



**HAL**  
open science

## Adaptive time splitting method for multi-scale evolutionary partial differential equations

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

► **To cite this version:**

Stéphane Descombes, Max Duarte, Thierry Dumont, Violaine Louvet, Marc Massot. Adaptive time splitting method for multi-scale evolutionary partial differential equations. *Confluentes Mathematici*, 2011, 3, pp.1-31. 10.1142/S1793744211000412 . hal-00587036v2

**HAL Id: hal-00587036**

**<https://hal.science/hal-00587036v2>**

Submitted on 31 Jul 2011

**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.

Confluentes Mathematici  
© World Scientific Publishing Company

## ADAPTIVE TIME SPLITTING METHOD FOR MULTI-SCALE EVOLUTIONARY PARTIAL DIFFERENTIAL EQUATIONS

STÉPHANE DESCOMBES

*Laboratoire J. A. Dieudonné - UMR CNRS 6621, Université Nice Sophia Antipolis, Parc  
Valrose, Nice Cedex 02, 06108, France  
sdescomb@unice.fr*

MAX DUARTE

*Laboratoire EM2C - UPR CNRS 288, Ecole Centrale Paris, Grande Voie des Vignes  
Chatenay-Malabry Cedex, 92295, France  
max.duarte@em2c.ecp.fr*

THIERRY DUMONT

*Institut Camille Jordan - UMR CNRS 5208, Université de Lyon, Université Lyon 1, INSA de  
Lyon 69621, Ecole Centrale de Lyon, 43 Boulevard du 11 novembre 1918,  
Villeurbanne Cedex, 69622, France  
tdumont@math.univ-lyon1.fr*

VIOLAINE LOUVET

*Institut Camille Jordan - UMR CNRS 5208, Université de Lyon, Université Lyon 1, INSA de  
Lyon 69621, Ecole Centrale de Lyon, 43 Boulevard du 11 novembre 1918,  
Villeurbanne Cedex, 69622, France,  
louv@math.univ-lyon1.fr*

MARC MASSOT

*Laboratoire EM2C - UPR CNRS 288, Ecole Centrale Paris, Grande Voie des Vignes,  
Chatenay-Malabry Cedex, 92295 France,  
marc.massot@em2c.ecp.fr*

Received Day Month Year

Revised Day Month Year

This paper introduces an adaptive time splitting technique for the solution of stiff evolutionary PDEs that guarantees an effective error control of the simulation, independent of the fastest physical time scale for highly unsteady problems. The strategy considers a second order Strang method and another lower order embedded splitting scheme that takes into account potential loss of order due to the stiffness featured by time-space multi-scale phenomena. The scheme is then built upon a precise numerical analysis of the method and a complementary numerical procedure, conceived to overcome classical restrictions of adaptive time stepping schemes based on lower order embedded methods, whenever asymptotic estimates fail to predict the dynamics of the problem. The performance of the method in terms of control of integration errors is evaluated by numerical simulations of stiff propagating waves coming from nonlinear chemical dynamics models

2 *Descombes, Duarte, Dumont, Louwet, Massot*

as well as highly multi-scale nanosecond repetitively pulsed gas discharges, which allow to illustrate the method capabilities to consistently describe a broad spectrum of time scales and different physical scenarios for consecutive discharge/post-discharge phases.

*Keywords:* Time adaptive integration; error control; operator splitting; reaction-diffusion; multi-scale reaction waves; multi-scale discharge.

AMS Subject Classification: 65G20, 65M15, 65Z05, 65L04, 35K57, 35A35, 35C07

### Dedication

Cet article est dédié à la mémoire de Michelle Schatzman. Spécialiste des méthodes de décomposition d'opérateur, sa grande clairvoyance scientifique lui a permis d'orienter plusieurs chercheurs débutants sur ce sujet à un moment où il pouvait sembler achevé. Michelle aimait dire qu'il n'y a pas de frontière entre les branches des mathématiques et que seule une grande culture permet de naviguer dans cette forêt et d'y trouver les bonnes techniques pour résoudre un problème. Ce travail est un hommage; à la croisée des mathématiques et de leur applications effectives, il tente d'illustrer cette assertion. Michelle, ton dynamisme, ton humour et ton plaisir à parler mathématiques nous manquent.

### 1. Introduction

Numerical simulations of multi-scale phenomena are commonly used for modeling purposes in many applications such as combustion, plasma discharges, chemical vapor deposition or air pollution modeling. In general, all these models raise several difficulties created by the high number of unknowns, the wide range of temporal scales due to large and detailed chemical kinetic mechanisms, as well as steep spatial gradients associated with localized fronts of high chemical activity. In this context, faced with the induced stiffness of these time dependent problems, a high performing numerical strategy for multidimensional simulations considers a time operator splitting with dedicated high order time integration methods for reaction and diffusion problems, in order to exploit efficiently the special features of each problem. Such a numerical strategy for time discretization has been presented in [9] and extended in [8] with multiresolution techniques for adaptive space discretization. The main idea is to use a second order Strang scheme to solve independently reaction and diffusion problems in three successive fractional steps, taking into account that for multi-scale phenomena better performances are usually expected while ending the splitting scheme by the part involving the fastest scales, as it has been proven in [5]. Therefore, based on these theoretical results and on the construction of the splitting solver, this strategy provides an accurate resolution of such stiff problems even for splitting time steps much larger than either the fastest time scales involved in the source terms or the time step restrictions related to spatial grid discretizations.

