



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. SMAI Journal of Computational Mathematics, 2017, pp.1-23. 10.5802/smai-jcm.19 . hal-01148617v3

HAL Id: hal-01148617

<https://hal.science/hal-01148617v3>

Submitted on 15 Oct 2016

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¶

October 15, 2016

Abstract

A new solver featuring time-space adaptation and error control has been recently introduced to tackle the numerical solution of stiff reaction-diffusion systems. Based on operator splitting, finite volume adaptive multiresolution and high order time integrators with specific stability properties for each operator, this strategy yields high computational efficiency for large multi-dimensional computations on standard architectures such as powerful workstations. However, the data structure of the original implementation, based on trees of pointers, provides limited opportunities for efficiency enhancements, while posing serious challenges in terms of parallel programming and load balancing. The present contribution proposes a new implementation of the whole set of numerical methods including RADAU5 and ROCK4, relying on a fully different data structure together with the use of a specific library, TBB, 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, stiff reaction-diffusion equations

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

1 Introduction

Stiff reaction-diffusion systems model various complex phenomena across different disciplines such as combustion, atmospheric sciences, plasma physics or biomedical engineering. Such models can also involve the dynamics of moving fronts, usually very localized in space (see, *e.g.*, [23] and references therein), and potentially entail a large number of unknowns. A general reaction-diffusion system can be written as follows, for $i = 1, 2, \dots, m$:

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

*Université Côte d'Azur, CNRS, Inria, LJAD, France (stephane.descombes@unice.fr).

†CCSE, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd. MS 50A-1148, Berkeley, CA 94720, USA and CD-adapco, 200 Shepherds Bush Road, London W6 7NL, UK (max.duarte@cd-adapco.com).

‡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).

§Intel, Les Montalets, 2 rue de Paris, 92196 Meudon, France and Exascale Computing Research, Campus Teratec, 2 rue de la Piquetterie, 91680 Bruyères-le-Châtel, France (thomas.guillet@intel.com).

¶CNRS UPR 288, Laboratoire EM2C, CentraleSupélec, Fédération de Mathématiques de l'École Centrale Paris, CNRS FR 3487, Grande Voie des Vignes, 92295 Chatenay-Malabry Cedex, France (marc.massot@centralesupelec.fr).

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.

Two major difficulties need to be addressed when solving numerically this kind of problem in two and three dimensions. First, a large spectrum of temporal scales in the nonlinear source terms yields highly stiff equations¹. Systems of stiff ordinary differential equations impose the use of numerical methods with specific stability properties in order to achieve accuracy and stability within reasonable memory and computing costs [36]. 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 [17].

We have recently introduced in [27] a tailored numerical strategy to cope with the latter difficulties using reasonable computing resources, that is, on a sufficiently powerful workstation, possibly exploiting shared-memory parallelism [26]. It relies on time operator splitting² with one-step, high order integration schemes, namely, RADAU5³ [36] and ROCK4⁴ [1] for the time integration of the reaction and diffusion sub-systems, respectively. The discretized equations are solved on a dynamically adapted grid generated by multiresolution analysis (MRA) in a finite volume multidimensional framework in the spirit of the original work of Harten [37] and then Cohen et al. and Müller [10, 46, 57]. MRA is based on a wavelet decomposition and a multiresolution transform, yielding both highly compressed representations for problems displaying localized fronts as well as a compression error control with respect to the solution on the full grid. Let us emphasize that this property is a major difference with an AMR (Adaptive Mesh Refinement) strategy where the refinement criterion rather relies on heuristics. This MRA-operator splitting numerical strategy was implemented in the MBARETE code [23], using a tree-structured data with pointers and recursive navigation. However, the latter implementation shows serious 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 is thus needed to achieve more efficient MRA-operator splitting implementations.

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 MRA, 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 as it provides a number of advantages well-adapted to MRA applications. Contrary to uniform Cartesian grids for which arrays can be generally accessed following regular patterns within long loops, adaptive meshing in general, that is both multiresolution and AMR codes, relies on fine-grained and dynamic data structures. The corresponding algorithms can thus have intricate data dependencies, especially for complex operations such as those associated with mesh adaptation. Hence, 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. As a consequence of the complex layouts of the resulting adapted meshes, an efficient parallel implementation requires non-trivial balancing of the computations across all available computing cores. Moreover, the load balancing needs to be frequently updated as the mesh structure evolves during the simulation. Actually, many existing mesh refinement algorithms and libraries address this issue in distributed-memory settings for both patch- and cell-based AMR [11, 50, 2, 53, 44, 64, 60, 13, 9, 14, 43, 22, 30].

¹The 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 or in the presence of strong attracting slow manifolds [36].

²High performance computing can be achieved by choosing dedicated schemes for each split sub-system [27, 28]. Operator splitting schemes also exhibit a large data-driven parallelism, and their mathematical analysis is well-established for relatively large splitting time steps even when stiffness is present [19, 17, 15].

³An implicit, fifth order Runge-Kutta scheme with A - and L -stability properties and time step adaptation.

⁴A stabilized, explicit Runge-Kutta method of order four and time step adaptation.

However, even if MRA leads to compression error control, the mesh adaptation relies on a multiresolution transform that involves recursive navigations throughout the data sets [7]. The latter makes it much more difficult to parallelize than classic AMR approaches. Some advances have been conducted recently into that direction but in a different framework [31, 42, 47].

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 [11, 4], several authors have proposed multithreaded implementations for adaptive gridding techniques [3, 20] with some focusing on specialized grid structures for multi-core processors [29, 58]. 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 and recursive navigation in the data set, such as MRA applications.

We have thus opted for a shared-memory parallel approach based on work-stealing, relying on the Intel Threading Building Blocks (TBB) library [52, 39]. While other runtime libraries supporting task parallelism are currently available, such as the OpenMP application programming interface (API), and could have been used for our purposes with equivalent functionality, they would have required a much more important programming effort. Shared-memory parallelism was hence introduced using the task-based Intel Threading Building Blocks runtime library. The latter turns out to be particularly well suited to the multiresolution algorithm in order to cope with nested parallel regions, which are directly related to the MRA-splitting strategy.

Additionally, a completely new implementation of the original numerical strategy of MBARETE code has been conducted, 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 work-stealing techniques and in particular the use of TBB have already been shown in [38, 54] for multiresolution schemes applied to non-stiff time-dependent PDEs, that is, using explicit time integration schemes⁵.

In order to assess the new implementation and code performance we have chosen three reaction-diffusion models with increasing complexity in two ($d = 2$) and three ($d = 3$) dimensions:

I. **NAGUMO.** A bistable, Nagumo-type reaction-diffusion equation⁶. Here we have $m = 1$ and $f_1(u_1) = ku_1^2(1 - u_1)$ in (1). We consider $k = 10$, $\varepsilon_1 = 0.1$, and an initial condition satisfying $0 \leq u_1^0(x) \leq 1$ everywhere⁷.

II. **BZ.** The Belusov-Zhabotinsky reaction [49, 40]. This is a system of three ($m = 3$) equations, where the reaction term is given by⁸ $f_1(u_1, u_2, u_3) = 10^5 (-0.02 u_1 - u_2 u_1 + 1.6 u_3)$, $f_2(u_1, u_2, u_3) = 10^2 (u_2 - u_2^2 - u_1(u_2 - 0.02))$, and $f_3(u_1, u_2, u_3) = u_2 - u_3$. For the diffusion part, we set $\varepsilon_1(x) = \varepsilon_2(x) = 2.5 \times 10^{-3}$ and $\varepsilon_3(x) = 1.5 \times 10^{-3}$ in (1)⁹.

⁵These authors developed a data structure based on the definition of *wavelet blocks*, specifically a tree of wavelet blocks, where each block contains a predefined number of cells at the same grid-level.

