

Improved accuracy for time-splitting methods for the numerical solution of parabolic equations

A. Arrarás^{a,*}, L. Portero^a

^a*Departamento de Ingeniería Matemática e Informática, Universidad Pública de Navarra, Edificio Las Encinas, Campus de Arrosadía, 31006 Pamplona (Spain)*

Abstract

In this work, we study time-splitting strategies for the numerical approximation of evolutionary reaction–diffusion problems. In particular, we formulate a family of domain decomposition splitting methods that overcomes some typical limitations of classical alternating direction implicit (ADI) schemes. The splitting error associated with such methods is observed to be $\mathcal{O}(\tau^2)$ in the time step τ . In order to decrease the size of this splitting error to $\mathcal{O}(\tau^3)$, we add a correction term to the right-hand side of the original formulation. This procedure is based on the improved initialization technique proposed by Douglas and Kim in the framework of ADI methods. The resulting non-iterative schemes reduce the global system to a collection of uncoupled subdomain problems that can be solved in parallel. Computational results comparing the newly derived algorithms with the Crank–Nicolson scheme and certain ADI methods are presented.

Keywords: alternating direction implicit, domain decomposition, partition of unity, splitting error, time-splitting method

2010 MSC: 35K57, 65M55, 68W10

1. Introduction

The numerical approximation of parabolic problems using time-splitting procedures has been a wide field of research since the pioneering works of

*Corresponding author.

Email addresses: andres.arraras@unavarra.es (A. Arrarás),
laura.portero@unavarra.es (L. Portero)

Douglas, Peaceman and Rachford in the decade of the 1950s (cf. [6, 10, 11, 26]). In such works, they introduced the so-called alternating direction implicit (ADI) methods by noting that, in a d -dimensional spatial domain Ω , the diffusion operator $-\nabla \cdot (a \nabla)$ can be expressed as the sum of d one-dimensional operators $\{\partial_k(a\partial_k)\}_{k=1,2,\dots,d}$, a being a uniformly positive function on $\overline{\Omega}$ and $\partial_k = \partial/\partial x_k$. Using this idea, multidimensional parabolic problems can be solved as a sequence of one-dimensional problems, each formulated for one of the spatial variables under consideration. Any time-stepping procedure based on a component-wise splitting of this kind is called a locally one-dimensional method (see, e.g., [15]). For an extensive study of ADI and related time-splitting methods, we refer to the monographs [13, 15, 20].

Significantly, an ADI method can be viewed as a perturbation of some underlying implicit scheme, such as the backward Euler or the Crank–Nicolson method. In this setting, the ADI method may be formulated as the corresponding implicit scheme plus a perturbation term called the splitting error. In general, this splitting error is of the same –or higher– order in the time step τ as the truncation error associated with the underlying unsplit method. As a result, the asymptotic rate of convergence for both the ADI and its underlying method should be the same; in practice, however, the actual errors associated with the former are much larger than those associated with the latter. This fact is due to the presence of the splitting error and, typically, it is considered to be the main drawback of time-splitting methods. In order to reduce the size of such an error, Douglas and Kim introduced in [9] the so-called alternating direction method with improved initialization (AD-II) (sometimes referred to as the modified alternating direction (AD-M) method, see [1]). Essentially, they proposed to add a correction term to the right-hand side of the original ADI scheme with the aim of reducing the splitting error from $\mathcal{O}(\tau^2)$ to $\mathcal{O}(\tau^3)$. This idea was used in [1] to formulate improved ADI methods for regular and mixed finite element discretizations, and further studied in [19] to derive linear multistep methods by the approximate factorization technique.

In this paper, we extend the improved initialization procedure of Douglas and Kim to the case of domain decomposition-based splittings. This kind of splittings was first introduced in [35, 36] to obtain the so-called regionally-additive schemes, and has been subsequently studied in [22, 29, 37] for the solution of evolutionary problems. In the context of linear and semilinear parabolic equations, it has been successfully used in combination with various spatial discretization techniques, such as mimetic finite differences (cf. [3,

28]), mixed finite elements (cf. [2]) or multipoint flux approximation methods (cf. [5]). Some additional results regarding nonlinear parabolic equations can be found in [4, 27]. The monographs [21, 31] show an overview of some recent contributions to the topic.

The key to the efficiency of domain decomposition splitting methods lies in reducing the system matrix to a collection of uncoupled submatrices of lower dimension. As compared to classical overlapping domain decomposition algorithms (cf. [30]), this approach does not involve any Schwarz iteration procedure, thus reducing the computational cost of the overall solution process. In addition, it overcomes two typical limitations of alternating direction splittings, namely: (a) their need to deal with rectangular or hexahedral spatial grids (in two- or three-dimensional problems, respectively); and (b) their difficulty to handle mixed derivative terms. In this respect, although several ADI methods have been specifically designed along the years to overcome this latter constraint (see, e.g., [8, 23, 24] or, more recently, [16, 17, 18]), no AD-II scheme with this property has been developed so far.

In order to introduce the improved time-splitting procedures, we consider a parabolic initial-boundary value problem of the form

$$u_t - \nabla \cdot (a \nabla u) + cu = f, \quad \mathbf{x} \in \Omega, \quad 0 < t \leq T, \quad (1a)$$

$$u = 0, \quad \mathbf{x} \in \partial\Omega, \quad 0 < t \leq T, \quad (1b)$$

$$u = u_0, \quad \mathbf{x} \in \Omega, \quad t = 0, \quad (1c)$$

where $\Omega \subset \mathbb{R}^2$ is a bounded open domain with boundary $\partial\Omega$, $a(\mathbf{x}) \in \mathbb{R}^{2 \times 2}$ is a symmetric positive definite matrix function, with elements $\{a_{ij}(\mathbf{x})\}_{i,j=1,2}$, $c(\mathbf{x})$ is a uniformly positive function on $\bar{\Omega}$, and the subscript t denotes partial differentiation with respect to time. The entries of $a(\mathbf{x})$ and the terms $c(\mathbf{x})$, $f(\mathbf{x}, t)$ and $u_0(\mathbf{x})$ are assumed to be sufficiently smooth. In the sequel, we denote by $Au = -\nabla \cdot (a \nabla u) + cu$ the elliptic operator applied to the exact solution $u(\mathbf{x}, t)$. For simplicity in the exposition, we consider homogeneous Dirichlet boundary conditions, although more general conditions can also be handled.