Up to our days, fixed splitting time step schemes have been largely used in the literature [16, 24, 22], and the relevance of our numerical strategy [9, 8] has been evaluated in the framework of stiff reaction waves for which a constant splitting

time step is more than reasonable to precisely describe the global coupling of the split phenomena. However, such a fixed time stepping strategy would surely lead to major difficulties and limitations for problems describing highly non stationary models with very different dynamics such as flame ignition and propagation or repetitively pulsed plasmas discharges [23], all the more in the framework of large scale simulations. It is thus essential to be able to dynamically adapt splitting time steps for the simulation of such multi-scale problems with strongly evolving dynamics.

In order to guarantee a precise description of the coupled multi-scale phenomenon, this splitting time step adaptation strategy must rely on a local error estimate, which can be obtained by considering a lower order embedded method. This is a common practice for ODEs numerical solution [13], which yields very efficient and eventually high order methods for which time steps can dynamically adapt according to a given tolerance, to sufficiently small values in order to cope with the fastest time scales of the problem. However, it is well known that for stiff problems and larger accuracy tolerances, the order of the methods can degenerate, yielding non reliable error estimates and possibly, much larger global errors than expected by the given tolerance. Such a scenario will be all the more valid in the framework of the resolution of PDEs where fine grid and large gradients coupled with stiff source terms lead to especially stiff problems. In particular, our numerical strategy [9, 8] is built in such a way that the main source of error is the splitting error, each building block relying on high order adaptive and dedicated numerical methods; therefore, it is essential not only to construct a reliable splitting error estimate, but also to guarantee an effective error control within the so claimed accuracy tolerance.

In this article, we present a novel strategy to control the local splitting error with two different splitting schemes, the first one is a second order Strang technique whereas the second one considers a shifted Strang formula, built with a  $\varepsilon$ -shift in time of the classical Strang formula. This second method is embedded because the first substep is common to both methods to reduce computational cost, and inherits from the Strang scheme, stability properties and the same numerical behavior in the context of stiff problems; nevertheless, it is only of order one due to the slightly lack of symmetry. In the first part of the paper, we conduct a complete error estimate of this new splitting method in order to characterize the local error estimate that will be computed out of first and second order splitting resolutions. We define then a domain of application of the adaptive method in which the local error estimates guarantee an effective error control of the solution according to the given tolerance. The key issue is related to the evaluation of a maximum splitting time step, called the critical splitting time step, as a function of  $\varepsilon$ , for which local error estimates are valid. A numerical validation of the theoretical estimates is performed in the framework of traveling reaction waves for a simple PDE, for which the threshold and critical time steps can be also theoretically estimated and compared with numerical results.

However, in order to extend the numerical strategy to more realistic configu-

4 *Descombes, Duarte, Dumont, Louwet, Massot*

rations, for which theoretical evaluation of critical time steps is out of reach, we develop a complementary and general numerical procedure based on numerical estimates, that allows to establish the domain of application of the method by simultaneously choosing the appropriate  $\varepsilon$  for a given tolerance. This procedure is tested in the framework of nonlinear chemical dynamics of Belousov-Zhabotinsky (BZ) reactions in a very stiff case in both time and space, yielding satisfactory results. As a consequence, a final numerical strategy is conceived that considers adaptive splitting time steps and that evaluates simultaneously critical time steps as well as best-suited  $\varepsilon$ , in order to guarantee error control for a given accuracy tolerance of the simulation with splitting time steps as large as possible. The relevance of the proposed strategy is first evaluated for the BZ reaction-diffusion equations, whereas a more complex problem issued from the simulation of multi-pulsed gas discharges involving several dynamics with very different typical time scales, constitutes the second test-case. It is shown that for this second very stiff reaction-diffusion system, splitting time steps can cover a range of three orders of magnitude and always guarantee a proper respect of the prescribed tolerance.

The paper is organized as follows: section 2 describes the adaptive time splitting strategy; in section 3, we perform the numerical analysis of the proposed method and identify the limit of validity of the local error estimate which is at the heart of the adapting procedure. Section 4 is devoted to the validation of the previous theoretical estimates and to a theoretical/numerical study of the critical splitting time steps in the context of a 1D reaction-diffusion problem featuring traveling wave solutions. In section 5 we present the final numerical strategy that includes an additional numerical procedure to evaluate critical time steps and suitable  $\varepsilon$ . The potential of the method is illustrated for the proposed two test-cases in section 6. We end in the last part with some concluding remarks.

