann solvers is that solutions are collapsed into piecewise-constant regions with discontinuous jumps. Inviscid rarefactions are therefore modelled as an artificial jump, introducing an entropy violation. By applying a StARS modification to existing solvers such as the Roe solver, it is possible to furnish a simple yet analytically correct entropy fix that does not require tuning of artificial diffusivity as with some traditional fixes.

The occurrence and magnitude of entropy violation errors are more prevalent under conditions of transcritical or supercritical flow with large gradients. Thus, entropy fixes are warranted. Using a rigorous analysis of the conditions under which transonic rarefactions occur, it is possible to derive and plot curves that define the thermodynamic state boundaries where entropy violations occur. These entropy errors accumulate and propagate through space and time, and lead to numerical instabilities or negative pressures, temperatures, and negative internal energies, thus heightening the importance of an accurate and stable Riemann solver. For fluid simulations involving high-speed non-ideal flows with large gradients, the computational cost of an entropy-stable Riemann solver (*e.g.* using a StARS fix) is justified.

In numerical simulations, step functions and regions of large gradients are

indistinguishable. It is therefore useful to observe the behaviour of numerical schemes using a *gradient* RP. In this problem setup, the initial step function is replaced by a gradient of varying slope. Entropy violations in Riemann solvers do not necessarily exhibit monotonic decrease with respect to decreasing magnitude of initial gradients. In other words, weaker shocks or shallower flow gradients do not always result in smaller entropy errors.

Learning-based Riemann solvers, or *FluxNets*, offer yet another option to increase the accuracy of numerical simulations at a runtime complexity in between that of exact and approximate solvers. Whereas the error in star-state primitive variables may reach as high as 9% in the case of entropy-fixed approximate solvers, FluxNet errors are typically well below 0.3%. However, FluxNets also possess around 13% greater time complexity than with the StARS approach assuming parallelized matrix computations. Because of this, FluxNets may be best suited to those flow problems where numerical stability is especially sensitive to noise and errors. Even so, alternate data-driven techniques such as look-up tables or interpolation curves may be advisable for computational efficiency.

The learning curves of FluxNets with different network depths seem to suggest that predictive accuracy on test data is only loosely correlated with network size and complexity. In fact, the largest neural network that was tested exhibited poor stability during the training process, due to competition between achieving low loss on the training data versus low loss on the physics-informed penalty.

Physics-informed loss functions provide smoother, more generalizable solu-

tions than traditional loss functions that are purely data-driven. The Rankine-Hugoniot jump conditions applied at the contact discontinuity suffice as a natural physical constraint. However, it may be advisable to avoid an energy conservation constraint with non-ideal thermodynamics, due to the significant non-linearities it can introduce. In this thesis, networks demonstrated poorer convergence and greater losses when the energy constraint was included.

7.2 Future Research

Fundamental ideas are developed and tested in this thesis, however their practical application and extension to more complex problems remain outstanding. Some future research directions include:

1. Explore characteristic curve domain mappings for other types of shock problems (*e.g.* detonation) and types of matter (*e.g.* solid, liquid) in which hyperbolic partial differential equations may permit analytical solutions to emerge.

2. Extend the study of non-ideal shock and expansion waves to include Lorentz transformations, which could be relevant for future astrophysical fluid dynamics.

3. Extend the exact, approximate, and ML Riemann solvers of this thesis to contexts where mixtures or other terms in the Navier-Stokes equations are relevant. Thermodynamic mixing rules, viscosity models, and additional heat transfer properties may need to be considered when adapting the approaches of this thesis.

4. Apply StARS to restore the expansion wave to other Riemann solvers, *e.g.* HLL, HLLC, or AUSM, and extend the numerical tests to include more practical 2D and 3D problems with higher-order numerical schemes. As part of this, compare low-order schemes with more advanced Riemann solvers against high-order schemes with simpler Riemann solvers, in terms of accuracy and computational cost.

5. Experiment with alternate implementations of StARS for the Roe solver, including designs that take advantage of massively parallelized matrix computations, in order to further reduce time complexity and accelerate practical computations.

6. Perform a broader sensitivity analysis of FluxNet performance to various factors, such as training data sampling techniques, breadth of training data, number of network nodes, and loss functions. Given the highly nonlinear thermodynamics around the critical point of gases, it is possible that alternate data generation and training techniques could significantly influence the minimum network complexity required to achieve a certain level of accuracy.

7. Test compact convolutional or recurrent networks as an alternative to traditional multi-layer perceptron approaches. Early work has been promising in the use of convolutional networks for estimating weights in the weighted essentially non-oscillatory scheme [106].

8. Test alternative physical constraints when developing loss functions for learning-based Riemann solvers, *e.g.* constraints based on EOS, boundary condition, or wavespeed. The use of different or additional physical constraints could further en-

hance accuracy, but also increases the risk of destabilizing the training process due to competing behaviour between the loss terms.

9. Conduct a performance analysis of all contemporary solvers against various test cases, plotting results on an accuracy vs. time complexity graph. There may be an empirical scaling law or clustering of solvers' performance capabilities that could provide CFD practitioners with a tool to select the right Riemann solver for the desired level of accuracy.

References

[1] A. Harten and J. M. Hyman, “Self-adjusting grid methods for one-dimensional hyperbolic conservation laws,” *Journal of Computational Physics*, vol. 50, pp. 235–269, 1983.

[2] D.-Y. Peng and D. B. Robinson, “A new two-constant equation of state,” *Industrial & Engineering Chemistry Fundamentals*, vol. 15, no. 1, pp. 59–64, 1976.

[3] K. S. Pitzer, “The volumetric and thermodynamic properties of fluids. i. theoretical basis and virial coefficients,” *Journal of the American Chemical Society*, vol. 77, no. 13, pp. 3427–3433, 1955.

[4] K. S. Pitzer, D. Z. Lippmann, R. F. C. Jr., C. M. Huggins, and D. E. Petersen, “The volumetric and thermodynamic properties of fluids. ii. compressibility factor, vapor pressure and entropy of vaporization,” *Journal of the American Chemical Society*, vol. 77, no. 13, pp. 3427–3433, 1955.

[5] P. C. Ma, Y. Lv, and M. Ihme, “An entropy-stable hybrid scheme for simulations of transcritical real-fluid flows,” *Journal of Computational Physics*, vol. 340, pp. 330–357, 2017.

[6] B. Riemann, “Ueber die fortpflanzung ebener luftwellen von endlicher schwingungsweite,” *Abhandlungen der Koeniglichen Gesellschaft der Wissenschaften zu Goettingen*, vol. 8, pp. 43–66, 1860.

[7] R. Leveque, *Finite-Volume Methods for Hyperbolic Problems.* Cambridge University Press, 2002.