For the sake of completeness, it must be mentioned that the solution of problem (1) can also be approximated with a high order of accuracy by means of unsplit methods. Among the time-stepping schemes that have been recently designed for this purpose, it is worth referring to the so-called ADER (Arbitrary DERivative in space and time) approach. This method was initially developed to provide high-order approximations to the solution

of hyperbolic conservation laws (see, e.g., [12, 32]), and was subsequently formulated for nonlinear reaction–diffusion problems (cf. [33]). More recently, it has been extended in [14] to nonlinear systems of advection–diffusion–reaction equations involving stiff source terms. Some novel contributions to the topic can be found in [25, 34].

The rest of the paper is organized as follows. In §2, we formulate some classical unsplit implicit methods for the numerical solution of problem (1). Two time-splitting strategies are described in §3: on one hand, the well-known Douglas (cf. [7]) and Douglas–Rachford (cf. [11]) alternating direction methods; on the other, a non-iterative domain decomposition splitting technique based on a family of partition of unity functions. The splitting error associated with the preceding schemes is studied in §4. In this section, we also introduce a correction procedure to reduce the size of such an error. Numerical experiments comparing the previous time-splitting schemes with the Crank–Nicolson method are reported in §5. The paper ends with some concluding remarks summarized in §6.

2. Classical implicit methods

Let us consider a suitable partition \mathcal{T}_h of the spatial domain Ω , where h denotes the maximal grid spacing. In principle, \mathcal{T}_h is an unstructured grid composed of either triangular or quadrilateral elements. We will see in the next section that, if an ADI method is used for solving (1), \mathcal{T}_h must be assumed to be a rectangular grid. Domain decomposition splittings, by contrast, will remain valid for unstructured partitions. On the other hand, let us consider a constant time step $\tau > 0$ and define the discrete times $t_n = n\tau$, for $n = 0, 1, \dots, N_T + 1$, with $N_T = \lceil T/\tau \rceil - 1$.

In this framework, let A_h be a symmetric positive definite matrix obtained from finite difference or finite element discretization of the elliptic operator A with order of accuracy $\mathcal{O}(h^s)$. Then, if U_h^n denotes the fully discrete solution at time t_n , the classical implicit time-stepping schemes can be written together, for $n = 0, 1, \dots, N_T$, in the form

$$\frac{U_h^{n+1} - U_h^n}{\tau} + A_h(\theta U_h^{n+1} + (1 - \theta)U_h^n) = F_h^{n+\theta}, \quad (2)$$

where $\theta = 1$ for backward Euler and $\theta = \frac{1}{2}$ for Crank–Nicolson. Note that $F_h^{n+\theta} = \theta F_h(t_{n+1}) + (1 - \theta)F_h(t_n)$, $F_h(t_n)$ being the discrete forcing term at time t_n . The initial condition is given by $U_h^0 = \mathcal{P}_h u_0$, where \mathcal{P}_h is a

suitable restriction or projection operator. It is well known that, for a spatial discretization of order $\mathcal{O}(h^s)$, the backward Euler solution converges to the true solution with order $\mathcal{O}(\tau + h^s)$; in turn, the Crank–Nicolson solution converges with order $\mathcal{O}(\tau^2 + h^s)$. In compact form, both results can be written together as $\mathcal{O}(\tau^{3-2\theta} + h^s)$.

3. Time-splitting methods

In this section, we introduce two classes of time-splitting procedures for the solution of problem (1), namely: (a) the classical alternating direction methods based on a component-wise splitting; and (b) a family of non-iterative schemes based on an overlapping domain decomposition splitting.

3.1. Alternating direction splitting

Let us assume that $a(\mathbf{x})$ in (1a) is a diagonal matrix with elements $a_{11}(\mathbf{x})$ and $a_{22}(\mathbf{x})$. This assumption avoids the existence of mixed derivative terms in the parabolic problem. In this way, the elliptic operator A can be expressed as the sum $A = A_1 + A_2$, where

$$A_1 u = -(a_{11} u_x)_x + \frac{1}{2} c u, \quad A_2 u = -(a_{22} u_y)_y + \frac{1}{2} c u. \quad (3)$$

If the spatial domain Ω admits a rectangular grid \mathcal{T}_h , problem (1) can be approximated, for $n = 0, 1, \dots, N_T$, by an ADI method of the form

$$\begin{aligned} \frac{V_h^{n,1} - V_h^n}{\tau} + A_{1h}(\theta V_h^{n,1} + (1-\theta)V_h^n) + A_{2h}V_h^n &= F_h^{n+\theta}, \\ \frac{V_h^{n+1} - V_h^n}{\tau} + A_{1h}(\theta V_h^{n,1} + (1-\theta)V_h^n) + A_{2h}(\theta V_h^{n+1} + (1-\theta)V_h^n) &= F_h^{n+\theta}. \end{aligned} \quad (4)$$

Here, A_{1h} and A_{2h} are suitable discretizations of A_1 and A_2 , respectively, such that $A_h = A_{1h} + A_{2h}$. Note that we introduce the notation V_h^n to distinguish the ADI discrete solution from the unsplit discrete solution U_h^n of the preceding section. In this case, the initial condition is also given by $V_h^0 = \mathcal{P}_h u_0$.

The previous discretization scheme yields some classical ADI methods for certain values of the parameter θ . More precisely, if $\theta = 1$, we recover the Douglas–Rachford method, first proposed in [11]; i.e., for $n = 0, 1, \dots, N_T$,

$$\begin{aligned} (I + \tau A_{1h}) V_h^{n,1} &= (I - \tau A_{2h}) V_h^n + \tau F_h^{n+1}, \\ (I + \tau A_{2h}) V_h^{n+1} &= V_h^{n,1} + \tau A_{2h} V_h^n, \end{aligned} \quad (5)$$

for a given V_h^0 , where I denotes the identity matrix. In turn, if $\theta = \frac{1}{2}$, the Douglas method (cf. [7]) is obtained; i.e., for $n = 0, 1, \dots, N_T$,

$$\begin{aligned} \left(I + \frac{\tau}{2}A_{1h}\right) V_h^{n,1} &= \left(I - \frac{\tau}{2}A_{1h} - \tau A_{2h}\right) V_h^n + \tau F_h^{n+\frac{1}{2}}, \\ \left(I + \frac{\tau}{2}A_{2h}\right) V_h^{n+1} &= V_h^{n,1} + \frac{\tau}{2}A_{2h}V_h^n, \end{aligned} \quad (6)$$

for a given V_h^0 .

