A Modular Parallel-Computing Strategy for Multiresolution Simulations of Complex Flows on High-Performance Computers

Nils Hoppe

Die Dissertation wurde am 15.09.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Engineering and Design am 27.02.2023 angenommen.
Sandra and Konrad: I love you to pieces, to you I dedicate this thesis.
Acknowledgement

Many people have contributed to this thesis, some directly with advise on its subject, others indirectly through continuous moral support. Some have helped knowingly, others may not even know me. No matter the way you influenced this thesis for the better, my deepest gratitude to you!

In particular, I want to thank my supervisor Prof. Nikolaus A. Adams for the opportunity to work in a stimulating environment with large thematic freedom. Thank you also for the support and guidance throughout the years and last but not least for the examination of this work.

I further want to express my gratitude towards Prof. Petros Koumoutsakos for taking the time and effort to evaluate this thesis. The same applies to Prof. Michael Bader, whom I also thank for his challenging but inspiring lectures and the good times at the Bavarian Graduate School of Computational Engineering. Similarly, I am thankful to Prof. Gee whose lectures sparked the interest for numerical mechanics in my first undergraduate semesters and who volunteered as chairmen of the examination board of this thesis.

My thanks expand to Dr. Stefan Adami an extraordinary supervisor and reviewer, who spent hours and days elevating my notes into reports, my slides into presentations and my drafts into papers. Thank you for keeping the Nanoshock group on track and for the memorable times at the workshops.

I thank Prof. Bungartz for his tremendous support during my master study program, which enabled and encouraged me to enroll in a doctoral program to begin with. Your taking on the mentorship for my dissertation project was much appreciated, and doing so even after I rejected your offer to work at your chair speaks for itself.

Thanks is due to Prof. Christian Stemmer not only for the countless technical assistance provided, but for the constructive work and the high levels of trust you put in me during the NFDI4Ing project.

I want to thank Angela Grygier and Hua Liu for their help with the bureaucracy. Even though I never manged to fill any form correctly on the first try, you never lost your patience nor your good attitude towards me.

I thank Josef for the fun and effort in our joint publication, his high work and coding ethics and his friendship outside the office. I appreciate that the words “No time” and “I cannot help” do not exist in your vocabulary. Further, I thank Vladimir for the many joyful hours in our office room. It was a pleasure learning about Russian and German language and culture together. Thanks go out to Nico for his physics and numerics explanations for dummies without which I’d still be lost in the forest of papers.
Although your mountaineering is on a different level than mine, you made sure everyone enjoyed the joint hikes. I thank you, Jakob, for sharing your TUM intel with us. For your sarcasm, cynicism and the vodka: thank you, Alexandr. Naeimeh, thanks for keeping up the good spirit despite being all alone - culture, gender and topic wise. I thank Thomas for his commitment to the shared code base and for keeping a smile on his face and a calm hand even in rough times. For keeping the ALPACA legacy alive and rejuvenating it, I thank you, Alexander.

My deep gratitude to Benjamin without you I would have constantly drowned in NFDI4Ing organizational tasks. Thanks also to Vasiliki for your good mood and for the constant plant-watering efforts. Further, I would like to thank Steffen and Marcus for intense physics debates after hours and the peeks behind the scenes of university politics. Christopher, Christian, Deniz, Rim, Fabian, Aaron, Ludger, and Vladislav, thanks for the great lunch times and the long nights at the Italian restaurant. Thanks to Patrick, Stefan, Aleksandra, Felix, Andrei and all the other colleagues who made the time in-between work more enjoyable.

Outside the chair I want to thank my family. Without the continuous support of my parents and my brother I would not have made it to the finish line. I want to thank my in-laws not only, but in particular for the countless hours of baby-sitting allowing me working on this very manuscript. Thank you Martin and Haosi for continuous friendship throughout the years. And thanks to Nicolas, Felix and Elena: it is a privilege calling you my friends.

And finally, I want to thank Scott Adams, Charles M. Schulz, Bill Watterson, Thomas ‘Tom’ Körner, and Jorge Cham for keeping me entertained throughout the years.
Abstract

Today, resolving all relevant scales and phenomena of practical compressible multiphase flow applications is infeasible for direct numerical simulation (DNS) without subgrid modelling. This is due to the complex flow features such as shock waves, the inherent three-dimensionality of the problem, and the need to resolve a wide range of physical scales. These hindrances stem from a lack of computational power as well as the complexity of the required numerical methods. Nevertheless, a decent range of scales are already representable if compression algorithms and high-performance computing (HPC) systems are utilized. Growing computational power translates hence directly to more powerful DNS as long as the numerical methods and algorithms use (all the features of) the provided hardware efficiently. This, however, is rarely the case.

This thesis presents concepts and implementations for the parallelization of a modular DNS finite-volume method (FVM) compressible multiphase flow solver. The parallelization strategy takes into account the requirements of multiresolution (MR) for spatial and adaptive local-time stepping (LTS) for temporal compression. Additionally, techniques to efficiently incorporate the level-set (LS) algorithm for the representation of multiple phases as well as free surfaces or complex solid geometries are given. The described concepts allow a modular implementation capable of maintaining the parallel performance across numerical schemes of different nominal convergence order. In turn, this modularity eases integration of new schemes speeding up method development and eases benchmarking new methods on large-scale three-dimensional (3D) problems. The performance portability is demonstrated on homogenous supercomputing hardware, with special focus on improved auto-vectorization. Besides thorough verification and validation, various large-scale single- and two-phase simulations of 3D flows demonstrate the capabilities of the solver.
Zusammenfassung


# Contents

List of Figures xi  
Acronyms xiii  
Nomenclature xvii  

1 Introduction 1  

2 Fundamentals 5  
2.1 Governing Equations ........................................ 5  
2.1.1 Multiphase Flows ........................................ 6  
2.2 Numerical Model ........................................... 8  
2.2.1 Flux Evaluations ........................................ 10  
2.2.2 Time Integration .......................................... 12  
2.2.3 Interface Discretization .................................. 13  
2.3 Compression Algorithms .................................... 13  
2.4 Build-up of Supercomputers ................................ 17  
2.5 Code Quality ................................................ 21  
2.6 Challenges for Numerical Software & Parallel Efficiency ........................................... 22  
2.7 Computational Frameworks for Complex Flows ........................................... 24  

3 Accomplishments 25  
3.1 Improved Multiresolution Compression Algorithm .................. 25  
3.1.1 Summary of Hoppe et al. 2022, CMAME ......................... 26  
3.2 Addressing Auto-Vectorization ................................ 27  
3.2.1 Summary of Hoppe et al. 2019, HPCS .......................... 27  
3.3 Large Scale Sharp-interface Simulations .......................... 28  
3.3.1 Summary of Hoppe et al., 2022 CPC .......................... 29  
3.4 Modular Framework ........................................... 30  

4 List of Peer-Reviewed Publications 31  
4.1 Publications comprising this Thesis ............................ 31  
4.2 Further Publications .......................................... 31  

5 Discussion and Concluding Remarks 33  

Bibliography 37
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Schematic of a Cut Cell</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Multiresolution: Averaging and Prediction Operators</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>Prediction Operator: Q-terms</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Schematic of the Interconnect between and within Compute Nodes</td>
<td>18</td>
</tr>
<tr>
<td>2.5</td>
<td>Memory Hierarchy in a Compute Node</td>
<td>19</td>
</tr>
<tr>
<td>2.6</td>
<td>Concept of the Internal Structure of a Core</td>
<td>20</td>
</tr>
<tr>
<td>2.7</td>
<td>Pipelining Concept within a Core</td>
<td>21</td>
</tr>
<tr>
<td>2.8</td>
<td>Vectorization versus Sequential Addition</td>
<td>22</td>
</tr>
<tr>
<td>2.9</td>
<td>Outline of the roofline model</td>
<td>23</td>
</tr>
</tbody>
</table>
Acronyms

3D three-dimensional

AGU address generation unit

ALPACA Adaptive Level-set PArallel Code ALPACA

ALTS adaptive LTS

ALU arithmetic logic unit

AMR adaptive mesh refinement

API application programming interface

CFD computational fluid dynamics

CFL Courant-Friedrichs-Lewy

CPU central processing unit

CRTP curiously recurring template pattern

DGM Discontinuous Galerkin method

DNS direct numerical simulation

DOF degree of freedom

ENO essentially non-oscillatory

EOS equation of state
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDM</td>
<td>finite-difference method</td>
</tr>
<tr>
<td>FEM</td>
<td>finite-element method</td>
</tr>
<tr>
<td>FLAC</td>
<td>Free Lossless Audio Codec</td>
</tr>
<tr>
<td>FLOP</td>
<td>floating-point operations</td>
</tr>
<tr>
<td>FPU</td>
<td>floating-point unit</td>
</tr>
<tr>
<td>FVM</td>
<td>finite-volume method</td>
</tr>
<tr>
<td>GFM</td>
<td>ghost-fluid method</td>
</tr>
<tr>
<td>GPU</td>
<td>graphics processing unit</td>
</tr>
<tr>
<td>HLLC</td>
<td>Harten-Lax-van Leer with restored contact surface</td>
</tr>
<tr>
<td>HPC</td>
<td>high-performance computing</td>
</tr>
<tr>
<td>I/O</td>
<td>input/output</td>
</tr>
<tr>
<td>ID</td>
<td>identifier</td>
</tr>
<tr>
<td>JPEG</td>
<td>Joint Photographic Experts Group</td>
</tr>
<tr>
<td>LAN</td>
<td>local area network</td>
</tr>
<tr>
<td>LRZ</td>
<td>Leibniz Supercomputing Centre</td>
</tr>
<tr>
<td>LS</td>
<td>level-set</td>
</tr>
<tr>
<td>LTS</td>
<td>local-time stepping</td>
</tr>
<tr>
<td>MP3</td>
<td>Moving Picture Experts Group Audio Layer III</td>
</tr>
<tr>
<td>MPI</td>
<td>Message Passing Interface</td>
</tr>
</tbody>
</table>
MR  multiresolution

MUSCL  Monotonic Upstream-centered Scheme for Conservation Laws

NSE  Navier-Stokes equations

NUMA  non-uniform memory access

OS  operating system

PC  personal computer

PNG  Portable Network Graphics

RAM  random-access memory

RANS  Reynolds-averaged Navier-Stokes

RK  Runge-Kutta

RTI  Rayleigh-Taylor instability

SFC  space-filling curve

SIMD  single instruction multiple data

SMT  simultaneous multithreading

TENO  targeted ENO

TPU  tensor processing unit

TVD  total variation diminishing

VOF  volume of fluid

WENO  weighted ENO
Nomenclature

Thermo-Fluid Dynamics

\( \rho \)  \hspace{1cm} \text{Density}  \hspace{1cm} \frac{\text{kg}}{\text{m}^3} \\
\( \mathbf{v} \)  \hspace{1cm} \text{Velocity vector}  \hspace{1cm} \frac{\text{m}}{\text{s}} \\
\( p \)  \hspace{1cm} \text{Pressure}  \hspace{1cm} \frac{\text{kg}}{\text{m} \cdot \text{s}^2} \\
\( e \)  \hspace{1cm} \text{Specific internal energy}  \hspace{1cm} \frac{\text{m}^2}{\text{s}^2} \\
\( E \)  \hspace{1cm} \text{Total specific energy}  \hspace{1cm} \frac{\text{m}^2}{\text{s}^2} \\
\( \Pi \)  \hspace{1cm} \text{Stress tensor}  \hspace{1cm} \frac{\text{kg}}{\text{m} \cdot \text{s}^2} \\
\( \Upsilon \)  \hspace{1cm} \text{Viscous stress tensor}  \hspace{1cm} \frac{\text{kg}}{\text{m} \cdot \text{s}^2} \\
\( q \)  \hspace{1cm} \text{Heat flux density}  \hspace{1cm} \frac{\text{kg}}{\text{s}^2} \\
\( g \)  \hspace{1cm} \text{Gravitational pull}  \hspace{1cm} \frac{\text{m}}{\text{s}^2} \\
\( T \)  \hspace{1cm} \text{Temperature}  \hspace{1cm} \text{K} \\
\( k \)  \hspace{1cm} \text{Thermal conductivity}  \hspace{1cm} \frac{\text{kg} \cdot \text{m}}{\text{s}^2 \cdot \text{K}} \\
\( \mu \)  \hspace{1cm} \text{Dynamic viscosity}  \hspace{1cm} \frac{\text{kg}}{\text{m} \cdot \text{s}} \\
\( \mathbf{U} \)  \hspace{1cm} \text{State vector}  \hspace{1cm} \begin{bmatrix} \frac{\text{kg}}{\text{m}^3}, \frac{\text{kg}}{\text{m}^2 \cdot \text{s}}, \frac{\text{kg}}{\text{m} \cdot \text{s}^2} \end{bmatrix}^T \\
\( \mathbf{f} \)  \hspace{1cm} \text{Volume-force vector}  \hspace{1cm} \begin{bmatrix} \frac{\text{kg}}{\text{m}^3}, \frac{\text{kg}}{\text{m}^2 \cdot \text{s}}, \frac{\text{kg}}{\text{m} \cdot \text{s}^2} \end{bmatrix}^T \\
\( \gamma \)  \hspace{1cm} \text{Heat capacity ratio}  \hspace{1cm} — \\
\( \sigma \)  \hspace{1cm} \text{Surface tension coefficient}  \hspace{1cm} \frac{\text{kg}}{\text{s}^2} \\
\( a \)  \hspace{1cm} \text{(Local) speed of sound}  \hspace{1cm} \frac{\text{m}}{\text{s}} \\