## 2. Adaptive Time Splitting Method

Let us first set the general mathematical framework of this work. A class of multi-scale phenomena can be modeled by general reaction-diffusion systems of type:

$$\left. \begin{aligned} \partial_t \mathbf{u} - \partial_{\mathbf{x}} (\mathbf{D}(\mathbf{u}) \partial_{\mathbf{x}} \mathbf{u}) &= \mathbf{f}(\mathbf{u}), & \mathbf{x} \in \mathbb{R}^d, t > 0, \\ \mathbf{u}(0, \mathbf{x}) &= \mathbf{u}_0(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^d, t = 0, \end{aligned} \right\} \quad (2.1)$$

where  $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  and  $\mathbf{u} : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ , with a tensor of order  $d \times d \times m$  as diffusion matrix  $\mathbf{D}(\mathbf{u})$ .

In the following we will focus on the simplified case of linear diagonal diffusion, for which the elements of the diffusion matrix are written as  $D_{i_1 i_2 i_3}(\mathbf{u}) = D_{i_3} \delta_{i_1 i_2}$  for some positive indices  $i_1, i_2, i_3$ , so that the diffusion operator reduces to the heat operator with some scalar diffusion coefficient  $D_{i_3}$  for component  $u_{i_3}$  of  $\mathbf{u}$ . A scalar one-dimensional model is considered in order to simplify the presentation, taking

into account that extension into higher dimensions of  $\mathbf{x}$  or  $\mathbf{u}$  is straightforward:

$$\left. \begin{aligned} \partial_t u - \partial_x^2 u &= f(u), & x \in \mathbb{R}, t > 0, \\ u(0, x) &= u_0(x), & x \in \mathbb{R}, t = 0, \end{aligned} \right\} \quad (2.2)$$

where  $f$  and  $u_0$  are smooth functions. We denote by  $T^t u_0$  the solution of (2.2).

Introducing standard decoupling of the diffusion and reaction parts of (2.2), we denote by  $X^t u_0$  the solution of the diffusion equation:

$$\partial_t u_D - \partial_x^2 u_D = 0, \quad x \in \mathbb{R}, t > 0, \quad (2.3)$$

with initial data  $u_D(0, \cdot) = u_0(\cdot)$  after some time  $t$ ; and by  $Y^t u_0$ , the solution of the reaction part where spatial coordinate  $x$  can be considered as a parameter:

$$\partial_t u_R = f(u_R), \quad x \in \mathbb{R}, t > 0, \quad (2.4)$$

with  $u_R(0, \cdot) = u_0(\cdot)$ .

The two Lie approximation formulae of the solution of system (2.2) are then defined by

$$L_1^t u_0 = X^t Y^t u_0, \quad L_2^t u_0 = Y^t X^t u_0, \quad (2.5)$$

whereas the two Strang approximation formulae [25, 26] are given by

$$S_1^t u_0 = X^{t/2} Y^t X^{t/2} u_0, \quad S_2^t u_0 = Y^{t/2} X^t Y^{t/2} u_0. \quad (2.6)$$

It is well known that Lie formulae (2.5) (resp. Strang formulae (2.6)) are an approximation of order 1 (resp. 2) of the exact solution of (2.2). Higher order splitting schemes are also possible. Nevertheless, the order conditions for such composition methods state that either negative time substeps or complex coefficients or non convex combinations are necessary [13]. The formers imply usually important stability restrictions and more sophisticated numerical implementations. In the particular case of negative time steps, they are completely undesirable for PDEs that are ill-posed for negative time progression.

An adaptive time stepping strategy is based on a local error estimate which can be obtained by using two schemes of different order, in this case  $S_1^t$  or  $S_2^t$ , locally of order 3, and  $L_1^t$  or  $L_2^t$ , locally of order 2. For instance, the Embedded Split-Step Formulae given in [17] consider  $S_1^t$  and  $L_2^t$  or  $S_2^t$  and  $L_1^t$ , noticing that

$$L_1^t u_0 = X^t Y^{t/2} Y^{t/2} u_0,$$

where  $Y^{t/2} u_0$  is also used to compute  $S_2^t u_0$ . Nevertheless, in the context of multi-scale phenomena, order reductions may appear due to short-life transients associated with the fastest variables when one considers splitting time steps larger than the fastest scales. It has been proved in [5] that better performances are expected while ending the splitting scheme by the part involving the fastest time scales of the phenomenon. In particular, in the case of linear diagonal diffusion problems, no order loss is expected for the  $L_2^t$  and  $S_2^t$  schemes when fast scales are present in

6 *Descombes, Duarte, Dumont, Louwet, Massot*

the reactive term. Therefore, the embedding procedure must be carefully conceived taking into consideration these theoretical studies.