3.2. Domain decomposition splitting

As an alternative to component-wise splitting methods, we present a time-stepping procedure based on a suitable decomposition of the spatial domain. Unlike the ADI methods introduced above, this technique permits to handle mixed derivative terms (i.e., full tensor coefficients $a(\mathbf{x})$ in (1a)), and is also valid for unstructured partitions \mathcal{T}_h of the spatial domain Ω . In this case, we consider the splitting $A = A_1 + A_2 + \dots + A_m$ into an arbitrary number m of split terms, where

$$A_k u = -\nabla \cdot (\rho_k a \nabla u) + \rho_k c u, \quad \text{for } k = 1, 2, \dots, m. \quad (7)$$

The family of functions $\{\rho_k(\mathbf{x})\}_{k=1,2,\dots,m}$ is subordinate to an overlapping decomposition of Ω and conforms a smooth partition of unity. The construction of such a partition is discussed in the sequel.

Let $\{\Omega_k^*\}_{k=1,2,\dots,m}$ form a non-overlapping decomposition of Ω into m subdomains. Such a decomposition fulfills the conditions $\overline{\Omega} = \bigcup_{k=1}^m \overline{\Omega_k^*}$ and $\Omega_k^* \cap \Omega_l^* = \emptyset$, for $k \neq l$. In turn, each $\Omega_k^* \subset \Omega$ is considered to be an open disconnected set involving m_k connected components, i.e.,

$$\Omega_k^* = \bigcup_{l=1}^{m_k} \Omega_{kl}^*, \quad \text{for } k = 1, 2, \dots, m.$$

Such components are pairwise disjoint (that is, $\Omega_{ki}^* \cap \Omega_{kj}^* = \emptyset$, for $i \neq j$) and typically chosen to be shape regular of diameter h_0 . For instance, the components Ω_{kl}^* may correspond to the elements in a coarse partition \mathcal{T}_{h_0} of Ω with mesh size h_0 . Let Ω_{kl} be the extension of Ω_{kl}^* obtained by suitably translating its internal boundaries, $\partial\Omega_{kl}^* \cap \Omega$, within a distance βh_0 in Ω . The parameter $\beta > 0$ is usually referred to as the overlapping factor and its value must be chosen in such a way that the extended components are also pairwise disjoint (i.e., $\Omega_{ki} \cap \Omega_{kj} = \emptyset$, for $i \neq j$). The distance $\xi = 2\beta h_0$ is

called the overlapping size. If we denote by $\Omega_k \subset \Omega$ the open disconnected set defined as

$$\Omega_k = \bigcup_{l=1}^{m_k} \Omega_{kl}, \quad \text{for } k = 1, 2, \dots, m, \quad (8)$$

then the collection $\{\Omega_k\}_{k=1,2,\dots,m}$ forms an overlapping decomposition of Ω into m subdomains, that is, $\Omega = \bigcup_{k=1}^m \Omega_k$.

Subordinate to this overlapping covering of Ω , we construct a smooth partition of unity consisting of a family of m non-negative and $\mathcal{C}^\infty(\Omega)$ functions $\{\rho_k(\mathbf{x})\}_{k=1,2,\dots,m}$. Each function $\rho_k : \bar{\Omega} \rightarrow [0, 1]$ is chosen to be

$$\rho_k(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \in \bar{\Omega} \setminus \bar{\Omega}_k, \\ h_k(\mathbf{x}), & \text{if } \mathbf{x} \in \bigcup_{l=1; l \neq k}^m (\bar{\Omega}_k \cap \bar{\Omega}_l), \\ 1, & \text{if } \mathbf{x} \in \bar{\Omega}_k \setminus \bigcup_{l=1; l \neq k}^m (\bar{\Omega}_k \cap \bar{\Omega}_l), \end{cases}$$

where $h_k(\mathbf{x})$ is $\mathcal{C}^\infty(\Omega)$ and satisfies the conditions

$$0 \leq h_k(\mathbf{x}) \leq 1, \quad \sum_{k=1}^m h_k(\mathbf{x}) = 1,$$

for any $\mathbf{x} \in \bigcup_{l=1; l \neq k}^m (\bar{\Omega}_k \cap \bar{\Omega}_l)$. The family of functions $\{\rho_k(\mathbf{x})\}_{k=1,2,\dots,m}$ is such that

$$\text{supp}(\rho_k(\mathbf{x})) \subset \bar{\Omega}_k, \quad 0 \leq \rho_k(\mathbf{x}) \leq 1, \quad \sum_{k=1}^m \rho_k(\mathbf{x}) = 1, \quad (9)$$

for any $\mathbf{x} \in \bar{\Omega}$. For numerical purposes, $h_k(\mathbf{x})$ may not necessarily be $\mathcal{C}^\infty(\Omega)$, but only a continuous and piecewise smooth function (cf. [22]); this fact will be illustrated in the numerical experiments provided below.

In component-wise splittings, the number m of split terms is equal to the dimension d of the spatial domain Ω . In particular, the alternating direction splitting corresponding to the two-dimensional problem (1) consists of two split terms given by (3). However, for domain decomposition splittings as that considered in (7), m is not necessarily equal to d . As a result, such splittings may require a generalization of the two-stage procedure (4) to an arbitrary number m of stages. This general formulation was first proposed in [8] in the context of ADI methods. In our case, it is given, for $n = 0, 1, \dots, N_T$ and $k = 1, 2, \dots, m$, by

$$\frac{W_h^{n,k} - W_h^n}{\tau} + \sum_{i=1}^k A_{ih}(\theta W_h^{n,i} + (1 - \theta)W_h^n) + \sum_{i=k+1}^m A_{ih}W_h^n = F_h^{n+\theta}, \quad (10)$$

$$W_h^{n+1} = W_h^{n,m},$$

where $W_h^0 = \mathcal{P}_h u_0$. Note that the domain decomposition discrete solution is denoted by W_h^n . Here, $\{A_{kh}\}_{k=1,2,\dots,m}$ are suitable discretizations of the split operators $\{A_k\}_{k=1,2,\dots,m}$ introduced in (7); such discretizations are constructed to satisfy $A_h = A_{1h} + A_{2h} + \dots + A_{mh}$. For convenience, the algorithm (10) is rewritten in the equivalent form