Numerics

\( F \)  \hspace{1cm} \text{Flux}  \hspace{1cm} \begin{bmatrix} \frac{\text{kg}}{\text{m}^3}, \frac{\text{kg}}{\text{m}^2 \cdot \text{s}}, \frac{\text{kg} \cdot \text{m}^3}{\text{s}^2} \end{bmatrix}^T \\
\( \mathbf{R} \)  \hspace{1cm} \text{Roe eigenvector}  \hspace{1cm} — \\
\( \Lambda \)  \hspace{1cm} \text{Roe eigenvalue matrix}  \hspace{1cm} —
### Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Signal speed estimate</td>
<td>$\text{m/s}$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Velocity component normal to Riemann problem</td>
<td>$\text{m/s}$</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Permutated velocity components of HLLC-input state</td>
<td>$\text{m/s}$</td>
</tr>
<tr>
<td>$\Theta$</td>
<td>Stencil</td>
<td>—</td>
</tr>
<tr>
<td>$w, a, b, d$</td>
<td>WENO weights</td>
<td>—</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Small positive quantity</td>
<td>—</td>
</tr>
</tbody>
</table>

### Time

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>Time</td>
<td>$\text{s}$</td>
</tr>
<tr>
<td>$\Delta t$</td>
<td>Time step</td>
<td>$\text{s}$</td>
</tr>
<tr>
<td>$C$</td>
<td>CFL constant</td>
<td>—</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Pseudo time</td>
<td>—</td>
</tr>
</tbody>
</table>

### Geometric Quantities

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V$</td>
<td>Volume</td>
<td>$\text{m}^3$</td>
</tr>
<tr>
<td>$A$</td>
<td>Aperture</td>
<td>$\text{m}$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Volume fraction</td>
<td>$\text{m}^3$</td>
</tr>
<tr>
<td>$n$</td>
<td>Normal vector</td>
<td>$\text{m}$</td>
</tr>
<tr>
<td>$\Delta x$</td>
<td>Cell size</td>
<td>$\text{m}$</td>
</tr>
</tbody>
</table>

### Multiphase

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>Level set</td>
<td>—</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Fluid identifier</td>
<td>—</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>Ideal-gas variable for diffusive interface model</td>
<td>—</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>Background-pressure variable for diffusive interface model</td>
<td>$\text{kg/m}^2$</td>
</tr>
</tbody>
</table>

### Multiresolution

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l$</td>
<td>Level of refinement</td>
<td>—</td>
</tr>
<tr>
<td>$m$</td>
<td>Level index</td>
<td>—</td>
</tr>
<tr>
<td>$c$</td>
<td>Scaling coefficient</td>
<td>—</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Scaling Function</td>
<td>—</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>$d$</td>
<td>Detail</td>
<td></td>
</tr>
<tr>
<td>$\Psi$</td>
<td>Wavelet function</td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>Averaging operator</td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>Prediction operator</td>
<td></td>
</tr>
<tr>
<td>$Q$</td>
<td>Q-terms of the prediction operator</td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>Prediction coefficients</td>
<td></td>
</tr>
</tbody>
</table>

**Mathematical Symbols**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\otimes$</td>
<td>Tensor product</td>
</tr>
<tr>
<td>$\cdot$</td>
<td>Dot product</td>
</tr>
<tr>
<td>$\nabla$</td>
<td>Gradient</td>
</tr>
<tr>
<td>$\partial$</td>
<td>Partial derivative</td>
</tr>
</tbody>
</table>

**Miscellaneous**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$</td>
<td>Identity matrix</td>
</tr>
<tr>
<td>$i, j, k$</td>
<td>Cell indices</td>
</tr>
<tr>
<td>$p, q, r, s$</td>
<td>Arbitrary indices</td>
</tr>
<tr>
<td>$N$</td>
<td>Arbitrary natural number</td>
</tr>
<tr>
<td>$\mathbb{N}$</td>
<td>Set of all natural numbers</td>
</tr>
<tr>
<td>$\mathbb{Z}$</td>
<td>Set of all integer numbers</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Compressible multiphase flows govern a wide variety of engineering applications. Cavitation bubbles damage, e.g., ship propellers [34] by emitting shockwaves when collapsing. Bubble dynamics are also relevant within nozzles such as fuel injectors or for flows around airfoils, submarines and alike [23, 169, 172, 173, 194, 195]. Breakup and atomization of liquid droplets is relevant for combustion engines [49, 151, 152, 196, 197, 237, 241], supersonic aircraft design [54, 111, 122, 217], and nuclear safety [36, 37]. Nanoparticles can be generated using plasma-based laser ablation [116, 153], but also biomedical procedures rely on compressible flow phenomena including shockwaves, phase-separation and phase change. Examples include sonoporation [164], lithotripsy [133, 147, 150, 176], histotripsy [131], artificial heart valves and pumps [30], blast trauma care [55, 135, 178], or manufacturing [235] and delivery [210, 233] of medical drugs. Further applications range from boiling flows [38, 138, 141] and liquid jets [132, 184, 215, 216, 244] to blast effects on structures [57, 75, 91, 92, 159, 160].

Despite successful practical application of complex flows in the aforementioned areas, not all underlying physical processes are understood. Hence, to make these methods effective or at least more efficient a better understanding is required. Experimental investigations have revealed many details of compressible multiphase flows: From singular-bubble collapse and rebound behavior [181] and jetting-behavior of bubbles [21, 32, 87, 143, 163, 193] to fragmentation and atomization of droplets [102, 235] to name, but a few. Yet, they cannot produce complete field solutions and are often limited by measurement equipment, e.g. exposure times of cameras or measurements probes altering the observed flow field. The more complex the investigated flow becomes the more cumbersome, expensive and potentially dangerous experiments become [228, 236].

Numerical investigations do not suffer from these shortcomings, but bring their own challenges. For the simulation of multiphase flows, one such challenge lies in correctly representing the fluid interface(s). While the Navier-Stokes equations (NSE) accurately describe the flow of a single phase, the interplay between two or more immiscible fluids is not described by them. Hence, one needs to introduce additional models for the fluid interface and the transport across it. Such interface models usually fall into the categories of interface capturing with smeared interfaces or interface tracking schemes with a sharp interface representation. The volume of fluid (VOF) methods [101, 161] or the diffuse interface method [1, 4] are prominent examples for the first and the free-Lagrange [11, 227], the front-tracking [78, 79, 225, 229], or the level-set (LS) algorithm [47, 167, 213] for the latter.

Numerical discretization schemes for (compressible) flows should ideally provide vanishing smearing of flow states, prevent oscillations near discontinuities, maintain conservation and ideally provide high-order convergence. The (cell-averaged) finite-volume method (FVM) Godunov’s approach [80] achieves these goals and is hence typically employed. As for flux limiting procedures [142, 221], it ensures discrete conservation. Similar to the Monotonic Upstream-centered Scheme for Conservation Laws (MUSCL) approach [230], it allows combinations of approximate Riemann solvers [64, 97, 186] but enables application of
higher-order methods than MUSCL. To reach these higher-orders, weighted ENO (WENO) type reconstruction stencils [118] are typically chosen [16, 119, 170, 219] as they allow straightforward construction of stencils with arbitrary order while introducing limited numerical diffusion. However, their computational cost is high and for certain problem types commonly selected combinations of Riemann solvers and high-order stencils may fail [65, 66, 70]. Hence, support for a wide range of numerical schemes is beneficial in a simulation framework for compressible flow research. It shall be noted, that also finite-difference method (FDM) [146] and Discontinuous Galerkin methods (DGMs) are applied to compressible flows [40]. For these methods, however, it is more challenging to meet all the mentioned desired properties [62, 158, 174, 182]. Hence, they are not further discussed within this thesis.

For Godunov-type schemes, in addition to the numerically expensive discretization, the wide range of scales poses high computational demands. Physically, all scales from the domain size to the Kolmogorov scale and in multiphase flows additionally the scale of the smallest interface waves need to be resolved or modelled appropriately. Although subgrid-modelling methods [73, 208, 223] are commonly used the focus here lies on direct numerical simulation (DNS), i.e. the aim to ultimately resolve all relevant scales. While this eliminates modelling errors it makes the application of compression algorithms and the usage of high-performance computing (HPC) systems inevitable. A variety of compression algorithms exists. While the most prominent ones stem from media compression such as the lossy Joint Photographic Experts Group (JPEG) [112] [112] or Moving Picture Experts Group Audio Layer III (MP3) [113] and the lossless Portable Network Graphics (PNG) [114] or Free Lossless Audio Codec (FLAC) [68] for image and audio compression, respectively. For the application of compression algorithms to complex flows, however, the aforementioned properties of the numerical methods must stay intact. Additionally, the compression algorithms should compress both number of floating-point operations (FLOP) and the memory footprint of the simulations. While multigrid [60], sparse grids [204], adaptive DGM and wavelet-collocation [231] have been applied to complex flows [5, 98, 232, 245], the commonly applied techniques within the simulations of physical application are the gradient-based adaptive mesh refinement (AMR) [17] and the wavelet-based multiresolution (MR) [95] algorithms [48, 90, 106, 162, 198, 207, 209, 239]. Utilization of such mesh compression, however, introduces additional algorithmic complexity and may affect the performance of the computation on HPC hardware. In order to efficiently use such hardware, all its layers of parallelism from distributed-memory to single instruction multiple data (SIMD) parallelism need to be addressed [89, 236].

This thesis, hence, addresses the aforementioned challenges for numerical simulation of compressible multiphase flows. To improve the parallel performance of MR compression on modern HPC systems a block-based variant is proposed. The scheme is the algorithmic backbone of the open-source C++ framework Adaptive Level-set PArallel Code ALPACA (ALPACA) [3]. For the algorithm design a modular built-up is proposed to allow exchanging numerical schemes and parameters, material models, equation of states (EOSs), and partitioning methods. The modular design even allows exchanging of the underlying equations at compile time, yielding straightforward extension to different governing equations. Beyond that, the narrowband algorithm [2, 39, 171] for LS simulations is modified to naturally fit into the block-based MR scheme allowing for large scale three-dimensional (3D) simulations using the sharp-interface method. The implementation focuses on modern programming patterns [71, 81, 149, 155] and the use of C++20 features to enable the algorithmic flexibility. For increased usability clean-coding techniques and high-level source code optimizations are employed. Therein, the flexibility and maintainability of the code base is also taken into account. The aim is to obtain steady parallel performance across the mod-
ules while providing a compute environment that fosters rapid large-scale prototyping of new numerical methods and novel physical setups alike allowing a more agile approach to computational fluid dynamics (CFD).