We introduce a shifted Strang formula

$$S_{2,\varepsilon}^t u_0 = Y^{(1/2-\varepsilon)t} X^t Y^{(1/2+\varepsilon)t} u_0, \quad (2.7)$$

locally of order 2, due to the lack of symmetry, for  $\varepsilon$  in  $[-1/2, 0) \cup (0, 1/2]$ . In this way, a local error estimate is computed based on two solutions for which orders are guaranteed and a potential loss of order is simultaneous, following

$$\begin{pmatrix} S_2^{\Delta t} u_0 \\ S_{2,\varepsilon}^{\Delta t} u_0 \end{pmatrix} = \begin{pmatrix} Y^{\Delta t/2} X^{\Delta t} Y^{\Delta t/2} u_0 \\ Y^{(1/2-\varepsilon)\Delta t} X^{\Delta t} Y^{(1/2+\varepsilon)\Delta t} u_0 \end{pmatrix}, \quad (2.8)$$

for some splitting time step  $\Delta t > 0$ . Embedding is accomplished as long as  $\varepsilon$  is different from  $-1/2$ , that is  $S_{2,\varepsilon}^{\Delta t} u_0$  different from  $L_2^t u_0$ . On the other hand, if  $\varepsilon$  is equal to  $1/2$ ,  $S_{2,\varepsilon}^{\Delta t} u_0$  is defined as  $L_1^t u_0$ , which it is not suitable for stiff configurations as it was previously discussed [5]. Therefore,  $\varepsilon$  should be contained in  $(-1/2, 0) \cup (0, 1/2)$ . Shifted  $S_{1,\varepsilon}^{\Delta t} u_0$  is defined in a similar way and depending on the multi-scale character of the problem, it might be the appropriate choice along with  $S_1^{\Delta t} u_0$ .

Taking into account that

$$\begin{aligned} S_2^{\Delta t} u_0 - S_{2,\varepsilon}^{\Delta t} u_0 &= S_2^{\Delta t} u_0 - T^{\Delta t} u_0 + T^{\Delta t} u_0 - S_{2,\varepsilon}^{\Delta t} u_0, \\ &= \mathcal{O}(\Delta t^3) + \mathcal{O}(\Delta t^2) \approx \mathcal{O}(\Delta t^2), \end{aligned} \quad (2.9)$$

for a given accuracy tolerance  $\eta$ ,

$$\|S_2^{\Delta t} u_0 - S_{2,\varepsilon}^{\Delta t} u_0\| < \eta \quad (2.10)$$

must be verified in order to accept current computation with  $\Delta t$ , while new time step is calculated by

$$\Delta t^{\text{new}} = v \Delta t \sqrt{\frac{\eta}{\|S_2^{\Delta t} u_0 - S_{2,\varepsilon}^{\Delta t} u_0\|}}, \quad (2.11)$$

with security factor  $0 < v \leq 1$  close to one. This comes from a classical adaptive time stepping procedure for stiff ODEs solution, for which more sophisticated formulae than (2.11) can be also considered, see [14] for example.

The error control of these adaptive methods is fully guaranteed as long as the orders of both, the main and the embedded integration methods, remains valid. This is the case for small enough time steps for which asymptotic theoretical estimates hold, but remains an open problem for larger time steps for which the validity of the formers is assumed. This is a key point in this work, because we propose not only a new splitting strategy with adaptive time steps as described in this section, but we aim also at applications for which splitting time steps may go beyond the fastest scales associated with each subproblem in order to obtain important computational savings. Therefore, a technique that guarantees consistently error control for all possible separation scales must be pursued, but first of all, a detailed numerical analysis of the method must be performed. This is the goal of the following part.

### 3. Numerical Analysis of the Method

In this part, we develop the numerical analysis of the proposed method. It is mainly based on the theoretical study of the introduced shifted Strang formula (2.7) and the domain of validity of the local error estimates. However, first of all, we introduce the Lie formalism which will be used as mathematical tool of analysis.

#### 3.1. The Lie operator formalism

We introduce the Lie operator formalism in order to generalize the exponential of a linear operator in the context of nonlinear operators. Let  $X$  be a Banach space,  $T_0 > 0$  and  $F$ , an unbounded nonlinear operator from  $D(F) \subset X$  to  $X$ , we consider the general autonomous equation:

$$\left. \begin{aligned} u'(t) &= F(u(t)), & 0 < t < T_0, \\ u(0) &= u_0, & t = 0. \end{aligned} \right\} \quad (3.1)$$