$$(I + \theta\tau A_{1h}) W_h^{n,1} = \left(I - (1 - \theta)\tau A_{1h} - \tau \sum_{i=2}^m A_{ih} \right) W_h^n + \tau F_h^{n+\theta}, \quad (11a)$$

$$(I + \theta\tau A_{kh}) W_h^{n,k} = W_h^{n,k-1} + \theta\tau A_{kh} W_h^n, \quad \text{for } k = 2, 3, \dots, m, \quad (11b)$$

$$W_h^{n+1} = W_h^{n,m}. \quad (11c)$$

Thus, the linear system to be solved at the k -th internal stage is of the form

$$(I + \theta\tau A_{kh}) W_h^{n,k} = Q_h^{n,k}, \quad \text{for } k = 1, 2, \dots, m, \quad (12)$$

where $Q_h^{n,k}$ stands for the corresponding right-hand side. As stated in (9), the function $\rho_k(\mathbf{x})$ has compact support on $\bar{\Omega}_k$ and, by construction, the entries of A_{kh} corresponding to the nodes that lie outside of this subdomain are zero. In addition, since Ω_k involves m_k disjoint connected components Ω_{kl} due to (8), the linear system (12) is indeed a collection of m_k uncoupled subsystems of the form

$$(I_{\Omega_{kl}} + \theta\tau A_{\Omega_{kl}h}) \mathcal{R}_{\Omega_{kl}} W_h^{n,k} = \mathcal{R}_{\Omega_{kl}} Q_h^{n,k}, \quad \text{for } k = 1, 2, \dots, m, \quad (13)$$

where $I_{\Omega_{kl}} = \mathcal{R}_{\Omega_{kl}} I \mathcal{R}_{\Omega_{kl}}^T$ and $A_{\Omega_{kl}h} = \mathcal{R}_{\Omega_{kl}} A_{kh} \mathcal{R}_{\Omega_{kl}}^T$. The rectangular matrices $\mathcal{R}_{\Omega_{kl}}$ and $\mathcal{R}_{\Omega_{kl}}^T$ are usually called restriction and extension matrices, respectively, and represent a type of domain decomposition preconditioners (cf. [21]). To obtain such matrices, we first order the nodes of \mathcal{T}_h in Ω_{kl} in some local ordering. Let us denote by N_S and N , respectively, the number of nodes of \mathcal{T}_h in Ω_{kl} and the total number of nodes of \mathcal{T}_h . Then, $\mathcal{R}_{\Omega_{kl}}$ is an $N_S \times N$ matrix defined as

$$(\mathcal{R}_{\Omega_{kl}})_{ij} = \begin{cases} 1, & \text{if } \text{index}(\Omega_{kl}, i) = j, \\ 0, & \text{if } \text{index}(\Omega_{kl}, i) \neq j, \end{cases}$$

where the index function $\text{index}(\Omega_{kl}, i)$ obtains the global index of the i -th local node of \mathcal{T}_h in Ω_{kl} , for $i = 1, 2, \dots, N_S$. In other words, this matrix maps

a vector in \mathbb{R}^N of nodal values in \mathcal{T}_h into a vector in \mathbb{R}^{N_S} of nodal values corresponding to the nodes of \mathcal{T}_h in Ω_{kl} (in the local ordering). Likewise, the $N \times N_S$ transpose matrix $\mathcal{R}_{\Omega_{kl}}^T$ maps a vector in \mathbb{R}^{N_S} into a vector in \mathbb{R}^N , inserting zero entries for the global indices which do not belong to Ω_{kl} . Finally, given a global matrix $M_h \in \mathbb{R}^{N \times N}$, the submatrix $M_{\Omega_{kl}h} \in \mathbb{R}^{N_S \times N_S}$ associated with the nodes in Ω_{kl} may be obtained from the restriction and extension mappings as $M_{\Omega_{kl}h} = \mathcal{R}_{\Omega_{kl}} M_h \mathcal{R}_{\Omega_{kl}}^T$.

From a computational viewpoint, this domain decomposition splitting procedure has three main parameters to adjust its efficiency, namely: the number m of subdomains, the number m_k of disjoint connected components inside a subdomain, and the overlapping size ξ . In order to achieve an efficient performance of the algorithm, m should be chosen as small as possible to minimize the number of sequential steps (i.e., internal stages) in (13), m_k should be chosen as large as possible to maximize the number of parallel components, and ξ should be chosen as small as possible to minimize the number of unknowns within the overlapping regions. In addition, to ensure that the loads assigned to each processor are balanced, there should be approximately the same number of disjoint connected components in each subdomain and, moreover, each such component should have approximately the same diameter. For instance, suppose that we have q available processors for parallel computing. Then, if m_k is approximately the same in each subdomain Ω_k and also a multiple of q , each subdomain can be partitioned into q groups of $\frac{m_k}{q}$ components and each group assigned to one of the processors. Remarkably, unlike most classical overlapping domain decomposition algorithms (cf. [30]), the solution to (13) does not require any Schwarz iteration procedure, since the internal stages in the algorithm are solved sequentially (i.e., interface conditions need not be imposed on subdomains during the solution process).

4. The splitting error

As noted in the introduction, the time-splitting methods presented above can be interpreted as perturbations of some unsplit implicit schemes, such as the backward Euler and Crank–Nicolson methods. Recall that the scheme (4) considers the alternating direction splitting (3), while the algorithm (10) makes use of the domain decomposition splitting (7). In essence, the former method is a particular instance of the latter for $m = 2$. In this section, we reformulate (10) as a classical implicit scheme (backward Euler, for $\theta = 1$, and

Crank–Nicolson, for $\theta = \frac{1}{2}$) plus a perturbation term, which will be referred to as the splitting error. As we will see, this splitting error is $\mathcal{O}(\tau^2)$, i.e., one order higher than the truncation error associated with the backward Euler method and of the same order as that associated with the Crank–Nicolson method. As a result, one would expect to obtain the same asymptotic rate of convergence for both the time-splitting and its underlying implicit method. In practice, however, the size of the splitting error can be much larger than that of the truncation error corresponding to the underlying scheme, thus degrading the efficiency of the resulting algorithm. In the sequel, we derive a general expression for the splitting error and present a strategy to reduce its actual size.