⁶A bistable case of type A according to [61], which is a limit case of the Nagumo function, similar to what is found in combustion models [65] or in nonlinear chemical dynamics when studying traveling waves due, for instance, to chemical reactions with cubic auto-catalysis [33].

⁷In 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.

⁸This set describes a chemical reaction between $HBrO_2$, $Ce(IV)$ and Br^- , also known as the Oregonator problem.

⁹The reaction term is stiff. The system of ordinary differential equations has a limit cycle along which the amplitude of the eigenvalues of the Jacobian attains values of the order of 10^5 . As a comparison, when computing in

III. **STROKE.** An ischemic stroke model [21, 16]. This is a system of 21 ($m = 21$) equations with a very stiff reaction term¹⁰. Its computation is performed by a quite complex and computationally expensive program of about 400 instructions, containing many log-function evaluations.

The paper is organized as follows. After a brief description of the key aspects of the numerical strategy in Section 2, we focus in Section 3 on the task-based parallel implementation developed in this work. A performance analysis is then presented in Section 4, to end with some concluding remarks and future developments related to this work.

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 standard centered second-order discretization on a Cartesian mesh, A_ε is given by a matrix with five (resp., seven) non-zero elements per line in two (resp., three) dimensions. Following [27] a dedicated splitting solver, briefly described in what follows, is implemented to advance system (2) in time.

2.1 Dedicated splitting solver

Considering independently the two sub-problems coming from (2):

$$\left. \begin{aligned} \frac{dV}{dt} &= A_\varepsilon V, \\ V(0) &= V_0, \end{aligned} \right\} \quad \left. \begin{aligned} \frac{dW}{dt} &= F(W), \\ W(0) &= W_0, \end{aligned} \right\}$$

we denote by $D_{\Delta t}V_0$ and $R_{\Delta t}W_0$ the solution of the first and second sub-problem, respectively, after a given *splitting time step* Δt . Assuming in general stiff reaction terms, we consider without any loss of generality the second-order Strang scheme¹¹ [59], $U_{n+1} = R_{\Delta t/2} \circ D_{\Delta t} \circ R_{\Delta t/2} U_n$.

The RADAU5 method is used to solve each of the independent initial value problems coming from the reaction term and defined at each grid-point, during the *Reaction Steps*, $R_{\Delta t/2}$; whereas the *Diffusion Step*, $D_{\Delta t}$, uses the explicit ROCK4 method to globally advance the discrete diffusion sub-problem. Notice that with the latter choice of solver, one only needs to evaluate matrix-vector products without any need of preconditioning nor factorizing the diffusion matrices when the mesh has changed. In this work we have chosen splitting time steps, Δt , equal to 10^{-2} , 10^{-3} , and 1 for the NAGUMO, BZ, and STROKE models, respectively.

Finally, in terms of 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. The situation is different

$\Omega = [0, 1]^d$ with a spatial discretization step of $h = 1/1024$, 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.

¹⁰The right-hand side, $f(u)$, incorporates an important amount of biological knowledge about ion channels (see [28] for the model here considered). A numerical computation of the eigenvalues of the Jacobian near a stable equilibrium state, $f(u) = 0$, shows real parts in the interval of $[-10^8, 0[$. 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 a given threshold. Gates are modeled through sigmoid functions, closely approximating a Heaviside function. Simulations show steep spatial gradients, as well.

¹¹It has been shown that better accuracies are expected by ending the splitting scheme with the substep involving the fastest time scales [18, 19, 15]. Both theoretical and numerical results show that this method performs very well for stiff reaction-diffusion systems [23, 27, 28].

for the $D_{\Delta t}$ -step, which involves the whole computational domain. A simple and straightforward parallelism consists in parallelizing the solution of the m independent parabolic linear equations in (1) [26]. However, an alternative that is not limited by the number of equations should be considered to enhance the parallel performance of the diffusion solver. The latter is much simpler to achieve in the case of an explicit scheme like ROCK4, since matrix-vector products are easier to parallelize than linear solvers.

2.2 Adaptive multiresolution method

For the kind of problems we are interested in, featuring propagating localized fronts, an adaptive meshing technique that reduces the number of grid-nodes over near-equilibrium zones can significantly reduce the CPU time¹² together with the required computer memory. In the following we describe some basic features of the multiresolution method introduced in [10], and implemented in [27] with the aforementioned dedicated splitting solver for reaction-diffusion problems¹³.

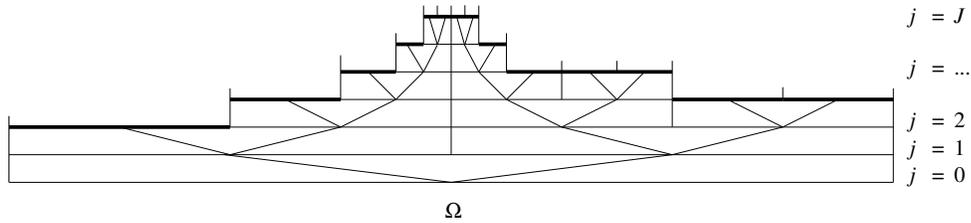


Figure 1: One-dimensional graded tree composed of nested grids. The resulting adapted mesh is given by the leaves of the tree, indicated in bold.

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 correspond, respectively, to the whole computational domain Ω and the mesh at the finest spatial resolution. An adapted spatial representation can be thus achieved by trimming the set of nested grids throughout the tree, as shown in Figure 1 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 only performed on the adapted mesh, that is on the leaf nodes, the solution is updated and stored at all nodes of the tree. The set of values at nodes of level j is denoted as \mathbf{U}_j . Data at different levels of discretization are related by two inter-level transformations defined by 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 computes the average of the 2^d nested nodes at the successive finer level j . The prediction operator, P_j^{j-1} , maps \mathbf{U}_{j-1} to an approximation $\hat{\mathbf{U}}_j$ of \mathbf{U}_j ; that is, the value of a given node at level j is approximated based on the surrounding coarser nodes at level $j-1$.

Grid adaptation is performed based on local estimators 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

$$d_{j,k} = u_{j,k} - \hat{u}_{j,k} = u_{j,k} - P_j^{j-1} \circ P_{j-1}^j u_{j,k}, \quad (3)$$

¹²Since this is a key aspect, a numerical experimentation on uniform meshes is provided in Appendix A for the three aforementioned models.

¹³More details on this particular multiresolution implementation can be found in [23].

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. From the theoretical point of view, the details correspond to the coefficients related to a given discrete function when represented on a wavelet basis. This wavelet decomposition is a general procedure, independent of any physical particularity of the problem. 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 [23]. Conversely, nodes whose details exceed ϵ_j must be refined. Grid adaptation is hence driven by the local value of the details, while the tolerance parameter defines the level of approximation errors introduced by the data compression [10]. This is the main difference between MRA and AMR strategies. In particular, while enabling compression error control, the multiresolution analysis also leads to recursive navigation in the data set.

In this work the approximated values, \hat{U}_j , generated with the prediction operator, are obtained using centered polynomial interpolations of accuracy order, $o = 2l + 1$, computed with the l nearest neighboring cells in each direction including the diagonals in multi-dimensional configurations. For a one-dimensional configuration with $l = 1$, the prediction operator is then given by [10]

$$\hat{u}_{j+1,2k} = u_{j,k} + \frac{1}{8}(u_{j,k-1} - u_{j,k+1}), \quad \hat{u}_{j+1,2k+1} = u_{j,k} + \frac{1}{8}(u_{j,k+1} - u_{j,k-1}). \quad (4)$$

Extension to multi-dimensional Cartesian grids is easily obtained by a tensor product of the one-dimensional operator (4) (see, *e.g.*, [55]). The interpolation stencil is thus given by $(2l + 1)^d$ cells at the coarser level. Here we will only consider third-order interpolations with $l = 1$. Notice that these local interpolation stencils must be available throughout the tree-structure during the simulation. A tree-structure that complies with such a constraint is called a *graded tree* [10]. Hence, a graded tree-structure must be always guaranteed after the refining and coarsening operations.