The remainder of this publication-based thesis is structured as follows. In chapter 2 the governing equations and their numerical discretization are stated. Details on mesh-compression algorithms in general and MR in particular are given. Furthermore, the parallelization concepts employed in HPC hardware are stated and their implications for software implementations are discussed. A discussion of competing compute frameworks for complex flows is given. Chapter 3 lists the main findings of this thesis and chapter 4 links the respective publications. A conclusion including a review on relevant literature is drawn in chapter 5.
Chapter 2

Fundamentals

2.1 Governing Equations

Under continuum assumptions the compressible NSE

\[
\begin{align*}
\frac{\partial \rho}{\partial t} &= -\nabla \rho \mathbf{v}, \quad (2.1) \\
\frac{\partial \rho \mathbf{v}}{\partial t} &= -\nabla (\rho \mathbf{v} \otimes \mathbf{v} + \Pi) + \rho \mathbf{g}, \quad (2.2) \\
\frac{\partial \rho E}{\partial t} &= -\nabla (\rho E \mathbf{v} + \Pi \mathbf{v} - \mathbf{q}) + \rho \mathbf{g} \cdot \mathbf{v}, \quad (2.3)
\end{align*}
\]

describe the flow of a fluid in 3D space \( \mathbf{x} = [x^1, x^2, x^3]^T \) and time \( t \) by stating the conservation of mass, momentum and energy, respectively. In the above equations, density \( \rho (\mathbf{x}, t) \), the velocity vector \( \mathbf{v} (\mathbf{x}, t) \), the gravitational acceleration vector \( \mathbf{g} (\mathbf{x}, t) \), and the total specific energy \( E = e + \mathbf{v}^2 / 2 \) which is the sum of the specific internal energy \( e (\mathbf{x}, t) \) and the specific kinetic energy are used. Note, throughout this thesis the spatial \( \mathbf{x} \) and temporal \( t \) variables are skipped unless they are crucial for the understanding of the described phenomenon. Heat fluxes are considered by

\[
\mathbf{q} = -k \nabla T \quad (2.4)
\]

with the thermal conductivity \( k (\mathbf{x}, t) \) and temperature \( T (\mathbf{x}, t) \). The stress tensor for Newtonian fluids is given by

\[
\Pi = -p \mathbf{I} + \Upsilon = -p \mathbf{I} + \mu \left( \nabla \otimes \mathbf{v} + (\nabla \otimes \mathbf{v})^T - \frac{2}{3} \text{div} (\mathbf{v}) \mathbf{I} \right) \quad (2.5)
\]

with pressure \( p (\mathbf{x}, t) \), dynamic viscosity \( \mu \), identity matrix \( \mathbf{I} \) and viscous stress tensor \( \Upsilon (\mathbf{x}, t) \). The integral formulation of eqs. (2.1) to (2.3) reads

\[
\begin{align*}
\int_V \frac{\partial \rho}{\partial t} dV &= -\oint_{\partial V} (\rho \mathbf{v}) \cdot \mathbf{n} d\partial V \quad (2.6) \\
\int_V \frac{\partial \rho \mathbf{v}}{\partial t} dV &= -\oint_{\partial V} (\rho \mathbf{v} \otimes \mathbf{v} - p \mathbf{I} + \Upsilon) \cdot \mathbf{n} d\partial V + \int_V \rho \mathbf{g} dV \quad (2.7) \\
\int_V \frac{\partial \rho E}{\partial t} dV &= -\oint_{\partial V} (\rho E \mathbf{v} + p \mathbf{I} \mathbf{v} + \Upsilon \mathbf{v} + \mathbf{q}) \cdot \mathbf{n} d\partial V + \int_V \rho \mathbf{g} \cdot \mathbf{v} dV, \quad (2.8)
\end{align*}
\]

wherein \( V \) denotes a material volume. By rearranging and summarizing terms in this set of equations the flux-based formulation gives

\[
\begin{align*}
\int_V \frac{\partial \mathbf{U}}{\partial t} dV &= -\oint_{\partial V} (\mathbf{F}^c + \mathbf{F}^\mu + \mathbf{F}^\Upsilon) d\partial V + \int_V \mathbf{f} dV \quad (2.9)
\end{align*}
\]
with the state vector \( \mathbf{U} = [\rho \quad \rho v \quad \rho E]^T \), the convective contribution

\[
\mathbf{F}^c = \begin{bmatrix}
(\rho v) \cdot \mathbf{n} \\
(\rho v \otimes \mathbf{v} - p \mathbf{I}) \cdot \mathbf{n} \\
(\rho \mathbf{E} \mathbf{v} - p \mathbf{I}) \cdot \mathbf{n}
\end{bmatrix},
\]

(2.10)

the viscous contribution

\[
\mathbf{F}^\mu = \begin{bmatrix}
0 \cdot \mathbf{n} \\
-\Upsilon \cdot \mathbf{n} \\
-\Upsilon \mathbf{v} \cdot \mathbf{n}
\end{bmatrix},
\]

(2.11)

heat-flux densities

\[
\mathbf{F}^q = \begin{bmatrix}
0 \\
0 \\
q \cdot \mathbf{n}
\end{bmatrix},
\]

(2.12)

and the volume-force vector

\[
\mathbf{f} = \begin{bmatrix}
0 \cdot \mathbf{n} \\
(\rho g) \cdot \mathbf{n} \\
(\rho g \cdot \mathbf{v}) \cdot \mathbf{n}
\end{bmatrix}.
\]

(2.13)

Thermodynamic closure of the NSE is obtained by an EOS in the form

\[
p = f(\rho, e),
\]

(2.14)

which relates pressure, density, and energy.

### 2.1.1 Multiphase Flows

The interaction between immiscible phases is typically modelled with either interface tracking or interface capturing schemes. The schemes relevant to this thesis of either group define the interface with respect to a material or numerical property. E. g. in the diffusive interface model \cite{1, 6} the ratio of the specific heats \( \gamma(x, t) \) and the background pressure \( p_{\infty}(x, t) \) of the stiffened gas EOS \cite{93} are used to define \( \Gamma(x, t) = \frac{1}{\gamma - 1} \) and \( \Omega(x, t) = \frac{1}{\gamma - 1} p_{\infty} \). For these two properties the additional advection equations

\[
\frac{\partial}{\partial t} \left( \begin{array}{c}
\Gamma \\
\Omega
\end{array} \right) + \frac{\partial}{\partial x} u \left( \begin{array}{c}
\Gamma \\
\Omega
\end{array} \right) + \frac{\partial}{\partial y} v \left( \begin{array}{c}
\Gamma \\
\Omega
\end{array} \right) + \frac{\partial}{\partial z} w \left( \begin{array}{c}
\Gamma \\
\Omega
\end{array} \right) = 0,
\]

(2.15)

are solved alongside eqs. (2.1) to (2.3). The interface position is then defined with respect to transition in \( \gamma \). Note, for an ideal gas law only the first line of eq. (2.15) is required. Consequently, more complex EOSs might require solving even more equations. In VOF formulations \cite{101, 161}, a volume fraction function \( \alpha(x, t) \) is advected in addition to the NSE, as

\[
\frac{\partial \alpha}{\partial t} + (\mathbf{v} \cdot \nabla) \alpha = 0.
\]

(2.16)

Therein, \( \alpha = 1 \) in one and \( \alpha = 0 \) in the other fluid, respectively. For immiscible fluids the interface is defined by the \( \alpha = 0.5 \) contour. Despite this precise definition for the interface location, the actual state change from one fluid to the other is smeared out over all cells in
2.1 Governing Equations

which \(0 < \alpha < 1\). The volume-fraction approach is typically coupled with mixture models for the fluid states [4, 130]. For example the mass conservation eq. (2.1) is split for each phase to read

\[
\begin{align*}
\frac{\partial \alpha \rho_1}{\partial t} &= -\text{div}(\alpha \rho_1 \mathbf{v}), \\
\frac{\partial (1-\alpha) \rho_2}{\partial t} &= -\text{div}((1-\alpha) \rho_2 \mathbf{v}).
\end{align*}
\]

(2.17) (2.18)

Therein, the overall density is given by the weighted sum of the partial densities. Similarly, eqs. (2.2) and (2.3) may be split and enriched by additional terms to model more complex thermodynamic phenomena [185].

In contrast, interface tracking schemes such as free-Lagrange [11, 227], front-tracking [78, 79, 225, 229], or the LS algorithm [47, 167, 213] represent the interface sharply. Here, the focus lies on the LS approach as it locates the interface accurately and simplifies imposing arbitrary coupling conditions between the phases. Furthermore, it handles topological changes [77] naturally and gives geometric quantities via straightforward calculations [201]. In addition, it allows prescribing interfacial physics directly [33]. By adjusting the exchange terms across the interface the LS model can also be used to model free surfaces or solid boundaries with complex geometries [154]. For a more extensive overview about LS methods and comparison to other interface modeling approaches the interested reader is referred to [167, 201].

In the LS method, also, the additional advection equation

\[
\frac{\partial \phi}{\partial \tau} = -u_S n_\xi \cdot \nabla \phi = -u_S |\nabla \phi|,
\]

(2.19)

with the interface velocity \(u_S(x,t)\) obtained from solving the two-material Riemann problem for the stress balance

\[
-\mathbf{n}_S \cdot \Pi (U^1_S) \cdot \mathbf{n}_S + \mathbf{n}_S \cdot \Pi (U^2_S) \cdot \mathbf{n}_S = -\sigma \mathbf{n}_S (\nabla^T \cdot \mathbf{n}_S) \cdot \mathbf{n}_S,
\]

(2.20)

is solved for the signed-distance function \(\phi(x,t)\). Therein,

\[
\mathbf{n}_\xi (x) = \frac{\nabla \phi (x)}{||\nabla \phi (x)||},
\]

(2.21)

is the interface normal pointing towards the respective fluid, indicated by \(\xi \in 1,2\). By defining the interface as

\[
S(x,t) = \{x : \phi(x,t) = 0\},
\]

(2.22)

an exact representation of the interface is obtained. Hence, this method is called a sharp-interface method. Away from the interface, the sign of \(\phi\) indicates which of the two fluids is present. Advecting the interface according to eq. (2.19) might, however, violate the signed-distance property of \(\phi\). To recover this property the steady-state solution of the reinitialization equation

\[
\frac{\partial \phi}{\partial \tau} = \text{sign}(\phi)(1 - |\nabla \phi|)
\]

(2.23)

needs to be found in pseudo-time \(\tau\) [213]. However, finding such a solution while maintaining conservation, i.e. leaving the zero contour intact, is challenging. Numerous types of reinitialization techniques have been developed addressing this issue [156, 166]. Evaluating fluxes at the interface across two phases may produce unphysical oscillations. Therefore,
the ghost-fluid method (GFM) [59] treats the phases separately. The states of each fluid are extrapolated by the steady-state solution of

$$\frac{\partial U(x)}{\partial \tau} = n_s(x) \cdot \nabla U(x).$$

(2.24)

Interactions between two fluids are considered by adding

$$- \oint_S \left( F^c_S + F^\mu_S + F^q_S \right) d \partial V$$

(2.25)

on the right-hand side of eq. (2.9). Where the interface-flux densities are defined as

$$F^c_S = \begin{bmatrix} (\rho v) \cdot n_s \\ (\rho v \otimes v - pI) \cdot n_s \\ (\rho e v - pI) \cdot n_s \end{bmatrix}$$

(2.26)

for convection,

$$F^\mu_S = \begin{bmatrix} 0 \cdot n_s \\ -\Upsilon \cdot n_s \\ -\Upsilon v \cdot n_s \end{bmatrix}$$

(2.27)

for viscosity and

$$F^q_S = \begin{bmatrix} 0 \\ 0 \\ q \cdot n_s \end{bmatrix},$$

(2.28)

for heat-transfer, respectively. The LS algorithm allows using distinct EOSs for each fluid as long as the interface states are compatible.

### 2.2 Numerical Model

The aforementioned equations are discretized to be solved numerically. A DNS approach is followed without any models for under-resolved physical effects in contrast to e.g. large-eddy or Reynolds-averaged Navier-Stokes (RANS) simulations. While this eliminates modeling errors it requires finer discretization and hence more compute power as discussed later in this thesis.

The domain $V$ is discretized by partitioning it into a disjunct set of cuboid finite volumes

$$V_{i,j,k} \subseteq V \mid \Sigma V_{i,j,k} = V$$

(2.29)

with cell size $\Delta x$. Note, throughout this work subscript $i,j,k \in \mathbb{N}$ are indices for the three spatial directions, respectively. Also, they indicate that the respective quantity is discretized. Interface capturing schemes can be computed directly on the thus defined mesh. Sharp-interface schemes, however, require additional attention as each cell may contain part of the fluid interface $S$.