Let us consider the formulation given by (11). In order to eliminate the intermediate values $W_h^{n,1}, W_h^{n,2}, \dots, W_h^{n,m-1}$, we multiply the equations (11b) by $(I + \theta\tau A_{1h})(I + \theta\tau A_{2h}) \cdots (I + \theta\tau A_{(k-1)h})$, sum on k and add the equation (11a) to obtain, for $n = 0, 1, \dots, N_T$,

$$\frac{W_h^{n+1} - W_h^n}{\tau} + A_h(\theta W_h^{n+1} + (1 - \theta)W_h^n) + B_h(W_h^{n+1} - W_h^n) = F_h^{n+\theta}, \quad (14)$$

where

$$\begin{aligned} B_h = & \theta^2 \tau \sum_{1 \leq i_1 < i_2 \leq m} A_{i_1 h} A_{i_2 h} + \theta^3 \tau^2 \sum_{1 \leq i_1 < i_2 < i_3 \leq m} A_{i_1 h} A_{i_2 h} A_{i_3 h} \\ & + \dots + \theta^m \tau^{m-1} A_{1h} A_{2h} \cdots A_{mh}. \end{aligned}$$

Note that, depending on the value of θ , (14) is, in fact, a perturbed backward Euler or Crank–Nicolson method (as compared with (2)). The perturbation term $B_h(W_h^{n+1} - W_h^n)$ is the aforementioned splitting error and satisfies

$$B_h(W_h^{n+1} - W_h^n) = \tau B_h \left(\frac{W_h^{n+1} - W_h^n}{\tau} \right) = \mathcal{O} \left(\tau^2 \sum_{i,j} \left| \frac{\partial}{\partial t} A_i A_j u \right| \right) = \mathcal{O}(\tau^2),$$

where the split operators A_k are given by either (3) or (7), provided that u , c and the entries of a are sufficiently smooth.

In [9], Douglas and Kim noted that, if we could add $B_h(W_h^{n+1} - W_h^n)$ to the right-hand side of (14), then the perturbation term would be cancelled. However, since W_h^{n+1} is not known at time t_n , this modification cannot be done. Taking into account that the best available approximation to the

difference $W_h^{n+1} - W_h^n$ at time t_n is $W_h^n - W_h^{n-1}$, the algorithm (11) can be modified, for $n = 1, 2, \dots, N_T$, to obtain

$$\begin{aligned}
\widehat{F}_h^{n+\theta} &= F_h^{n+\theta} + B_h(Z_h^n - Z_h^{n-1}), \\
(I + \theta\tau A_{1h})Z_h^{n,1} &= \left(I - (1 - \theta)\tau A_{1h} - \tau \sum_{i=2}^m A_{ih} \right) Z_h^n + \tau \widehat{F}_h^{n+\theta} \\
(I + \theta\tau A_{kh})Z_h^{n,k} &= Z_h^{n,k-1} + \theta\tau A_{kh} Z_h^n, \quad \text{for } k = 2, 3, \dots, m, \\
Z_h^{n+1} &= Z_h^{n,m},
\end{aligned} \tag{15}$$

where $Z_h^0 = \mathcal{P}_h u_0$ and Z_h^1 must be suitably approximated. In this case, eliminating the intermediate values $Z_h^{n,1}, Z_h^{n,2}, \dots, Z_h^{n,m-1}$, or referring to (14), we get, for $n = 1, 2, \dots, N_T$,

$$\frac{Z_h^{n+1} - Z_h^n}{\tau} + A_h(\theta Z_h^{n+1} + (1 - \theta)Z_h^n) + B_h(Z_h^{n+1} - 2Z_h^n + Z_h^{n-1}) = F_h^{n+\theta}.$$

Now, the perturbation term $B_h(Z_h^{n+1} - 2Z_h^n + Z_h^{n-1})$ satisfies

$$\begin{aligned}
B_h(Z_h^{n+1} - 2Z_h^n + Z_h^{n-1}) &= \tau^2 B_h \left(\frac{Z_h^{n+1} - 2Z_h^n + Z_h^{n-1}}{\tau^2} \right) \\
&= \mathcal{O} \left(\tau^3 \sum_{i,j} \left| \frac{\partial^2}{\partial t^2} A_i A_j u \right| \right) = \mathcal{O}(\tau^3),
\end{aligned}$$

provided that u, c and the entries of a are sufficiently smooth. In other words, the corrected right-hand side $\widehat{F}_h^{n+\theta}$ introduced in (15) yields a reduction in the order of the splitting error from $\mathcal{O}(\tau^2)$ to $\mathcal{O}(\tau^3)$. As a result, the actual discretization error of the modified algorithm (15) is of the same size as that associated with its underlying unsplit method. Remarkably, if the domain decomposition splitting (7) is used to obtain A_k , the operator B_h will be non-zero only within the overlapping regions. Thus, the newly added term $B_h(Z_h^n - Z_h^{n-1})$ will have a very low extra computational cost.

The convergence analysis for the improved accuracy method (15) must take into account the splitting strategy under consideration. The case in which the split operators A_k are obtained via the alternating direction splitting (3) is discussed in [1] and [9], and error estimates of order $\mathcal{O}(\tau^{3-2\theta} + h^s)$ are derived. In the latter work, the discrete split operators are assumed to

commute pairwise, while the former analyzes the non-commuting case. As for the domain decomposition splitting (7), a convergence analysis of this kind is lacking so far. A similar analysis for the unimproved algorithm (11), considering a finite difference spatial discretization, is carried out in [22]. The derivation of error estimates for the newly proposed method (15) is a topic of current research.

5. Numerical experiments

The numerical experiments performed in this section are partly inspired by those contained in [1, 9]. Let us consider the parabolic initial-boundary value problem (1), where $\Omega = (0, 1) \times (0, 1)$, $T = 1$ and $c(\mathbf{x}) \equiv 0$. Given a certain diffusion term $a(\mathbf{x})$ (to be specified below), the functions $f(\mathbf{x}, t)$ and $u_0(\mathbf{x})$ are defined in such a way that the exact solution is

$$u(x, y, t) = \sin(2\pi t) \sin(2\pi x) \sin(2\pi y).$$

Throughout this section, a finite difference spatial discretization is combined with several second-order time integrators (i.e., we fix $\theta = \frac{1}{2}$), considering the discretization parameters $h = \tau = \frac{1}{M}$. In particular, we compare the accuracy of the Crank–Nicolson (CN) method (2) to that of the Douglas–Gunn scheme (11) and the Douglas–Kim method (15), combined with either the ADI splitting (3) (DG_{ADI} and DK_{ADI}, respectively) or the domain decomposition splitting (7) (DG_{DD} and DK_{DD}, respectively). For the Douglas–Kim procedures, DK_{ADI} and DK_{DD}, the approximation of the exact solution at time $t_1 = \tau$ is obtained by running a single step of the Crank–Nicolson scheme. The computed errors are measured in the discrete $\ell^\infty(\ell^2)$ -norm, that is,

$$\|e\|_{\ell^\infty(\ell^2)} = \max_{1 \leq n \leq N_T} \left(\sum_{i=1}^{M-1} \sum_{j=1}^{M-1} (e_{i,j}^n)^2 h^2 \right)^{1/2},$$

where $e_{i,j}^n$ denotes the error at the (i, j) -th node and time t_n , obtained as the difference between the exact solution at (ih, jh, t_n) and the corresponding numerical solution.