3 Task-based parallelism for MRA-splitting applications

We have introduced shared-memory parallelism using the task-based Intel Threading Building Blocks runtime library [52, 39]. It turns out to be particularly well suited to the multiresolution algorithm in order to cope with nested parallel regions, which are directly related to the MRA-splitting strategy. Other than TBB, several runtime libraries supporting task parallelism are currently available, such as the OpenMP API [51], XKaapi [32], among others. They could potentially have been used for our purposes with equivalent functionality, but requiring a much more important programming effort. A discussion clarifying the reasons why we have opted for TBB, in particular over OpenMP API, is provided in Appendix B.

3.1 Parallel implementation of the multiresolution algorithm

One of the main constraints on parallel implementations of the multiresolution algorithm has to do with the data locality required to efficiently perform the inter-level operations throughout the tree data structure (quadtrees or octrees in two or three dimensions, respectively). More specifically, the prediction operator needs to access all nodes in the interpolation stencil, meaning that for any given node, one must access the parent-node at the successive coarser level and all its neighboring nodes, as shown in Figure 2 in the case of quadtrees. On the other hand, the projection operator needs to access the child-nodes of any given node that is not a leaf. Notice that it is very unlikely to achieve data locality in standard pointer-based tree data structures, like the one considered in the MBARETE code. In the latter case nodes are generated and linked recursively and even though pointers to the neighboring nodes are introduced [23], nodes belonging to a given interpolation stencil will hardly be stored contiguously in memory. Additionally, other the projection and prediction operators, the refinement and coarsening operations, also performed on trees, should be executed in parallel as well, in order to maximize parallel coverage and minimize the impact of Amdahl's law.

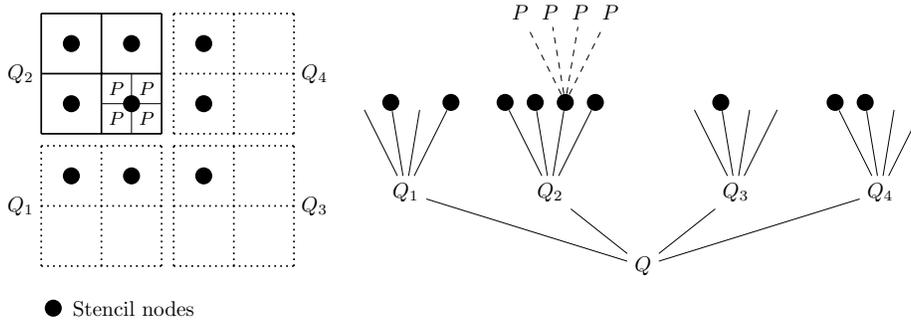


Figure 2: Prediction stencil nodes to compute values at nodes P within a given quadtree. The stencil comprises nodes belonging to four different quadtrees: Q_i , $i = 1, 2, 3, 4$.

The data structure implemented in this work to support quadtrees and octrees, as well as the computation of the key multiresolution operations are detailed in what follows.

3.1.1 Data structure and implementation of trees

Without loss of generality we consider again the computational domain $\Omega = [0, 1]^d$, $d = 2, 3$. The dyadic refinement of cells allows us to define the coordinates of a given node, (x_1, \dots, x_i) , $i = 1, d$, using the binary numeral system. That is, for each dimension i , the x_i -coordinate of a given node P at the grid-level j is given by $x_i = 0.x_{i,1}x_{i,2} \dots x_{i,j-1}x_{i,j}$, where $x_{i,j-1}$ is equal to 0 or 1 depending on which of the two possible nodes at grid-level $j - 1$ generated the node P ; and the same follows for $x_{i,j-2}$, up to $x_{i,1}$. Using this convention one can easily define a space-filling curve [56] throughout all nodes of the tree; a powerful tool already employed in multiresolution [8, 12] and AMR [9, 62] applications. In our particular case we use a *Morton order* space-filling curve [45], also known as *Z-order* or *Morton code*. For instance, in two dimensions, for a given node $(x_1 = 0.x_{1,1}x_{1,2} \dots x_{1,j}, x_2 = 0.x_{2,1}x_{2,2} \dots x_{2,j})$, its Morton abscissa is constructed by alternating the digits of x_1 and x_2 :

$$s = x_{1,1}x_{2,1}x_{1,2}x_{2,2} \dots x_{1,j}x_{2,j};$$

and the same follows in three dimensions. In this way each node of the tree is uniquely defined by an integer. For each node of the tree we can then store in a 64-bit integer, its abscissa s and its corresponding grid-level j . In our implementation the first 48 bits are intended to contain the abscissa and the following four, the grid-level, leaving the remaining bits available for tagging purposes during the multiresolution operations, as shown in Figure 3. In the case of three-dimensional problems, the latter choice allows one to encode nodes of trees with up to 16 grid-levels.



Figure 3: Representation of a node using a 64-bit integer.

This representation of nodes is similar to the so-called CSAMR data structure introduced in [41], which then uses hash tables for efficient node lookup. In our implementation we also employ hash tables, but over clustered nodes, in what follows denoted as *blocks*, which enables further control on lookup granularity, and eases the operations for shared-memory parallelism. Taking into

account that $r = 0.s$ belongs to $[0, 1[$, we partition $[0, 1[$ into intervals. The data structure is thus given by *blocks* and *collections of blocks*. A block is then a structure that contains data related to a given interval, namely, the interval bounds: s_{\min} and s_{\max} ; a vector containing all nodes such that s belongs to $[s_{\min}, s_{\max}[$; and the size of the vector. A collection of blocks is given by pointers to the blocks, ordered according to their s_{\min} values. In practice, based on preliminary tests, we operate with two collections of blocks, one for the leaf nodes and another one for the non-leaf nodes.

Abscissas are not necessary ordered inside blocks, and binary search is used to locate the block where the abscissa of a given node is possibly stored, followed by a sequential search within that block. In our implementation the latter strategy has proven to be the most efficient, in particular compared with one based on a fully ordered storage. A system of caches keeps track of the last blocks accessed, effectively decreasing the number of binary searches. These caches are local to TBB tasks.

Additionally, five auxiliary operations are implemented to properly handle this data structure. Two of them aim at maintaining the size of the blocks within a predefined range, by splitting into two a relatively large block or merging two neighboring blocks of relatively small size. The complexity of the search is hence of $O(\log_2 N_b)$, where N_b is the number of blocks in the collection. The next two operations are used to enlarge or reduce the size of the vectors contained in the blocks. These are performed by allocating a new vector of suitable size, copying the content of the original vector, and finally, deleting the original vector. The fifth operation is implemented as a garbage collector, for instance, to effectively delete previously tagged nodes. Notice that all these five operations are such that can be performed in parallel, by using the `parallel_for` structure of the TBB library while looping over blocks.

Finally, the refinement process entails adding the abscissas of the new nodes at the end of the vector of the block that contains the abscissa of the original node. On the contrary, the coarsening process is performed using a *lazy* approach; that is, all nodes that need to be deleted are first tagged, using the free space in their 64-bit representation, while in a second stage the garbage collector effectively remove them from their corresponding blocks. Further details are given in the following.

3.1.2 Parallel implementation of the key multiresolution operations

Grid adaptation in the multiresolution algorithm is essentially driven by the value of details (3), which in turn is computed using the projection and prediction operators. In practice, details are computed in two stages, as follows.