The exact solution of this evolutionary equation is (formally) given by

$$u(t) = T^t u_0, \quad 0 \leq t \leq T_0 \quad (3.2)$$

where  $T^t$  is the semiflow associated with (3.1); in particular we can set  $F(u) = \partial_x^2 u + f(u)$  as in (2.2). The Lie operator  $D_F$  associated with  $F$  is then a linear operator acting on the space of operators defined in  $X$  [13, 6]. More precisely, for any unbounded nonlinear operator  $G$  from  $D(G) \subset X$  to  $X$  with Fréchet derivative  $G'$ ,  $D_F$  maps  $G$  into a new operator  $D_F G$ , such that for any  $v$  in  $X$ :

$$(D_F G)(v) = G'(v)F(v). \quad (3.3)$$

Hence, by induction on  $n$  with solution  $u$  of (3.1), we obtain

$$\frac{\partial^n}{\partial t^n} G(u(t)) = (D_F^n G)(u(t)),$$

and a formal Taylor expansion yields

$$G(u(t)) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} \left( \frac{\partial^n}{\partial t^n} G(u(t)) \right) \Big|_{t=0} = \left( \sum_{n=0}^{+\infty} \frac{t^n}{n!} D_F^n G \right) u_0 = (e^{tD_F} G) u_0. \quad (3.4)$$

If we now assume that  $G$  is the identity operator  $\text{Id}$ , we obtain

$$u(t) = T^t u_0 = (e^{tD_F} \text{Id}) u_0.$$

Therefore, the Lie operator is indeed a way to write the solution of a nonlinear equation in terms of a linear but differential operator. Following (3.4), an important result obtained by Gröbner in 1960 [12], considers the composition of two semiflows  $T_1^t$  and  $T_2^s$  associated with  $F_1$  and  $F_2$  for any  $v$  in  $X$ :

$$T_1^t T_2^s v = (e^{sD_{F_2}} T_1^t) v = (e^{sD_{F_2}} e^{tD_{F_1}} \text{Id}) v.$$



### 3.2. Error analysis

In this paragraph, we conduct the error analysis of the approximation of  $T^t$  by  $S_{2,\varepsilon}^t$  in a linear framework. Then, we extend these results to a general nonlinear configuration given by problem (2.2), using the Lie operator formalism. General estimates for the approximation of  $T^t$  by  $S_2^t$  are also drawn. We end in the last part with a mathematical study that shows the domain of application of the method described in § 2, for which an effective error control is guaranteed within an accuracy tolerance. To simplify the notations in what follows, we will denote  $S_2^t$  by  $S^t$  and  $S_{2,\varepsilon}^t$  by  $S_\varepsilon^t$ .

Assume that  $A$  and  $B$  are linear bounded operators and define

$$S_\varepsilon^t u_0 = e^{(1/2-\varepsilon)tA} e^{tB} e^{(1/2+\varepsilon)tA} u_0$$

as an approximation of  $e^{t(A+B)}$ . The following theorem gives the expansion in powers of  $t$  of the difference between  $e^{t(A+B)}$  and  $S_\varepsilon^t$ . We recall the definition of the brackets between  $A$  and  $B$ :  $[A, B] = AB - BA$ .

**Theorem 3.1.** *Assume that  $A$  and  $B$  are linear bounded operators, for  $t$  and  $\varepsilon$  small enough, the following asymptotic holds*

$$e^{t(A+B)} u_0 - S_\varepsilon^t u_0 = -\varepsilon t^2 [A, B] u_0 + \frac{t^3}{24} ([A, [A, B]] + 2[B, [A, B]]) u_0 + \mathcal{O}(\varepsilon t^3) + \mathcal{O}(t^4).$$

**Proof.** Proof is straightforward by using the Taylor formula with integral remainder for a linear bounded operator  $A$ :

$$e^{tA} = \text{Id} + tA + \frac{t^2 A^2}{2} + \frac{t^3 A^3}{6} + \int_0^t \frac{(t-s)^3}{6} A^4 e^{sA} ds. \quad \square$$

We extend now the previous theorem to our nonlinear framework given by (2.2). In order to do this, we introduce the spaces  $C^\infty(\mathbb{R})$  of functions of class  $C^\infty$  on  $\mathbb{R}$ , and  $C_b^\infty(\mathbb{R})$  of functions of class  $C^\infty$  on  $\mathbb{R}$  and bounded over  $\mathbb{R}$ . We consider also the Schwartz space  $\mathcal{S}(\mathbb{R})$  defined by

$$\mathcal{S}(\mathbb{R}) = \{g \in C^\infty(\mathbb{R}) \mid \sup_{v \in \mathbb{R}} |v^{\alpha_1} \partial_v^{\alpha_2} g(v)| < \infty \text{ for all integers } \alpha_1, \alpha_2\};$$

and we define the space  $\mathcal{S}_1(\mathbb{R})$ , made out of functions  $v$  belonging to  $C_b^\infty(\mathbb{R})$  such that  $v'$  belongs to  $\mathcal{S}(\mathbb{R})$ . Let us consider now equation (2.2) and give the expansion in powers of  $t$  of the difference between  $T^t$  and  $S_\varepsilon^t$ , given by (2.7).