Unless otherwise stated, both the DG_{DD} and the DK_{DD} methods consider two overlapping subdomains, each consisting of four non-overlapping

connected components. More precisely, let $I = (0, 1)$ and define the intervals

$$I_1 = \left(0, \frac{1}{8} + \frac{\xi}{2}\right) \cup \left(\frac{1}{4} - \frac{\xi}{2}, \frac{3}{8} + \frac{\xi}{2}\right) \cup \left(\frac{1}{2} - \frac{\xi}{2}, \frac{5}{8} + \frac{\xi}{2}\right) \cup \left(\frac{3}{4} - \frac{\xi}{2}, \frac{7}{8} + \frac{\xi}{2}\right),$$

$$I_2 = \left(\frac{1}{8} - \frac{\xi}{2}, \frac{1}{4} + \frac{\xi}{2}\right) \cup \left(\frac{3}{8} - \frac{\xi}{2}, \frac{1}{2} + \frac{\xi}{2}\right) \cup \left(\frac{5}{8} - \frac{\xi}{2}, \frac{3}{4} + \frac{\xi}{2}\right) \cup \left(\frac{7}{8} - \frac{\xi}{2}, 1\right),$$

where ξ denotes the overlapping size (to be specified below). Then, we can define the subdomains $\Omega_1 = I_1 \times I$ and $\Omega_2 = I_2 \times I$. Figure 1 shows a domain decomposition of this kind on the unit square, for an arbitrary number q of connected components per subdomain.

According to [22], we construct a piecewise smooth partition of unity as follows. The closure of subdomain Ω_k can be expressed as the union of four closed disjoint rectangles given by $R_k^i = [a_k^i, b_k^i] \times [0, 1]$, for $i = 1, 2, 3, 4$ and $k = 1, 2$. Then, for each rectangle R_k^i , we define a function of the form

$$w_k^i(x, y) = \begin{cases} \sin\left(\frac{\pi(x - a_k^i)}{b_k^i - a_k^i}\right), & \text{if } (x, y) \in R_k^i, \\ 0, & \text{if } (x, y) \in \bar{\Omega} \setminus R_k^i. \end{cases}$$

The partition of unity functions are thus obtained as

$$\rho_k(x, y) = \frac{\sum_{i=1}^4 w_k^i(x, y)}{\sum_{l=1}^2 \sum_{i=1}^4 w_l^i(x, y)}, \quad \text{for } k = 1, 2. \quad (16)$$

Table 1 shows the global errors and mean rates of convergence obtained for different values of the discretization parameter M , with the aforementioned methods. In this experiment, the diffusion coefficient is $a(x, y) = 1$, and the overlapping size is $\xi = \frac{1}{8}$. We observe that the global errors corresponding to the DG_{ADI} and the DG_{DD} methods are about 10 and 14 times larger, respectively, than the CN errors. The DK_{ADI} errors are larger but comparable to the CN errors, since the splitting error for the former method is much smaller than that for the DG_{ADI} scheme. In this case, the mean rate of convergence is greater than 2, because the splitting error is being removed at a rate of $\mathcal{O}(\tau^3)$. Accordingly, the DK_{DD} errors are much smaller than the DG_{DD} errors, and very similar to those obtained for the CN scheme. In fact, due to the observed higher rate of convergence, the DK_{DD} errors are indeed smaller than the CN errors, for $M = 160$ and $M = 320$.

In Table 2, we fix the values $M = 160$ and $\xi = \frac{1}{8}$, and compute the global errors corresponding to the previous schemes, for five different choices of the diffusion coefficient $a(x, y)$. Besides the one tested in the preceding table, $a_1(x, y) = 1$, we also consider

$$a_2(x, y) = \frac{1}{2 + \cos(3\pi x) \cos(2\pi y)},$$

$$a_3(x, y) = \begin{cases} 1 + 0.5 \sin(5\pi x) + y^3, & \text{if } x \leq 0.5, \\ \frac{1.5}{1 + (x - 0.5)^2} + y^3, & \text{otherwise,} \end{cases}$$

$$a_4(x, y) = \begin{pmatrix} a_2(x, y) & 0 \\ 0 & a_3(x, y) \end{pmatrix},$$

$$a_5(x, y) = \begin{pmatrix} a_2(x, y) & 1/4 \\ 1/4 & a_2(x, y) \end{pmatrix}.$$

If we take $a(x, y) = a_i(x, y)$, for $i = 1, 2, 3, 4$, the computed DG_{ADI} errors are much larger than the corresponding CN errors. In turn, the errors associated with the DK_{ADI} method are comparable to, and only slightly larger than, those obtained with the CN scheme. On the other hand, if the diffusion coefficient is chosen to be $a_5(x, y)$, none of these two ADI methods can be applied, due to the presence of mixed derivative terms. As for the domain decomposition splitting techniques, in all five cases, the DK_{DD} scheme is shown to be comparable to the CN scheme and superior to the DG_{DD} method.

Table 3 shows the effect of the overlapping size ξ on the accuracy of the implemented methods. In particular, we consider $a(x, y) = a_2(x, y)$ and $M = 160$, and compute the global errors corresponding to the CN, the DG_{DD} and the DK_{DD} schemes, for different values of ξ . The observed DG_{DD} errors increase as ξ becomes smaller, due to the presence of negative powers of ξ in the splitting error (cf. [22]). This undesirable behaviour is avoided by the DK_{DD} scheme, whose global errors get slightly smaller as ξ decreases.