[8] S. K. Godunov, “Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics,” *Matematicheskii Sbornik*, vol. 47, no. 89-3, 1959.

[9] G. A. Sod, "A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws," *Journal of Computational Physics*, vol. 27, pp. 1–31, 1978.

[10] B. Einfeldt, C. D. Munz, P. L. Roe, and B. Sjoegreen, "On godunov-type methods near low densities," *Journal of Computational Physics*, vol. 92, pp. 273–295, 1991.

[11] G. Kogekar, C. Karakaya, G. J. Liskovich, M. A. Oehlschlaeger, S. C. DeCaluwe, and R. J. Kee, "Impact of non-ideal behavior on ignition delay and chemical kinetics in high-pressure shock tube reactors," *Combustion and Flame*, vol. 189, pp. 1–11, 2018.

[12] P. A. Davidson, *An Introduction to Magnetohydrodynamics.* Cambridge University Press, 2001.

[13] C. Clarke and R. Carswell, *Principles of Astrophysical Fluid Dynamics.* Cambridge University Press, 2007.

[14] A. Martin, M. Reggio, and J.-Y. Trepanier, "Numerical solution of axisymmetric multi-species compressible gas flow: towards improved circuit breaker simulation," *International Journal of Computational Fluid Dynaimcs*, vol. 22, no. 4, pp. 259–271, 2008.

[15] W. Merzkirch and K. Bracht, "The erosion of dust by a shock wave in air: initial stages with laminar flow," *International Journal of Multiphase Flow*, vol. 4, no. 1, pp. 89–95, 1978.

[16] J. H. Geng and H. Groenig, "Dust suspensions accelerated by shock waves," *Experiments in Fluids*, vol. 28, pp. 360–367, 2000.

[17] R. R. Parker, J. F. Klausner, and R. Mei, "Supersonic two-phase impinging jet heat transfer," *Journal of Heat Transfer*, vol. 135, 2013.

[18] J. W. Hargis and E. L. Petersen, "Shock-tube boundary-layer effects on reflected-shock conditions with and without co2," *AIAA Journal*, vol. 55, no. 3, pp. 902–912, 2017.

[19] K. Grogan and M. Ihme, "Stanshock: a gas-dynamic model for shock tube simulations with non-ideal effects and chemical kinetics," *Shock Waves*, vol. 30, pp. 425–438, 2020.

[20] H. Hugoniot, “Memoire sur la propagation des mouvements dans les corps et specialement dans les gaz parfaits,” *Journal de l’Ecole Polytechnique*, vol. 58, pp. 1–125, 1887.

[21] A. Harten, P. D. Lax, and B. V. Leer, “On upstream differencing and godunov-type schemes for hyperbolic conservation laws,” *SIAM Review*, vol. 25, no. 1, 1983.

[22] B. Einfeldt, “On godunov-type methods for gas dynamics,” *SIAM Journal of Numerical Analysis*, vol. 25, no. 2, 1988.

[23] E. F. Toro, M. Spruce, and W. Speares, “Restoration of the contact surface in the hll-riemann solver,” *Shock Waves*, vol. 4, pp. 25–34, 1994.

[24] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics.* Springer, 3rd ed., 2009.

[25] B. E. Poling, J. M. Prausnitz, and J. P. O’Connell, *The Properties of Gases and Liquids.* McGraw-Hill, 2001.

[26] R. Arina, “Numerical simulation of near-critical fluids,” *Applied Numerical Mathematics*, vol. 51, no. 4, pp. 409–426, 2004.

[27] H. Terashima, S. Kawai, and N. Yamanishi, “High-resolution numerical method for supercritical flows with large density variations,” *AIAA Journal*, vol. 49, no. 12, pp. 2658–2672, 2011.

[28] W. J. M. Rankine, “On the thermodynamic theory of waves of finite longitudinal disturbances,” *Philosophical Transactions of the Royal Society of London*, vol. 160, pp. 277–288, 1870.

[29] J. D. Anderson, *Modern Compressible Flow: with Historical Perspective.* New York, NY: McGraw-Hill, 3rd ed., 2003.

[30] J.-P. Hickey and M. Ihme, “Supercritical mixing and combustion in rocket propulsion,” *Annual Research Briefs, Stanford Centre for Turbulence*, pp. 21–36, 2013.

[31] R. Saurel, M. Larini, and J. C. Loraud, “Exact and approximate riemann solvers for real gases,” *Journal of Computational Physics*, vol. 112, pp. 126–127, 1994.

[32] C. Pantano, R. Saurel, and T. Schmitt, “An oscillation free shock-capturing method for compressiblevan der waals supercritical fluid flows,” *Journal of Computational Physics*, vol. 335, pp. 780–811, 2017.

[33] C. Rodriguez, A. Vidal, P. Koukouvinis, M. Gavaises, and M. A. McHugh, "Simulation of transcritical fluid jets using the pc-saft eos," *Journal of Computational Physics*, vol. 374, pp. 444–468, 2018.

[34] C. Rodriguez, P. Koukouvinis, and M. Gavaises, "Simulation of supercritical diesel jets using the pc-saft eos," *Journal of Supercritical Fluids*, vol. 145, pp. 48–65, 2019.

[35] M. T. Migliorino and C. Scalo, "Heat-induced planar shock waves in supercritical fluids," *Shock Waves*, vol. 30, pp. 153–167, 2019.

[36] P. Colella and H. M. Glaz, "Efficient solution algorithms for the riemann problem for real gases," *Journal of Computational Physics*, vol. 59, pp. 264–289, 1985.

[37] D. A. Kouremenos, "The normal shock waves of real gases and the generalized isentropic exponents," *Forschung im Ingenieurwesen*, vol. 52, no. 1, pp. 23–31, 1986.

[38] D. A. Kouremenos and K. A. Antonopoulos, "Real gas normal shock waves with the redlich-kwong equation of state," *Acta Mechanica*, vol. 76, pp. 223–233, 1989.

[39] J. W. Banks, "On exact conservation for the euler equations with complex equations of state," *Communications in Computational Physics*, vol. 8, no. 5, pp. 995–1015, 2010.

[40] J. R. Kam, "An exact, compressible one-dimensional riemann solver for general, convex equations of state," tech. rep., Los Alamos National Laboratory, 2015.

[41] M. Passmann, S. aus der Wiesche, and F. Joos, "A one-dimensional analytical calcluation method for obtaining normal shock losses in supersonic real gas flows," *Journal of Physics Conference Series*, vol. 821, no. 012004, 2017.

[42] W. A. Sirignano, "Normal shocks with high upstream pressure," *Physical Review Fluids*, vol. 3, no. 093401, 2018.

