



HAL
open science

Task-based adaptive multiresolution for time-space multi-scale reaction-diffusion systems on multi-core architectures

Stéphane Descombes, Max Duarte, Thierry Dumont, Thomas Guillet,
Violaine Louvet, Marc Massot

► To cite this version:

Stéphane Descombes, Max Duarte, Thierry Dumont, Thomas Guillet, Violaine Louvet, et al.. Task-based adaptive multiresolution for time-space multi-scale reaction-diffusion systems on multi-core architectures. 2015. hal-01148617v1

HAL Id: hal-01148617

<https://hal.science/hal-01148617v1>

Preprint submitted on 5 May 2015 (v1), last revised 15 Oct 2016 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Task-based adaptive multiresolution for time-space multi-scale reaction-diffusion systems on multi-core architectures

Stéphane Descombes^{*†} Max Duarte[‡] Thierry Dumont[§] Thomas Guillet^{††‡‡}
Violaine Louvet[§] Marc Massot^{¶||**}

May 5, 2015

Abstract

Reaction-diffusion systems involving a large number of unknowns and a wide spectrum of scales in space and time model various phenomena across disciplines such as combustion dynamics, atmospheric science, or biomedical engineering. The numerical solution of these multi-scale systems of partial differential equations entails specific challenges due to the potentially large stiffness stemming from the broad range of temporal scales in the nonlinear source term or from the presence of steep spatial gradients at the localized reaction fronts. A new generation of techniques featuring adaptation in space and time as well as error control has been introduced recently. Based on operator splitting, finite volume adaptive multiresolution and high order time integrators with specific stability properties for each operator, these methods yield a high computational efficiency for stiff reaction-diffusion problems. While demonstrating the potential of such techniques the data structures of the original implementation, based on trees of pointers, provided limited opportunities for computational efficiency optimizations and posed challenges in terms of parallel programming and load balancing. The present contribution proposes a new implementation of the whole set of numerical methods relying on fully different data structures together with the use of a specific library for shared-memory, task-based parallelism with work stealing. The performance of our implementation is assessed in a series of test-cases of increasing difficulty in two and three dimensions on multi-core and many-core architectures, demonstrating high scalability.

Keywords: Task-based parallelism, multi-core architectures, multiresolution, adaptive grid, reaction-diffusion equations

AMS subject classifications: 65Y05, 65T60, 65M50, 65L04, 35K57

^{*}Université Nice Sophia Antipolis, CNRS, LJAD, UMR 7351, 06100 Nice, France (stephane.descombes@unice.fr).

[†]INRIA Sophia Antipolis-Méditerranée Research Center, Nachos project-team, 06902 Sophia Antipolis Cedex, France).

[‡]CCSE, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd. MS 50A-1148, Berkeley, CA 94720, USA (MDGonzalez@lbl.gov).

[§]Université de Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne Cedex, France (tdumont@math.univ-lyon1.fr, louvet@math.univ-lyon1.fr).

[¶]CNRS UPR 288, Laboratoire EM2C, Grande Voie des Vignes, 92295 Chatenay-Malabry Cedex, France (marc.massot@centralesupelec.fr).

^{||}CentraleSupélec, Grande Voie des Vignes, 92295 Chatenay-Malabry Cedex, France

^{**}Fédération de Mathématiques de l'École Centrale Paris, CNRS FR 3487, Grande Voie des Vignes, 92295 Chatenay-Malabry Cedex, France

^{††}Intel, Les Montalets, 2 rue de Paris, 92196 Meudon, France (thomas.guillet@intel.com).

^{‡‡}Exascale Computing Research, Campus Teratec, 2 rue de la Piquetterie, 91680 Bruyères-le-Châtel, France.

1 Introduction

Reaction-diffusion systems play an important role in many research domains such as chemistry, combustion, air pollution modeling or biomedical engineering applications involving the dynamics of moving fronts, usually very localized in space (see, e.g., [25] and references therein). They can model complex phenomena and potentially involve a large number of unknowns as well as a large spectrum of spatial and temporal scales. A general reaction-diffusion system can be written as follows, for $i = 1, 2, \dots, m$:

$$\left. \begin{aligned} \frac{\partial u_i}{\partial t}(x, t) - \operatorname{div}(\varepsilon_i(x) \operatorname{grad} u_i(x, t)) &= f_i(u(x, t)), & x \in \Omega \subset \mathbb{R}^d, t > 0, \\ u_i(x, 0) &= u_i^0(x), & x \in \Omega; \end{aligned} \right\} \quad (1)$$

with the compact notation: $u = (u_1, \dots, u_m)^t$. In particular we denote $f(u) = (f_1(u), \dots, f_m(u))^t$. With no loss of generality we restrict our presentation to homogeneous Neumann boundary conditions. Such systems can be also seen as the building block of more complex models including additional physical phenomena like convection.

Solving numerically this kind of problem involving a large number of nonlinear coupled equations in two and three dimensions is a challenging task. In particular two major difficulties need to be addressed. First, a large spectrum of temporal scales in the nonlinear source terms yields highly stiff equations. The latter can be expected when the Jacobian matrices, $(\partial f_i / \partial u_j)_{1 \leq i, j \leq m}$, have eigenvalues whose real part varies within a large negative interval. Systems of stiff ordinary differential equations impose the use of dedicated numerical methods in order to achieve accuracy and stability within reasonable memory and computing costs [38]. Secondly, steep fronts require a very fine discretization mesh, at least locally, which leads to problems of large size if no mesh adaptation is used. Additionally spatial stiffness may arise as a consequence of these steep spatial gradients even with non-stiff source terms and diffusion coefficients of relatively small value [20].

1.1 A new numerical strategy

We have recently introduced in [30] a tailored numerical strategy to cope with the latter difficulties using reasonable computing resources, that is, on a sufficiently powerful workstation possibly with a shared-memory architecture [29]. It relies on the following key ingredients:

- Time operator splitting: high performance computing can be achieved in the numerical solution of reaction-diffusion systems by choosing dedicated schemes for each split sub-system [30, 31]. Additionally these methods exhibit a large data-driven parallelism, and their mathematical analysis is well-established for relatively large splitting time steps even when stiffness is present [22, 20, 17]. Dynamically adapted splitting time steps can be also considered in the case of strongly varying dynamics [18].
- Multiresolution finite volume scheme: based on wavelet decomposition multiresolution schemes yield highly compressed representations for problems displaying localized fronts [12, 46]. In particular grid adaptation is performed dynamically between time steps according to an accuracy tolerance that tracks the approximation errors coming from data compression.
- One-step high order integration schemes: RADAU5 [38], an implicit, fifth order Runge–Kutta scheme with A - and L -stability properties for the time integration of the reaction terms given by stiff systems of ordinary differential equations; and ROCK4 [1], a stabilized, explicit Runge–Kutta method of order four for the diffusion problems. Both schemes feature time-stepping strategies that guarantee numerical approximations within a user-defined accuracy.

All these numerical techniques were implemented in the MBARETE code [25], using a tree-structured data with pointers and recursive navigation. While showing a great potential in several

cases, such a multiresolution implementation relied on a data structure that offered reduced optimization opportunities as well as limitations in terms of parallel programming due to the lack of data locality and load balancing. A new paradigm of parallelism together with a different and customized data structure needed to be sought for more performing splitting-multiresolution implementations.

1.2 A new paradigm of parallelism for shared-memory architectures

Until about 2005, continuous increases in both CPU clock frequency and transistor density had been driving the increment of computing performance from one processor generation to the next [7]. However, CPU frequencies have reached a plateau in the end of the last decade. This can be jointly attributed to the rise of leakage currents with transistor density, and the fact that the CPU total power budget has hit practical limits [8]. Still, improvements in manufacturing processes have been delivering ever-increasing transistor densities following Moore's law. As modern multi- and many-core CPUs use the increasing transistor budget from Moore's law to provide a growing number of cores, the bulk of performance gains is due today to parallelism.

A wide spectrum of techniques and runtime implementations are available for application developers to express parallelism. For the problems we are interested in and using the multiresolution approach, one single modern compute node provides enough memory and computing power to run simulations with a reasonable time-to-solution. Therefore, we have so far focused on parallelism over shared-memory architectures. Shared-memory parallelism provides a number of advantages well-adapted to multiresolution applications. Contrary to uniform Cartesian grids for which arrays can be generally accessed following regular patterns within long loops, adaptive meshing in multiresolution and AMR codes rely on fine-grained and dynamic data structures. The corresponding algorithms can have intricate data dependencies, especially for complex operations such as those associated with mesh adaptation. Therefore, exposing parallelism in these methods is best done using programming techniques that combine both coarse- and fine-grained parallel constructs, keeping in mind that maximizing parallel coverage is crucial to limiting the impact of Amdahl's law on scalability. Another key characteristic of multiresolution applications is their large dynamic range in both spatial and temporal scales. Because the resulting adapted meshes have complex geometries, an efficient parallel implementation requires non-trivial balancing of the computations across all available computing cores. Since the mesh structure evolves during a simulation, the load balancing needs to be frequently updated. Many existing mesh refinement algorithms and libraries address this issue in distributed-memory settings for both patch- and cell-based AMR [13, 49, 2, 53, 44, 64, 60, 15, 10, 16, 43].