**Theorem 3.2.** *Assume that  $u_0$  belongs to  $\mathcal{S}_1(\mathbb{R})$  and that  $f$  belongs to  $C^\infty(\mathbb{R})$ .*

For  $t$  and  $\varepsilon$  small enough, the following asymptotic holds

$$\begin{aligned}
T^t u_0 - S_\varepsilon^t u_0 &= -\varepsilon t^2 f''(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 \\
&+ \frac{t^3}{24} (f'(u_0) f''(u_0) + f(u_0) f^{(3)}(u_0)) \left( \frac{\partial u_0}{\partial x} \right)^2 \\
&- \frac{t^3}{12} f^{(4)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^4 - \frac{t^3}{3} f^{(3)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 \frac{\partial^2 u_0}{\partial x^2} \\
&- \frac{t^3}{6} f''(u_0) \left( \frac{\partial^2 u_0}{\partial x^2} \right)^2 + \mathcal{O}(\varepsilon t^3) + \mathcal{O}(t^4). \tag{3.5}
\end{aligned}$$

**Proof.** We introduce the two Lie operators  $D_\Delta$  and  $D_f$  associated with  $\partial_x^2$  and  $f$  and write

$$T^t u_0 - S_\varepsilon^t u_0 = \left( e^{t(D_\Delta + D_f)} \text{Id} \right) u_0 - \left( e^{(1/2+\varepsilon)t D_f} e^{t D_\Delta} e^{(1/2-\varepsilon)t D_f} \text{Id} \right) u_0.$$

With Theorem 3.1 we can deduce that

$$\begin{aligned}
T^t u_0 - S_\varepsilon^t u_0 &= -\varepsilon t^2 ([D_f, D_\Delta] \text{Id}) u_0 + \frac{t^3}{24} ([D_f, [D_f, D_\Delta]] \text{Id}) u_0 \\
&+ \frac{t^3}{12} ([D_\Delta, [D_f, D_\Delta]] \text{Id}) u_0 + \mathcal{O}(\varepsilon t^3) + \mathcal{O}(t^4). \tag{3.6}
\end{aligned}$$

We are not interested in giving the exact form of the terms  $\mathcal{O}(\varepsilon t^3)$  and  $\mathcal{O}(t^4)$ , but these terms can be computed following the same technique developed in [6]. For the term in  $\mathcal{O}(t^2)$ , we have by definition and with (3.3),

$$\begin{aligned}
([D_f, D_\Delta] \text{Id}) u_0 &= (D_f(D_\Delta \text{Id}) - D_\Delta(D_f \text{Id})) u_0, \\
&= (D_\Delta \text{Id})'(u_0) f(u_0) - (D_f \text{Id})'(u_0) \frac{\partial^2 u_0}{\partial x^2}, \\
&= \frac{\partial^2}{\partial x^2} (f(u_0)) - f'(u_0) \frac{\partial^2 u_0}{\partial x^2}.
\end{aligned}$$

The last term is by definition the Lie bracket between  $\partial_x^2$  and  $f$ , a simple computation shows that

$$\begin{aligned}
\frac{\partial^2 f(u_0)}{\partial x^2} - f'(u_0) \frac{\partial^2 u_0}{\partial x^2} &= f''(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 + f'(u_0) \frac{\partial^2 u_0}{\partial x^2} - f'(u_0) \frac{\partial^2 u_0}{\partial x^2}, \\
&= f''(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2.
\end{aligned}$$

Furthermore,

$$([D_f, [D_f, D_\Delta]] \text{Id}) (u_0) = (f'(u_0) f''(u_0) + f(u_0) f^{(3)}(u_0)) \left( \frac{\partial u_0}{\partial x} \right)^2$$

10 *Descombes, Duarte, Dumont, Louwet, Massot*

and

$$\begin{aligned} ([D_\Delta, [D_f, D_\Delta]]\text{Id}) u_0 &= -f^{(4)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^4 - 4f^{(3)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 \frac{\partial^2 u_0}{\partial x^2} \\ &\quad - 2f''(u_0) \left( \frac{\partial^2 u_0}{\partial x^2} \right)^2. \end{aligned}$$

All the terms are now computed and this concludes the proof of Theorem 3.2.  $\square$

For  $\varepsilon = 0$ , the next corollary follows directly.