[43] W. A. Sirignano, "Compressible flow at high pressure with linear equation of state," *Journal of Fluid Mechanics*, vol. 843, pp. 244–292, 2018.

[44] P. A. Thompson, "A fundamental derivative in gas dynamics," *Physics of Fluids*, vol. 14, pp. 1843–1849, 1971.

[45] K. C. Lambrakis and P. A. Thompson, "Existence of real fluids with a negative fundamental derivative," *Physics of Fluids*, vol. 15, pp. 933–935, 1972.

[46] R. Menikoff and B. J. Plohr, “The riemann problem for fluid flow of real gases,” *Reviews of Modern Physics*, vol. 61, no. 1, pp. 75–130, 1989.

[47] T. Hitz, M. Heinen, J. Vrabec, and C.-D. Munz, “Comparison of macro- and microscopic solutions of the riemann problem i. supercritical shock tube and expansion into vacuum,” *Journal of Computational Physics*, vol. 402, no. 109077, 2020.

[48] P. J. Milan, J.-P. Hickey, X. Wang, and V. Yang, “Deep-learning accelerated calculation of real-fluid properties in numerical simulation of complex flowfields,” *Journal of Computational Physics*, vol. 444, no. 110567, 2021.

[49] P. L. Roe, “Approximate riemann solvers, parameter vectors, and difference schemes,” *Journal of Computational Physics*, vol. 43, pp. 357–372, 1981.

[50] B. Engquist and S. Osher, “One-sided difference approximations for nonlinear conservation laws,” *Mathematics of Computation*, vol. 36, no. 154, 1981.

[51] E. F. Toro, “A linearised riemann solver for the time–dependent euler equations of gas dynamics,” *Proceedings of the Royal Society of London*, vol. 434, pp. 683–693, 1991.

[52] S. F. Davis, “Simplified second-order godunov methods,” *SIAM Journal of Scientific and Statistical Computing*, vol. 9, no. 3, 1988.

[53] E. F. Toro, “The hllc riemann solver,” *Shock Waves*, vol. 29, pp. 1065–1082, 2019.

[54] E. F. Toro, “Riemann problems and the waf method for solving two–dimensional shallow water equations,” *Philosophical Transactions of the Royal Society A*, vol. 338, pp. 43–68, 1992.

[55] S. Li, “An hllc riemann solver for magneto-hydrodynamics,” *Journal of Computational Physics*, vol. 203, pp. 344–357, 2005.

[56] A. Mignone and G. Bodo, “An hllc riemann solver for relativistic flows – i. hydrodynamics,” *Monthly Notices of the Royal Astronomical Society*, vol. 364, pp. 126–136, 2005.

[57] S. Osher, “Riemann solvers, the entropy condition, and difference approximations,” *SIAM Journal of Numerical Analysis*, vol. 21, no. 2, pp. 217–235, 1984.

[58] J. J. Quirk, “A contribution to the great riemann solver debate,” *International Journal for Numerical Methods in Fluids*, vol. 18, pp. 555–574, 1994.

[59] F. Qu, D. Sun, Q. Liu, and J. Bai, “A review of riemann solvers for hypersonic flows,” *Archives of Computational Methods in Engineering*, vol. 29, pp. 1771–1800, 2021.

[60] F. Dubois and G. Mehlman, “A non-parameterized entropy correction for roe’s approximate riemann solver,” *Numerische Mathematik*, vol. 73, pp. 169–208, 1996.

[61] P. L. Roe, “Sonic flux formulae,” *SIAM J. Sci. Stat. Comput.*, vol. 13, no. 2, pp. 611–630, 1992.

[62] M. Svard, “Entropy stable bounadry conditions for the euler equations,” *Journal of Computational Physics*, vol. 426, no. 109947, 2021.

[63] X.-S. Li, X.-D. Ren, C.-W. Gu, and Y.-H. Li, “Shock-stable roe scheme combining entropy fix and rotated riemann solver,” *AIAA Journal*, vol. 58, no. 2, 2020.

[64] H. Chizari, V. Singh, and F. Ismail, “Cell-vertex entropy-stable finite volume methods for the system of euler equations on unstructured grids,” *Computers and Mathematics with Applications*, vol. 98, pp. 261–279, 2021.

[65] A. Gouasmi, K. Duraisamy, and S. M. Murman, “Formulation of entropy-stable schemes for the multicomponent compressible euler equations,” *Computer Methods in Applied Mechanics and Engineering*, vol. 363, no. 112912, 2020.

[66] A. Gouasmi, S. M. Murman, and K. Duraisamy, “Entropy-stable schemes in the low-mach-number regime: flux-precondition, entropy breakdowns, and entropy transfers,” *Journal of Computational Physics*, vol. 456, no. 111036, 2022.

[67] P. Helluy, J.-M. Herard, H. Mathis, and S. Mueller, “A simple parameter-free entropy correction for approximate riemann solvers,” *Comptes Rendus Mecanique*, vol. 338, no. 9, pp. 493–498, 2010.

[68] B. Schmidtmann and A. R. Winters, “Hybrid entropy stable hll-type riemann solvers for hyperbolic conservation laws,” *Journal of Computational Physics*, vol. 330, pp. 566–570, 2017.

[69] A. Colombo, A. Crivellini, and A. Nigro, “On the entropy conserving/stable implicit dg linearization of the euler equations in entropy variables,” *Computers and Fluids*, vol. 232, no. 105198, 2021.

[70] F. Renac, “Entropy stable, robust and high-order dgsem for the compressible multicomponent euler equations,” *Journal of Computational Physics*, vol. 445, no. 110584, 2021.

[71] X. Wu, E. J. Kubatko, and J. Chan, “High-order entropy stable discontinuous galerkin methods for the shallow water equations: curved triangular meshes and gpu acceleration,” *Computers and Mathematics with Applications*, vol. 82, pp. 172–199, 2021.

[72] S. Brull, B. Dubroca, and X. Lhebrard, “Modelling and entropy satisfying relaxation scheme for the nonconservative bitemperature euler system with transverse magnetic field,” *Computers and Fluids*, vol. 214, no. 104743, 2021.

[73] A. Chan, G. Gallice, R. Loubere, and P.-H. Maire, “Positivity preserving and entropy consistent approximate riemann solvers dedicated to the high-order mood-based finite volume discretization of langrangian and eulerian gas dynamics,” *Computers and Fluids*, vol. 229, no. 105056, 2021.

[74] J. Duan and H. Tang, “Entropy stable adaptive moving mesh schemes for 2d and 3d special relativistic hydrodynamics,” *Journal of Computational Physics*, vol. 426, no. 109949, 2021.

[75] P. Cinella, “Roe-type schemes for dense gas flow computations,” *Computers and Fluids*, vol. 35, pp. 1264–1281, 2006.