In shared-memory architectures, the common memory address space enables dynamic load balancing while avoiding data transfers, but it also requires careful synchronizations between threads. Starting from the original implementations on a single computer [13, 4], several authors have proposed multithreaded implementations for adaptive gridding techniques [3, 23] with some focusing on specialized grid structures for multi-core processors [32, 57]. Task-based parallelism provides an attractive shared-memory solution to both granularity and dynamic load balancing requirements. In a task-based approach the programmer introduces parallelism by specifying computations that can be carried out in parallel. Expressing tasks can be done using different techniques, for example, relying on compiler directives or through library calls. Scheduling of the tasks is determined at runtime based on the available computing resources, yielding dynamic load balancing using techniques such as work-stealing [5, 6]. A key feature associated with tasks is that parallelism is not limited to flat parallel iteration constructs, but can be introduced recursively by having tasks create other tasks. This makes the task concept particularly suited to codes with complex hierarchical operations, such as multiresolution applications.

We have thus opted for a shared-memory parallel approach based on task-stealing, relying in particular on the Intel® Threading Building Blocks (TBB) library, very well suited to our case. Such a choice involved a completely new implementation of the MBARETE code, where a different and more efficient data structure has been introduced as an alternative to the original pointer-based

approach. Important performance enhancements, due to task-stealing techniques and in particular the use of the TBB library, have been already demonstrated in [39, 54] for multiresolution schemes applied to non-stiff time-dependent PDEs, that is, using explicit time integration schemes. These authors developed a different data structure based on the definition of *wavelet blocks* and thus a tree of wavelet blocks, where each block contains a predefined number of cells at the same grid-level. Here we focus on the numerical solution of stiff reaction-diffusion systems using the numerical strategy developed in [30] and summarized in § 1.1. In order to assess the new implementation and code performance we have chosen three reaction-diffusion models with increasing complexity.

1.3 Benchmark problems

In this work we consider three reaction-diffusion problems in dimension $d = 2$ and $d = 3$, with increasing complexity:

1. A bistable, Nagumo-type reaction-diffusion equation¹ (NAGUMO), similar to what can be found in combustion models [65] or in nonlinear chemical dynamics when studying traveling waves due, for instance, to chemical reactions with cubic auto-catalysis for which analytic solutions are available [35]. Here we have in (1), $m = 1$ and $f_1(u_1) = ku_1^2(1 - u_1)$. We consider $k = 10$, $\varepsilon_1 = 0.1$, and an initial condition verifying $0 \leq u_1^0(x) \leq 1$ everywhere. In general the reaction term is not stiff as $|df_1(u_1)/du_1| \leq 3k$. However, with this setting, traveling waves develop with a sharp spatial gradient yielding a space multi-scale configuration.
2. The Belusov-Zhabotinsky reaction (BZ). This is a system of $m = 3$ equations. The reaction term reads

$$\begin{aligned} f_1(u_1, u_2, u_3) &= 10^5 (-0.02 u_1 - u_2 u_1 + 1.6 u_3), \\ f_2(u_1, u_2, u_3) &= 10^2 (u_2 - u_2^2 - u_1(u_2 - 0.02)), \\ f_3(u_1, u_2, u_3) &= u_2 - u_3. \end{aligned}$$

This set describes a chemical reaction between $HBrO_2$, $Ce(IV)$ and Br^- [48, 41], also known as the Oregonator problem. For the diffusion part, we set $\varepsilon_1(x) = 2.5 \times 10^{-3}$, $\varepsilon_2 = \varepsilon_1$, and $\varepsilon_3(x) = 1.5 \times 10^{-3}$ in (1). The reaction term is stiff. Actually, this system of ordinary differential equations has a limit cycle along which the amplitude of the eigenvalues² of the Jacobian attains values of the order of 10^5 . As a comparison, when computing in $\Omega = [0, 1]^d$ with a spatial discretization step $h = 1/1024$, the largest negative eigenvalues in the diffusion term are about -2×10^4 . Propagating fronts with steep spatial gradients are also developed in this case yielding a time-space multi-scale configuration.

3. An ischemic stroke model (STROKE) [24, 19, 31]. This is a system of $m = 21$ equations. The reaction term is very stiff. A numerical computation of the eigenvalues of the Jacobian near a stable equilibrium state, $f(u) = 0$, shows real parts in the interval of $[-10^8, 0[$. This stiffness is due in part to the fact that $f(u)$ models voltage-gated ion channels that open or close when the difference of potential between a cell and the surrounding media attains some threshold. Gates are modeled through sigmoid functions, closely approximating a Heaviside function. Simulations show steep spatial gradients, as well. The right-hand side, $f(u)$, incorporates an important amount of biological knowledge about ion channels and consequently, its computation is performed using a quite complex program of about 400 instructions, computationally expensive, containing many log-function evaluations.

¹Type A bistable case in [61] is a limit case of the Nagumo function.

²In such systems, the eigenvalues can be positive, even with large amplitude, in some regions of the limit cycle.

1.4 Paper outline

The paper is organized as follows: in a second section, the operator splitting strategy is presented on the systems semi-discretized in space, and we describe the resolution of the split sub-steps, that is, parabolic diffusion equations and stiff systems of ordinary differential equations. Then we justify the use of an adaptive mesh strategy and summarize briefly the finite volume multiresolution approach in order to have a comprehensive picture of the numerical strategy. The main part of the paper is devoted to our choice of parallel runtime and the description of our parallel implementation of multiresolution with a specific data structure, as well as the new parallel implementation of ROCK4 and RADAU5 solvers for diffusion and reaction. The last part is devoted to a detailed performance analysis of our new software applied to reaction-diffusion systems. Potential improvements and future developments are highlighted in conclusion.

2 Numerical strategy

We are interested in two- and three-dimensional simulations. Discretizing (1) in space yields a large and stiff system of ordinary differential equations given by

$$\frac{dU}{dt} = A_\varepsilon U + F(U), \quad (2)$$

where A_ε is a matrix corresponding to the discretization of the diffusion operator. Using a Cartesian mesh A_ε is in general a block-diagonal matrix with 5 (resp., 7) non-zero elements per line in dimension 2 (resp., 3). In order to avoid the drawbacks of purely explicit or implicit methods a decoupling technique could be implemented instead. Here we consider operator splitting methods, briefly described in what follows in the context of the dedicated splitting solver introduced in [30].

2.1 Dedicated splitting solver

We consider the two sub-problems:

$$\frac{dV}{dt} = A_\varepsilon V, \quad \frac{dW}{dt} = F(W),$$

and we denote $D_{\Delta t}V_0$ and $R_{\Delta t}W_0$ the solution of the first and second sub-problem, respectively, after a time step Δt with the initial conditions $V = V_0$ and $W = W_0$. One can define the first order Lie schemes: $U_{n+1} = D_{\Delta t} \circ R_{\Delta t} U_n$ and $U_{n+1} = R_{\Delta t} \circ D_{\Delta t} U_n$, as well as the second order Strang schemes: $U_{n+1} = R_{\Delta t/2} \circ D_{\Delta t} \circ R_{\Delta t/2} U_n$ and $U_{n+1} = D_{\Delta t/2} \circ R_{\Delta t} \circ D_{\Delta t/2} U_n$ [58]. The practical implementation of both Lie and Strang schemes is straightforward once the substeps $R_{\Delta t}$ and $D_{\Delta t}$ have been defined. In this work we focus on the Strang scheme. In particular it has been shown that better accuracies are expected by ending the splitting scheme with the substep involving the fastest time scales [21, 22, 17]. Therefore, assuming in general stiff reaction terms, we consider without any loss of generality the scheme, $U_{n+1} = R_{\Delta t/2} \circ D_{\Delta t} \circ R_{\Delta t/2} U_n$. Both theoretical and practical results show that this method performs very well for stiff reaction-diffusion systems [29, 30, 31]. One of the most interesting properties is that the *splitting time step*, Δt , can be chosen independently of standard stability constraints associated with mesh size or stiff source terms, as long as the sub-problems are solved either exactly or with dedicated solvers with appropriate stability properties and time-stepping features [30]. This property allows us to choose Δt some orders of magnitude larger than the fastest time scales of the system. Additionally, an adaptive splitting scheme can be easily implemented within this framework [18] to cope with solutions displaying different dynamic behaviors in time.

2.1.1 General considerations on the splitting implementation

The following remarks can be made in terms of the solvers implemented for the numerical time integration of the split sub-systems and their capabilities for parallelism in the context of the present work.

Reaction step $R_{\Delta t/2}$: after spatial discretization by finite volumes, finite differences or finite elements one obtains a large set of independent Cauchy problems given by systems of ordinary differential equations. As previously mentioned, when these systems are stiff we use the RADAU5 method [38] which has experimentally proved to be highly efficient in terms of computational costs, accuracy and robustness. When the systems are not stiff one can consider any explicit Runge–Kutta method such as the well-known and simple RK4 explicit method.

Diffusion step $D_{\Delta t}$: one could consider, for instance, diagonally implicit Runge–Kutta methods like SDIRK schemes [38], for which we would need to solve several linear systems successively. However, when the diffusion problem is mildly stiff, one can consider instead a stabilized, explicit Runge–Kutta scheme like the ROCK4 method [1]. The main advantage of such an alternative is that one only needs to evaluate matrix-vector products³ given its explicitness. The number of Runge–Kutta stages, s , can be larger than the order of the method, but the stability domain grows like s^2 . For system (1) the stiffness of the diffusion sub-systems, given by $\varepsilon_i(x)\Delta t$, must be contained within the stability domain. In particular we have chosen $\Delta t = 10^{-2}$, 10^{-3} , and 1 for the (NAGUMO), (BZ), and (STROKE) models, respectively. ROCK4 is well-suited to all these cases. An additional advantage in the context of dynamic meshing strategies is that an explicit approach does not need to compute any preconditioner nor perform any factorization when the mesh has changed.