**Corollary 3.1.** *Assume that  $u_0$  belongs to  $\mathcal{S}_1(\mathbb{R})$  and that  $f$  belongs to  $C^\infty(\mathbb{R})$ . For  $t$  small enough, the following asymptotic holds*

$$\begin{aligned} T^t u_0 - S^t u_0 &= \frac{t^3}{24} (f'(u_0)f''(u_0) + f(u_0)f^{(3)}(u_0)) \left( \frac{\partial u_0}{\partial x} \right)^2 \\ &\quad - \frac{t^3}{12} f^{(4)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^4 - \frac{t^3}{3} f^{(3)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 \frac{\partial^2 u_0}{\partial x^2} \\ &\quad - \frac{t^3}{6} f''(u_0) \left( \frac{\partial^2 u_0}{\partial x^2} \right)^2 + \mathcal{O}(t^4). \end{aligned} \quad (3.7)$$

From (3.5) and (3.7), we can see that

$$S^t u_0 - S_\varepsilon^t u_0 = \varepsilon t^2 f''(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 + \mathcal{O}(\varepsilon t^3), \quad (3.8)$$

and thus,

$$T^t u_0 - S_\varepsilon^t u_0 = \underbrace{T^t u_0 - S^t u_0}_{\mathcal{O}(t^3)} + \underbrace{S^t u_0 - S_\varepsilon^t u_0}_{\mathcal{O}(\varepsilon t^2)}. \quad (3.9)$$

Therefore, we are sure that the real local error of the method,  $T^t u_0 - S^t u_0$ , will be bounded by the local error estimate,  $err = S^t u_0 - S_\varepsilon^t u_0$ , when for a given  $\varepsilon$ ,

$$T^t u_0 - S_\varepsilon^t u_0 \approx \mathcal{O}(t^2) \quad (3.10)$$

is verified into (3.9); that is, when the embedded method is really of lower order as it was assumed in (2.9). This will be always verified for small enough time steps  $t$ , for which  $T^t u_0 - S^t u_0 \approx \mathcal{O}(t^3) < err \approx \mathcal{O}(\varepsilon t^2)$  is guaranteed. Nevertheless, for larger time steps,  $err$  will fail to properly predict  $T^t u_0 - S^t u_0$  since we will eventually have  $T^t u_0 - S^t u_0 \approx \mathcal{O}(t^3) > err \approx \mathcal{O}(\varepsilon t^2)$ . When this happens, (3.10) is no longer true and the previous estimates show that we will rather have  $T^t u_0 - S_\varepsilon^t u_0 \approx \mathcal{O}(t^3)$ , and assumption (2.9) will no longer hold.

In order to overcome this difficulty, we must therefore estimate a critical time step  $t^* > 0$  such that for all  $t$  in  $[0, t^*]$ , (3.10) is guaranteed for a given  $\varepsilon$ . This will imply that Strang local error,  $T^t u_0 - S^t u_0$ , will be indeed bounded by the local error estimate,  $err$ , and that an effective error control will be achieved for  $err$  smaller than a given accuracy tolerance  $\eta$ . Finally, a suitable choice of  $\varepsilon$  can be also made since  $t^*$  is related to  $\varepsilon$  following (3.9).

A natural strategy to predict this critical  $t^*$  will rely on the previous theoretical estimates and on a more precise knowledge of the structure of the solutions of the PDEs; this is for instance illustrated in the next part in the context of traveling wave solutions.

#### 4. Application to Reaction Traveling Waves

In this part, we will confront the previous theoretical study to a simple reaction diffusion problem that admits self-similar traveling wave solutions such as the KPP equation [18]. The main advantages of considering this kind of problems are that analytic solutions exist and that the featured stiffness can be tuned using a space-time scaling. Therefore, it provides a first numerical validation of the numerical estimates of the method and an evaluation of its domain of application; and on the other hand, a detailed study can be conducted on the impact of the stiffness featured by propagating fronts with steep spatial gradients.

In what follows, we recast previous estimates in the context of these reaction traveling waves, to then deduce an estimate of the time step  $t^*$  that defines the limit of application of the method for which local error estimates yield effective error control. We end with a numerical validation of the theoretical results in the context of the resolution of KPP model.

##### 4.1. Numerical estimates

We are interested in the propagation of self-similar waves modeled by parabolic PDEs of type:

$$\left. \begin{aligned} \partial_t u - D \partial_x^2 u &= kf(u), & x \in \mathbb{R}, t > 0, \\ u(0, x) &= u_0(x), & x \in \mathbb{R}, t = 0, \end{aligned} \right\} \quad (4.1)$$

with solution  $u(x, t) = u_0(x - ct)$ , where  $c$  is the steady speed of the wavefront, and  $D$  and  $k$  stand respectively for diffusion and reaction coefficients.

Considering Theorem 3.2 we obtain the following estimate for system (4.1).