- I. **Projection.** Data is propagated from the leaf nodes to the root of the tree, an operation easily cast as a recursive computation. Based on preliminary tests, we came out with the following implementation. Identifying a relatively low, fully populated level, j_{\min} , two recursive computation are performed. First, for every node P at grid-level j_{\min} , data is recursively transferred from the corresponding leaves down to P . These are independent computations, hence performed in parallel by using recursive `parallel_for` constructions. In a second step, a sequential computation propagates recursively data from level j_{\min} to the root of the tree. In this work, j_{\min} was taken equal to 3 and 5, respectively, for three- and two-dimensional applications. It is worth noticing that often the sequential computational, although inexpensive, needs not be performed, as in general no grid adaptation is present at those low grid-levels.
- II. **Prediction and detail computation.** Details are computed for every node P according to (3), after applying the prediction operator by means of interpolation, in the following schematic manner.

```

for  $P \in \{\text{Nodes}\}$  do
    Define the list  $L$  of Nodes in the prediction stencil associated with  $P$  (see, e.g., Figure 2);
    Get values of Nodes in  $L$ ;

```

```

    Compute the predicted value and the detail;
    Store the detail.

```

```

end for

```

All these computations are easily implemented within a `parallel_for` structure, looping over blocks.

Once the details have been computed, the adapted grid is updated, also in parallel, in two stages by means of the refinement and coarsening procedures, detailed in what follows. Notice that in all cases a lazy strategy was again implemented.

- I. **Refinement.** Based on the detail values, nodes that need to be refined are first tagged in parallel. Then, all tagged cells are refined. The resulting tree will not necessarily be a graded one. Therefore, additional nodes must be tagged for refinement in order to guarantee a graded tree-structure. The process of tagging and refining is thus repeated until a graded tree-structure is achieved.

The task-based refinement operation is implemented in the following way. Each available task considers a given set of consecutive blocks, that is blocks of nodes within a given abscissa interval, say $[s_i, s_f]$. Whenever a node is created by refinement, it is directly stored in the corresponding block that contains its abscissa, as previously explained. However, if necessary, two new blocks are created by each task to temporarily store those new nodes whose abscissas are not contained in $[s_i, s_f]$ (one for abscissas smaller than s_i , and another one for those larger than or equal to s_f). These operations can be safely done in parallel because no task will be writing into any set or temporary block belonging to another task. Once all tagged nodes were refined, a new task-based operation starts, again over consecutive sets of blocks, to move the previously created nodes from the temporary blocks to the corresponding block actually containing its abscissa. The latter process can be also safely performed in parallel because it only considers read-only accesses.

- II. **Coarsening.** At this point, leaf nodes that are no longer necessary can be eliminated. The latter applies to leaf nodes whose parent-nodes were not tagged for refinement in the previous stage, and that can be deleted without compromising the graded tree-structure.

The task-based coarsening operation is implemented in the following way. Each available task considers a given set of consecutive blocks, and tags the leaf nodes that need to be deleted. Once the tagging phase is over, parent-nodes of tagged nodes are moved to the collection corresponding to leaf nodes, using the same task-based techniques described for the refinement. Finally, a parallel call to the garbage collector effectively deletes the tagged nodes. These three operations are repeated until no node deletion is needed. Again these computations are safely performed in parallel using `parallel_for` over blocks.

3.2 Parallel implementation of the splitting solver

Before considering the parallel implementation itself, we have to define the way of storing the unknowns of system (1) in memory. This set of unknowns needs to be solved in every leaf node of the tree, and there are basically two options to store them in memory, following either an *array of structures* or *structure of arrays* approach. The first option consists in storing all data related to a given node, in particular the unknowns, in a common structure, thus achieving node-wise memory contiguity. On the other hand, the second option consists in storing each unknown in its own array, thus achieving variable-wise memory contiguity.

In general, within a splitting framework, the first layout turns out to be the most appropriate for the reaction solver, as it computes local reaction rates using all the unknowns at the current node. The diffusion problem, on the other hand, is solved for each unknown independently, hence a structure of arrays would provide better data locality. Nevertheless, the reaction solver is highly

compute-intensive and therefore, largely insensitive to the memory layout, whereas the diffusion step displays very low compute-intensity. Consequently, we have opted to store each unknown in its own contiguous array. As we will later see, a large fraction of the runtime is spent in the reaction and diffusion steps; having such a memory layout proves then to be very relevant, especially when the variables are not coupled in the diffusion operator. The Morton order defines the order of the unknowns, which are stored in vectors. Additionally, the sparse matrices used during the diffusion steps can be stored according to a classical CSR storage [48]. However, an even more compact format was developed in this work for diffusion problems with constant diffusion coefficients (see details in Appendix C).

As previously mentioned, the parallel implementation of the reaction solver is quite straightforward. Once the access to the values of the unknowns u at a given node is defined, the reaction solver operates over the N leaf nodes present in the adapted grid as follows. The reaction solver class is such that computes the solution for a given number of nodes, defined as a `blocked_range<size_t>` within the TBB framework. Then, a call to `parallel_for(blocked_range<size_t>(0,N),Reaction)` parallelizes the computation, where `Reaction` corresponds to the solver class.

For the diffusion solver, two levels of parallelism is implemented. Other the aforementioned parallel computation of the m unknowns in system (1), the matrix-vector products performed by ROCK4 are also parallelized, using two `parallel_for` constructions. Further details on this ROCK4 implementation is described in what follows

3.2.1 Implementation of ROCK4

ROCK4 is an explicit time integration method, which can be viewed as the composition of two methods that are successively performed [1]. The first one is based on orthogonal polynomials and uses a three-term recurrence formula where the number of function evaluations, hence stages of the method, depends on the spectral radius of the Jacobian of the system. The second one is an explicit four-stage Runge-Kutta scheme. For a system $dy/dt = f(y)$ and a time step δt , the latter is given by

$$\left. \begin{aligned} k_i &= f \left(y_n + \delta t \sum_{j=1}^{i-1} a_{i,j} k_j \right), \quad i = 1, 2, 3, 4, \\ y_{n+1} &= y_n + \delta t \sum_{i=1}^4 b_i k_i, \end{aligned} \right\} \quad (5)$$

where $a_{i,j}$ and b_i correspond to the coefficients of the Runge-Kutta scheme.

Computing linear combinations of large vectors of `double` type, as in (5), is characterized by a rather low *arithmetic intensity*¹⁴, less than 1/4, because of the multiple memory accesses associated with reading and writing temporary arrays in memory. Nevertheless, achieving a high arithmetic intensity is fundamental to obtain a high FLOP/s efficiency (see, *e.g.*, [63]). In our implementation, the arithmetic intensity is largely improved by exploiting the fact that an explicit Runge-Kutta scheme applied to a linear problem $dy/dt = Ay$ is given by $y_{n+1} = \pi(\delta t A)y_n$, where π is a polynomial whose coefficients are a function of the $a_{i,j}$'s and k_i 's coefficients [35], which are known and precomputed. In this way, we use the Horner's method to compute the four-stage Runge-Kutta formula, thus reducing considerably the aforementioned memory traffic. Notice that a factorization of π is not feasible because this polynomial has complex roots. Finally, the spectral radius $\rho(A)$ of A , used to determine the number of stages of the method, is estimated using the Gershgorin circle theorem.

¹⁴The arithmetic intensity is a measure of floating-point operations (FLOPs) performed by a given code (or code section) relative to the amount of memory accesses (Bytes) that are required to support those operations.

4 Code performance and scalability

We have developed a C++ code, named *Z-code*, which operates on multi-dimensional configurations; the dimension d is a template parameter for most classes and methods. It runs successfully on modern Linux systems, and has been tested with both `gcc` (version 4.8 and higher) and the Intel C++ (version 14.0.1 and higher) compilers. Figure 4 shows some results for the BZ problem.