Parallelism: the $R_{\Delta t/2}$ step exhibits a high data-driven parallelism for one has as many independent problems as nodes in the spatial discretization. However, for the $D_{\Delta t}$ step one can only achieve a *poor man's parallelism* based on the solution of m independent parabolic linear equations [29]. Better alternatives should be thus considered to enhance the parallel performance of the diffusion solver. The latter can be much simpler to achieve in the case of an explicit scheme like ROCK4 since matrix-vector product are easier to parallelize than linear solvers.

2.1.2 Numerical experimentation on uniform Cartesian meshes

Let us consider a numerical implementation of the Strang scheme to solve (1) on a uniform Cartesian grid. ROCK4 and RADAU5 are considered to solve the diffusion and reaction sub-systems, respectively. With this setting the following table provides an estimate of the percentage of total CPU time spent in the diffusion step of the splitting strategy:

Grid size	(NAGUMO)	(BZ)	(STROKE)
512^3	92.5	34.5	1.6
1024^2	88.5	20.1	1.0

Notice that for the models of higher complexity, that is, (BZ) and (STROKE), most of the simulation time is spent in the reaction step. Consequently, as a measure of the computational complexity, we use the number of evaluations of the right-hand side performed by the RADAU5 program on a given grid-node during the $R_{\Delta t/2}$ step. However, this complexity measure can vary considerably throughout the computational domain given that the splitting time step $\Delta t/2$ can be further split

³Additionally, one scalar product per time step must be performed but only when the time-stepping technique is enabled.

according to the time-stepping procedure and that an iterative, simplified Newton method is implemented to solve the nonlinear systems. In particular the minimum complexity corresponds to grid-nodes where only one reaction sub-step is required during a given $\Delta t/2$, as well as one iteration of the Newton method.

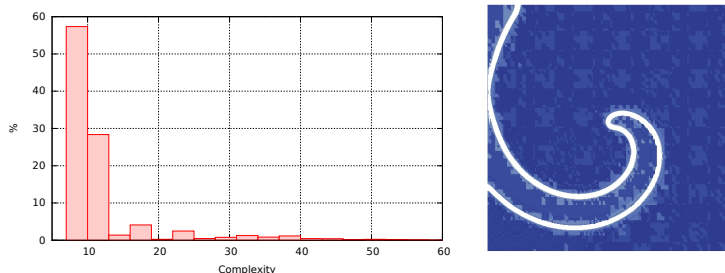


Figure 1: BZ problem: histogram of complexity (left). In the right picture, white zones correspond to points where the complexity is greater than 17.

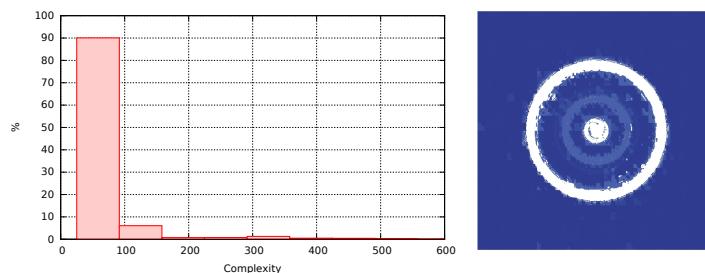


Figure 2: STROKE problem: histogram of complexity (left). In the right picture, white zones correspond to points where the complexity is greater than 80.

Considering a uniform mesh of 1024^2 grid-nodes, Figures 1 and 2 show the computational complexity measured for a given $\Delta t/2$ for the (BZ) and (STROKE) models, respectively. For the (BZ) problem the Jacobian is computed analytically involving three right-hand side evaluations. In this case 87% of the reaction CPU time is spent over regions where the complexity is less than 17. These are the regions where the solution is close to the reaction equilibrium and therefore, the complexity remains also close to its minimum value. Similarly, about 90% of the reaction CPU time is spent in regions where the complexity is less than 60 for the (STROKE) model. Here the Jacobian is computed numerically. One can conclude that an adaptive meshing technique that reduces the number of grid-nodes over near-equilibrium zones would significantly reduce the CPU time together with the required computer memory.

2.2 Adaptive multiresolution method

Grid adaptation must be achieved automatically by the software. We describe here some basic features of the multiresolution method introduced in [12], applied to system (1). For a comprehensive overview on adaptive multiresolution techniques, we refer to the books of Cohen [11] and Müller [46]. More details on this particular multiresolution implementation can be found in [25].

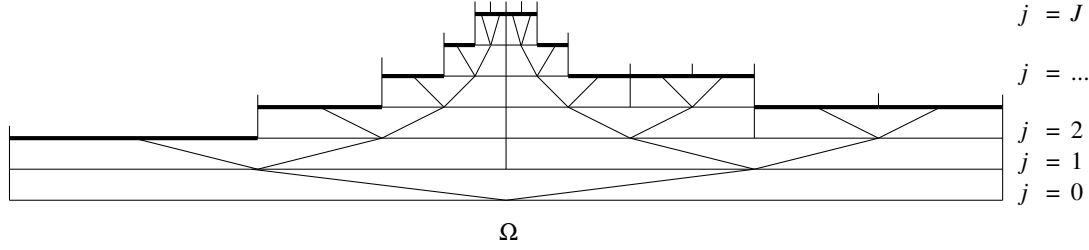


Figure 3: One-dimensional graded tree composed of nested grids. The resulting adapted mesh is given by the leaves of the tree, indicated in bold.

Without loss of generality we consider the computational domain $\Omega = [0, 1]^d$, $d = 1, 2, 3$. A recursive dyadic subdivision of Ω is then performed yielding 2^{dj} cells (segments, squares or cubes) of equal size at each grid-level j , from $j = 1$ to $j = J$. Levels 0 and J stand, respectively, for the whole computational domain Ω and the mesh at the finest spatial resolution. An adapted spatial discretization can be thus achieved by truncating the set of nested grids throughout the tree as shown in Figure 3 for a one-dimensional case. In particular the adapted grid is given by the *leaves* of the tree. Finite volumes are associated with all cells of the tree and throughout this paper we will refer to all cells, including the leaves, and their corresponding finite volumes as *nodes*. The *root* of the tree is the entire domain Ω . Even though the solution of problem (1) is performed on the adapted mesh, that is on the leaf nodes, the solution is updated and stored at all nodes of the tree. Hence, we denote \mathbf{U}_j as the set of values at nodes of level j . Data at different levels of discretization are related by two inter-level transformations performed by means of the *projection* and *prediction* operators. The projection operator, P_{j-1}^j , which maps \mathbf{U}_j to \mathbf{U}_{j-1} , is obtained through exact averages computed at the finer level; that is, for a given node at level $j-1$ one takes the average of values associated with its children at the successive level j . The prediction operator, P_j^{j-1} , which maps \mathbf{U}_{j-1} to an approximation $\hat{\mathbf{U}}_j$ of \mathbf{U}_j . The following two constraints must be observed when defining P_j^{j-1} [12]:

1. The prediction must be local, that is, it must depend on cells contained in a finite neighborhood.
2. The prediction must be consistent with the projection in the sense that $P_{j-1}^j \circ P_j^{j-1} = Id$.

2.2.1 Compression on graded tree-structured data

Data compression and thus grid adaptation are achieved by introducing a local estimator of the spatial regularity of the solution at a given simulation time. These local estimators are known as *details*, and for a given node at level j its detail is defined as

$$d_{j,k} = u_{j,k} - \hat{u}_{j,k} = u_{j,k} - P_j^{j-1} \circ P_{j-1}^j u_{j,k},$$

where $k \in \mathbb{Z}^d$ accounts for the location of the node at level j and $u_{j,k}$ represents the cell-average of $u(x, t)$ there. Introducing a tolerance parameter, $\epsilon > 0$, threshold values are defined level-wise as $\epsilon_j = 2^{\frac{d}{2}(j-J)}\epsilon$, $j = 0, 1, \dots, J$. Data compression is thus achieved by discarding nodes whose *details* are smaller than ϵ_j in a given norm; here we consider an L_2 -norm [25]. Conversely, nodes whose *details* exceed ϵ_j must be refined. However, a graded tree-structure must be guaranteed during the coarsening/refining procedure, meaning that all cells required to compute P_j^{j-1} must always be available (see [12] for more details).

In this work the approximated values, $\hat{\mathbf{U}}_j$, generated with the P_j^{j-1} -operator are obtained using centered polynomial interpolations of accuracy order $r = 2l + 1$, computed with the l nearest

neighboring cells in each direction including the diagonals in multi-dimensional configurations. This procedure is exact for solutions given by polynomials of degree $2l$. For a one-dimensional configuration with $l = 1$, the prediction operator is explicitly given by [12]

$$\hat{u}_{j+1,2k} = u_{j,k} + \frac{1}{8}(u_{j,k-1} - u_{j,k+1}), \quad \hat{u}_{j+1,2k+1} = u_{j,k} + \frac{1}{8}(u_{j,k+1} - u_{j,k-1}). \quad (3)$$

Higher order formulas can be found in [46], while extension to multi-dimensional Cartesian grids is easily obtained by a tensor product of the one-dimensional operator (3) (see, e.g., [55]). In general the interpolation stencil is given by $(2l + 1)^d$ cells at the coarser level. Here we will only consider third order interpolations with $l = 1$; the impact of higher order approximations has been investigated, for instance, in [59, 30].