Finally, we study the influence of the number of disjoint connected components per subdomain on the accuracy of the domain decomposition procedures. For that purpose, we consider the same number q of components in Ω_1 and Ω_2 ; i.e., recalling the notation introduced in §3, $m_1 = m_2 = q$. Table 4 shows the global errors corresponding to the CN, the DG_{DD} and the DK_{DD} schemes, for different values of q , when $a(x, y) = a_2(x, y)$, $M = 160$

and $\xi = \frac{1}{16}$. Note that, if $q = 2$ or $q = 8$, the partition of unity functions (16) must be modified accordingly. As in the preceding experiments, the DG_{DD} errors are larger than the CN errors; moreover, they even grow as q increases, which represents a serious drawback from a computational viewpoint. By contrast, the errors for the DK_{DD} scheme are comparable to, and even smaller than, those associated with the CN method. In this case, such errors decrease as q increases, thus permitting to exploit the parallel capabilities of the new algorithms.

6. Conclusions

A family of improved domain decomposition splitting methods for solving parabolic equations has been presented. These methods can be formulated as unsplit implicit schemes (such as backward Euler or Crank–Nicolson) perturbed by a term that is $\mathcal{O}(\tau^3)$ in the time step τ . This perturbation term, also known as splitting error, is one order higher than its $\mathcal{O}(\tau^2)$ analogue stemming from the unimproved formulation. Such an increase in the order of the splitting error is achieved by adding a correction term to the right-hand side of the original scheme. This technique was proposed in [9] in the context of ADI methods, and has been extended here to domain decomposition-based splittings for the first time. The newly derived methods can be implemented with a very low additional cost, since the correction term only requires certain computations within the overlapping regions. In addition, the resulting subdomain problems can be solved in parallel with no need for Schwarz-type iteration procedures. Numerical results show that the proposed modification leads to a significant reduction in the splitting error. Moreover, the global discretization error is comparable to that associated with the Crank–Nicolson method, independently of the number of connected components per subdomain and the overlapping size.

References

- [1] T. Arbogast, C.-S. Huang, S.-M. Yang, Improved accuracy for alternating-direction methods for parabolic equations based on regular and mixed finite elements, *Math. Models Methods Appl. Sci.* 17 (2007), 1279–1305.

- [2] A. Arrarás, L. Portero, Expanded mixed finite element domain decomposition methods on triangular grids, *Int. J. Numer. Anal. Model.* 11 (2014), 255–270.
- [3] A. Arrarás, L. Portero, J.C. Jorge, Convergence of fractional step mimetic finite difference discretizations for semilinear parabolic problems, *Appl. Numer. Math.* 60 (2010), 473–485.
- [4] A. Arrarás, L. Portero, J.C. Jorge, Locally linearized fractional step methods for nonlinear parabolic problems, *J. Comput. Appl. Math.* 234 (2010), 1117–1128.
- [5] A. Arrarás, L. Portero, I. Yotov, Error analysis of multipoint flux domain decomposition methods for evolutionary diffusion problems, *J. Comput. Phys.* 257 (2014), part B, 1321–1351.
- [6] J. Douglas, Jr., On the numerical integration of $\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = \partial u / \partial t$ by implicit methods, *J. Soc. Indust. Appl. Math.* 3 (1955), 42–65.
- [7] J. Douglas, Jr., Alternating direction methods for three space variables, *Numer. Math.* 4 (1962), 41–63.
- [8] J. Douglas, Jr., J.E. Gunn, A general formulation of alternating direction methods. I. Parabolic and hyperbolic problems, *Numer. Math.* 6 (1964), 428–453.
- [9] J. Douglas, Jr., S. Kim, Improved accuracy for locally one-dimensional methods for parabolic equations, *Math. Models Methods Appl. Sci.* 11 (2001), 1563–1579.
- [10] J. Douglas, Jr., D.W. Peaceman, Numerical solution of two-dimensional heat-flow problems, *AIChE J.* 1 (1955), 505–512.
- [11] J. Douglas, Jr., H.H. Rachford, Jr., On the numerical solution of heat conduction problems in two and three space variables, *Trans. Amer. Math. Soc.* 82 (1956), 421–439.
- [12] M. Dumbser, C. Enaux, E.F. Toro, Finite volume schemes of very high order of accuracy for stiff hyperbolic balance laws, *J. Comput. Phys.* 227 (2008), 3971–4001.

- [13] I. Faragó, Á. Havasi, *Operator Splittings and their Applications*, Mathematics Research Developments, Nova Science Publishers, Inc., New York, 2009.
- [14] A. Hidalgo, M. Dumbser, ADER schemes for nonlinear systems of stiff advection–diffusion–reaction equations, *J. Sci. Comput.* 48 (2011), 173–189.
- [15] W. Hundsdorfer, J.G. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer Ser. Comput. Math. 33, Springer-Verlag, Berlin, 2003.
- [16] K.J. in 't Hout, C. Mishra, Stability of ADI schemes for multidimensional diffusion equations with mixed derivative terms, *Appl. Numer. Math.* 74 (2013), 83–94.
- [17] K.J. in 't Hout, B.D. Welfert, Stability of ADI schemes applied to convection–diffusion equations with mixed derivative terms, *Appl. Numer. Math.* 57 (2007), 19–35.
- [18] K.J. in 't Hout, B.D. Welfert, Unconditional stability of second-order ADI schemes applied to multi-dimensional diffusion equations with mixed derivative terms, *Appl. Numer. Math.* 59 (2009), 677–692.
- [19] S. Karaa, Improved accuracy for the approximate factorization of parabolic equations, *Computing* 86 (2009), 23–36.
- [20] G.I. Marchuk, Splitting and alternating direction methods, in: P.G. Ciarlet, J.L. Lions (editors), *Finite Difference Methods. Part 1. Solution of Equations in \mathbb{R}^n . Part 1*, *Handb. Numer. Anal. I*, North-Holland, Amsterdam, 1990, pp. 197–462.
- [21] T.P.A. Mathew, *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*, *Lect. Notes Comput. Sci. Eng.* 61, Springer-Verlag, Berlin, 2008.
- [22] T.P. Mathew, P.L. Polyakov, G. Russo, J. Wang, Domain decomposition operator splittings for the solution of parabolic equations, *SIAM J. Sci. Comput.* 19 (1998), 912–932.