[76] A. Guardone, C. Zamfirescu, and P. Colonna, “Maximum intensity of rarefaction shock waves for dense gases,” *Journal of Fluid Mechanics*, vol. 642, pp. 127–146, 2010.

[77] N. R. Nannan, A. Guardone, and P. Colonna, “Critical point anomalies include expansion shock waves,” *Physics of Fluids*, vol. 26, no. 021701, 2014.

[78] N. R. Nannan, C. Sirianni, T. Mathijssen, A. Guardone, and P. Colonna, “The admissibility domain of rarefaction shock waves in the near-critical vapour-liquid equilibrium region of pure typical fluids,” *Journal of Fluid Mechanics*, vol. 795, pp. 241–261, 2016.

[79] A. Giaque, C. Corre, and A. Vadrot, “Direct numerical simulations of forced homogeneous isotropic turbulence in a dense gas,” *Journal of Turbulence*, vol. 21, no. 3, pp. 186–208, 2020.

[80] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, “Machine learning for fluid mechanics,” *Annual Review of Fluid Mechanics*, vol. 52, pp. 477–508, 2020.

[81] A. Komolgrov, “The local structure of turbulence in incompressible viscous fluid for very large reynolds number,” *Cr. Acad. Sci. USSR*, vol. 30, pp. 301–305, 1941.

[82] L. Sirovich, “Turbulence and the dynamics of coherent structures, parts i-iii,” *Quarterly of Applied Mathematics*, vol. 45, pp. 561–590, 1987.

[83] L. Sirovich and M. Kirby, “Low-dimensional procedure for the characterization of human faces,” *Journal of the Optical Society of America*, vol. 4, no. 3, pp. 519–524, 1987.

[84] K. Jambunathan, S. Hartle, S. Ashforth-Frost, and V. Fontama, “Evaluating convective heat transfer coefficients using neural networks,” *International Journal of Heat and Mass Transfer*, vol. 39, pp. 2329–2332, 1996.

[85] M. Milano and P. Koumoutsakos, “Neural network modeling for near-wall turbulent flow,” *Journal of Computational Physics*, vol. 182, pp. 1–26, 2002.

[86] M. Ma, J. Lu, and G. Tryggvason, “Using statistical learning to close two-fluid multiphase flow equations for a simple bubbly system,” *Physics of Fluids*, vol. 27, no. 9, 2015.

[87] B. Lusch, J. N. Kutz, and S. L. Brunton, “Deep learning for universal linear embeddings of nonlinear dynamics,” *Nature Communications*, vol. 9, no. 1, 2018.

[88] O. San, R. Maulik, and M. Ahmed, “An artificial neural network framework for reduced order modeling of transient flows,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 77, pp. 271–287, 2019.

[89] S. Pawar, S. E. Ahmed, O. San, and A. Rasheed, “An evole-then-correct reduced order model for hidden fluid dynamics,” *Mathematics*, vol. 8, no. 570, 2020.

[90] K. Duraisamy, G. Iaccarino, and H. Xiao, “Turbulence modeling in the age of data,” *Annual Review of Fluid Mechanics*, vol. 51, no. 1, pp. 357–377, 2019.

[91] W. E. Faller and S. J. Schreck, “Neural networks: applications and opportunities in aeronautics,” *Progress in Aerospace Sciences*, vol. 32, pp. 433–456, 1996.

[92] C. Lee, J. Kim, D. Babcock, and R. Goodman, “Application of neural networks to turbulence control for drag reduction,” *Physics of Fluids*, vol. 9, pp. 1740–1747, 1997.

[93] A. T. Mohan and D. V. Gaitonde, “A deep learning based approach to reduced order modelingfor turbulent flow control using lstm neural networks,” *arXiV:Computational Physics*, 2018.

[94] N. Benard, J. Pons-Prats, J. Periaux, G. Bugeda, and P. Braud, “Turbulent separated shear flow control bysurface plasma actuator: experimental optimization by genetic algorithm approach,” *Experiments in Fluids*, vol. 57, no. 22, 2016.

[95] S. Pierret and R. Van Den Braembussche, “Turbomachinery blade design using a navier-stokes solver and artificial neural network,” *Journal of Turbomachinery*, vol. 121, pp. 326–332, 1999.

[96] D. Loucks, E. van Beek, J. Stedinger, J. Dijkman, and M. Villars, *Water Resources Systems Planning and Management: An Introduction to Methods*. Springer, 2005.

[97] F. Gueniat, L. Mathelin, and M. Y. Hussaini, “A statistical learning strategy for closed-loop control of fluid flows,” *Theory of Computational Fluid Dynamics*, vol. 30, pp. 497–510, 2016.

[98] M. Gazzola, A. Tchieu, D. Alexeev, A. D. Brauer, and P. Koumoutsakos, “Learning to school in the presence ofhydrodynamic interactions,” *Journal of Fluid Mechanics*, vol. 789, pp. 726–749, 2016.

[99] G. Reddy, A. Celani, T. J. Sejnowski, and M. Vergassola, “Learning to soar in turbulent environments,” *Proceedings of the National Academy of Science*, vol. 113, pp. 4877–4884, 2016.

[100] H. J. Kim, M. I. Jordan, S. Sastry, and A. Y. Ng, “Autonomous helicopter flight via reinforcement learning,” *Advances in Neural Information Processing Systems 17*, p. 799–806, 2004.

[101] M. Dissanayake and N. Phan-Thien, “Neural-network-based approximations for solving partial differential equations,” *Communications in Numerial Methods in Engineering*, vol. 10, pp. 195–201, 1994.

[102] R. Gonzalez-Garcia, R. Rico-Martinez, and I. Kevrekidis, “Identification of distributed paramater systems: a neural net based approach,” *Computers and Chemical Engineering*, vol. 22, pp. 965–968, 1998.

[103] I. E. Lagaris, A. Likas, and D. I. Fotiadis, “Artificial neural networks for solving ordinary and partial differential equations,” *IEEE Transactions on Neural Networks*, vol. 9, pp. 987–1000, 1998.

[104] G. E. K. M. Raissi, P. Perdikaris, “Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[105] G. E. K. Zhiping Mao, Ameya D. Jagtap, “Physics-informed neural networks for high-speed flows,” *Computer Methods in Applied Mechanics and Engineering*, vol. 360, no. 112789, 2020.

[106] D. A. Bezgin, S. J. Schmidt, and N. A. Adams, “A data-driven physics-informed finite volume scheme for nonclassical undercompressive shocks,” *Journal of Computational Physics*, vol. 437, no. 110324, 2021.

[107] J. Magiera, D. Ray, J. S. Hesthaven, and C. Rohde, “Constraint-aware neural networks for riemannproblems,” *Journal of Computational Physics*, vol. 409, no. 109345, 2019.