3 Shared memory task-based parallelism for multiresolution problems

In our application, we have chosen to introduce shared memory parallelism using the task-based Intel[®] Threading Building Blocks runtime, which we found to be particularly suited to our multiresolution implementation. We now discuss parallel runtimes in the context of multiresolution codes, and motivate our choice of TBB, in particular instead of the OpenMP[®] application programming interface (API).

There is a number of runtimes available today supporting task parallelism, such as Intel[®] Threading Building Blocks [52, 40], the OpenMP API [50], XKaapi [33], among others. We have considered the OpenMP API and Intel[®] Threading Building Blocks, both well-established and proven options. The OpenMP standard is a widely supported and popular option for shared memory parallelism, in particular in HPC codes. It relies on compiler directives to introduce parallel regions, work sharing constructs, and tasks, and interfaces with a runtime library component. However, the OpenMP API relies on a thread-based (instead of task-based) approach to parallelism; among important consequences are:

- OpenMP parallelism is strongly tied to explicitly defined parallel regions: a parallel region enforces parallel execution, and no code outside parallel regions will be executed in a thread-parallel way.
- OpenMP worksharing constructs such as `omp for` decompose the work across worker threads, not in terms of tasks. But since all threads of a parallel region are required to enter a worksharing construct, a consequence is that it is not legal to arbitrarily nest OpenMP tasks and worksharing constructs. In particular, one cannot introduce parallelism by nesting `omp for` constructs within `task` constructs.

Because of the above points, OpenMP parallelism is not *composable*: it is not possible to expose parallelism in an opportunistic way, regardless of the outer context from which a piece of code may be called. We believe that composability is a desirable property for parallel runtimes, in particular for C++ multiresolution codes which combine complex call graphs with code reuse through classes.

The OpenMP standard provides some support for nested parallel regions which can help resolve this composability issue, but which also introduce problems of their own. Depending on the OpenMP implementation, nested regions can lead to exponential creation of OpenMP worker threads, easily resulting in oversubscription. More fundamentally, dynamic load balancing cannot occur between two concurrently executing parallel regions, because the standard requires that the worker threads participating in a region are determined when entering the region.

Intel[®] Threading Building Blocks (TBB) relies on a different approach to parallel programming. TBB is an open-source C++ template library for task parallelism, and requires no special support in the compiler. In addition to task-based parallel constructs, TBB also provides concurrent data

structures and memory allocators. The library implements a composable task-based runtime: any C++ function can expose parallelism and combine tasks with constructs such as `parallel_for` or `parallel_reduce`, irrespective of the calling context. Dynamic load balancing is achieved through work stealing. Since both tasks and constructs such as `parallel_for` rely on the same task-based scheduler, TBB's parallel idioms may be combined and nested arbitrarily, as long of course as the programmer avoids races through appropriate synchronizations.

In addition to a composable task runtime, TBB provides a parallel algorithm library, with support in particular for parallel reductions and parallel sort.

We chose TBB over the OpenMP model because the former provides an efficient task runtime with dynamic load balancing, composability of parallel code, and an array of parallel data structures and algorithms. Equivalent functionality could likely have been achieved using OpenMP tasks, but with more effort to circumvent some of the shortcomings of the OpenMP model we discussed in this section.

4 Sketch of the implementation

- For the Reaction step $R_{\Delta t/2}$, the implementation is very simple, once the leaf nodes of the spatial discretization are numbered from 0 to $M - 1$, and an access to the values of u at nodes is provided: one must write a class `R` which computes the solution on an interval of nodes (a `blocked_range<size_t>` for TBB); then, a call to `parallel_for(blocked_range<size_t>(0,M),R)` parallelizes the computation.
- For the Diffusion step $D_{\Delta t}$, one can use 2 levels of parallelism: a parallelization on the m unknowns of (1), and a parallelization of the matrix-vector products, using here two `parallel_for` constructions.

5 Parallel implementation of the multiresolution algorithm

The implementation of multiresolution faces two problems:

1. Data locality: projection and prediction operators are defined on trees (quad-trees in dimension 2, oct-trees in dimension 3). Moreover the prediction operator is based on a stencil, and, for any given node, one must access not only the father and the brothers of the node in the tree but also its uncles and cousins (see Figure 4).

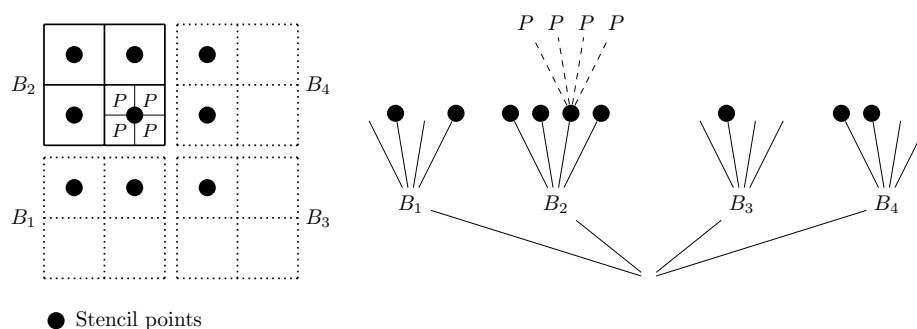


Figure 4: Computing values at nodes marked p ; stencil extends on 4 brotherhoods $B_i, i = 1, 4$.

The projection operator is much less demanding, as it needs only to access the sons of the non leaf nodes. An efficient implementation of quad-trees and oct-trees which insures as much data

locality as possible is necessary: classical implementation of trees using pointers is probably not efficient, since it is certainly difficult to ensure any data locality.

2. Parallelism: the projection and prediction operators, refinement and coarsening operations, all defined on trees, must be parallelized in order to maximize parallel coverage and minimize the impact of Amdahl's law.

In the following, we explain our implementation of quad-trees and oct-trees, and describe the parallel implementation of the most important parts of the multiresolution.

5.1 Implementation of trees

Now, Ω will be $]0, 1[^d$ with $d = 2$ or $d = 3$. In a refinement process, we first note that the coordinates $x_i, i = 1, d$ of the centers of the volumes created have all a base 2 development limited to n digits: $x_i = 0.x_{i,1}x_{i,2} \dots x_{i,n}$, n being the level of the volume in the tree. The idea, which has already been used in multiresolution [9, 14] and AMR [10, 62] contexts, is to index all nodes in the tree by a space-filling curve [56]. We have used a Morton order indexation [45] (also known as Z-Order curve or Morton code): in dimension 2, for a given node $(x_1 = 0.x_{1,1}x_{1,2} \dots x_{1,n}, x_2 = 0.x_{2,1}x_{2,2} \dots x_{2,n})$, we define his Morton abscissa by

$$s = x_{1,1}x_{2,1}x_{1,2}x_{2,2} \dots x_{1,n}x_{2,n},$$

(that is to say by interlacing the digits of x_1 and x_2). This generalizes immediately in dimension 3. For all nodes of the tree, we store s and the level of the node in the tree in a 64 digits (unsigned) integer: in dimension 3, this allows to code all nodes of a 16 levels tree, and to add some tag marks during the computation (see Figure 5). This representation of nodes by unique integers, instead of an explicit tree in memory, is similar to the so-called CSAMR data structure introduced by [42]. Even though we also rely on hash tables for efficient node lookup, we use a slightly more complicated data structure based on *blocks* described below, which allows fine control over lookup granularity, and makes it easier to implement shared memory parallelism.

Data structure We define 2 levels:

1. *Blocks*: as $r = 0.s \in [0, 1[$ we partition $[0, 1[$ into intervals; a *block* is then a structure which contains data related to given interval, that is to say:
 - The bounds of the interval s_{\min} and s_{\max} .
 - A vector, where all the nodes with abscissa s for $s \in [s_{\min}, s_{\max}[$ are stored.
 - The size of the vector.
2. *Collections of blocks*: We store the pointers to the *blocks* ordered by the s_{\min} values in vectors. Actually, we manage 2 *collections of blocks*: one for leaf nodes and one for non leaf nodes: test have shown that transfers between these 2 structures is more efficient than using a unique collection.

Abscissa are not necessary ordered in a *block*, and the strategy for searching a given node is to find by a dichotomic search the *block* where his abscissa is (possibly) stored, and then do a sequential search in the *block*. This strategy has proven to be the most efficient, in particular compared to a fully ordered storage. A system of caches get track of the last *blocks* accessed, allowing to decrease the number of dichotomic searches. Caches are local to TBB tasks.

Associated to this data-structure, we just need to implement few methods:

1. cut a *block* in 2 *blocks* when the size of the *block* exceeds some threshold,
2. fusion of two neighbor *blocks* when the sum of their size is less than some other threshold.

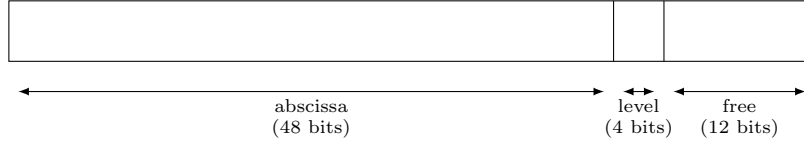


Figure 5: Representation of a node using 64-bit integers.