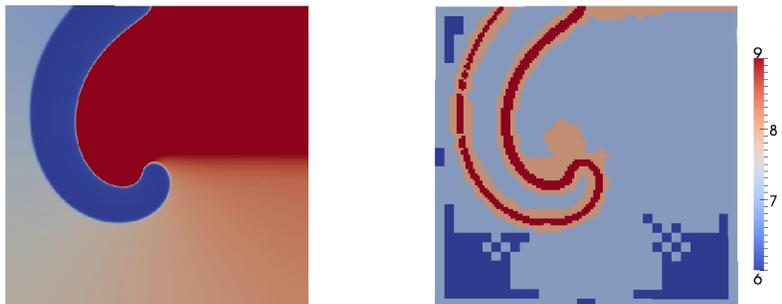


Figure 4: BZ model, two-dimensional simulation over 9 grid-levels. Variable u_1 (left) with the corresponding adapted grid defined at different grid-levels (right).

Regarding the methods considered for the dedicated splitting solver, codes for the RADAU5 and ROCK4 methods are publicly available on the Internet [34]. These codes, albeit very efficiently coded in Fortran 77, can be substantially improved. For RADAU5 we have carefully rewritten its implementation in C++, thus removing a significant amount of branching. The latter yields an enhanced performance between 20 to 30%. In the case of ROCK4, we have developed a completely new implementation, efficiently adapted to linear problems, that reduces memory accesses, as described before¹⁵.

Numerical performance can be measured from many points of view. In what follows we first present some comparisons between computations on uniform Cartesian and multiresolution adapted meshes. We then focus on the scalability attained for a given problem with increasing thread count (strong scaling). The following tests were performed on an Intel E5-2650 v3 platform (2 Haswell CPUs with 20 cores in total and 40 threads with hyperthreading) and on a Xeon Phi (MIC) 5110P (60 cores, 240 threads with simultaneous multithreading, and 8 GB of main memory) using the Intel `icpc` compiler.

4.1 Uniform Cartesian meshes versus multiresolution computations

Grid adaptation results in significant data compression. However, it is worth verifying whether the latter effectively leads to performance enhancement in terms of runtime, given the additional computations required to dynamically adapt the grid and the more elaborate implementation.

Nevertheless, in order to guarantee a fair comparison, we have replaced the diffusion implementation in the *Z-code* with one more suited to uniform meshes, for all runs performed on a uniform grid. Specifically, the matrix-vector products were replaced with a more efficient procedure for five- and seven-point stencils, using two- and three-dimensional arrays, respectively, and cache blocking. The latter avoids the additional cost of manipulating more complicated data structures and reduces the bandwidth of the computation. The following results were obtained on the *Haswell* machine, using 40 threads over 20 cores.

¹⁵The sources of *Z-code* and related software can be provided by contacting T. Dumont, the corresponding author of the present contribution. All codes are distributable under CeCILL-B licence (<http://www.cecill.info/licences.en.html>).

Recalling that J stands for the maximum grid-level allowed for mesh refinement, Tables 1 and 2 gather the wall-clock computing times in seconds, for various grid-levels J and the aforementioned three problems in two and three dimensions. Additionally, we can define a compression ratio in percentage, CR , as

$$CR = \left(1 - \frac{N}{N_J}\right) \times 100,$$

where N and N_J correspond to the number of nodes in the adapted grid and at the finest level, respectively. In particular N_J is given by 2^{dJ} and is equal to the number of nodes in the corresponding uniform Cartesian mesh. During all these computations, the compression ratio remained between 80% and 83%.

		J		
		8	9	10
NAGUMO	MR	2.51×10^{-3} (1.7)	3.02×10^{-3}	4.97×10^{-3}
	CM	1.49×10^{-3}	8.27×10^{-3} (2.7)	2.09×10^{-2} (4.2)
BZ	MR	1.14×10^{-2}	2.51×10^{-2}	4.06×10^{-2}
	CM	1.31×10^{-2} (1.2)	4.06×10^{-2} (1.6)	1.78×10^{-1} (4.4)
STROKE	MR	2.4×10^{-2}	4.03×10^{-2}	1.03×10^{-1}
	CM	3.2×10^{-1} (13.3)	1.21 (30.0)	4.80 (46.6)

Table 1: Wall-clock times in seconds for two-dimensional multiresolution (MR) and uniform Cartesian mesh (CM) computations. Largest values are indicated in bold with the corresponding ratio in parenthesis.

		J	
		8	9
NAGUMO	MR	0.13	0.31
	CM	0.54 (2.5)	2.47 (4.5)
BZ	MR	0.97	5.75
	CM	3.05 (3.1)	23.60 (4.1)
STROKE	MR	1.53	10.81
	CM	76.68 (50.0)	610.50 (56.5)

Table 2: Wall-clock times in seconds for three-dimensional multiresolution (MR) and uniform Cartesian mesh (CM) computations. Largest values are indicated in bold with the corresponding ratio in parenthesis.

These results clearly highlight the advantage of using the multiresolution algorithm to solve stiff problems disclosing localized reaction fronts with an appropriate grid resolution. As a matter of fact, looking at the previous results we see that computations on a uniform grid are more efficient only for the two-dimensional, non-stiff NAGUMO problem on a rather coarse grid. In general the additional cost related to mesh adaptation and more complex data structures is counterbalanced by the computing savings achieved with a compressed representation. The latter is even more relevant for problems with very stiff reaction terms, confirming the findings and forecasts previously described in Appendix A.

4.2 Analysis of scalability

We investigate the strong scaling behavior of Z-code by timing one single simulation time step, and breaking down the total step time (S) into the mesh adaptation (A), reaction solver (R), and diffusion solver (D). Notice that the adapted grid is updated at every time step, the detail computation is hence included in the mesh adaptation process, and the diffusion solver must also construct the matrices used in the linear systems.

The three reference problems probe different regimes of compute-intensity. The NAGUMO problem is characterized by a rather inexpensive, non-stiff reaction term, whereas the computing effort for the BZ problem is roughly balanced between the reaction solver and the other two procedures; that for the STROKE problem, on the other hand, is completely dominated by a very expensive, stiff reaction term. Table 3 shows the percentages of wall-clock time spent on each of the three major code phases, for three-dimensional simulations performed over 10 grid-levels ($J = 10$) and 40 threads on the Haswell platform.

	NAGUMO	BZ	STROKE
Mesh Adaptation	68.00	19.23	7.20
Reaction Solver	0.92	56.23	88.89
Diffusion Solver	31.08	24.54	3.91

Table 3: Percentages of wall-clock time for three-dimensional simulations with maximum grid-level $J = 10$ and 40 threads on the Haswell platform. Dominant contributions are indicated in bold.

We now discuss scalability in terms of parallel efficiency. Let T_k be the wall-clock computing time in seconds when running on k threads. We define the parallel efficiency on k threads by

$$s_k = \frac{T_1}{n_k T_k},$$

where n_k corresponds to the number of cores used when computing over k threads. In particular because of the *simultaneous multithreading*, that is, running more than one thread per core, n_k is taken either as

$$n_k = \min(k, 20),$$

for the Haswell platform, or

$$n_k = \min(k, 60),$$

for the Xeon Phi. In this way we compute the efficiency based on the number of CPU *cores*, not the number of available hardware threads. The motivation for this choice is that cores, unlike hardware threads, constitute independent compute units. The scalability plots therefore present scaling results as a function of independent hardware resources, *i.e.* cores, participating in the computation. An important consequence is that, when using hyperthreading, efficiencies may well exceed one. This simply means that running more than one thread per core may result in a more efficient execution at the core level than what can be achieved with a single thread. We will discuss hyperthreading in more detail in the following. Figure 5 illustrates the parallel efficiency, s_k , for all three problems, on Haswell and Xeon Phi, in three-dimensional simulations.

Overall, the code achieves good shared-memory scalability on compute-intensive problems (BZ and STROKE), with efficiencies over 80 % over 2 Haswell sockets. The scaling of the reaction step is very good mainly because of its embarrassingly parallel nature and the dynamic load balancing provided by TBB; there is actually no shortage of parallelism in this part of the time step. In addition, since the reaction step is strongly compute bound, cores do not compete for shared resources, allowing for linear scaling. Only for the NAGUMO problem, whose reaction computation is quite inexpensive, does the reaction efficiency drop below 90 %.