[108] V. Gyrya, M. J. Shashkov, A. N. Skurikhin, and S. Tokareva, “Machine learning approaches for the solution of the riemann problem influid dynamics: a case study (with reviewers),” *Journal of Computational Physics*, 2019.

[109] O. Fuks and H. A. Tchelepi, “Limitations of physics-informed machine learning for nonlinear two-phase transport in porous media,” *Journal of Machine Learning for Modeling and Computing*, vol. 1, no. 1, pp. 19–37, 2022.

[110] T. Dieselhorst, W. Cook, S. Bernuzzi, and D. Radice, “Machine learning for conservative-to-primitive in relativistic hydrodynamics,” *Symmetry*, vol. 13, no. 11, 2021.

[111] M. M. Abbott, “Cubic equations of state: an interpretive review,” *Advances in Chemistry*, vol. 182, pp. 47–70, 1979.

[112] O. Redlich and J. N. S. Kwong, “On the thermodynamics of solutions. v. an equation of state. fugacities of gaseous solutions.,” *Chemical Reviews*, vol. 44, no. 1, pp. 233–244, 1949.

[113] G. Soave, “Equilibrium constants from a modified redlich-kwong equation of state,” *Chemical Engineering Science*, vol. 27, pp. 1197–1203, 1972.

[114] B. J. McBride, S. Gordon, and M. A. Reno, *Coefficients for Calculating Thermodynamic and Transport Properties of Individual Species.* National Aeronautics and Space Administration, 1993.

[115] N. J. Hurst, “Determination of differences and ratios of specific heats and change in entropy as applied to five equations of state for gases,” *US Army Missile Command, Redstone Arsenal*, 1963.

[116] J. R. Andrews and O. Biblarz, “Temperature dependence of gases in polynomial form,” *Naval Postgraduate School*, 1981.

[117] B. V. Leer, “Towards the ultimate conservative difference scheme v. a second-order sequel to godunov’s metod,” *Journal of Computational Physics*, vol. 32, pp. 101–136, 1979.

[118] M.-S. Liou and C. S. Jr, “A new flux splitting scheme,” *Journal of Computational Physics*, vol. 107, no. 23-39, 1993.

[119] M.-S. Liou, “A sequel to ausm: Ausm+,” *Journal of Computational Physics*, vol. 129, no. 364-382, 1996.

[120] X.-D. Liu, S. Osher, and T. Chan, “Weighted essentially non-oscillatory schemes,” *Journal of Computational Physics*, vol. 115, pp. 200–212, 1994.

[121] S. Gottlieb, C.-W. Shu, and E. Tadmor, “Strong stability-preserving high-order time discretization methods,” *SIAM Review*, vol. 43, no. 1, pp. 89–112, 2001.

[122] T. M. Mitchell, *Machine Learning.* McGraw-Hill, 1997.

[123] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach.* Prentice Hall, 3rd ed., 2010.

[124] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* MIT Press, 2016.

[125] M. Ihme, C. Schmitt, and H. Pitsch, “Optimal artificial neural networks and tabulation methods for chemistry representation in les of a bluff-body swirl-stabilized flame,” *Proceedings of the Combustion Institute*, 2009.

[126] S. Bhalla, M. Yao, J.-P. Hickey, and M. Crowley, “Compact representation of a multidimensional combustion manifold using deep neural networks,” *European Conference on Machine Learning*, 2019.

[127] P. J. Milan, X. Wang, J.-P. Hickey, Y. Li, and V. Yang, “Accelerating numerical simulations of supercritical fluid flows using deep neural networks,” *AIAA ScieTech Forum*, 2020.

[128] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning.* Springer, 2nd (12th printing) ed., 2017.

[129] G. Hinton, N. Srivastava, and K. Swersky, “Neural networks for machine learning: Lecture 6,” *Coursera*, 2012.

[130] D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” in *International Conference on Learning Representations*, 2015.

[131] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of Controls, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, 1989.

[132] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991.

[133] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, “Mutilayer feedforward networks with a nonpolynomial activation function can approximate any function,” *Neural Networks*, vol. 6, no. 6, pp. 861–867, 1993.

[134] Y.-X. Yuan, “Recent advances in trust region algorithms,” *Mathematical Programming: Series A and B*, vol. 151, no. 1, 2015.

[135] C. L. Yaws, *Chemical Properties Handbook: Physical, Thermodynamic, Environmental, Transport, Safety, and Health Related Properties for Organic and Inorganic Chemicals.* McGraw-Hill, 1999.

[136] D. T. Banuti, M. Raju, P. C. Ma, M. Ihme, and J.-P. Hickey, “Seven questions about supercritical fluids towards a new fluid state diagram,” *AIAA SciTech Forum*, 2017.

[137] J. C.-H. Wang and J.-P. Hickey, “Analytical solutions to shock and expansion waves for non-ideal equations of state,” *Physics of Fluids*, vol. 32, p. 086105, 2020.

[138] Q. Wang, J. S. Hesthaven, and D. Ray, “Non-intrusive reduced order modeling of unsteady flows using artificial neural networkswith application to a combustion problem,” *Journal of Computational Physics*, vol. 384, pp. 289–307, 2019.

[139] M. S. B. Coleman, “An extension of the athena++ framework for general equations of state,” *The Astrophysics Journal*, vol. 248, no. 7, 2020.

[140] R. Abgrall and S. Karni, “Computations of compressible multifluids,” *Journal of Computational Physics*, vol. 169, pp. 594–623, 2001.

[141] J. Glimm, C. Klingenberg, O. McGryan, B. Plohr, D. Sharp, and S. Yaniv, "Front tracking and two-dimensional riemann problems," *Advances in Applied Mathematics*, vol. 6, pp. 259–290, 1985.

[142] T. Zhang and Y. Zheng, "Conjecture on the structure of solutions of the riemann problem for two-dimensional gas dynamics systems," *SIAM Journal of Mathematical Analysis*, vol. 21, no. 3, pp. 593–630, 1990.

[143] C. W. Schulz-Rinne, "Classification of the riemann problem for two-dimensional gas dynamics," *SIAM Journal of Mathematical Analysis*, vol. 24, no. 1, pp. 76–88, 1993.

[144] C. W. Schulz-Rinne, J. P. Collins, and H. M. Glaz, "Numerical solution of the riemann problem for two-dimensional gas dynamics," *SIAM Journal of Scientific Computing*, vol. 14, no. 6, pp. 1394–1414, 1993.

[145] P. D. Lax and X.-D. Liu, "Solution of two-dimensional riemann problems of gas dynamics by positive schemes," *SIAM Journal of Scientific Computing*, vol. 19, no. 2, pp. 319–340, 1998.