3. enlarge the vector contained in *blocks* (this is done by allocating a vector of double size, copying and deleting the old vector).
4. decrease the size of the vector contained in *blocks* when it contains too few data (here also by allocating, copying deleting).
5. some garbage collecting operations (for example suppress nodes marked for deletion).

With these methods, we maintain the size of individual *blocks* between some fixed limits: thus the complexity of a search is $O(\log_2 N_b) + \text{Constant}$ where N_b is the number of *blocks* in the collection.

New nodes, in a refinement process, are just added at the end of the vector of the *block* which contains their abscissa. To delete nodes in the coarsening process, we use a *lazy* strategy: first a digit is positioned in the free digits part of the 64 digits word which represent the nodes we want to delete; in a second step, the garbage collector really suppresses marked nodes. Note that the 5 operations described above can all be done in parallel, using the `parallel_for` structure of the TBB library, looping on *blocks*.

5.2 Parallel implementation of some multiresolution algorithms

Computing the “details” Here, we have two steps: 1) *projection*, then 2) *prediction* and *Details* computation.

1. *projection*: we must propagate information from the leaves to the root of the tree, which can be seen as a recursive computation. The most efficient implementation we have tested is the following:
 - (a) We cut the tree at a low, fully populated level J_{\min} : for each node N of level J_{\min} we have an independent computation to perform, recursively transferring data from leaves to the node N . These computations are done using recursive `parallel_for` constructions.
 - (b) A sequential (recursive) computation propagates data from level J_{\min} to the root of the tree.

In practice, we choose $J_{\min} = 3$ in dimension $d = 3$ and $J_{\min} = 5$ in dimension $d = 2$. Note also that, often, the second sequential phase of the computation does not need to be performed, as the leaves remain everywhere at a level $\geq J_{\min}$; anyway, phase (b) is an inexpensive task.

2. *prediction* and *Details* computation:

```

for  $N \in \{\text{Non leaf Nodes}\}$  do
    Make the list  $L$  of Nodes connected to  $N$  by the prediction stencil 3;
    Get values associated to Nodes in  $L$ 
    Compute prediction and details, and store details.
end for

```

These computations are easily be implemented using a `parallel_for` structure, iterating on *blocks*.

Mesh management Here also, *lazy* strategies are the key to parallelism.

1. *Refinement*: Once the “details” have been computed, we first tag in parallel the nodes which must be refined. Then the idea is to loop on the following two steps (1) and (2) until a fixed mesh is obtained:

- (1) Refine the Nodes which are tagged. Note that, after this step, the tree is no more graded.
- (2) Attempt to make the tree graded: for this, some nodes must be refined, and thus we tag them.

For (1) we proceed in 2 sub steps (see Figure 6):

- (1a) Each task operates on a consecutive set S of *blocks*, thus on an interval $[s_0, s_1[$. The created nodes which belong to $[s_0, s_1[$ are directly stored in S ; other nodes (which are actually very few) are temporally stored in new created *blocks* (2 blocks by task). All of these operations can be safely done in parallel, because no task will be writing into any other’s blocks or temporary storage.
- (1b) Each task works on a consecutive set T of blocks; the *blocks* created in (1a) only need to be accessed read-only to move the stored nodes to leaves, and thus this can be done in parallel.

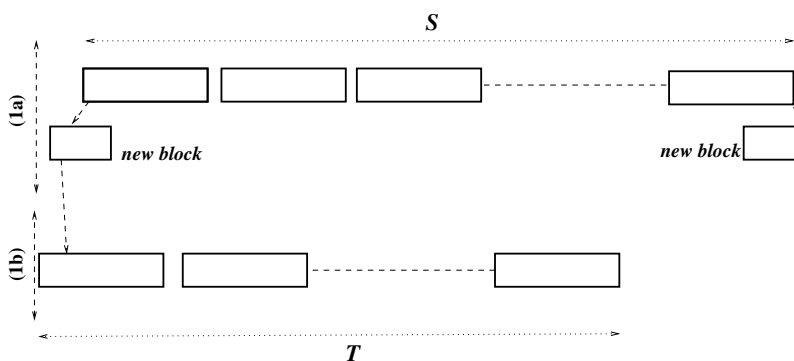


Figure 6: Refinement, phases (1a) and (1b).

2. *Coarsening*: we iterate on the following steps, as long as some nodes can be suppressed:
 - (a) Tasks work on consecutive sets of *blocks*, searching for the leaf nodes which can be deleted (that is to say the leaf nodes which have not been created in the refinement phase, and which can be suppressed keeping the graded structure of the tree) and marking them as deleted.
 - (b) Fathers of suppressed nodes are moved in the leave collection (using the technique described for Refinement, see (1a), (1b) above).
 - (c) A parallel call to the garbage collector really suppress the nodes deleted.

Each of these operations can be parallelized using `parallel_for` on blocks, since they involve either changes to blocks local to the task, or in the case of (b), two passes as described in the refinement steps (1a) and (1b) above.

Data storage The values of the computed unknowns are associated to nodes of the tree. We have two options for the storage memory layout of the unknowns: either to store the data related to each node in a common structure (in the so-called “array of structures” scheme), or to store each unknown in its own contiguous array (as a “structure of arrays”). As we shall see later, a large fraction of the run time is spent in the reaction and diffusion steps. The reaction step computes local reaction rates from all the unknowns at the current point, and would therefore in principle benefit from the array of structures scheme for data locality. However, the reaction solver is extremely compute intensive, and therefore largely insensitive to the memory layout. The diffusion step, on the other hand, has very low compute intensity, which makes it sensitive to the memory layout. Since the diffusion step is computed on each unknown independently, the “structure of arrays” scheme provides better data locality. We have therefore opted to store each unknown in its own contiguous array, which will prove to be a relevant choice when the variables are not coupled in the diffusion operator. The Morton order defines the order of the unknowns, which are stored in vectors. For the sparse matrices used in the diffusion steps, we use a classical CSR storage [47] or a more compact storage as described in the appendix when the diffusion coefficients ε_i are constant.

5.3 Implementation of Rock4 for the heat equation

ROCK4 [1] is an explicit method; it is the composition of two methods. The first one is based on orthogonal polynomials: thus its implementation uses a three-term recurrence formula (whose length depends on the spectral radius of the Jacobian of the system). The second one is an explicit 4-stage Runge–Kutta formula; for a system $dy/dt = f(y)$ and a time step Δt it is given by:

$$\text{Compute } k_i = f(x_n + \Delta t \sum_{j=1}^{i-1} a_{i,j} k_j), i = 1, 4, \text{ and then: } x_{n+1} = x_n + \Delta t \sum_{i=1}^4 b_i k_i.$$

($a_{i,j}$ and b_i being the coefficients of the formula). One must remark that in a naive implementation of the second method, we would compute linear combinations of (large) vectors (from 1 to 4 vectors). Let us recall that the *arithmetic intensity* of an algorithm is the ratio of total floating-point operations to total data movement (bytes). Achieving a high arithmetic intensity is a key for obtaining high FLOP/s efficiency (see [63]). For linear combinations of large vectors of *double*, this intensity is very low (less than 1/4). Here, the arithmetic intensity can be improved using the well known fact [37] that an explicit Runge–Kutta formula applied to a linear problem $dy/dt = Ay$ is given by $x_{n+1} = P(\Delta t A)x_n$ where P is a polynomial whose coefficients depends of the $a_{i,j}$ and k_i Runge–Kutta coefficients. An implementation based on Horner formula avoids expensive linear combinations (note that P has complex roots, thus an evaluation based on a factorization of P is not feasible). The spectral radius R_A of A is estimated using the Gershgorin theorem; R_A is used to determine the number s of stages of the method.

6 The computing code

We have developed a pure C++ code, named **Z-code**, where the most part uses the dimension d as template parameter. This program runs on recent Linux systems, and as been tested with `gcc` (version ≥ 4.8) and the Intel C++ (version $\geq 14.0.1$) compilers. Figure 7 shows some results of computations.

Related developments

Codes for RADAU5 and ROCK4 methods are publicly available on the Internet[36]. These codes, albeit very efficiently coded in Fortran 77 can be improved; for RADAU5 a careful rewriting in C++ allows to remove a lot of branching; this result in a limited but not negligible improvement

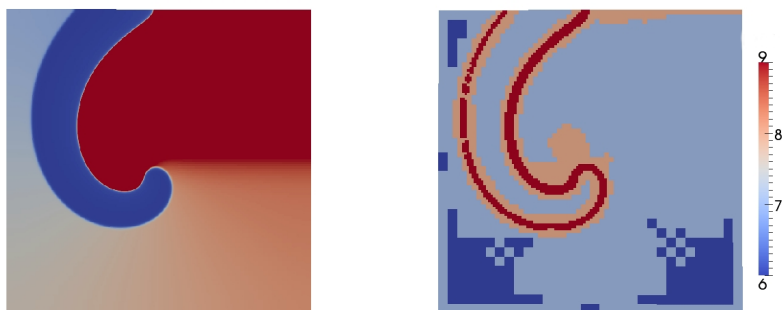


Figure 7: (BZ) model, a simulation in dimension $d = 2$ with 9 levels. left: u_1 , right: levels.

of the performances (20 to 30 %). For ROCK4, we have developed a complete alternative avoiding unnecessary vector copies and a version adapted to linear problems, as described above⁴.