- [23] S. McKee, A.R. Mitchell, Alternating direction methods for parabolic equations in two space dimensions with a mixed derivative, *Comput. J.* 13 (1970), 81–86.
- [24] S. McKee, D.P. Wall, S.K. Wilson, An alternating direction implicit scheme for parabolic equations with mixed derivative and convective terms, *J. Comput. Phys.* 126 (1996), 64–76.
- [25] G.I. Montecinos, E.F. Toro, Reformulations for general advection–diffusion–reaction equations and locally implicit ADER schemes, *J. Comput. Phys.* 275 (2014), 415–442.
- [26] D.W. Peaceman, H.H. Rachford, Jr., The numerical solution of parabolic and elliptic differential equations, *J. Soc. Indust. Appl. Math.* 3 (1955), 28–41.
- [27] L. Portero, A. Arrarás, J.C. Jorge, Contractivity of domain decomposition splitting methods for nonlinear parabolic problems, *J. Comput. Appl. Math.* 234 (2010), 1078–1087.
- [28] L. Portero, A. Arrarás, J.C. Jorge, Domain decomposition operator splitting for mimetic finite difference discretizations of non-stationary problems, *Comput. & Fluids* 46 (2011), 398–403.
- [29] L. Portero, B. Bujanda, J.C. Jorge, A combined fractional step domain decomposition method for the numerical integration of parabolic problems, in: R. Wyrzykowski, J. Dongarra, M. Paprzycki, J. Wasniewski (editors), *Parallel Processing and Applied Mathematics, Lecture Notes in Comput. Sci.* 3019, Springer-Verlag, Berlin, 2004, pp. 1034–1041.
- [30] A. Quarteroni, A. Valli, *Domain Decomposition Methods for Partial Differential Equations*, *Numer. Math. Sci. Comput.*, Oxford University Press, New York, 1999.
- [31] A.A. Samarskiĭ, P.P. Matus, P.N. Vabishchevich, *Difference Schemes with Operator Factors*, *Mathematics and Its Applications* 546, Kluwer Academic Publishers, Dordrecht, 2002.
- [32] V.A. Titarev, E.F. Toro, ADER: arbitrary high order Godunov approach, *J. Sci. Comput.* 17 (2002), 609–618.

- [33] E.F. Toro, A. Hidalgo, ADER finite volume schemes for nonlinear reaction–diffusion equations, *Appl. Numer. Math.* 59 (2009), 73–100.
- [34] E.F. Toro, G.I. Montecinos, Advection-diffusion-reaction equations: hyperbolization and high-order ADER discretizations, *SIAM J. Sci. Comput.* 36 (2014), A2423–A2457.
- [35] P.N. Vabishchevich, Difference schemes with domain decomposition for solving non-stationary problems, *U.S.S.R. Comput. Math. Math. Phys.* 29 (1989), 155–160.
- [36] P.N. Vabishchevich, Parallel domain decomposition algorithms for time-dependent problems of mathematical physics, in: I.T. Dimov, B. Sendov, P.S. Vassilevski (editors), *Advances in Numerical Methods and Applications*, World Scientific, River Edge, 1994, pp. 293–299.
- [37] P.N. Vabishchevich, Domain decomposition methods with overlapping subdomains for the time-dependent problems of mathematical physics, *Comput. Methods Appl. Math.* 8 (2008), 393–405.

Method	$M = 40$	$M = 80$	$M = 160$	$M = 320$	Rate
CN	1.029e-03	2.571e-04	6.426e-05	1.606e-05	2.000
DG _{ADI}	9.780e-03	2.457e-03	6.244e-04	1.658e-04	1.961
DK _{ADI}	2.558e-03	4.560e-04	9.920e-05	3.121e-05	2.119
DG _{DD}	1.444e-02	3.026e-03	8.488e-04	2.252e-04	2.001
DK _{DD}	2.180e-03	2.933e-04	6.079e-05	1.494e-05	2.396

Table 1: Global errors and mean rates of convergence for different values of the discretization parameter M , with $a(x, y) = 1$ and $\xi = \frac{1}{8}$.

Method	$a = a_1$	$a = a_2$	$a = a_3$	$a = a_4$	$a = a_5$
CN	6.426e-05	6.179e-05	7.456e-05	6.160e-05	9.339e-05
DG _{ADI}	6.244e-04	3.541e-04	9.391e-04	4.863e-04	–
DK _{ADI}	9.920e-05	7.732e-05	1.116e-04	8.308e-05	–
DG _{DD}	8.488e-04	4.427e-04	1.464e-03	5.313e-04	5.333e-04
DK _{DD}	6.079e-05	5.905e-05	8.327e-05	6.481e-05	9.593e-05

Table 2: Global errors for different choices of the diffusion coefficient $a(x, y)$, with $M = 160$ and $\xi = \frac{1}{8}$.

Method	$\xi = \frac{1}{8}$	$\xi = \frac{1}{16}$	$\xi = \frac{1}{32}$
CN	6.179e-05	6.179e-05	6.179e-05
DG _{DD}	4.427e-04	8.148e-04	1.747e-03
DK _{DD}	5.905e-05	4.600e-05	4.418e-05

Table 3: Global errors for different values of the overlapping size ξ , with $a(x, y) = a_2(x, y)$ and $M = 160$.

Method	$q = 2$	$q = 4$	$q = 8$
CN	6.179e-05	6.179e-05	6.179e-05
DG _{DD}	4.407e-04	8.148e-04	1.475e-03
DK _{DD}	5.717e-05	4.600e-05	3.742e-05

Table 4: Global errors for different number q of connected components per subdomain, with $a(x, y) = a_2(x, y)$, $M = 160$ and $\xi = \frac{1}{16}$.

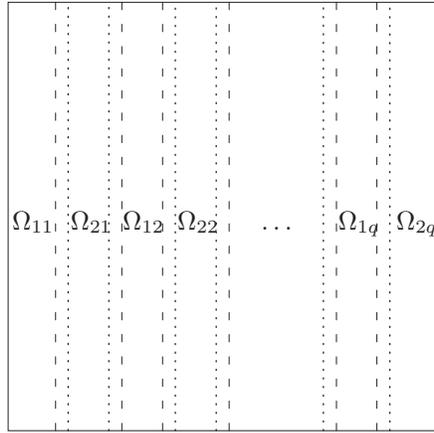


Figure 1: Decomposition of Ω into two subdomains, Ω_1 and Ω_2 , each consisting of q disjoint connected components, $\{\Omega_{1l}\}_{l=1,2,\dots,q}$ and $\{\Omega_{2l}\}_{l=1,2,\dots,q}$. The internal boundaries of such components are represented by dotted and dashed lines, respectively.