[146] D. Tan and T. Zhang, "Two-dimensional riemann problem for a hyperbolic system of nonlinear conservation laws i. four j-cases," *Journal of Differential Equations*, vol. 111, pp. 203–254, 1994.

[147] T. Zhang, G.-Q. Chen, and S. Yang, "On the 2d riemann problem for the compressible euler equations i. interaction of shocks and rarefaction waves," *Discrete and Continuous Dynamical Systems*, vol. 1, no. 4, pp. 555–584, 1995.

[148] T. Zhang, G.-Q. Chen, and S. Yang, "On the 2d riemann problem for the compressible euler equations ii. interaction of contact discontinuities," *Discrete and Continuous Dynamical Systems*, vol. 6, no. 2, pp. 419–430, 2000.

[149] J. Glimm, X. Ji, J. Li, X. Li, P. Zhang, T. Zhang, and Y. Zheng, "Transonic shock formation in a rarefaction riemann problem for the 2d compressible euler equations," *SIAM Journal on Applied Mathematics*, vol. 69, no. 3, pp. 720–742, 2008.

[150] L. Gosse, "A two-dimensional version of the godunov scheme for scalar balance laws," *SIAM Journal on Numerical Analysis*, vol. 52, no. 2, pp. 626–652, 2014.

[151] E. Dekel, D. Nassimi, and S. Sahni, "Parallel matrix and graph algorithms," *SIAM Journal on Computing*, vol. 10, no. 4, pp. 657–675, 1981.

[152] J. C. H. Wang and J.-P. Hickey, “A class of structurally complete approximate riemann solvers for trans- and supercritical flows with large gradients,” *Journal of Computational Physics*, vol. 468, no. 111521, 2022.

[153] S. J. Billett and E. F. Toro, “On the accuracy and stability of explicit schemesfor multidimensional linear homogeneous advection equations,” *Journal of Computational Physics*, vol. 131, pp. 247–250, 1997.

APPENDICES

Appendix A

Integral and differential forms of the governing equations

Anderson [29] provides a thorough derivation of the integral and differential forms of the governing equations of continuum fluid mechanics. In a generic three-dimensional flowfield, suppose there is a control volume of volume $\mathcal{V}$ and surface area $\mathcal{S}$. If the control volume is of infinitesimal size, then let its volume and surface area be denoted $d\mathcal{V}$ and $d\mathcal{S}$, respectively.

The integral form of mass conservation is:

$$\frac{\partial}{\partial t}\iiint_{\mathcal{V}}\rho d\mathcal{V}+\oiint_{\mathcal{S}}\rho\mathbf{V}\cdot\hat{n}d\mathcal{S}=0 \tag{A.1}$$

where the $\mathbf{V}$ is the velocity vector and $\hat{n}$ is the surface normal unit vector pointing outward. The left-hand side represents the rate of change of mass within the volume and the mass flow rate exiting the volume.

The integral form of momentum conservation is:

$$\iiint_{\mathcal{V}}\frac{\partial(\rho\mathbf{V})}{\partial t}d\mathcal{V}+\oiint_{\mathcal{S}}(\rho\mathbf{V}\cdot\hat{n})\mathbf{V}d\mathcal{S}=\iiint_{\mathcal{V}}\rho\mathbf{f}d\mathcal{V}+\oiint_{\mathcal{S}}-p\hat{n}d\mathcal{S}+\mathbf{f}_{viscous} \tag{A.2}$$

where the left-hand terms are the rate of change of momentum within the volume and the momentum flow rate exiting the volume; and the right-hand terms are the body forces $\mathbf{f}$, pressure forces, and viscous forces $\mathbf{f}_{viscous}$ (*i.e.* friction forces that act tangential to the surface) acting on the control volume. Since this thesis is concerned with inviscid flows, we generally assume that body and viscous forces are zero. It should be noted that $\mathbf{f}_{viscous}$

is commonly modelled as a Newtonian fluid whereby surface viscous stresses need to be estimated, such as with Reynolds-averaging.

The integral form of energy conservation is:

$$\begin{aligned} &\iiint_{\mathcal{V}} \frac{\partial}{\partial t}\left(\rho\left(\frac{e}{M}+\frac{1}{2}\mathbf{V}\cdot\mathbf{V}\right)\right) d\mathcal{V} + \oint\!\!\!\oint_{\mathcal{S}} \rho\left(\frac{e}{M}+\frac{1}{2}\mathbf{V}\cdot\mathbf{V}\right)\mathbf{V}\cdot\hat{n}d\mathcal{S} \\ &= \iiint_{\mathcal{V}} \dot{q}\rho d\mathcal{V} + \oint\!\!\!\oint_{\mathcal{S}} -p\mathbf{V}\cdot\hat{n}d\mathcal{S} + \iiint_{\mathcal{V}} \rho(\mathbf{f}\cdot\mathbf{V})d\mathcal{V} \end{aligned} \tag{A.3}$$

where the left-hand terms are the rate of change of total energy within the volume and the total energy flow rate out of the volume; and the right-hand terms are the rate of heat addition to the volume, the rate of work done on the volume by pressure forces, and the rate of work done on the volume by body forces.

In order to derive the differential form of the governing equations, it is helpful to use the divergence theorem for vector and scalar-valued functions:

$$\oint\!\!\!\oint_{\mathcal{S}} \mathbf{A}\cdot\hat{n}d\mathcal{S} = \iiint_{\mathcal{V}} (\nabla\cdot\mathbf{A})d\mathcal{V} \tag{A.4}$$

$$\oint\!\!\!\oint_{\mathcal{S}} A\hat{n}d\mathcal{S} = \iiint_{\mathcal{V}} (\nabla A)d\mathcal{V} \tag{A.5}$$

where **A** and A are continuously differentiable functions that are vectors or scalars, respectively. Then, the differential forms of mass, momentum, and energy conservation are:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{V}) = 0 \tag{A.6}$$

$$\frac{\partial}{\partial t}(\rho\mathbf{V}) + \nabla\cdot(\rho\mathbf{V}\otimes\mathbf{V}) = -\nabla p + \rho\mathbf{f} \tag{A.7}$$

$$\frac{\partial E}{\partial t} + \nabla\cdot E\mathbf{V} = \rho\dot{q} - \nabla\cdot(\rho\mathbf{V}) + \rho(\mathbf{f}\cdot\mathbf{V}) \tag{A.8}$$

where $\otimes$ is the outer product of two vectors. It is thus possible to consider only the x or x, y dimensions as required in this thesis.