7 Performance

Application performance can be measured from many points of view; we present comparisons between Cartesian meshes and multiresolution computations and also focus on scalability of given problem with increasing thread counts (strong scaling). The tests have been performed on an Intel Xeon E5-2680 platform (2 *Sandy Bridge* CPUs, 16 cores total, 32 threads with hyperthreading), and on a Xeon Phi (MIC) 5110P (60 cores, 240 threads with simultaneous multithreading, 8 GB main memory) using the Intel `icpc` compiler.

7.1 Uniform Cartesian meshes versus multiresolution

We present comparisons between fixed mesh (full Cartesian) and multiresolution computations. The question is whether the reduction of computing time using multiresolution is counterbalanced or not by the cost of mesh adaptation. To obtain a correct comparison, we have replaced the matrix vector products of the $D_{\Delta t}$ sub steps implemented in our code by an implementation adapted to regular 5 and 7 points stencils in Cartesian meshes, using 2D or 3D arrays and cache blocking. This avoids the extra cost of manipulating complicated data structures and also reduces the bandwidth of the computation.

Results presented were obtained on the *Sandy Bridge* machine, using 32 threads (on 16 cores).

Results In the following figures, *level* is the maximum level authorized for mesh refinement. The compression ratio is defined as $cr = 1 - nnodes/nlevel$ where *nnodes* is the number of nodes used for the solution (number of leaf nodes in the tree) and *nlevel* is the number of nodes in the full grid at the maximum authorized level ($nlevel = 2^{d \cdot level}$). Figure 8 shows the computing time (wall clock, in seconds) for different levels *lev* of discretization in dimension 2 and 3 for our 3 model problems. During all the computations, we had $80\% \leq cr \leq 83\%$.

From these experiments, we conclude that the multiresolution approach is more efficient than the Cartesian one, except for very small and easy computing cases (in dimension 2 with 8 levels, for problems where the reaction is inexpensive to solve). In most cases, the extra cost of the mesh adaptation, and of the manipulation of complex data structures for matrices is counter balanced by

⁴Sources for Z-code and related software can be obtained by contacting T. Dumont, the corresponding author of the present contribution. All codes are under Cecill-B licence (<http://www.cecill.info/licences.en.html>).

		$d = 2$			$d = 3$	
lev		8	9	10	8	9
(NAGUMO)	MR	$2.45 \cdot 10^{-3}$	$2.50 \cdot 10^{-3}$	$3.50 \cdot 10^{-3}$	0.12	0.43
	Cartesian	$1.23 \cdot 10^{-3}$	$3.33 \cdot 10^{-3}$	$1.82 \cdot 10^{-2}$	0.21	1.78
(BZ)	MR	$9.10 \cdot 10^{-3}$	$1.44 \cdot 10^{-2}$	$3.22 \cdot 10^{-2}$	1.27	7.86
	Cartesian	$1.35 \cdot 10^{-2}$	$4.88 \cdot 10^{-2}$	0.21	3.35	36.46
(STROKE)	MR	0.54	1.91	6.80	22.51	175.4
	Cartesian	1.02	3.38	13.1	146.5	1194.

Figure 8: Multiresolution (MR) versus Cartesian computations: computing time in seconds.

the cost of the reaction sub step. One must also keep in mind that these results, advantageous for multiresolution, are obtained only for problems where the solution exhibits localized large gradients.

7.2 Scalability

We explore the strong scalability (fixed problem size, increasing thread count) of the different parts of the computing process: Mesh adaptation (A) (including *details* computation), Reaction steps (R), Diffusion step (D) (building matrices and resolution) and the entirety of the step computation (S). Let T_k the computing time (wall clock, in seconds) using k threads. We define the parallel efficiency on k threads by $s_k = T_1/n_k T_k$, where n_k is the number of *cores* used when computing on k threads. Because of simultaneous multithreading, $n_k = \min(k, 16)$ on the Xeon, and $n_k = \min(k, 60)$ on the Xeon Phi. Note that, using \min , we compute the efficiency based on the number of CPU *cores*, not the number of available hardware threads. The motivation is that the cores are what constitute independent compute units: while using more than one thread per core can help hide various latencies, one cannot expect performance to scale linearly with the number of threads per core because these threads will compete for the same compute resources.

The following arrays give s_k for different configurations:

- (NAGUMO) problem, in dimension $d = 3$, on the *Sandy Bridge*:

		$J = 9$					$J = 10$				
k		2	4	8	16	32	2	4	8	16	32
A		0.98	0.95	0.88	0.82	0.96	0.99	0.95	0.89	0.71	0.91
R		0.98	0.99	0.97	0.96	1.09	0.98	1.00	0.99	0.93	1.12
D		0.98	0.97	0.91	0.83	0.95	0.98	0.95	0.91	0.81	1.00
S		0.98	0.96	0.89	0.83	0.96	0.98	0.95	0.90	0.75	0.95

On the *Xeon Phi* we obtain, for $d = 3$ and $J = 10$:

k	8	30	60	120	240
A	0.96	0.74	0.60	0.73	0.72
R	1.01	1.01	1.00	1.53	1.39
D	0.96	0.84	0.71	0.90	0.94
S	0.96	0.79	0.65	0.80	0.81

- (BZ) problem in dimension $d = 3$, with J levels of grids and for a typical step:

1. On the *Sandy-Bridge*:

	$J = 9$					$J = 10$				
k	2	4	8	16	32	2	4	8	16	32
A	0.97	0.95	0.89	0.85	1.00	1.01	0.99	0.94	0.90	1.09
R	1.06	1.03	1.13	1.15	1.39	1.01	1.12	1.15	1.13	1.43
D	0.99	0.95	0.89	0.87	1.01	0.99	0.95	0.90	0.88	1.01
S	1.03	1.01	1.05	1.04	1.24	1.01	1.07	1.06	1.04	1.28

2. On the *Xeon-Phi*, we are limited to 9 grid levels in dimension $d = 3$ due to the 8 GB of main memory on our model. We obtain:

k	2	4	8	15	30	60	120	240
A	1.08	1.10	1.08	1.06	0.95	0.78	1.06	1.17
R	1.06	1.07	1.12	1.12	1.14	1.16	1.90	2.72
D	1.00	0.97	0.97	0.94	0.89	0.79	1.11	1.22
S	1.05	1.06	1.09	1.08	1.07	1.03	1.59	2.03

- (STROKE) problem in dimension $d = 3$, with $J = 9$ levels:

	<i>Sandy Bridge</i>					<i>Xeon Phi</i>				
k	2	4	8	16	32	8	30	60	120	240
A	0.91	0.81	0.67	0.72	0.72	0.91	0.85	0.69	0.88	0.88
R	1.00	1.01	1.04	1.07	1.29	1.04	1.04	1.06	1.54	1.95
D	0.92	0.84	0.78	0.71	0.80	0.94	0.89	0.81	1.05	0.93
S	1.00	1.01	1.03	1.06	1.27	1.03	1.04	1.05	1.52	1.93

These results show that the full computation (S) scales correctly, even if the sub steps (A) and (D) do not scale perfectly. This is better understood by looking at the percentage of time spent in different steps of the computation (here, all computations were done in 3d, using 10 levels, with 32 threads on the *Sandy Bridge*):

	(BZ)	(STROKE)	(NAGUMO)
Adaptation	17.17	1.64	61.00
Reaction	60.86	97.02	1.32
Diffusion	21.97	1.34	37.68

The scalability of (STROKE) closely follows the scalability of the reaction step, but for (NAGUMO) the time is dominated by the adaptation and diffusion steps because the reaction term is much less expensive to compute for this problem.

Some comments:

- As could be expected, the Reaction step scales very well on both computers, because the arithmetic intensity is very high and it is embarrassingly parallel across mesh points.
- The Diffusion step scales correctly: here we have used the compact structure described in the appendix; using a classical CSR structure results in lower performance. Note that because the Diffusion step is memory-intensive, one cannot expect it to scale better than the memory bandwidth with the number or cores. Because the available memory bandwidth per core decreases with core count on common architectures, it will not be possible to achieve perfect scaling on the diffusion step.
- All the adaptation process also scales quite correctly but not ideally, and the same arguments discussed for the diffusion also apply.

- With our code, simultaneous multithreading (running more than one thread per core) noticeably improves performance, in particular on Xeon Phi (where maximum performance cannot be achieved with one thread per core alone). Performance gains are greatest for the reaction solver; we speculate that simultaneous multithreading helps hiding the latencies related to the complex control flow in RADAU5.

8 Conclusions and prospects

The purpose of the present contribution was to investigate the capabilities of a new numerical strategy (involving operator splitting with specific high order numerical integrators coupled to adaptive multiresolution in space) in the framework of shared-memory architectures by introducing new data structure and algorithms. Before getting into the summary of the achievements, let us underline that we have tackled three model reaction-diffusion problems. They are representative building blocks of more complex and realistic applications encountered in biomedical engineering [31], combustion [28] even in the low Mach number approximation with detailed transport and complex chemistry [17], plasma physics [27], and which can be coupled to a Poisson solver if needed [26]. By representative, we mean that the level of stiffness reached in these three models is of the same order compared to the ones observed in these realistic applications. It is also important to notice that multiresolution methods are by no means universal: the solution must exhibit localized sharp structures to get a significant performance advantage. Nevertheless, even if restricted to specific problems, the proposed strategy can be the key to solve some multiscale problems in two or three dimensions, which would be out of reach on standard computing resources using classical approaches because of the induced stiffness and memory requirements (see [31] for example in biomedical applications).