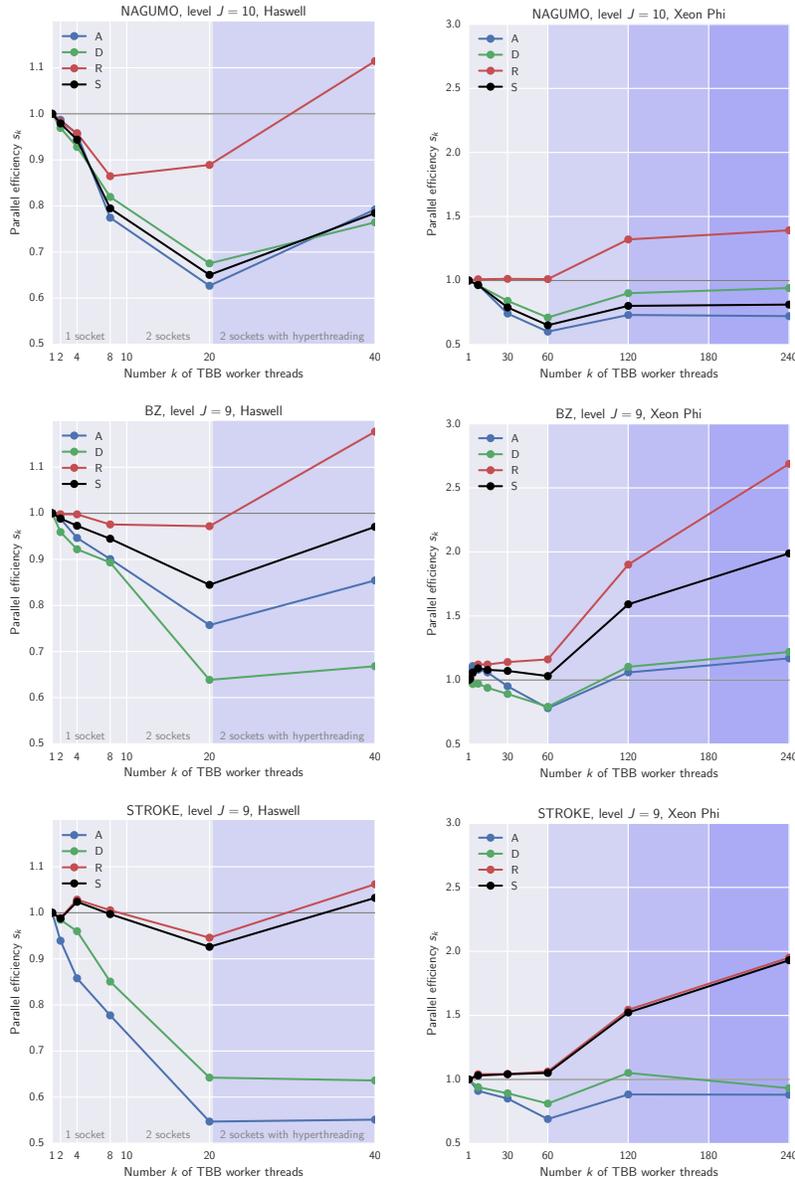


Figure 5: Parallel efficiency s_k in terms of number k of TBB worker threads for the NAGUMO (top), BZ (middle), and STROKE (bottom) problems on Haswell (left) and Xeon Phi (right) architectures, for mesh adaptation (A), diffusion (D), reaction (R), and total time step (S). The blue-shaded areas highlight thread counts with more than one TBB worker thread per CPU core (2 threads/core on Xeon Haswell and 2–4 threads/core on Xeon Phi).

The diffusion and adaptation steps achieve lower parallel efficiencies than the reaction one. A first reason for this is that both phases are more memory-intensive than the reaction; therefore, more pressure is put on the memory controllers, which are shared between cores of a given socket. This results in less-than-ideal scaling, as memory accesses become a bottleneck. Note that for the diffusion operator, we have used the compact structure described in Appendix C, which effectively improved performance over a classical CSR structure for the problems studied here. Other the

considerations on memory access, parallelization of the diffusion solver and multiresolution operations is intrinsically more complex than that of the reaction solver, yielding slightly less exposed parallelism and a degradation of parallel performance due to Amdahl’s law. This highlights the importance of exposing as much parallelism as possible, in particular in complicated algorithms such as mesh adaptation via multiresolution.

We now discuss the impact of simultaneous multithreading (hyperthreading) on the performance of *Z-Code* on the Haswell Xeon and Xeon Phi platforms. Hyperthreading allows the cores’ execution units to be used by a second application thread whenever the core stalls executing an instruction stream from the first thread. Typically, it is beneficial for applications that have good thread scalability, and where the execution is frequently stalled by latency-related events such as cache misses, branches, or bad pipelining.

For *Z-Code* on Haswell, moving from 1 to 2 threads per core (20 to 40 TBB worker threads) provides significant performance improvements, especially in the reaction solver where hyperthreading gains up to 1.25 speedup, with a geomean across problems of 1.20. The reaction solver features a very complex control flow in the RADAU5 algorithm, together with BLAS function calls and operations on small matrices. All these contribute to core stalls due to bad instruction pipelining and complex branching. These latencies, together with the good thread scalability of the reaction solver, explain the significant performance gains attained by hyperthreading.

By contrast, still on Haswell, the gains are more modest for the multiresolution algorithm (geomean 1.13) and even more so for the diffusion solver (geomean 1.05). The diffusion step features simpler control flow and, as previously discussed, its performance is dominated by memory controller bottlenecks, which hyperthreading cannot alleviate because they lie outside of the CPU core.

In regard to the Xeon Phi, on our Knights Corner coprocessor, simultaneous multithreading is crucial to enhancing performance. Specifically, it allows hiding some of the execution latencies that cannot be overlapped by the in-order architecture. In addition, because the Xeon Phi features two pipelines per core, it cannot reach its peak performance at less than 2 threads per core. Just like on the Haswell architecture, we find the largest performance gains on the reaction solver, with a geomean speedup of 1.80 across problems between 60 and 240 threads (1 and 4 threads/core). On Xeon Phi, however, hyperthreading is also beneficial to multiresolution mesh adaptation (geomean 1.32) and diffusion (geomean 1.33), highlighting the importance of also overlapping the latencies of memory accesses.

5 Concluding remarks and prospects

The main purpose of the present work was to introduce shared-memory parallelism in a thoughtful way, to a tailored numerical strategy previously developed to tackle the simulation of stiff reaction-diffusion models disclosing propagating fronts [27]. In order to achieve this, a new data structure was conceived along with parallel-friendly implementations of the key algorithms into play, including the ROCK4 and RADAU5 solvers. Shared-memory parallelism was introduced using the task-based TBB runtime library, which turned out to be particularly well suited to our purposes.

The numerical performance of the resulting implementation was assessed for three reaction-diffusion models, reaching a very satisfactory level of efficiency on shared-memory architectures such as the Intel E5-2650 and the Xeon Phi (MIC) 5110P. For models involving a certain level of modeling complexity such as BZ or STROKE, high scalability is achieved mainly because of the very efficient performance of the reaction solver within a splitting context. In addition to that, the present parallel implementation of the diffusion solver and the multiresolution algorithm shows also to perform reasonably well, given the limitations commonly encountered in standard computing architectures, as previously discussed. Finally, simultaneous multithreading is shown to improve considerably the computational performance, in particular on many-core architectures.