In the LS approach, the interface is linearized inside such cells, and cell-face apertures $A^\xi_{i,j+1/2,k+1/2}(t)$, $A^\xi_{i+1/2,j,k+1/2}(t)$, and $A^\xi_{i+1/2,j+1/2,k}(t)$ are defined, which describe the cell-face fraction
covered by fluid $\xi$. From them, the computation of the volume fraction $\alpha_{i,j,k}^\xi$ is straightforward. The volume fraction denotes the volumetric portion of $V_{i,j,k}$ which is covered by the respective fluid. Figure 2.1 visualizes the geometric properties in a cell cut by the interface. From this interface discretization two partially-overlapping sets of cells are defined as

\begin{align}
\mathcal{V}^\xi &= \left\{ V_{i,j,k} | \alpha_{i,j,k}^\xi \neq 0 \right\}, \text{ and} \\
S &= \left\{ V_{i,j,k} | \alpha_{i,j,k}^1 \neq 0 \text{ and } \alpha_{i,j,k}^2 \neq 0 \right\}.
\end{align}

Using volume-averaged conservative quantities indicated by the over bar

\[ \bar{U} = \frac{1}{\alpha_{i,j,k}} \int_{\Delta_{i,j,k}} U \, dV \]

together with eqs. (2.30) and (2.31), the discretized form of eq. (2.9) becomes

\[ \frac{\partial \left( \alpha_{i,j,k}^\xi \bar{U} \right)}{\partial t} = - \oint_{\partial \mathcal{V}_{i,j,k}^\xi \setminus S_{i,j,k}} \mathbf{F} + \mathbf{F}^\mu + \mathbf{F}^q d\partial V - \oint_{S_{i,j,k}} \mathbf{F}_S^c + \mathbf{F}_S^\mu + \mathbf{F}_S^q d\partial V + \int_{\mathcal{V}_{i,j,k}^\xi} f \, dV. \]

Where the integrals on the right-hand side account for fluxes over the cell faces, the interaction between the fluids, and the volume forces acting on the fluids.
2.2.1 Flux Evaluations

Equation (2.33) is solved by introducing numerical flux functions for each of the fluxes and the volume force term. With these numerical flux functions the conservative balance equations for each term on the right-hand side of eq. (2.33) is defined. E.g., the convective flux integral becomes

\[
- \int_{\partial V_{i,j,k} \setminus S_{i,j,k}} F^c \, d\partial V = \Delta x^2 \left( A^\xi_{i-\frac{1}{2},j,k}^F F^c_{i-\frac{1}{2},j,k} - A^\xi_{i+\frac{1}{2},j,k}^F F^c_{i+\frac{1}{2},j,k} + A^\xi_{i,j-\frac{1}{2},k}^F F^c_{i,j-\frac{1}{2},k} - A^\xi_{i,j+\frac{1}{2},k}^F F^c_{i,j+\frac{1}{2},k} + A^\xi_{i,j,k-\frac{1}{2}} F^c_{i,j,k-\frac{1}{2}} - A^\xi_{i,j,k+\frac{1}{2}} F^c_{i,j,k+\frac{1}{2}} \right),
\]

(2.34)

Here the number in the superscript indicates the component of the Flux vector \( F = [F^1, F^2, F^3]^T \) for each spatial dimension, respectively. The remaining integrals on the right-hand side are treated similarly.

Computation of \( F^u \) and \( F^d \) respectively is done by evaluating the gradients in eqs. (2.4) and (2.5) with discretization stencils of desired convergence order at the respective cell faces. Gravitational volume forces are trivially computed evaluating eq. (2.13) at cell centers.

The numerical cell-face flux \( F^c \) is determined by solving a Riemann problem between neighboring cells. While exact iterative solvers exists, using explicit approximate Riemann solver has shown to provide sufficient accuracy at a fraction of the computational cost. A variety of such approximate Riemann solvers is found in literature [97, 168, 186, 191, 222]. The solver of Roe [186] and the Harten-Lax-van Leer with restored contact surface (HLLC) [222] are the dominant representatives of two major groups of Riemann solvers commonly used in CFD.

The Roe flux reads

\[
F_{i+\frac{1}{2},j,k}^{\text{Roe}} = \frac{1}{2} \left( F_{i+1,j,k} - F_{i,j,k} \right) - \frac{1}{2} R_{i+\frac{1}{2},j,k} \Lambda_{i+\frac{1}{2},j,k} R^{-1}_{i+\frac{1}{2},j,k} \left( \bar{U}_{i+1,j,k} - \bar{U}_{i,j,k} \right),
\]

(2.35)

with the right eigenvector \( R \) and the diagonal eigenvalue matrix \( \Lambda \) of the Jacobian \( \partial F / \partial U \). If the solver is directly applied to cell-averaged states as described above the scheme is first-order accurate. Yet, higher convergence orders are achievable by using the flux-vector splitting form of eq. (2.35)

\[
F^+_{i+\frac{1}{2},j,k} = F_{i,j,k} + R_{i+\frac{1}{2},j,k} \Lambda_{i+\frac{1}{2},j,k} R^{-1}_{i+\frac{1}{2},j,k} \bar{U}_{i,j,k} \quad \text{and} \quad F^-_{i+\frac{1}{2},j,k} = F_{i+1,j,k} - R_{i+\frac{1}{2},j,k} \Lambda_{i+\frac{1}{2},j,k} R^{-1}_{i+\frac{1}{2},j,k} \bar{U}_{i+1,j,k}
\]

(2.36)

and therein employing higher-order reconstructions of \( F^+_{i+\frac{1}{2},j,k} \) and \( F^-_{i+\frac{1}{2},j,k} \) denoted by a hat.

The higher-order flux then reads

\[
F_{i+\frac{1}{2},j,k}^{\text{Roe}} = \frac{1}{2} \left( \hat{F}^+_{i+\frac{1}{2},j,k} + \hat{F}^-_{i+\frac{1}{2},j,k} \right)
\]

(2.38)

For the HLLC solver [222], the flux is given by

\[
F_{\text{HLLC}} = \frac{1 + \text{sign}(S^+)}{2} \left[ F_L + S^- (U_L^* - \bar{U}_L) \right] + \frac{1 + \text{sign}(S^-)}{2} \left[ F_R + S^+ (U_R^* - \bar{U}_R) \right],
\]

(2.39)

with

\[
S^- = \min(S_L, 0), \quad \text{and} \quad S^+ = \max(S_R, 0)
\]

(2.40)

(2.41)
for the left and right states denoted by subscripts $L$ and $R$, respectively. For the signal speeds estimates $S_L, S_R, S^*$ various competing definitions exist \[15, 44, 53, 86, 220\]. The fluxes are given by $F_{L/R} = F(U_{L/R})$

$$U_{L/R}^* = \frac{S_{L/R} - \beta}{S_{L/R} - S^*} \left( \begin{array}{c} \rho_{L/R} \\ \rho_{L/R} \zeta_1 \\ \rho_{L/R} \zeta_2 \\ \rho_{L/R} \zeta_3 \\ E_{L/R} + (S^* - \beta)(\rho_{L/R} S^* + \frac{p_{L/R}}{S_{L/R} - \beta}) \end{array} \right).$$  \hspace{1cm} (2.42)

If the flux is computed in $x^1$-direction, $\beta = v^1$, $\zeta_1 = S^*$, $\zeta_2 = v^2$, and $\zeta_3 = v^3$ for the respective components of the velocity vector $v$. Along the $x^1$-direction and $x^2$-direction, respectively $\beta = v^2$ and $\beta = v^3$. Accordingly, $\zeta_1 = v^1$, $\zeta_2 = S^*$, and $\zeta_3 = v^3$ or $\zeta_1 = v^1$, $\zeta_2 = v^2$, and $\zeta_3 = S^*$. To obtain higher orders when using the HLLC solver, the input left and right states are determined from a high-order reconstruction.

For the reconstruction in both solver types, typically, essentially non-oscillatory (ENO) \[96\] type stencils are chosen as they minimize spurious oscillations. The most prominent such stencil is the (fifth-order) WENO stencil by Jiang and Shu \[118\]. It uses a nonlinear weighting of three second-order three-point sub-stencils $\Theta_p, p = 1, 2, 3$. For each stencil an accompanying smoothness measure is computed. If the input function is sufficiently smooth all stencils are optimally weighted and recover the nominal fifth-order convergence rate. Otherwise, the contribution of the sub-stencil containing the non-smooth region of the function is reduced accordingly. The reconstructed value along the first spatial dimension is obtained from

$$U_{i+\frac{1}{2},j,k} = w_1 \bar{U}_{\Theta_1} + w_2 \bar{U}_{\Theta_2} + w_3 \bar{U}_{\Theta_3},$$  \hspace{1cm} (2.43)

where

$$\bar{U}_{\Theta_1} = \frac{1}{3} \bar{U}_{i-2,j,k} - \frac{7}{6} \bar{U}_{i-1,j,k} + \frac{11}{6} \bar{U}_{i,j,k},$$ \hspace{1cm} (2.44)
$$\bar{U}_{\Theta_2} = -\frac{1}{6} \bar{U}_{i-1,j,k} + \frac{5}{6} \bar{U}_{i,j,k} + \frac{1}{3} \bar{U}_{i+1,j,k},$$ \hspace{1cm} (2.45)
$$\bar{U}_{\Theta_3} = \frac{1}{3} \bar{U}_{i,j,k} + \frac{5}{6} \bar{U}_{i+1,j,k} - \frac{1}{6} \bar{U}_{i+2,j,k},$$ \hspace{1cm} (2.46)

with weights

$$w_p = \frac{a_p}{\sum_{q=0}^{3} a_q},$$ \hspace{1cm} (2.47)
$$a_p = \frac{d_p}{(\epsilon + b_p)^2},$$ \hspace{1cm} (2.48)

with coefficients $d_1 = 1/10$, $d_2 = 6/10$, and $d_3 = 3/10$. The value $\epsilon = 10^{-6}$ was introduced solely to prohibit division by zero, but has proven to affect the order \[26\]. Therefore, $\epsilon$ is set to machine precision in this work. The missing smoothness indicators read

$$b_1 = \frac{13}{12} \left( \bar{U}_{i-2,j,k} - 2 \bar{U}_{i-1,j,k} + \bar{U}_{i,j,k} \right)^2 + \frac{1}{4} \left( \bar{U}_{i-2,j,k} - 4 \bar{U}_{i-1,j,k} + 3 \bar{U}_{i,j,k} \right)^2,$$ \hspace{1cm} (2.49)
$$b_2 = \frac{13}{12} \left( \bar{U}_{i-1,j,k} - 2 \bar{U}_{i,j,k} + \bar{U}_{i+1,j,k} \right)^2 + \frac{1}{4} \left( \bar{U}_{i-1,j,k} - \bar{U}_{i+1,j,k} \right)^2,$$ \hspace{1cm} (2.50)
$$b_3 = \frac{13}{12} \left( \bar{U}_{i,j,k} - 2 \bar{U}_{i+1,j,k} + \bar{U}_{i+2,j,k} \right)^2 + \frac{1}{4} \left( \bar{U}_{i,j,k} - \bar{U}_{i+1,j,k} + \bar{U}_{i+2,j,k} \right)^2.$$ \hspace{1cm} (2.51)
The presented reconstruction is the one from the left. The one from the right is symmetric and obtained accordingly.

This initial scheme has not only been extended to higher \cite{13} and even adaptive convergence orders \cite{12}, but has been extended and modified hundreds of times. E.g., the weights have been adjusted to reduce dissipation \cite{72}, the direction of the reconstruction has been adjusted creating a blend of upwind and central schemes \cite{108}, and in targeted ENO (TENO) sub-stencils containing discontinuities are ignored altogether \cite{70}.

The possibility to combine multiple Riemann solvers with multiple reconstruction stencils creates a huge variety of numerical schemes. These schemes provide different benefits such as reduced numerical dissipation or dispersion, but may be vulnerable to unphysical asymmetries or instabilities \cite{64, 66}. Hence, it is beneficial to choose the best-fitting scheme for a particular problem.