The main achievement of the paper is the introduction of a new data structure and algorithm implementation focusing on the parallelism for finite volumes mesh adaptation for such stiff systems, which allows reaching a very satisfactory level of efficiency on shared memory architectures such as the Intel Xeon E5-2680 and the Xeon Phi (MIC) 5110P. This is conducted through a complete new coding of ROCK4 and RADAU5 solvers associated with the use of the TBB library which seems to be a good alternative in many cases to the OPENMP API, offering a different way of thinking the parallelism. In particular, it is shown that, depending on the complexity of the problem, the conclusions may differ. However, for the models involving some level of modeling complexity such as (BZ) or (STROKE), high scalability is achieved for several reasons. First, the reaction step constitutes an important part of the simulation time and it scales very well; second, the diffusion and adaptation scale correctly; third, simultaneous multithreading improves the performance, in particular on many-core architecture.

Beyond the fact that the proposed strategy is evaluated on a building block of more complex simulation and should allow tackling more complex systems, future work should address the following aspects:

- The extension to arbitrary domains, possibly by some immersed boundary method [51]. Let us recall that the domain can be as complicated as a human brain [31].
- The diffusion operator could involve some level of coupling of the variable in models involving detailed transport for example in multicomponent reactive flows and flames [17, 34]. In such a situation, the scope of the ROCK4 solver has to be extended and the data structure should be optimized.
- The implementation of hybrid parallelism including distributed memory architecture.

In terms of the last item, even though we have only covered shared memory considerations, we believe the scope of this work is relevant to distributed memory codes as well. Achieving strong scaling is particularly challenging for multiresolution and AMR codes, because of load imbalance

and fine-grain communications. For some applications, shared memory parallelism with dynamic load balancing may allow distributing work evenly between the CPU cores within an MPI rank, thereby helping reduce total imbalance. We expect shared memory parallelism to only gain in relevance for complex and imbalanced MPI applications.

Acknowledgments

This development was funded by grants from the ANR project (French National Research Agency - ANR Blancs) S echelles (2009-2013) and from the French *AMIES Peps* program. Part of this work was conducted by the Exascale Computing Research laboratory, thanks to the support of CEA, GENCI, Intel, UVSQ. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of CEA, GENCI, Intel or UVSQ.

A A compact data structure for the sparse matrices associated to diffusion

When the coefficients $\varepsilon_i(x)$ in (1) are constant, it is possible to reduce the amount of memory necessary to store the matrix associated to diffusion, and lower the memory bandwidth requirements of the program. For that, we remark that the coefficients fall into 3 types, describing either a link between 2 same level nodes, a link to upper level or to lower level. Given the types of the unknowns linked to a given node, and the level l of the node, one immediately knows the coefficient associated up to a factor $1/2^l$. We use the data structure shown at Figure 9: the array P contains pointers to the beginning of lines in the matrix ; the array K contains lines descriptions. The first integer of a line description is a 32 bits integer word, which contains the level l of the node, and the number of connections to level l , $l + 1$ and $l - 1$. The following integers give the connections. Figure 10 shows a typical line description, followed by the corresponding ranks of the nodes.

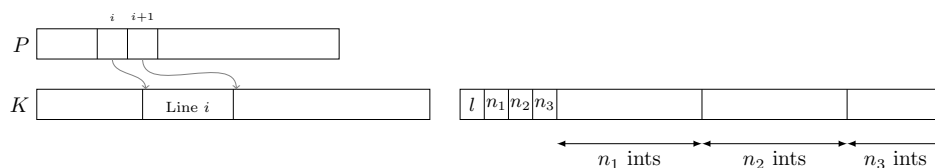


Figure 9: Data structure. The right part shows the structure of a line.

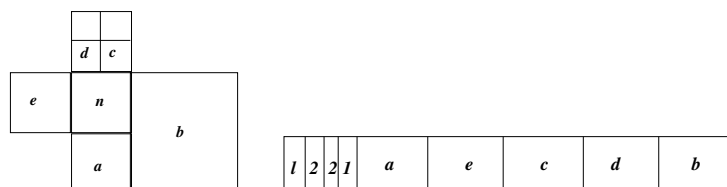


Figure 10: Example of line description.

Comparison with CSR matrices We compare the performance of the solution of $D_{\Delta t}$ using ROCK4 with the classical CSR data structure and our compact storage, applied to the (BZ) problem (Figure 11). In 2 dimensional computations, matrices are quite small, and the gain using compact storage is negligible; unlike dimension 3 where the compact storage structure is clearly beneficial.

	SB, $d = 2, J = 10$	SB, $d = 3, J = 10$	XPhi, $d = 2, J = 10$	XPhi $d = 3, J = 9$
Ratio	0.98	0.67	0.97	0.86

Figure 11: Ratio= cpu time with compact storage / cpu time with csr storage.

SB: Dual-socket Sandy Bridge processor using 16 cores and 32 threads, XPhi: Xeon Phi coprocessor using 60 cores and 240 threads. d : spatial dimension, J : maximum level in grids.

References

- [1] A. Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput.*, 23(6):2041–2054 (electronic), 2002.
- [2] S.B. Baden, N.P. Chrisochoides, D.B. Gannon, and M.L. Norman, editors. *Structured adaptive mesh refinement (SAMR) grid methods*, volume 117 of *The IMA Volumes in Mathematics and its Applications*. Springer-Verlag, New York, 2000. Papers from the IMA Workshop held at the University of Minnesota, Minneapolis, MN, March 12–13, 1997.
- [3] D.S. Balsara and C.D. Norton. Highly parallel structured adaptive mesh refinement using parallel language-based approaches. *Parallel Computing*, 27(12):37–70, January 2001.
- [4] J. Bell, M. Berger, J. Saltzman, and M. Welcome. Three-dimensional adaptive mesh refinement for hyperbolic conservation laws. *SIAM J. Sci. Comput.*, 15(1):127–138, 1994.
- [5] R.D. Blumofe, C.F. Joerg, B.C. Kuszmaul, C.E. Leiserson, K.H. Randall, and Y. Zhou. Cilk: An efficient multithreaded runtime system. *Journal of parallel and distributed computing*, 37(1):55–69, 1996.
- [6] R.D. Blumofe and C.E. Leiserson. Scheduling multithreaded computations by work stealing. *J. ACM*, 46(5):720–748, September 1999.
- [7] S. Borkar. Design challenges of technology scaling. *Micro, IEEE*, 19(4):23–29, Jul 1999.
- [8] S. Borkar and A.A. Chien. The future of microprocessors. *Commun. ACM*, 54(5):67–77, May 2011.
- [9] K. Brix, S.S. Melian, S. Müller, and G. Schieffer. Parallelisation of multiscale-based grid adaptation using space-filling curves. In *ESAIM: Proceedings*, volume 29, pages 108–129. EDP Sciences, 2009.
- [10] C. Burstedde, L. Wilcox, and O. Ghattas. p4est: Scalable algorithms for parallel adaptive mesh refinement on forests of octrees. *SIAM Journal on Scientific Computing*, 33(3):1103–1133, January 2011.
- [11] A. Cohen. *Wavelet methods in numerical analysis*, volume 7. Elsevier, Amsterdam, 2000.
- [12] A. Cohen, S.M. Kaber, S. Müller, and M. Postel. Fully adaptive multiresolution finite volume schemes for conservation laws. *Mathematics of Computation*, 72:183–225, 2003.
- [13] W. Crutchfield and M.L. Welcome. Object-oriented implementation of adaptive mesh refinement algorithms. *Journal of Scientific Programming*, pages 145–156, 1993.
- [14] W. Dahmen, T. Gotzen, S.S. Melian-Flamand, and S. Müller. *Numerical simulation of cooling gas injection using adaptive multiscale techniques*. Inst. für Geometrie und Praktische Mathematik, 2011.