Appendix B

Multi-dimensional Godunov scheme

Multi-dimensional Godunov schemes are generally derived through dimensional splitting (also called the method of fractional steps) or through an unsplit finite volume approach [24]. In the fractional step method, one applies 1D methods in each dimension. For each dimension, the fluxes are computed in that direction and all cells are updated before proceeding to the next dimension. In the unsplit method, the flux contributions from all dimensions are solved simultaneously. In this thesis, the unsplit finite volume method technique is used as it avoids the potential complexities of intermediate states during each time step, permitting analysis that is more closely linked to the Riemann solver itself.

Similar to Eq. (2.8), it may be shown that the 2D Euler equations are of the form:

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = 0 \tag{B.1}$$

where $\mathbf{G}$ is the flux in the y direction and depends on the conservative variables $\mathbf{U}$. Now, the conservative variables and flux vectors read:

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho w \\ E \end{bmatrix}; \qquad \mathbf{F} = \begin{bmatrix} \rho u \\ p + \rho u^2 \\ \rho u w \\ u(E+p) \end{bmatrix}; \qquad \mathbf{G} = \begin{bmatrix} \rho w \\ \rho w u \\ p + \rho w^2 \\ w(E+p) \end{bmatrix} \tag{B.2}$$

where w is the velocity in the y direction, and the total energy must now be computed as:

$$E = \frac{e}{v} + \frac{1}{2}\rho\left(u^2 + w^2\right) \tag{B.3}$$

Applying an explicit cell-centred finite volume scheme such as in Toro [24] §16.4.1, the resulting first-order Godunov scheme is:

$$\mathbf{U}_{i,j}^{n+1} = \mathbf{U}_{i,j}^{n} + \frac{\Delta t}{\Delta x}\left(\mathbf{F}_{i-\frac{1}{2},j} - \mathbf{F}_{i+\frac{1}{2},j}\right) + \frac{\Delta t}{\Delta y}\left(\mathbf{G}_{i,j-\frac{1}{2}} - \mathbf{G}_{i,j+\frac{1}{2}}\right) \tag{B.4}$$

which is analogous to Eq. (2.40). It is evident that all flux contributions are considered at each time step, as required by the unsplit finite volume method. The final consideration when migrating to higher dimensions is the CFL condition. Let $S_{i,j}^{n,x}, S_{i,j}^{n,y}$ be the fastest-moving wavespeeds for the i, j cell in the x and y directions, respectively. Then the CFL condition becomes:

$$\Delta t = \text{CFL} \min_{i,j}\left(\frac{\Delta x}{S_{i,j}^{n,x}}, \frac{\Delta y}{S_{i,j}^{n,y}}\right) \tag{B.5}$$

That is, the CFL number must now be chosen considering the fastest wavespeeds in either Cartesian direction. Numerical stability analysis is also much more challenging due to the extra degree of freedom. Details on accuracy and stability in multiple dimensions can be found in Billett & Toro [153].

Appendix C

Speed of a moving normal shock

Earlier, the pre-shock and post-shock conditions of a stationary normal shock were denoted 1 and 2, respectively. Suppose that instead of a stationary shock, the shock is moving at a speed u_s with respect to the environment. Then, the velocities relative to the moving shock are:

$$\hat{u}_1 = u_1 - u_s \tag{C.1}$$

$$\hat{u}_2 = u_2 - u_s \tag{C.2}$$

where the circumflex (hat) symbol indicates that the quantity is measured in the frame where the moving shock appears to be stationary. The governing equations Eqs. (2.1) to (2.3) become:

$$\rho_2 \hat{u}_2 = \rho_1 \hat{u}_1 \tag{C.3}$$

$$\rho_2 \hat{u}_2^2 + p_2 = \rho_1 \hat{u}_1^2 + p_1 \tag{C.4}$$

$$\hat{u}_2 \rho_2 (\frac{1}{2}\hat{u}_2^2 + \frac{\hat{e}_2}{M} + \frac{p_2}{\rho_2}) = \hat{u}_1 \rho_1 (\frac{1}{2}\hat{u}_1^2 + \frac{\hat{e}_1}{M} + \frac{p_1}{\rho_1}) \tag{C.5}$$

Substituting Eq. (C.3) into Eq. (C.4):

$$\rho_2 \hat{u}_2^2 = \frac{\rho_2 \hat{u}_2}{\rho_1} \rho_2 \hat{u}_2 + p_1 - p_2 \tag{C.6}$$

which can be rearranged to obtain:

$$\hat{u}_2^2 = \frac{\rho_1}{\rho_2} \left(\frac{p_1 - p_2}{\rho_1 - \rho_2} \right) \tag{C.7}$$

or similarly:

$$\hat{u}_1^2 = \frac{\rho_2}{\rho_1}\left(\frac{p_1 - p_2}{\rho_1 - \rho_2}\right) \tag{C.8}$$

Finally, combining Eq. (C.7) or Eq. (C.8) with either Eq. (C.1) or Eq. (C.2), respectively:

$$u_s = u_1 - \sqrt{\frac{\rho_2}{\rho_1}\left(\frac{p_1 - p_2}{\rho_1 - \rho_2}\right)} \tag{C.9}$$

which can be used to calculate the speed of the moving shock in either exact-iterative or learning-based Riemann solvers (which iterate or estimate star-state, *i.e.* post-shock, conditions).

Appendix D

Restoring a right rarefaction wave

In a right rarefaction wave, the form of Eqs. 4.8 through 4.11 changes slightly:

$$v(x,t) = \exp\Big(- A_1 \frac{x}{t} - A_2 \Big) \tag{D.1}$$

$$p(x,t) = \exp\Big(A_3 \frac{x}{t} - A_4 \Big) \tag{D.2}$$

$$u(x,t) = A_5 - A_6 \frac{x}{t} \tag{D.3}$$

where A_i are again the constants determined by substituting the known conditions at the head and tail of the expansion wave. The head and tail of a right rarefaction are located along the C_+ rather than C_- characteristics in left rarefactions. Thus, the primitives evaluated at the cell interface for a right transonic rarefaction are given by:

$$v(0,t) = \exp\Big(\frac{1}{2} \ln v_R v_{*R} - \frac{u_R + c_R + u_* + c_{*R}}{2(u_R + c_R - u_* - c_{*R})} \ln \frac{v_R}{v_{*R}} \Big) \tag{D.4}$$

$$p(0,t) = \exp\Big(\frac{1}{2} \ln p_R p_* - \frac{u_R + c_R + u_* + c_{*R}}{2(u_R + c_R - u_* - c_{*R})} \ln \frac{p_R}{p_*} \Big) \tag{D.5}$$

$$u(0,t) = \frac{1}{2}(u_R + u_*) - \frac{(u_R - u_*)(u_R + c_R + u_* + c_{*R})}{2(u_R + c_R - u_* - c_{*R})} \tag{D.6}$$

from which it is possible to calculate the flux as described earlier.