### 2.2.2 Time Integration

The left-hand side of eq. (2.33) is discretized using a time-step size $\Delta t$ to advance from time $t^n$ to $t^{n+1} = t^n + \Delta t$ with an explicit total variation diminishing (TVD) \cite{94} Runge-Kutta (RK) \cite{134, 190} time integration scheme of second order

$$
\bar{U}_{i,j,k}^{n+1} = \frac{1}{2} \left( \bar{U}_{i,j,k}^{n} + U_{i,j,k}^{*} \right) + \frac{1}{2} \Delta t \mathcal{L} \left( \bar{U}_{i,j,k}^{n} \right) \tag{2.52}
$$

or third order

$$
\bar{U}_{i,j,k}^{n+1} = \frac{1}{3} \left( \bar{U}_{i,j,k}^{n} + 2 U_{i,j,k}^{**} \right) + \frac{2}{3} \Delta t \mathcal{L} \left( \bar{U}_{i,j,k}^{n+1} \right) \tag{2.53}
$$

The stability criterion for the explicit time integration reads

$$
\Delta t_{\text{NSE}} = C \min \left( \Delta t_a, \Delta t_\mu, \Delta t_q, \Delta t_s, \Delta t_S \right), \tag{2.57}
$$

with time step limitations according to acoustics \cite{128}, viscous shear \cite{212, 213}, thermal conductivity \cite{106}, volume forces \cite{212, 213} and interface terms \cite{29}, respectively. If $\Delta t$ is constant throughout the simulation, the Courant-Friedrichs-Lewy (CFL) constant has to be reduced from the theoretical stability limit $C < 1$ to account for a large safety factor.
Otherwise, $\Delta t$ is recalculated every time step according to eq. (2.57) where

$$
\Delta t_a = \max_{\Delta i,j,k \in V_1 \cup \Delta i,j,k \in V_2} \left( \frac{\Sigma ||v|| + a}{\Delta x} \right)^{-1},
$$

$$
\Delta t_{\mu} = \max_{\Delta i,j,k \in V_1 \cup \Delta i,j,k \in V_2} \left( \frac{14\mu}{3\rho \Delta x^2} \right)^{-1},
$$

$$
\Delta t_s = \max_{\Delta i,j,k \in V_1 \cup \Delta i,j,k \in V_2} \left( \frac{\Sigma ||v|| + a + \sqrt{\Sigma ||v|| + a} + 4|g|\Delta x}{2\Delta x} \right)^{-1},
$$

$$
\Delta t_q = \max_{\Delta i,j,k \in V_1 \cup \Delta i,j,k \in V_2} \left( \frac{14k}{3\rho c_p \Delta x^2} \right)^{-1}, \text{ and}
$$

$$
\Delta t_S = \max_{\Delta i,j,k} \left( \frac{8\pi \sigma}{\Delta x^3 \sum_i \rho^i \xi} \right)^{-1},
$$

with the cell-local speed of sound $a$.

### 2.2.3 Interface Discretization

Depending on the interface representation, cf. section 2.1.1, different numerical schemes need to be considered for the discretization and computation of interface position and the exchange terms across it. In diffusive interface and VOF schemes interface smearing needs to be prevented, e.g. by introducing anti-diffusion [205] or reverting to more complex geometric VOF formulations [129]. The interface position, its velocity and curvature are obtained from reconstruction algorithms [9]. Furthermore, special care has to be taken to prevent spurious oscillations at the interface [1], in particular when using high-order methods like WENO-type reconstruction [119, 121]. In the LS method, the interface position is trivially obtained from eq. (2.22). Solving eq. (2.21) via standard finite-difference schemes of desired order further allows to a straightforward computation of its curvature.

Similarly, the right-hand sides of eqs. (2.19), (2.23) and (2.24) can be discretized using finite-differences. However, the application of reconstruction stencils may be advantageous [117, 156, 166]. The left-hand side of eq. (2.19) is integrated with the same scheme as the fluid states, cf. section 2.2.2. For the integration of eqs. (2.23) and (2.24) in pseudo-time, however, a one-step Euler integration [56] is sufficient. In order to solve the Riemann problem eq. (2.20) a two-material Riemann solver needs to be applied [107, 144, 221].

Note, cells with small volume fractions $\alpha_{i,j,k}^\xi \ll 1$ require lowering the time-step size drastically or risk arising of instabilities. To prevent either of these a conservative mixing procedure is applied to clear the cell of the respective phase [109, 137].

### 2.3 Compression Algorithms

The numerical methods introduced above are computationally demanding. The required compute power is further increased as the studied physical effects such as compressible turbulence or shock-bubble/drop interactions are innately 3D [175, 179, 199, 218] and very high-resolutions are required to resolve enough physical scales. Hence, simulations with billions of degrees of freedom (DOFs) have to be conducted. Luckily, often the predominant physical processes only occupy a fraction of the computational domain [56, 120, 126, 224,
This allows mimicking the effective separation of scales of dominating physical effects by investing more resolution in these areas. Thereby, reducing the computational load compared to a homogeneously discretized domain.

A commonly used technique in compressible CFD is the truncation-error based AMR [17]. It offers the possibility of a straightforward integration of the aforementioned numerical schemes, a regular mesh, and local scale support, i.e. the refinement criterion is local and if information is exchanged between the grids of different resolution only cells in the proximity are involved. Since its refinement is sensitive to the steepness of flow-field gradients, however, regions of interest without strong gradients, e.g. contact discontinuities or smooth rarefaction waves, may not refine properly [45, 46, 199]. A variety of other compression schemes exists such as adaptive (possibly wavelet-based) DGM [40, 243], multigrid [60], sparse grids [204] or wavelet collocation methods [231]. However, combining the previously introduced numerical schemes within these algorithms is challenging due to a lack of discrete conservation, heterogeneous meshes, or wide scale support.

In this regard, the MR scheme [95] is advantageous as it creates regular meshes, allows straightforward application of the numerical schemes presented in section 2.2, has local scale support, and adapts also on relevant but relatively smooth flow features, typically even with higher compression rates than AMR [45, 46, 199]. MR uses dyadic grids which naturally results in octree data structures in 3D. In the tree nodes may be leaves if they are not further refined with children or parents otherwise. Same-level neighboring nodes are called brothers or cousins depending on whether they share a parent. If a node does not have a cousin on a side, this side makes up a resolution jump. The original flux-compression [95] MR scheme is natively conservative, but requires more memory than a comparable homogeneous grid. Its fully-adaptive form [41, 189] also compresses memory, yet it requires a flux adjustment at resolution jumps [189].

The MR method is based on the wavelet-form of the solution function $U(x, t)$

$$U(x, t) = \sum_{i} \sum_{j} \sum_{k} c_{i,j,k} \theta_{i,j,k}(x, t) + \sum_{m} \sum_{i} \sum_{j} \sum_{k} d_{i,j,k}^{m} \psi_{i,j,k}^{m}(x, t), \quad (2.63)$$

where the subscript $i, j, k$ sample the function at grid point $x_{i,j,k}^{m} = [2^{-m}i, 2^{-m}j, 2^{-m}k]^T$, $(i, j, k) \in \mathbb{Z}$ on $m$-th refinement level $l_{m}$, $m \in \mathbb{N}$. The first sum over the products of scaling coefficients $c_{k}$ and compact scaling functions $\theta_{k}$ defines the general shape of the function. The details $d_{i,j,k}^{m}$ and the wavelet function $\psi_{i,j,k}^{m}$ define the local fluctuations on the respective level. Hence, the details indicate the smoothness of the solution. Compression is thus achieved by setting vanishing details to zero. By using interpolation wavelets the details can be computed without a full wavelet transformation by a simple difference between the solution function on different levels.

$$d_{i,j,k}^{m} = \sum_{i} \sum_{j} \sum_{k} d_{i,j,k}^{m} \psi_{i,j,k}^{m} = \bar{U}_{i,j,k}^{m+1} - \bar{U}_{i,j,k}^{m}, \quad (2.64)$$

The solution can be propagated to higher, i.e. finer, or lower levels using the consistently defined averaging and prediction operators [95, 103], illustrated in fig. 2.2. Mathematically, the averaging operator is defined as

$$A(U^{m+1}_{i,j,k}) \rightarrow \bar{U}^{m}_{i,j,k} : \bar{U}^{m}_{i,j,k} = \frac{1}{N^2} \sum_{p} \sum_{q} \sum_{r} U^{m+1}_{p,q,r}, \quad (2.65)$$
Figure 2.2: Averaging (left) and third-order prediction (right) operators in 3D. The cells used in the computation of the respective operator are colored. The coarse grid $l_m$ is on the bottom with the once refined grid atop. Note, in the finer mesh eight cells are highlighted in both operators.
and the prediction operator as

\[
P(\bar{U}^m) \to \bar{U}^{m+1}: \]

\[
\bar{U}^{m+1}_{i,j,k} = \bar{U}^m_{i,j,k} + (-1)^i Q^x_{i,j,k} + (-1)^j Q^y_{i,j,k} + (-1)^k Q^z_{i,j,k}
\]

\[
+ (-1)^i (-1)^j Q^{xy}_{i,j,k} + (-1)^k (-1)^j Q^{xz}_{i,j,k} + (-1)^i (-1)^k Q^{yz}_{i,j,k},
\]

with the contributions of each separate dimension

\[
Q^x_{i,j,k} = \sum_{p=1}^{s} \omega_p \left( \bar{U}^m_{i+p,j,k} - \bar{U}^m_{i-p,j,k} \right), \tag{2.67}
\]

\[
Q^y_{i,j,k} = \sum_{p=1}^{s} \omega_p \left( \bar{U}^m_{i,j+p,k} - \bar{U}^m_{i,j-p,k} \right), \tag{2.68}
\]

\[
Q^z_{i,j,k} = \sum_{p=1}^{s} \omega_p \left( \bar{U}^m_{i,j,k+p} - \bar{U}^m_{i,j,k-p} \right), \tag{2.69}
\]

of two-dimensional cross interactions

\[
Q^{xy}_{i,j,k} = \sum_{p=1}^{s} \omega_p \sum_{q=1}^{s} \omega_q \left( \bar{U}^m_{i+p,j+q,k} \bar{U}^m_{i-p,j,q,k} - \bar{U}^m_{i+p,j,q,k} \bar{U}^m_{i-p,j,q,k} \right), \tag{2.70}
\]

\[
Q^{xz}_{i,j,k} = \sum_{p=1}^{s} \omega_p \sum_{r=1}^{s} \omega_r \left( \bar{U}^m_{i+p,j+k+r} \bar{U}^m_{i-p,j,k+r} - \bar{U}^m_{i+p,j+k+r} \bar{U}^m_{i-p,j,k+r} \right), \tag{2.71}
\]

\[
Q^{yz}_{i,j,k} = \sum_{q=1}^{s} \omega_q \sum_{r=1}^{s} \omega_r \left( \bar{U}^m_{i,j+q+k,r} \bar{U}^m_{i,j-q+k,r} - \bar{U}^m_{i,j+q+k,r} \bar{U}^m_{i,j-q+k,r} \right), \tag{2.72}
\]

and of 3D cross interactions

\[
Q^{xyz}_{i,j,k} = \sum_{p=1}^{s} \omega_p \sum_{q=1}^{s} \omega_q \sum_{r=1}^{s} \omega_r \left( \bar{U}^m_{i+p,j+q,k+r} \bar{U}^m_{i-p,j+q,k+r} - \bar{U}^m_{i+p,j+q,k+r} \bar{U}^m_{i-p,j+q,k+r} \right) \tag{2.73}
\]

\[
+ \bar{U}^m_{i+p,j,q,k+r} \bar{U}^m_{i-p,j,q,k+r} + \bar{U}^m_{i+p,j+q,k+r} + u_{i+p,j,q+k+r}^m \bar{U}^m_{i-p,j+q,k+r} + u_{i+p,j+q,k+r}^m \bar{U}^m_{i-p,j+q,k+r} + u_{i+p,j+q,k+r}^m \bar{U}^m_{i-p,j+q,k+r}.
\]

The resulting MR scheme is of order \(O(2s + 1)\) and the coefficients \(\omega_p\) read

\[
\omega_1 = \frac{1}{8} \quad \text{for } s = 1, \quad \omega_2 = \frac{3}{128} \quad \text{for } s = 2. \tag{2.74}
\]

The tensor structure of the \(Q\)-terms is illustrated in fig. 2.3.

The spatial compression allows for additional compression in time. Therein, coarser regions are advanced with a proportionally larger time step than in the finer regions. For the dyadic meshes of MR and AMR schemes the time step is simply scaled as

\[
\Delta t_{im} = 2^{l_{max-l_m}} \Delta t_{l_{max}} \tag{2.76}
\]

relative to the time step \(\Delta t_{l_{max}}\) on the finest level computed from eq. (2.57). In a straightforward implementation such local-time stepping (LTS) schemes synchronize the time-step sizes after the lowest level has been advanced once [51].
2.4 Build-up of Supercomputers

Despite application of the before mentioned compression algorithms, the numerical load of 3D compressible flows is still too high for average personal computers (PCs). Rather, HPC systems need to be employed. Writing efficient software for these systems, requires a thorough understanding of its hardware and the computational concepts employed in them. Until the late 1990s it was of little importance for the programmer to know these details. All programs benefited equally from the rapidly and continuously growing clock frequency, i. e. the speed at which the chip conducts one operation. Around the change of the century, however, the high clock frequency together with the shrinking size of the transistors lead to higher power dissipation in form of heat [27]. Hence, current and future (silicon-based) chips employ multiple levels of parallelism to increase the performance without increasing (electrical) power demands [28]. A wide variety of special-purpose architectures such as graphics processing units (GPUs) or tensor processing units (TPUs) has emerged with different emphasis on the respective levels of parallelism. Hence, the actual implementation must explicitly exploit each level of parallelism to yield good performance on the respective architecture [214].