The proposed strategy is thus suitable for the simulation of multi-dimensional time-space multi-scale problems, which would be out of reach using classical approaches on standard computing

resources. In particular these reaction-diffusion problems are representative building blocks of more complex and realistic applications, encountered, for instance, in biomedical engineering [28], combustion [25], or plasma physics [24]. In this context ongoing work encompasses the extension to arbitrary domains, taking into account, for example, that the STROKE model applies to the human brain [28]; and to more general diffusion problems, as used, for example, in multi-component transport modeling for combustion problems [15]. A key extension of the present work deals with the implementation of hybrid parallelism using both shared- and distributed-memory formalisms. In this regard we believe the implementations and findings of this work are 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 thus expect shared-memory parallelism to gain in relevance for complex and imbalanced MPI applications. These are topics of our current research.

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 Numerical experimentation on uniform Cartesian meshes

Let us consider the numerical performance of the dedicated Strang scheme introduced in §2.1, to solve (1) on a uniform Cartesian grid. Table 4 shows the percentages of total wall-clock 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

Table 4: Percentages of wall-clock time employed to solve the diffusion problem.

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. This complexity measure may considerably vary throughout the computational domain, given that the splitting time step $\Delta t/2$ can be further split 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.

Considering a uniform mesh of 1024^2 grid-nodes, Figures 6 and 7 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 CPU time is spent over regions where the complexity is less than 17. These are

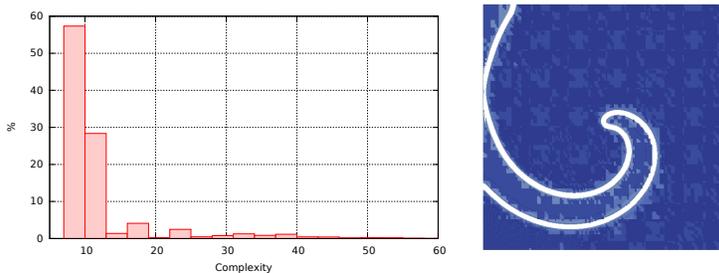


Figure 6: BZ problem. Histogram of computational complexity (left). In the right picture, white zones correspond to points where the complexity is greater than 17.

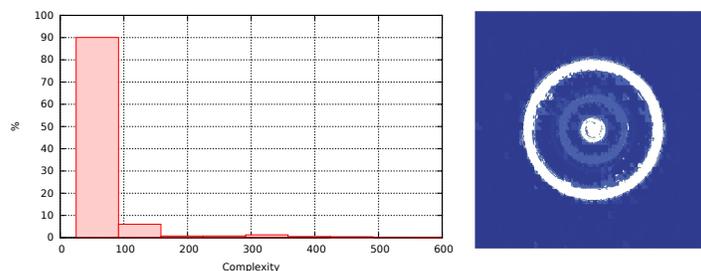


Figure 7: STROKE problem. Histogram of computational complexity (left). In the right picture, white zones correspond to points where the complexity is greater than 80.

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.

B Choice of TBB for multiresolution applications

In this work we have initially focused our attention on OpenMP API and TBB, both well-established options for shared-memory parallelism.

The OpenMP standard is a widely supported and popular choice, especially for high-performance computing (HPC) codes. It relies on compiler directives to introduce parallel regions and work-sharing constructs and tasks, which are executed in parallel by a runtime library component. OpenMP has been supporting tasking since version 3, and has recently introduced more sophisticated task constructs with version 4, with the `taskgroup` and `depend` constructs.

A key design characteristic of the OpenMP API is the reliance on a thread-based approach. That is, parallelism is defined explicitly by parallel regions which enforce parallel execution, and work and tasks are scheduled across threads participating in their containing parallel region. The set of threads taking part in a region is determined when entering the region, and fixed for its entire duration. Unfortunately, this severely limits opportunities for exploiting nested parallelism with OpenMP, even though the standard supports nested parallel regions. Actually, due to the fact that threads are assigned to a region for its whole duration, there can be no dynamic load balancing or work-sharing *across* parallel regions. In addition, creating an OpenMP region requires knowing

about the outside parallel execution context. That is, since OpenMP parallelism is mandatory, nesting parallel regions may lead to exponential creation of OpenMP worker threads for some runtime implementations, resulting in oversubscription. Finally, since all threads of a parallel region are required to enter an OpenMP work-sharing construct, one cannot arbitrarily nest OpenMP tasks and work-sharing constructs. In particular one cannot introduce parallelism by nesting `omp for` constructs within `task` constructs.

Because of the aforementioned points, OpenMP parallelism is not *composable*, that is, 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. For the development of a C++ multiresolution code, we found composability to be a desirable property, as it allows introducing parallelism at any level in the code, without knowledge or constraint on the outside calling context. This is particularly important for codes with complex call graphs and multiple call paths, as often found in C++ when using classes for code reuse.

Intel TBB, on the other hand, relies on a different approach to parallel programming. It 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, meaning that any C++ function can expose parallelism with (potentially arbitrarily nested) constructs such as `parallel_for` or `parallel_reduce`, irrespective of the calling context. Dynamic load balancing is achieved through a work-stealing runtime. Since both tasks and constructs such as `parallel_for` rely on the same task-based scheduler, TBB parallel idioms may be combined and nested arbitrarily, as long as the programmer avoids races through appropriate synchronizations. In particular, TBB’s composability allows us to achieve good scalability with the adaptive multiresolution algorithm (see § 3.1.2) by recursively nesting `parallel_for` constructs to expose parallelism in a way that follows the tree structure. In addition to a composable task runtime, TBB provides a parallel algorithm library with support for parallel reduction and sorting.

We chose TBB over an OpenMP implementation mainly because the former provides an efficient task runtime with dynamic load balancing, composability of parallel codes, and a set of parallel data structures and algorithms. Equivalent functionality could have been achieved using OpenMP tasks, but more effort would have been necessary to circumvent some of the aforementioned shortcomings.

C A compact data structure for diffusion sparse matrices

When the diffusion coefficients, $\varepsilon_i(x)$ in (1), are constant, it is possible to further reduce the amount of memory necessary to store the matrix entries associated with the diffusion problem, leading to lower memory bandwidth requirements for the program. Notice that for a given node, the matrix entries defined by the spatial discretization of the Laplace operator can have three possible values, depending on whether the neighboring nodes belong to the same, upper or lower grid-level. We therefore use the data structure shown in Figure 8. It consists of two arrays; the array A contains pointers to the beginning of lines in the second array K , where each line corresponds to one node in the adapted grid. The array K contains *line descriptors*, one for each node in the adapted grid, defined as follows. The first integer of a line descriptor is a 32-bit integer, containing the grid-level j of the current node, followed by the number of neighboring nodes at grid-level j , $j + 1$, and $j - 1$, as shown in Figure 8. The following integers then indicate the corresponding neighboring nodes. As an illustration, Figure 9 shows a typical line descriptor.

Table 5 compares the wall-clock times obtained when computing one diffusion step with the classical CSR data structure and the aforementioned compact storage, for the BZ model. Simulations on the Haswell platform were performed using 20 cores and 40 threads, whereas for Xeon Phi, 60 cores and 240 threads were considered. Computational gains are rather weak for two-dimensional simulations, for which the matrices remain relatively small. However, better performances are obtained with the compact structure when dealing with three-dimensional matrices.

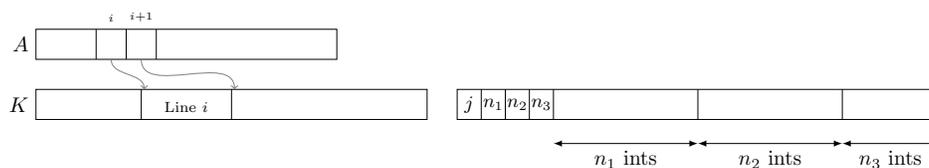


Figure 8: Compact data structure for diffusion matrices. The right part shows the structure of a line descriptor.