Appendix E

Roe solver for gases with non-ideal thermodynamics

The governing equations (2.1) to (2.3) may be cast into matrix form:

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = 0 \tag{E.1}$$

as shown in (2.8). By introducing the Jacobian:

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \tag{E.2}$$

and using the chain rule, the governing equations may be restated as:

$$\mathbf{U}_t + \mathbf{A}(\mathbf{U})\mathbf{U}_x = 0 \tag{E.3}$$

Roe [49] assumed that the Jacobian could be approximated by a constant matrix $\tilde{\mathbf{A}}$ of the form:

$$\tilde{\mathbf{A}} = \tilde{\mathbf{A}}(\mathbf{U}_L, \mathbf{U}_R) \tag{E.4}$$

that depends only on the initial conditions to the left and right of the cell interface. Roe therefore chose to solve the approximate, linearized Riemann problem:

$$\mathbf{U}_t + \tilde{\mathbf{A}}\mathbf{U}_x = 0 \tag{E.5}$$

subject to the initial conditions $\mathbf{U}(x < 0, t = 0) = U_L$ and $\mathbf{U}(x > 0, t = 0) = U_R$. Roe further imposed that the Jacobian matrix $\tilde{\mathbf{A}}$ must satisfy the properties of hyperbolicity, consistency, and conservation across discontinuities.

The resulting system may be solved analytically using standard techniques for linear hyperbolic equations. By decomposing this system in terms of eigenvalues, eigenvectors, and wave strengths, and with some algebraic manipulation, Roe showed that the intercell flux becomes:

$$\mathbf{F} = \frac{1}{2}(\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2}\sum_{i=1}^{m} \tilde{\alpha}_i |\tilde{\lambda}_i| \tilde{\mathbf{K}}^{(i)} \tag{E.6}$$

where $\tilde{\alpha}_i$ are wave strengths, $\tilde{\lambda}_i$ are eigenvalues, and $\tilde{\mathbf{K}}_i$ are right eigenvectors computed from so-called Roe-averaged physical quantities. The index $i = 1$ extends to m where m is the number of spatial dimensions plus 2. Solving the Riemann problem thus reduces to selecting a methodology for finding the Roe-averages, the above mentioned terms, and finally the flux. If deriving a Roe solver for 2D flow, then an analogous result is obtained for fluxes in the y direction, that is, $\mathbf{G}$ if using the same notation as Appendix B.

For the case of one-dimensional flow of non-ideal gases, where thermodynamic quantities are expressed in terms of v and T, the relevant Roe averages are:

$$\tilde{u} = \frac{\sqrt{\rho_L}u_L + \sqrt{\rho_R}u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \tag{E.7}$$

$$\tilde{H} = \frac{\sqrt{\rho_L}H_L + \sqrt{\rho_R}H_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \tag{E.8}$$

$$\tilde{v} = \frac{\sqrt{\rho_L}v_L + \sqrt{\rho_R}v_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \tag{E.9}$$

$$\tilde{T} = \frac{\sqrt{\rho_L}T_L + \sqrt{\rho_R}T_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \tag{E.10}$$

$$\tilde{c} = c(\tilde{v}, \tilde{T}) \tag{E.11}$$

where c is the real-gas speed of sound from (2.25). In the case of ideal gases, these equations may be simplified further. The required eigenvalues are thus:

$$\tilde{\lambda}_1 = \tilde{u} - \tilde{c}; \tilde{\lambda}_2 = \tilde{u}; \tilde{\lambda}_3 = \tilde{u} + \tilde{c} \tag{E.12}$$

the corresponding right eigenvectors are:

$$\tilde{\mathbf{K}}^{(1)} = \begin{bmatrix} 1 \\ \tilde{u} - \tilde{c} \\ \tilde{H} - \tilde{u}\tilde{c} \end{bmatrix}; \tilde{\mathbf{K}}^{(2)} = \begin{bmatrix} 1 \\ \tilde{u} \\ \frac{1}{2}\tilde{u}^2 \end{bmatrix}; \tilde{\mathbf{K}}^{(3)} = \begin{bmatrix} 1 \\ \tilde{u} + \tilde{c} \\ \tilde{H} + \tilde{u}\tilde{c} \end{bmatrix} \tag{E.13}$$

and the corresponding wave strengths are:

$$\tilde{\alpha}_1 = \frac{\Delta p - \tilde{\rho}\tilde{c}\Delta u}{2\tilde{c}^2}; \tilde{\alpha}_2 = -\left(\frac{\Delta p}{\tilde{c}^2} - \Delta\rho\right); \tilde{\alpha}_3 = \frac{\Delta p + \tilde{\rho}\tilde{c}\Delta u}{2\tilde{c}^2} \tag{E.14}$$

where $\Delta p = p_R - p_L$, $\Delta u = u_R - u_L$, and $\Delta\rho = \rho_R - \rho_L$.

In the case of two-dimensional flow, then solving for **F** requires an additional Roe average:

$$\tilde{w} = \frac{\sqrt{\rho_L}w_L + \sqrt{\rho_R}w_R}{\sqrt{\rho_L} + \sqrt{\rho_R}} \tag{E.15}$$

along with modified right eigenvectors:

$$\tilde{\mathbf{K}}^{(1)} = \begin{bmatrix} 1 \\ \tilde{u} - \tilde{c} \\ \tilde{w} \\ \tilde{H} - \tilde{u}\tilde{c} \end{bmatrix}; \tilde{\mathbf{K}}^{(2)} = \begin{bmatrix} 1 \\ \tilde{u} \\ \tilde{w} \\ \frac{1}{2}\tilde{u}^2 \end{bmatrix}; \tilde{\mathbf{K}}^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \tilde{w} \end{bmatrix}; \tilde{\mathbf{K}}^{(3)} = \begin{bmatrix} 1 \\ \tilde{u} + \tilde{c} \\ \tilde{w} \\ \tilde{H} + \tilde{u}\tilde{c} \end{bmatrix} \tag{E.16}$$

and modified wave strengths:

$$\tilde{\alpha}_1 = \frac{\Delta p - \tilde{\rho}\tilde{c}\Delta u}{2\tilde{c}^2}; \tilde{\alpha}_2 = -\left(\frac{\Delta p}{\tilde{c}^2} - \Delta\rho\right); \tilde{\alpha}_3 = \tilde{\rho}\Delta w; \tilde{\alpha}_4 = \frac{\Delta p + \tilde{\rho}\tilde{c}\Delta u}{2\tilde{c}^2} \tag{E.17}$$

Additionally, these calculations must be repeated for **G** except with the positions and ordering of u and w swapped accordingly. For details, see Toro [24] §11.3.3.