Within an HPC system the obvious level of parallelism is the distributed-memory parallelism. Comparably to a manifold of PC connected via a local area network (LAN), a specialized interconnect links the compute nodes of a super computer together. The nodes typically hold two or four sockets in which the respective central processing unit (CPU) and its main memory is located. Even though the random-access memory (RAM) is shared between the CPUs of a node the access times depend on the physical location of data, i. e. to which socket it is associated. This non-uniform memory access (NUMA) design brings forth possible locality and contention problems. If data needs to be fetched from a remote location the application...
performance may degrade. Furthermore, the bandwidth of the interconnect may be depleted if many remote accesses to the same location happen simultaneously. Figure 2.4 shows the NUMA domains for an exemplary two-node machine. The main memory of the whole HPC machine is sum of the main memory available on all its compute nodes. No central intelligence keeps track of in which node’s RAM a certain piece of information is stored. Hence, the application has to actively manage data transfer between the nodes.

A variety of different libraries for this communication between nodes exists, cf. [25, 123, 203], but the Message Passing Interface (MPI) [67] is the de-facto standard software running on distributed-memory machines. It is based on point-to-point communication. The MPI runtime starts multiple ranks. The ranks are pinned to cores according to user input; over-subscription of cores is possible. All ranks individually start the same program in parallel. Although MPI offers a wide variety of communication routines including collective and one-sided operations, data transfer between ranks is typically triggered via explicit MPI_Send and MPI_Receive calls from within the application. The asynchronous variants of these functions allow for intermediate work to be done while the messages are being send. Note, an MPI implementation may optimize the transfer based on the network topology and, e.g., replace network transfer with simple load instruction if the ranks are on cores of the same NUMA domain [43]. Such optimizations, however, cannot be actively exploited by developers as they are transparent to the application.

The next level of parallelism is the shared memory parallelism. All the cores within a CPU and typically even all cores within the same compute node have direct access to the same ’shared’ RAM. Physically close cores may in addition share further parts of the memory hierarchy as depicted in fig. 2.5. In this example, the level three ‘L3’ cache is shared within a socket and the faster yet smaller L2-cache is only shared within a CPU. The smallest and
fastest L1-Cache is then exclusively used by the respective core. From thereon data is read into the registers on which the bit-wise operations specified by the program are performed. When data is loaded into the caches, always a whole cache line is handled. Such a cache line comprises multiple bytes which are continuously located in a higher memory level. To keep the CPU fed an ideal program ensures that relevant data is grouped physically together in memory and the program flow is predictable. The load and stores are performed coherently, i.e. all loaded cache lines get invalidated if any core writes to this line. Nevertheless, cores can independently execute their computation and, hence, access the memory differently. This is indicated in the figure by the relative position of the colored loaded cache lines within the respective cache. This independence is typically exploited by running threads. These threads execute (part of) a program in parallel and may be spawned or joined arbitrarily. The shared-memory eliminates the need of explicit data transfer between threads. Nevertheless, access to the shared data must be coordinated to prevent race conditions and deadlocks. Numerous concepts and libraries for shared-memory parallelization exist, e.g. operating system (OS) level threads [110] are relatively slow to spin-up and typically cumbersome to program, but they provide full OS control to the thread creating application. In contrast, designated HPC threading libraries [123, 165, 183] provide less powerful threads which in turn are much faster created and retired, allowing their programming with simple instructions. It is also possible to use MPI within a shared-memory context as ranks can be pinned to individual cores. Furthermore, the MPI standard version 3 defines a separate shared-memory model [67] and certain implementations may provide additional shared-memory optimization [43].

Further parallelism is provided within the cores of a CPU. Modern cores use different types of registers for the different compute purposes. From only 16 bit-wide registers
Figure 2.6: Concept of the internal structure of a core. The Level 1 Cache is split into an instruction cache (L1I) and a data cache (L1D). Instruction flow is indicated by thick orange arrows, while data flow is shown by thin green ones. The address and general purposes registers abbreviated Adr and GP, respectively. Instructions executed in the units are abbreviated according to assembly language.

for program instructions and arithmetic operations up to 512 bit-wide vector registers for floating-point computations. The different registers are processed in different units of the processor. For example the integer registers are handled by an arithmetic logic unit (ALU) or address generation unit (AGU), whereas floating-point registers are handled in a floating-point unit (FPU). A schematic of the internals of a core is shown in fig. 2.6. Note, registers, decoders and processing units are duplicated. This superscalar architecture allows for multiple instructions being executed simultaneously. This first allows simultaneous multithreading (SMT), i.e. running two different threads on the same core. While this might be beneficial in the context of server farms where diverse applications run concurrently, the scheduling overhead often outweighs the performance gained by the increased throughput for numerical computations [83, 89], and for CFD application in particular [14]. An alternative is to write a single application that fully exploit the superscalar capabilities of the core. This, however, is nearly impossible for sensible applications that where translated to machine code via a compiler [89]. Nevertheless, the abundance of units also enables out-of-order execution. Therein, upcoming instructions are saved in a separate buffer. If the instruction is executable with the current register states and the required unit is free in the current cycle, then the instruction is executed on the spot ahead of its ordinary execution time. Although speed-ups are mild, applications benefit automatically. Out-of-order scheduling is done by the hardware and modern compilers optimize for it.

The fine granularity of units, however, allows pipelining parallelism. In the pipeline, program operations are split into their instructions which are then executed by the designated unit. If an instruction is repeated multiple times, e.g. in a loop, the throughput of instructions per clock cycle is increased. Exemplary, fig. 2.7 shows an idealized pipeline for one operation consisting of just three instructions “fetch data”, “process data”, “write data”. It is assumed
that each instruction is handled by a corresponding unit and the execution of the instruction in these units takes exactly one clock cycle. After an upwind phase of two cycles, one whole operation is finished every cycle. Without pipelining, the completion of one operation, however, requires three cycles. In actual cores, a reduced instruction set is employed. This reduced set requires to split the operations into tiny instructions. Hence, operations are broken down into many such operations and modern pipelines contain more than 30 stages. This in turn, makes the up- and downwind phase relatively costly. The application’s performance is thus increased by long branch-free loops. To keep the pipeline fed, the data processed within such loops needs to be limited. It further helps to have the data located continuously in memory, to benefit from the cache-line loading mechanism described above.

The final level of parallelism discussed here is SIMD parallelism, also called vectorization. By using wide registers with up to 512 bit the same floating-point operation can be executed on all elements in the register. The speedup depends on the register width and on the precision, e.g. for single-precision computations with 512 bit-wide registers a speedup of 16 over the sequential version is theoretically achievable. A comparison of vectorized and sequential computation of the summation of two vectors for four clock cycles is given fig. 2.8. Similarly to pipelining, vectorization only works on streams of data. Applications need to ensure branch-free loops with predictable memory access patterns in order to benefit from vectorization capabilities of the core. Furthermore, exceptions, data misalignment and pointer aliasing must be prohibited.

2.5 Code Quality

Naturally, it is in the interest of every programmer to write effective and efficient code. I.e. the compiled executable computes the correct answer to the stated problem at the cheapest cost. Note, cost is loosely defined here. Depending on the context it can be a measure of runtime, development time, memory usage, CPU power requirements, monetary cost, etc. Besides these obvious (software) design goals, the quality of the code base is an important metric. The way the source code is written and structured strongly affects its readability, understand-ability, (re-)usability, maintainability, flexibility, portability, testability, and adaptability. This code cleanliness becomes even more important for larger code bases since it...
Figure 2.8: Comparison of sequential and four-way vectorized addition of two arrays $A$, and $B$. At the top sequential and in the bottom vectorized addition is performed, respectively. The clock cycles are indicated by the vertical lines. An addition is assumed to take one such cycle.

affects the productivity of more developers and users. Even though this software quality is often not incentivized directly, the increased development speed outweighs the added software-cleaning costs. Furthermore, clean code often automatically benefits from improvements in compilers or added libraries.

Although the quality of written code cannot be assessed by a (single) discrete measure, certain rules and guidelines became consensus as best-practice when developing software [84, 148, 149]. These include among others: selection of descriptive variable and function names [58, 69, 149], presence of an automated testing pyramid [136, 155], and adherence to language specific guidelines [81, 115]. For scientific codes in particular, it was found that solely providing automated testing increased the correctness, efficiency, reliability, integrity, accuracy, and robustness of the used software [22].

2.6 Challenges for Numerical Software & Parallel Efficiency

Naturally, the aim for scientific software is to produce the physically correct result as cheaply as possible. As discussed in section 2.2, however, the employed numerical schemes are computationally demanding. Hence, simulations are conducted on HPC systems. For apparent reasons, access to such systems is only granted if an application runs with sufficient (parallel) performance. A simple, yet powerful tool to assess an application’s performance is measuring the amount of FLOP it performs and put it in relation to the processed amount of memory. The thus obtained value can be plotted into a roofline model [240] to assess its performance. As visualized in fig. 2.9, the roofline model indicates innate performance bounds of the used hardware, such as memory bandwidth limits of the caches in the memory hierarchy and the maximum number of computations an ideally fed unit can compute. The limiting factor is stated atop the respective bound in the figure. An ideal application will appear in the top right corner running at peak performance of the chip even if new data for computation has to be
2.6 Challenges for Numerical Software & Parallel Efficiency

Fig. 2.9: Outline of the roofline model in log-log scale. The limiting memory- or computation-type is given by the labeled lines. The bound-regions with respect to the peak performance are indicated by double-sided arrows.

retrieved from slower parts of the memory hierarchy. Such programs are said to be compute bound otherwise they are memory bound. Raising the memory bandwidth is a greater challenge than to increase the compute capabilities [89]. Therefore, compute-bound applications typically benefit most from newer hardware.

Naturally, the implemented algorithm and the employed numerical schemes have the largest impact on the application’s performance. Nevertheless, modifications within the frame of the original algorithm, such as blocking strategies in matrix-matrix multiplications may provide a significant speedup. On a lower level, pure software improvements, such as loop reordering may help to move the application towards the top-right corner in the roofline model. However, extremely trimming the source code for highest performance may lower the code quality and contradict the guidelines stated in section 2.5. As an example, the MPI application programming interface (API) is defined on a low-level of abstraction, giving the application fine-grained control over the communication but hindering the usage of high-level programming paradigms.

The MR compression, the interface discretization, and numerical methods employed in this thesis allow for clean implementations straightforwardly. Nevertheless, incorporating them in a 3D framework with high parallel performance poses some challenges. So does e.g. the mapping of the 3D data to a flat memory complicates streamed data access in all spatial directions, the huge number of computations per value in the reconstruction stencils impede vectorization of the kernels, and nontrivial determination of neighbor values hinders parallelization of the MR algorithm.
2.7 Computational Frameworks for Complex Flows

A handful of CFD frameworks are generally suitable for the study of the complex flows phenomena stated in chapter 1. Theses include the commercial tool suites Ansys [7], Solid Works Flow [200] and Comsol [42], the semicommercial open-source openFOAM [85], and the research codebases Cubism(-Nova) [99, 239], MFC [33], UCNS3D [8] and ECOGEN solver [198]. While the (semi-)commercial frameworks have a wide area of application and provide additional plugin and modules, the research codes typically focus on certain problem types. So does ECOGEN focus on bubble dynamics, MFC on ultrasound and shock-induced bubble collapse, UCNS3D on turbulence effects for aviation, and Cubism on (clouds of) cavitation bubbles. A comparison of the implemented algorithms, the numerical models and the performance of each framework and the one developed within this thesis is given in chapter 5.
Chapter 3

Accomplishments

The main accomplishments of this work cover four categories: Improved distributed-memory parallelization of compression algorithms for compressible finite-volume method direct numerical simulation computational fluid dynamics simulations, improving the vectorization of the aforementioned schemes, enabling large-scale LS simulation by integration of the LS into the improved compression algorithms, and the modularity and cleanliness of the created software framework. The achievements comprising this thesis were published in three first author publications [103, 104, 106]. Note, alongside the manuscript the generated research data was published under permissive license allowing interested researchers to re-use the obtained data [105].