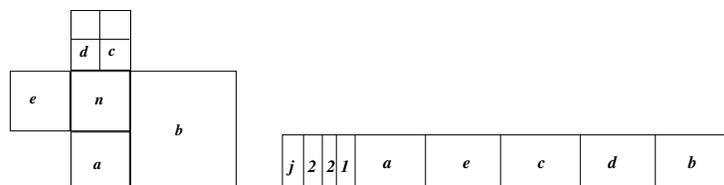


Figure 9: Illustration of a line descriptor for a given node n at grid-level j .

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 Comput.*, 27(12):37–70, 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. *J. Parallel Distr. Com.*, 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, 1999.
- [7] K. Brix, S. Melian, S. Müller, and M. Bachmann. Adaptive multiresolution methods: Practical issues on data structures, implementation and parallelization. In *Summer School on Multiresolution and Adaptive Mesh Refinement Methods*, volume 34 of *ESAIM Proc.*, pages 151–183. EDP Sci., Les Ulis, 2011.
- [8] K. Brix, S.S. Melian, S. Müller, and G. Schieffer. Parallelisation of multiscale-based grid adaptation using space-filling curves. In *ESAIM: Proc.*, volume 29, pages 108–129. EDP Sciences, 2009.
- [9] C. Burstedde, L. Wilcox, and O. Ghattas. p4est: Scalable algorithms for parallel adaptive mesh refinement on forests of octrees. *SIAM J. Sci. Comput.*, 33(3):1103–1133, 2011.
- [10] A. Cohen, S. M. Kaber, S. Müller, and M. Postel. Fully adaptive multiresolution finite volume schemes for conservation laws. *Math. Comput.*, 72:183–225, 2003.

	Haswell		Xeon Phi	
	$d = 2, J = 10$	$d = 3, J = 10$	$d = 2, J = 10$	$d = 3, J = 9$
Ratio	0.98	0.67	0.97	0.86

Table 5: Ratio between wall-clock times using the compact and CSR storages. BZ model simulated in dimension d and with maximum grid-level J

- [11] W. Crutchfield and M.L. Welcome. Object-oriented implementation of adaptive mesh refinement algorithms. *J. Sci. Program.*, pages 145–156, 1993.
- [12] 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.
- [13] R. Deiterding. A generic framework for block-structured adaptive mesh refinement in object-oriented C++. Technical report, available at <http://amroc.sourceforge.net>, 2003.
- [14] 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.
- [15] 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.
- [16] S. Descombes and T. Dumont. Numerical simulation of a stroke: Computational problems and methodology. *Prog. Biophys. Mol. Bio.*, 97(1):40–53, 2008.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] J. Dreher and R. Grauer. Racoon: A parallel mesh-adaptive framework for hyperbolic conservation laws. *Parallel Comput.*, 31(89):913–932, August 2005.
- [21] M.-A. Dronne, J.-P. Boissel, and E. Grenier. A mathematical model of ion movements in grey matter during a stroke. *J. Theor. Biol.*, 240(4):599–615, 2006.
- [22] F. Drui, A. Fikl, P. Kestener, S. Kokh, A. Larat, V. Le Chenadec, and M. Massot. Experimenting with the p4est library for AMR simulations of two-phase flows. *ESAIM: Proc. Surv.*, 53:232–247, 2016.
- [23] M. Duarte. *Adaptive numerical methods in time and space for the simulation of multi-scale reaction fronts*. PhD thesis, Ecole Centrale Paris, 2011.

- [24] 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.
- [25] M. Duarte, S. Descombes, C. Tenaud, S. Candel, and M. Massot. Time-space adaptive numerical methods for the simulation of combustion fronts. *Combust. Flame*, 160(6):1083–1101, 2013.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] W. Eckhardt and T. Weinzierl. A Blocking Strategy on Multicore Architectures for Dynamically Adaptive PDE Solvers. In D. Hutchison et al., editors, *Parallel Processing and Applied Mathematics*, volume 6067, pages 567–575. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [30] M. Essadki, S. de Chaisemartin, M. Massot, F. Laurent, and A. Larat. Adaptive mesh refinement and high order geometrical moment method for the simulation of polydisperse evaporating sprays. *Oil Gas Sci. Technol.*, 71:1–25, 2016.
- [31] C.J. Forster. *Parallel wavelet-adaptive direct numerical simulation of multiphase flows with phase change*. PhD thesis, Georgia Institute of Technology, 2016.
- [32] 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.
- [33] 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.
- [34] E. Hairer. Fortran and Matlab Codes. <http://www.unige.ch/~hairer/software.html>.
- [35] 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.
- [36] 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.
- [37] A. Harten. Multiresolution algorithms for the numerical solution of hyperbolic conservation laws. *Comm. Pure Applied Math.*, 48:1305–1342, 1995.

- [38] 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.
- [39] Intel Corporation. Intel Threading Building Blocks. <https://www.threadingbuildingblocks.org/>.
- [40] W. Jahnke, W.E. Skaggs, and A.T. Winfree. Chemical vortex dynamics in the Belousov-Zhabotinskii reaction and in the two-variable Oregonator model. *J. Phys. Chem.*, 93(2):740–749, 1989.
- [41] H. Ji, F.-S. Lien, and E. Yee. A new adaptive mesh refinement data structure with an application to detonation. *J. Comput. Phys.*, 229(23):8981–8993, 2010.
- [42] S. Jones and A. Lichtl. GPUs to Mars, full scale simulation of SpaceX’s Mars Rocket Engine. In *Proceedings of the GPU Technology Conference*, GPU Tech Conferences on demand, Silicon Valley, 2015. <http://on-demand.gputechconf.com/gtc/2015/video/S5398.html>.
- [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. *J. Comput. Phys.*, 231(3):718–744, 2012.
- [44] P. MacNeice, K.M. Olson, C. Mobarry, R. de Fainchtein, and C. Packer. PARAMESH: A parallel adaptive mesh refinement community toolkit. *Comput. Phys. Commun.*, 126(3):330–354, 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] A. Nejadmalayeri, A. Vezielainen, E. Brown-Dymkoski, and O.V. Vasilyev. Parallel adaptive wavelet collocation method for PDEs. *J. Comput. Phys.*, 298:237–253, 2015.
- [48] Netlib. Compressed row storage (crs). http://netlib.org/linalg/html_templates/node91.html.
- [49] R.M. Noyes, R. Field, and E. Koros. Oscillations in chemical systems. I. Detailed mechanism in a system showing temporal oscillations. *J. Ame. Chem. Soc.*, 94(4):1394–1395, 1972.
- [50] L. Oliker and R. Biswas. PLUM: Parallel load balancing for adaptive unstructured meshes. *J. Parallel Distr. Com.*, 52(2):150–177, 1998.
- [51] OpenMP Architecture Review Board. OpenMP Application Program Interface. Version 3.1., July 2011. Available at: <http://www.openmp.org>.
- [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. *Comput. Vis. Sci.*, 3(3):147–157, 2000.
- [54] D. Rossinelli, B. Hejazialhosseini, M. Bergdorf, and P. Koumoutsakos. Wavelet-adaptive solvers on multi-core architectures for the simulation of complex systems. *Concurr. Comput.: 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] K. Schneider and O.V. Vasilyev. Wavelet methods in computational fluid dynamics. *Annu. Rev. Fluid Mech.*, 42(1):473–503, 2010.
- [58] M. Schreiber, T. Weinzierl, and H.-J. Bungartz. Cluster Optimization and Parallelization of Simulations with Dynamically Adaptive Grids. In D. Hutchison et al., editors, *Euro-Par 2013 Parallel Processing*, volume 8097, pages 484–496. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [59] G. Strang. On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.*, 5:506–517, 1968.
- [60] R. Teyssier. Cosmological hydrodynamics with adaptive mesh refinement: A new high resolution code called RAMSES. *Astron. Astrophys.*, 385(1):337–364, 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 J. Sci. Comput.*, 33(5):2732–2760, 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] Ya.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.