- [15] R. Deiterding. A generic framework for blockstructured adaptive mesh refinement in object-oriented c++. Technical report, available at <http://amroc.sourceforge.net>, 2003.
- [16] R. Deiterding. Block-structured adaptive mesh refinement—theory, implementation and application. In *Summer School on Multiresolution and Adaptive Mesh Refinement Methods*, volume 34 of *ESAIM Proc.*, pages 97–150. EDP Sci., Les Ulis, 2011.
- [17] S. Descombes, M. Duarte, T. Dumont, F. Laurent, V. Louvet, and M. Massot. Analysis of operator splitting in the nonasymptotic regime for nonlinear reaction-diffusion equations. Application to the dynamics of premixed flames. *SIAM J. Numer. Anal.*, 52(3):1311–1334, 2014.
- [18] S. Descombes, M. Duarte, T. Dumont, V. Louvet, and M. Massot. Adaptive time splitting method for multi-scale evolutionary partial differential equations. *Confluentes Math.*, 3(3):413–443, 2011.
- [19] S. Descombes and T. Dumont. Numerical simulation of a stroke: Computational problems and methodology. *Progress in Biophysics & Molecular Biology*, 97(1):40–53, 2008.
- [20] S. Descombes, T. Dumont, V. Louvet, and M. Massot. On the local and global errors of splitting approximations of reaction-diffusion equations with high spatial gradients. *Int. J. Comput. Math.*, 84(6):749–765, 2007.
- [21] S. Descombes, T. Dumont, and M. Massot. Operator splitting for stiff nonlinear reaction-diffusion systems: order reduction and application to spiral waves. In *Patterns and waves (Saint Petersburg, 2002)*, pages 386–482. AkademPrint, St. Petersburg, 2003.
- [22] S. Descombes and M. Massot. Operator splitting for nonlinear reaction-diffusion systems with an entropic structure: singular perturbation and order reduction. *Numer. Math.*, 97(4):667–698, 2004.
- [23] J. Dreher and R. Grauer. Racocon: A parallel mesh-adaptive framework for hyperbolic conservation laws. *Parallel Computing*, 31(89):913–932, August 2005.
- [24] M.-A. Dronne, J.-P. Boissel, and E. Grenier. A mathematical model of ion movements in grey matter during a stroke. *J. of Theoretical Biology*, 240(4):599–615, 2006.
- [25] M. Duarte. *Adaptive numerical methods in time and space for the simulation of multi-scale reaction fronts*. PhD thesis, Ecole Centrale Paris, December 2011. <https://tel.archives-ouvertes.fr/tel-00667857/en>.
- [26] M. Duarte, Z. Bonaventura, M. Massot, and A. Bourdon. A numerical strategy to discretize and solve the Poisson equation on dynamically adapted multiresolution grids for time-dependent streamer discharge simulations. *J. Comput. Phys.*, 289:129–148, 2015.
- [27] M. Duarte, Z. Bonaventura, M. Massot, A. Bourdon, S. Descombes, and T. Dumont. A new numerical strategy with space-time adaptivity and error control for multi-scale streamer discharge simulations. *J. Comput. Phys.*, 231(3):1002–1019, 2012.
- [28] M. Duarte, S. Descombes, C. Tenaud, S. Candel, and M. Massot. Time-space adaptive numerical methods for the simulation of combustion fronts. *Combustion and Flame*, 160(6):1083–1101, 2013.
- [29] M. Duarte, M. Massot, S. Descombes, C. Tenaud, T. Dumont, V. Louvet, and F. Laurent. New resolution strategy for multi-scale reaction waves using time operator splitting and space adaptive multiresolution: application to human ischemic stroke. In *Summer School on Multiresolution and Adaptive Mesh Refinement Methods*, volume 34 of *ESAIM Proc.*, pages 277–290. EDP Sci., Les Ulis, 2011.

- [30] M. Duarte, M. Massot, S. Descombes, C. Tenaud, T. Dumont, V. Louvet, and F. Laurent. New resolution strategy for multiscale reaction waves using time operator splitting, space adaptive multiresolution, and dedicated high order implicit/explicit time integrators. *SIAM J. Sci. Comput.*, 34(1):A76–A104, 2012.
- [31] T. Dumont, M. Duarte, S. Descombes, M.-A. Dronne, M. Massot, and V. Louvet. Simulation of human ischemic stroke in realistic 3D geometry. *Commun. Nonlinear Sci. Numer. Simul.*, 18(6):1539–1557, 2013.
- [32] W. Eckhardt and T. Weinzierl. A Blocking Strategy on Multicore Architectures for Dynamically Adaptive PDE Solvers. In *Parallel Processing and Applied Mathematics*, volume 6067, pages 567–575. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [33] T. Gautier, J.V.F. Lima, N. Maillard, and B. Raffin. XKaapi: A Runtime System for Data-Flow Task Programming on Heterogeneous Architectures. In *Parallel Distributed Processing (IPDPS), 2013 IEEE 27th International Symposium on*, pages 1299–1308, May 2013.
- [34] V. Giovangigli and M. Massot. Asymptotic stability of equilibrium states for multicomponent reactive flows. *Math. Models Methods Appl. Sci.*, 8(2):251–297, 1998.
- [35] P. Gray and S.K. Scott. *Chemical Oscillations and Instabilities: Non-linear Chemical Kinetics*, volume 21 of *International Series of Monographs on Chemistry*. Oxford University Press, 1994.
- [36] E. Hairer. Fortran and matlab codes. <http://www.unige.ch/~hairer/software.html>.
- [37] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations I*, volume 8 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1993. Nonstiff problems.
- [38] E. Hairer and G. Wanner. *Solving ordinary differential equations II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1996. Stiff and differential-algebraic problems.
- [39] B. Hejazialhosseini, D. Rossinelli, M. Bergdorf, and P. Koumoutsakos. High order finite volume methods on wavelet-adapted grids with local time-stepping on multicore architectures for the simulation of shock-bubble interactions. *J. Comput. Phys.*, 229(22):8364–8383, 2010.
- [40] Intel Corp. Intel Threading Building Blocks. <https://www.threadingbuildingblocks.org/>.
- [41] W. Jahnke, W.E. Skaggs, and A.T. Winfree. Chemical vortex dynamics in the belousov-zhabotinskii reaction and in the two-variable oregonator model. *The Journal of Physical Chemistry*, 93(2):740–749, 1989.
- [42] H. Ji, F.-S. Lien, and E. Yee. A new adaptive mesh refinement data structure with an application to detonation. *Journal of Computational Physics*, 229(23):8981–8993, November 2010.
- [43] R. Keppens, Z. Meliani, A. J. van Marle, P. Delmont, A. Vlasis, and B. van der Holst. Parallel, grid-adaptive approaches for relativistic hydro and magnetohydrodynamics. *Journal of Computational Physics*, 231(3):718–744, February 2012.
- [44] P. MacNeice, K.M. Olson, C. Mobarry, R. de Fainchtein, and C. Packer. PARAMESH: A parallel adaptive mesh refinement community toolkit. *Computer Physics Communications*, 126(3):330–354, April 2000.
- [45] G.M. Morton. A computer oriented geodetic data base and a new technique in file sequencing. Technical report, IBM, Ottawa, Canada, 1966.

- [46] S. Müller. *Adaptive multiscale schemes for conservation laws*, volume 27. Springer-Verlag, Heidelberg, 2003.
- [47] Netlib. Compressed row storage (crs). http://netlib.org/linalg/html_templates/node91.html.
- [48] R.M. Noyes, R. Field, and E. Koros. Oscillations in chemical systems. i. detailed mechanism in a system showing temporal oscillations. *Journal of the American Chemical Society*, 94(4):1394–1395, 1972.
- [49] L. Oliker and R. Biswas. PLUM : Parallel Load Balancing for Adaptive Unstructured Meshes. *Journal of Parallel and Distributed Computing*, 52(2):150–177, August 1998.
- [50] OpenMP Architecture Review Board. OpenMP Application Program Interface. Version 3.1., July 2011. available at: <http://www.openmp.org>.
- [51] C.S. Peskin. The immersed boundary method. *Acta Numer.*, 11:479–517, 2002.
- [52] J. Reinders. *Intel Threading Building Blocks*. O’Reilly & Associates, Inc., Sebastopol, CA, USA, first edition, 2007.
- [53] C.A. Rendleman, V.E. Beckner, M. Lijewski, W. Crutchfield, and J.B. Bell. Parallelization of structured, hierarchical adaptive mesh refinement algorithms. *Computing and Visualization in Science*, 3(3):147–157, August 2000.
- [54] D. Rossinelli, B. Hejazialhosseini, M. Bergdorf, and P. Koumoutsakos. Wavelet-adaptive solvers on multi-core architectures for the simulation of complex systems. *Concurrency Computat.: Pract. Exper.*, 23(2):172–186, 2011.
- [55] O. Roussel, K. Schneider, A. Tsigulin, and H. Bockhorn. A conservative fully adaptive multiresolution algorithm for parabolic PDEs. *J. Comput. Phys.*, 188(2):493–523, 2003.
- [56] H. Sagan. *Space-filling curves*. Universitext. Springer-Verlag, New York, 1994.
- [57] M. Schreiber, T. Weinzierl, and H.-J. Bungartz. Cluster Optimization and Parallelization of Simulations with Dynamically Adaptive Grids. In *Euro-Par 2013 Parallel Processing*, volume 8097, pages 484–496. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [58] G. Strang. On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.*, 5:506–517, 1968.
- [59] C. Tenaud and M. Duarte. Tutorials on adaptive multiresolution for mesh refinement applied to fluid dynamics and reactive media problems. In *Summer School on Multiresolution and Adaptive Mesh Refinement Methods*, volume 34 of *ESAIM Proc.*, pages 184–239. EDP Sci., Les Ulis, 2011.
- [60] R. Teyssier. Cosmological hydrodynamics with adaptive mesh refinement: A new high resolution code called RAMSES. *Astronomy and Astrophysics*, 385(1):337–364, April 2002.
- [61] A.I. Volpert, V.A. Volpert, and V.A. Volpert. *Traveling wave solutions of parabolic systems*. American Mathematical Society, Providence, RI, 1994. Translated from the Russian manuscript by James F. Heyda.
- [62] T. Weinzierl and M. Mehl. Peano–A traversal and storage scheme for octree-like adaptive cartesian multiscale grids. *SIAM Journal on Scientific Computing*, 33(5):2732–2760, January 2011.

- [63] S. Williams, A. Waterman, and D. Patterson. Roofline: An insightful visual performance model for multicore architectures. *Commun. ACM*, 52(4):65–76, April 2009.
- [64] A.M. Wissink, R.D. Hornung, S.R. Kohn, S.S. Smith, and N. Elliott. Large Scale Parallel Structured AMR Calculations Using the SAMRAI Framework. In *Supercomputing, ACM/IEEE 2001 Conference*, pages 22–22, November 2001.
- [65] Y.B. Zel'dovich, G.I. Barenblatt, V.B. Librovich, and G.M. Makhviladze. *The mathematical theory of combustion and explosions*. Consultants Bureau [Plenum], New York, 1985. Translated from the Russian by Donald H. McNeill.