3.1 Improved Multiresolution Compression Algorithm

Here, only MR compression is considered for the reasons stated in section 2.3. A straightforward 3D MR implementation uses an octree to represent the computational mesh [41, 189]. Each node in the tree holds one cell or scale. Numerous such implementations exist [31, 48, 90, 207, 209]. When evaluating their capabilities, however, reveals their lack of support for high-order methods and limited parallel performance. This performance drawback typically stems from the difficulty of obtaining the value of the cell's neighbor states during (cell-face) flux computation. Such values have to be retrieved by a slow tree traversal. Worse, yet if the neighbor is missing on the given refinement level, the value has to be predicted from coarser levels first. Inevitably, this leads to non-regular memory access patterns, which should be avoided [31], cf. section 2.4. For broad high-order stencils this challenge is magnified as multiple neighboring cells are required.

Instead of following this approach, here, a predefined number of cells is packed into a block. Each octree node holds one such block instead of a single cell. This block-based approach is motivated by the remarkable performances reported for similar concepts in AMR and wavelet-collocation methods [61, 188]. This configuration also allows for streamlined and SIMD-optimized flux computations discussed in section 3.2. Equipping each block with its own set of halo cells also reduces communication overhead in distributed-memory environments.

The performance of the block-based MR algorithm is further improved by assigning each node an identifier (ID), which is based on a modified Morton order. This allows neighbor lookup by fast bit-shift operations, creation of tiled geometries and eases assignment of nodes to cores via space-filling curves (SFCs). In the implementation, this load balancing is further modified to account for the level of refinement when assigning nodes along the SFC. Although this leads to a more scattered distribution of the nodes, it is crucial to prevent idle cores if LTS is enabled. Details on the block-based MR algorithm, the modified Morton order and the adjusted SFC are published in [103], and summarized below.
Furthermore, an adaptive LTS (ALTS) scheme for temporal compression was developed that is integrated into the block-based MR algorithm. In contrast to previous LTS schemes, the time-step size can be adjusted after every time step on the finest level. This is achieved by projecting and averaging the right-hand side of eq. (2.9) instead of the integrated conservative quantities. The higher frequency of time-step size adjustments allows using larger time steps while maintaining numerical stability. The achieved results are published in [128]. My contribution to the latter lies in guidance on integration into the MR algorithm and the implementation in C++ as well as conducting the profiling of the parallel execution of the code.

3.1.1 Summary of: “A parallel modular computing environment for three-dimensional multiresolution simulations of compressible flows”

The manuscript discusses the details of the block-based MR algorithm and its modular implementation in detail. Concepts to enable the desired level of parallelism include a modified Morton order indexing of nodes in the MR-tree and traversing of SFCs in a level-wise manner. A wide range of test cases are simulated to assess the parallel performance across different computational modules such as the type of SFC, the MR-thresholding error norm, the Riemann solver, the reconstruction and derivative stencils, the number of cells per block, and more. The achieved compression is thoroughly analyzed and strong and weak scalings on up to twenty-four thousand cores are performed.

The paper starts with a thorough review of compression algorithms applied to compressible multiphase flows. The differences between AMR and wavelet-based methods are given and the choice of MR is motivated. Publications based on single-scale MR algorithms are examined and their shortcomings with respect to large-scale parallel computations are discussed.

Next, the governing equations and their numerical discretization are presented in detail. Throughout the work, different combinations of the Roe [186], the Rusanov [191] and the HLLC [222] and first-, third-, fifth-, seventh- [118], central upwind sixth-order [108] or adaptive fifth-third-order [12] WENO as well as targeted ENO [70] reconstruction stencils are used.

In the following, details on the genuine MR algorithm and LTS are stated. Then the block-based variant of the MR algorithm with activated ALTS is discussed. The special load-balancing requirements implied by the ALTS scheme are addressed by level-wise modification of SFCs. A modified Morton order is proposed to assign unique IDs to nodes in the MR-octree, allowing neighbor lookup via bit-shift operations from a hash map. This modification allows for straightforward implementation of periodic boundary conditions and non-cubic domains while requiring less memory than previously published indexing schemes [48]. The underlying algorithm is stated in detail. Remarks on the initialization routine and the time-step size restriction are made. Implementation specific details and enforced programming patterns which enable the modularity of the solver are given. The grain size of the parallelization and the MPI-only approach are discussed.

Finally, the correctness of the acoustic and source term solvers are shown by comparison with analytic solutions for the Sod shock tube [206] and the growth-rates of the Rayleigh-Taylor instability (RTI) [35], respectively. A convergence analysis is conducted verifying the nominal convergence orders of the mentioned schemes. It is shown, that the compression of the presented block-based MR variant is comparable to those of ‘traditional’ single-cell versions, in particular for 3D cases. For numerous test cases, it is shown that the parallel performance is maintained if various of the modular building-blocks are exchanged. This is done by conducting weak and strong scalings for each case on the Leibniz Supercomputing
3.2 Addressing Auto-Vectorization

As lead author of this publication, I wrote the script and addressed the reviewer comments. I have designed, implemented and optimized the modular code framework and in particular the parallelization routines. I have conducted the validation, numerical and physical test cases. Furthermore, I set up and evaluated the scaling runs and have created the shown visualizations. My co-authors have coordinated, supervised, and secured funding for the project. They have also reviewed the manuscript and responses to the reviewers.

3.2 Addressing Auto-Vectorization

The previously introduced block-based MR algorithm allows exploiting the SIMD capabilities of the CPU, cf. section 2.4. It is known from profiling runs that the flux computation accounts for roughly 60 - 75% of the overall runtime. Hence, vectorizing the respective compute kernels yields the highest benefit. However, the numerical scheme to determine the flux does not natively support vectorization. Worse yet, depending on the choice of solvers the processing steps differ. E.g., in the HLLC [222] solver the reconstruction stencil is applied directly on the conservative values whereas in the Roe Riemann solver [186] the stencil operates on characteristic values. Since experimenting with and exchanging these types of compute kernels is the norm in a scientific code base manual vectorization of each routine is destined to fail. Rather than trying manual insertion of intrinsic into the code base, restructured compute routines were employed. While still giving enough flexibility to the developer, the new structure is (more) transparent to the compiler. This results in auto-vectorization of the compute kernels. The details on the revised code structure are presented in [104] and summarized below. Note, the reported conflict between numerical symmetry and vectorization performance have since been resolved. Tests on LRZ’s SuperMUC, which also provides 512 bit wide registers, recover fully symmetric flow solutions.

3.2.1 Summary of: “Node-Level Optimization of a 3D Block-Based Multiresolution Compressible Flow Solver with Emphasis on Performance Portability”

This publication starts with listing challenges to SIMD vectorization in flux-based Riemann solvers with high-order reconstruction stencils. These challenges are then addressed by general and clean standard-compliant code improvements, rather than hard-coded intrinsic operations. These improvements result in compiler-generated vectorization and thus an increased performance across compute kernels.

The paper presents the governing NSE and the numerical discretization by the FVM, Riemann solvers and WENO-type reconstruction stencils. Throughout the work, different combinations of the Roe [186] and the HLLC [222] Riemann solver and the fifth- [118] and sixth-order central upwind [108] WENO stencils are used.

Then, the block-based MR algorithm is presented and its potential for efficient use of wide vector registers including its MPI-parallization is discussed. The modular C++ implementation is described and the need for multipurpose optimizations is stated. First performance-baselines are defined on two different Intel compute hardware with 256 bit and 512 bit wide vector registers, respectively. Baselines are defined with respect to the combinations of compute kernels with an additional variation in the number of cells per block.

Hindrances to the efficient use of the vector registers are listed next. They include unfavorable loop-nest structures and a lack of exception-free guarantees. The issues are subse-

---

1https://doku.lrz.de/display/PUBLIC/SuperMUC-NG
quentiy addressed. First, the loops are reorganized, as it was found that pre-computation and storage of intermediate values like the eigenvalues and -vectors is beneficial to the original computation on the fly. Also, array-index offsets stemming from the shift of total cells in a block to internal cells and from the mapping of cell-face fluxes to cell values now heavily use constexpr-features such that the compiler recognizes the constancy of array access patterns. Finally, potential exception throwing functions have been reformulated and marked noexcept. No non-standard intrinsic or even #pragma s have been employed.

The improved code version is then compared against the baseline on three 3D test cases: A Sod shock tube [206], an RTI [202] and the diffusion of a Gaussian pulse. Effective resolutions between the cases increase from one hundred thousand to one billion cells, respectively. It was found, that all cases benefit from the code changes. The hardware with the wider vector registers shows higher performance gains as do the configurations with more cells per block. As desired, all employed combinations of Riemann solver and reconstruction stencil benefit equally. Additional tuning of compiler flags shows to be only beneficial on the hardware with the wider vector registers. Yet, this compiler-flag optimization resulted in numerical symmetry breaking and reduces the program’s portability.

My contributions to the work were the initial initiative to address vectorization hindrances by portable standard-compliant source code modifications. In close collaboration with the second author, I performed the profiles and benchmarks. The new code blocks are provided by me as well as the setup of the presented test cases. I was responsible for the creation of the visualizations and tables. I have predominantly written the final manuscripts and incorporated feedback from the reviewers. My co-authors have coordinated, supervised, and secured funding for the project. They have also reviewed the manuscript and responses to the reviewers.

3.3 Large Scale Sharp-interface Simulations

The simulation of multiphase flows requires additional mathematical and numerical modeling to represent the fluid interface(s), cf. section 2.1.1. The LS with its sharp-interface representation is a natural choice for this discretization, in particular for interface-driven flows and the investigation of interface instabilities. However, to the best of the author’s knowledge, no simulation software exists that provides the LS algorithm within a parallelized framework that is capable to provide the needed resolution for 3D simulations.

This limitation is overcome by inserting the LS algorithm into the MR compressible CFD framework ALPACA [3]. Thereby, two key challenges for efficiently running the large scale LS simulations are addressed. First, in order to always fully refine the interface region, the narrowband [2, 39, 171] approach is picked-up on by defining interface tags representing the band structure alongside with prediction and averaging operators. This way the MR algorithm can react, i.e. refine or coarsen to approaching or vanishing interfaces, respectively. Without this tagging approach, the MR algorithm would coarsen on smooth \( \phi(x,t) \) and \( U(x,t) \) as e.g. present in a pure advection case. Second, since the GFM requires information from both phases at both sides of the interface, the storage must be managed. Each phase is stored in a separate block within the MR node. The block structure stays unchanged to section 3.1. However, nodes close to the interface allocate a second block allowing the (unchanged) solver to process both blocks independently. The implemented interface interaction allows for a straightforward representation of complex solid geometries or free surfaces by the LS. For these purposes, simply the two-material interface-interaction Riemann solver needs to be modified, cf. section 2.2. The capabilities of the implementation have been demonstrated on representable large scale two-phase computations. Details of these
improvements are published in [106] and summarized below.

3.3.1 Summary of: “ALPACA - a level-set based sharp-interface multiresolution solver for conservation laws”

In this paper, the adaptation of the LS method into the block-based MR algorithm is presented. In particular, a tagging system is proposed to replicate the narrowband approach [2, 39, 171] across all levels of refinement. The modular concept of the single-phase implementation is extended by the LS building-blocks, cf. section 2.1.1. In addition, competing schemes for the discretization of interface interaction, the GFM extrapolation and the reinitialization are provided.

At first, an overview of applications for compressible flows, the variety of numerical schemes, and the competing interface and compression algorithms is given. The challenges to efficiently run these algorithms and schemes on HPC clusters are stated. Furthermore, comments on general code features concerning the build process, test coverage, and input/output (I/O) are made. The mathematical formulation to extend the single-phase NSE to multiple phases using the LS method is stated alongside its numerical discretization. Therein, the LS re-initialization procedure, the GFM extrapolation, and the two-material Riemann-problem based interface interaction terms are highlighted. Further remarks are made concerning time-step size restrictions and the handling of marginally filled cut cells.

Then, the modular block-based MR algorithm is reviewed before the newly proposed interface tagging system is introduced. Its purpose is three-fold. First, it gives a memory-saving representation of the interface structure across all refinement levels, allowing the MR to respectively refine or coarsen on approaching or vanishing interfaces. Second, it allows limiting allocation of multiple fluid blocks to the narrowband region on all levels. Finally, it simplifies the restriction of LS related computations to the narrowband region. The algorithmic extension compared to the single-phase block-based MR algorithm are given alongside parallelization and implementation details, which enable the modular setup of the solver.

Next, a wide variety of test cases is simulated. First, the convergence order of the employed schemes is verified together with the correctness of the surface tension and viscous forces’ implementation. Subsequently, complex shock-bubble configurations with ‘heavy’ and ‘light’ bubbles are analyzed. Therein, complex material properties and high resolutions of up to four billion effective cells are used. In the following, the parallel performance of the multiphase implementation is evaluated on up to twenty-four thousand cores and shows high efficiencies across the different chosen compute kernels. The achieved compression over all conducted 3D simulations is determined. Four compression metrics have been proposed, revealing savings in terms of compute time and memory consumption. They are defined such that the MR compression and the enhanced narrowband-related compression are separately evaluable. High compression values are observed if the MR is allowed to provide high numbers of refinement levels.

For this publication, I have implemented the parallel framework and the interface tagging mechanism. I have conducted the convergence study, the scaling runs, and have defined and evaluated the compression rates. I have predominantly written sections 4 to 7, 8.1, 9 and 10. Furthermore, I have equally contributed to sections 1 and 11 as well as to the abstract and program summary. I have extensively reviewed the remaining sections. The visualizations in the respective sections were created by me. My first co-author has conducted the works presented in the remaining sections, which he has predominantly written and has equally contributed to sections 1 and 11 as well as to the abstract and program summary. He has extensively reviewed the sections written by me. Reviewer feedback was handled by both of us in a fair and equal manner. The remaining co-authors have coordinated, supervised, and
secured funding for the project. They have also reviewed the manuscript and responses to the reviewers.

### 3.4 Modular Framework

The presented software ALPACA [3] allows comparisons of competing numerical schemes for the simulation of compressible multiphase flows in high resolution and is available under open-source license. The clean code structure, the automated software testing, and the focus on high-level optimizations ease integration of new schemes and reduce development times. Therefore, data storage and data modification are strictly separated from data manipulation routines. Performance relevant settings such as e. g. the number of cells per block, the Riemann solver, the reconstruction stencil, etc. are determined at compile time to leverage maximum compiler support. Modern C++ 20 features are used. So do for example `if constexpr` clauses accompany the heavily used curiously recurring template pattern (CRTP) for adaptation of different numerical methods to the algorithmic API. Further settings e. g. material parameters, domain sizes, and boundary condition are choosable at runtime. Therein, the concept stays open for extension. We demonstrate the effectiveness of this approach in the before mentioned publications [103, 104, 106]. Therein, the scaling tests have always been conducted with different combinations of numerical schemes and settings and meaningful physical examples have been studied. Even exchanging schemes with different nominal convergence order yields similar parallel performance. The code setup even allows exchanging of the underlying system of equations. E. g. the five-equation NSE may be replaced by their four-equation isentropic counterpart. In this way, also the diffusive interface model is integrated by adding eq. (2.15) [3]. This modularity concept has already enabled a vast variety of research conducted by other authors [24, 63–66, 124–128, 241, 242].
Chapter 4

List of Peer-Reviewed Publications

4.1 Publications comprising this Thesis


4.2 Further Publications


- Benedikt Biller, Nils Hoppe, Stefan Adami, Nikolaus A. Adams: Jetting mechanisms in bubble-pair interactions [20].
Chapter 5

Discussion and Concluding Remarks

Numerical investigation of compressible multiphase flows poses multiple challenges. Firstly, the large range of physically relevant scales requires high resolutions in particular for realistic 3D problem setups. Secondly, the representation of the interface and the interaction between the fluids adds additional complexity. Furthermore, sophisticated high-order methods are required to avoid unphysical effects while recovering most flow details. A wide variety of these numerical methods exist whose strengths and drawbacks are problem dependent.

The work presented in this thesis addresses these problems. It presents a block-based MR algorithm with ALTS for efficient computations on HPC systems. The sharp-interface LS method has been adjusted to leverage this block-based MR algorithm. The algorithm has been implemented in the open-source framework ALPACA written in C++ which offers a modular structure to exchange the (high-order) numerical methods without loss of parallel performance. Even a seamless exchange of the underlying set of equations is possible, e.g. changing to the four-equation isentropic NSE or including additional advection equations for integration of diffusive interface models [3].

The presented MR algorithm follows the design of the single-scale MR algorithm [18, 19, 41, 95, 189]. In addition, however, it leverages the parallel capabilities of block-based approaches within the AMR [61] and wavelet-collocation [188] context. The scheme allows for straightforward integration of ALTS allowing for larger time-step sizes than in previously presented LTS [51] approaches.

For efficient domain decomposition with respect to the ALTS and MR requirements, a level-wise modification of well-established SFCs [10] is proposed. Therein, the position on a Hilbert [100] or Lebesque curve [139] can be computed directly from the introduced node IDs, which are generated from a modified Morton Order [157]. The ID mechanism allows neighbor lookup based on bit-shift operations. Given 64 bit per ID the same amount of refinement levels as in [48] is achievable. In contrast to many published MR schemes [48, 51, 189, 209], however, the presented ID scheme allows representation of non-cubic domains.

The LS sharp-interface method [47, 167, 213] is integrated into the block-based MR algorithm. Therefore, the efficient narrowband approach [2, 39, 171] is implemented based on an interface-tagging approach. On the finest resolution level, the original narrowband technique is used. In addition, the bands are represented via interface tags. By defining an average operator for interface tags, the evolution of the interface is propagated through the octree structure. This allows the tree to adjust for approaching or vanishing interfaces by refining or coarsening the appropriate leaves, respectively. In the LS approach the GFM method [59] is used to obtain conservative interface interactions [107, 109, 144, 221]. This method relies on information of all fluids in every cell close to the interface. Hence, memory for each fluid needs to be allocated. Here, the interface tagging mechanism is used again, allowing to invest this additional memory only close to the interface. In particular for 3D simulations, the overall memory-usage is drastically reduced over the implementation, as presented by Han et al. [90].

Highly resolved physical simulations on shock-bubble, shock-droplet, and bubble collapse
scenarios have been carried-out with the presented framework within this thesis [106]. The code has additionally been applied for further studies by other authors [64, 124, 126, 241]. The implemented parallelization strategy of the LS and MR algorithm allowed higher resolved simulations, larger time steps employing higher-order methods than presented in previous publications [52, 152]. The presented resolutions are comparable with large petascale simulations conducted on homogenous meshes [180, 238].

The simulation framework ALPACA developed within this thesis is offered under open-source license for re-use by the scientific community and to foster exchange within the field. It enables interested researchers to jump-start their own simulation work at a very high level in terms of computational efficiency on massively parallel architectures and accessible model complexity. The provided range of high-order numerical schemes for compressible multiphase flows is unmatched by commercial codes based on FVM [7, 200] and finite-element method (FEM) [42]. Similarly, the semicommercial open-source OpenFOAM solver provides only low-order discretization diffusive interface schemes [85]. However, its wide use in academia, results in frequent publication of additional libraries. E.g., a third-order WENO [74] and incompressible LS approaches [50, 145] have been implemented, but not yet reached the main releases of the framework.

Turning to research codes, a wide variety of parallelized compressible multiphase flow solvers exists. Solvers based on FEM or FVM, with different interface representations, with AMR or MR mesh compression, with low- or high-order methods can be found. Yet, no relevant parallel framework was found employing LTS. Reinarz et al. [182] provide a high-order discontinuous Galerkin framework with high parallel performance on $O(10^4)$ cores [192]. It uses VOF for the interface representation and AMR for mesh compression and is available under open-source license. Similarly, the open-source Cubism(-Nova) project [88, 99, 239] also uses VOF for the interface representation and is based on a high-order FVM. Although, the code base provides AMR mesh compression the published computations [180, 211, 234, 238] ran on homogenous (stretched) meshes. It has won the Gordon Bell Price 2013 with its 11 PFLOPs simulation on over one million compute cores [187]. Likewise, the MFC codebase [33] provides a high-order FVM implementation with a diffusive interface model on homogenous (stretched) meshes under open-source license. High parallel efficiency is recorded for scalings in $O(10^3)$ cores. Another open-source framework running high-order FVM with diffusive interface model on homogenous meshes is UCNS3D [8, 177, 226]. The ECOGEN solver [198] is also built on a FVM but provides at most second-order methods. It does, however, include AMR mesh compression. Scalings are presented in $O(10^2)$ and $O(10^3)$ cores for calculations on homogenous and AMR meshes, respectively. Numerous computations using the sharp interface level-set method for computation of compressible multiphase flows exist: With AMR [162] or MR [90] mesh compression as well as on homogenous meshes [76, 140]. However, none of the underlying code bases is freely available and details on their parallel performance are incomplete or missing completely. A handful of codes using MR mesh compression report their parallel performance solving the compressible NSE. E.g., Soni et al. ’s [207] implementation includes high-order methods and shows scalings in $O(10^2)$ cores. Sroka et al. [209] and Descombes et al. [48] release their respective low-order method codes under open-source license and present scalings in $O(10^2)$ cores.

The presented open-source C++20 codebase [3, 103, 104, 106] is hence unique as it provides a modular FVM framework for compressible multiphase simulations with high-order methods, the level-set method, MR mesh and ALTS temporal compression. Scalings in $O(10^4)$ cores conducted with temporal and spatial adaptivity activated show comparable parallel performance across methods of varying nominal convergence order. The modular structure further allows substitution of the LS interface representation with a diffusive or VOF-type representations.
In conclusion, the continuing growth of compute power benefits the usage of DNS simulations for the study of compressible multiphase flows. Nevertheless, fully resolved 3D simulations will require compression algorithms and efficiently used HPC infrastructure for years to come. The parallelization strategy presented in this thesis and its implementation in ALPACA combines these demands with parallelization-optimized MR, ALTS and level-set algorithms. The capabilities of the presented strategy and implementation have been shown in a range of simulations of physically challenging problems with high (effective) resolution. Although successful, the current MPI-only parallelization may be improved upon by introducing a hybrid parallelization with an additional threading concept. This could reduce the domain-partitioning overhead for multi-node and pure shared-memory computations. The block-based structure of the presented algorithms allows adaptation of the approach by Wermelinger et al. [239], which may also help to exploit the heterogeneous hardware of newer supercomputers.
Bibliography


[13] Balsara, D. S. and Shu, C.-W. “Monotonicity Preserving Weighted Essentially Non-
oscillatory Schemes with Increasingly High Order of Accuracy”. In: Journal of Com-
2000.6443.

of State-of-the-Art HPC Clusters with a Production CFD Code”. In: Proceedings of the
Platform for Advanced Scientific Computing Conference. PASC ’20. Geneva, Switzerland:
1145/3394277.3401847.

for the HLLC Riemann Solver”. In: SIAM Journal on Scientific Computing 18.6 (1997),

general systems of hyperbolic conservation laws”. In: Journal of Computational Physics

In: Journal of Computational Physics 82.1 (1989), pp. 64–84. ISSN: 0021-9991. DOI:
10.1016/0021-9991(89)90035-1.

[18] Bihari, B. L. “Multiresolution schemes for conservation laws with viscosity”. In: Journal
of Computational Physics 123.1 (Jan. 1996), pp. 207–225. ISSN: 00219991. DOI:
10.1006/jcph.1996.0017.

[19] Bihari, B. L. and Harten, A. “Multiresolution Schemes for the Numerical Solution
of 2-D Conservation Laws I”. In: SIAM Journal on Scientific Computing 18.2 (1997),

interactions”. In: Physics of Fluids 34.7 (2022), p. 076111. DOI: 10.1063/5.0097039.


[22] Bissi, W., Serra Seca Neto, A. G., and Emer, M. C. F. P. “The effects of test driven develop-
ment on internal quality, external quality and productivity: A systematic review”. In:
Information and Software Technology 74 (2016), pp. 45–54. ISSN: 0950-5849. DOI:
10.1016/j.infsof.2016.02.004.

[23] Bode, M., Lehn, F. vom, Maeda, K., Colonius, T., and Petsch, H. “Numerical investiga-
tion of the effect of dissolved non-condensable gases on hydraulic flip in cavitating
nozzles”. In: 10th International Cavitation Symposium, Baltimore, USA. 2018.

based sharp-interface method for moving contact lines”. In: Journal of Computational

[25] Bonachea, D. and Hargrove, P. H. “GASNet-EX: A High-Performance, Portable Com-
munication Library for Exascale”. In: Languages and Compilers for Parallel Computing.
Ed. by Hall, M. and Sundar, H. Cham: Springer International Publishing, 2019,

[26] Borges, R., Carmona, M., Costa, B., and Don, W. S. “An improved weighted essen-
tially non-oscillatory scheme for hyperbolic conservation laws”. In: Journal of Com-
putational Physics 227.6 (2008), pp. 3191–3211. ISSN: 0021-9991. DOI: 10.1016/j.
jcp.2007.11.038.

pp. 23–29. DOI: 10.1109/40.782564.