**Corollary 4.1.** *Assume that  $u_0$  belongs to  $\mathcal{S}_1(\mathbb{R})$  and that  $f$  belongs to  $C^\infty(\mathbb{R})$ . For  $t$  and  $\varepsilon$  small enough, the following asymptotic holds*

$$\begin{aligned} T^t u_0 - S_\varepsilon^t u_0 &= -\varepsilon k D t^2 f''(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 \\ &+ \frac{k^2 D t^3}{24} (f'(u_0) f''(u_0) + f(u_0) f^{(3)}(u_0)) \left( \frac{\partial u_0}{\partial x} \right)^2 \\ &- \frac{k D^2 t^3}{12} f^{(4)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^4 - \frac{k D^2 t^3}{3} f^{(3)}(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 \frac{\partial^2 u_0}{\partial x^2} \\ &- \frac{k D^2 t^3}{6} f''(u_0) \left( \frac{\partial^2 u_0}{\partial x^2} \right)^2 + \mathcal{O}(\varepsilon t^3) + \mathcal{O}(t^4). \end{aligned} \quad (4.2)$$

12 *Descombes, Duarte, Dumont, Louwet, Massot*

**Proof.** Proof follows directly from demonstration of Theorem 3.2, using (3.6) and considering that

$$\begin{aligned} [D_{kf}, D_{D\Delta}] &= kD[D_f, D_\Delta], \\ [[D_{kf}, D_{D\Delta}], D_{D\Delta}] &= kD^2[[D_f, D_\Delta], D_\Delta], \\ [[D_{kf}, D_{D\Delta}], D_{kf}] &= k^2D[[D_f, D_\Delta], D_f], \end{aligned}$$

where  $D_{D\Delta}$  and  $D_{kf}$  are the Lie operators associated with  $D\partial_x^2$  and  $kf$ .  $\square$

On the other hand, if we now consider system (4.1) with  $k = 1$  and  $D = 1$ , the following corollary establishes  $t^* > 0$  such that for all  $t$  in  $[0, t^*]$  (3.10) is guaranteed for a given  $\varepsilon$ .

**Corollary 4.2.** *Assume that  $u_0$  belongs to  $\mathcal{S}_1(\mathbb{R})$  and that  $f$  belongs to  $C^\infty(\mathbb{R})$ . For a given  $\varepsilon$  small enough, define*

$$M_1 = \left\| f''(u_0) \left( \frac{\partial u_0}{\partial x} \right)^2 \right\|_{L^2} \quad (4.3)$$

and

$$\begin{aligned} M_2 = \left\| \frac{f'(u_0)f''(u_0) + f(u_0)f^{(3)}(u_0)}{24} \left( \frac{\partial u_0}{\partial x} \right)^2 - \frac{f^{(4)}(u_0)}{12} \left( \frac{\partial u_0}{\partial x} \right)^4 \right. \\ \left. - \frac{f^{(3)}(u_0)}{3} \left( \frac{\partial u_0}{\partial x} \right)^2 \frac{\partial^2 u_0}{\partial x^2} - \frac{f''(u_0)}{6} \left( \frac{\partial^2 u_0}{\partial x^2} \right)^2 \right\|_{L^2}, \quad (4.4) \end{aligned}$$

define  $t^*$  by

$$t^* M_2 = \varepsilon M_1. \quad (4.5)$$

For all  $t$  such that  $0 < t \leq t^*$  then

$$\|T^t u_0 - S_\varepsilon^t u_0\|_{L^2} \approx \mathcal{O}(t^2).$$

In a general case, if evaluation of the derivatives of  $u_0$  and  $f$  is feasible, it is then possible to predict the domain of application of the method,  $[0, t^*]$ , for a given  $\varepsilon$  based on the previous result. In the particular case of traveling wave solutions for (4.1), diffusion and reaction coefficients,  $D$  and  $k$ , might be seen as scaling coefficients in time and space. A dimensionless analysis of a traveling wave, as shown in [11], can be then conducted considering a dimensionless time  $\tau$  and a dimensionless space  $r$  with

$$\tau = kt \quad \text{and} \quad r = (k/D)^{1/2}x.$$

This analysis allows to find a steady velocity of the wavefront,

$$c = x_t \propto (Dk)^{1/2}, \quad (4.6)$$

whereas the sharpness of the wave profile is measured by

$$u_x|_{\max} \propto (k/D)^{1/2}. \quad (4.7)$$

Therefore, condition  $Dk = 1$  implies constant velocity for all  $k = 1/D$  but greater  $k$  (or smaller  $D$ ) implies higher spatial gradients, and thus, stiffer configurations.

This study gives complementary information on the solution of (4.1) and in particular, when condition  $Dk = 1$  is satisfied, it allows to deduce from Corollary 4.2:

$$kt^*M_2 = \varepsilon M_1, \quad (4.8)$$

with  $M_1$  and  $M_2$  given by (4.3) and (4.4). Therefore, stiffer configurations given by the presence of steeper spatial gradients will restrain the application domain of the method, according to (4.8). Nevertheless, for larger gradients, smaller time steps are also required for a given level of accuracy and hence, we can expect a simultaneous reduction of both critical and accurate splitting time steps.

#### 4.2. Numerical illustration: KPP equation

Let us recall the Kolmogorov-Petrovskii-Piskunov model. In their original paper [18], these authors introduced a model describing the propagation of a virus and the first rigorous analysis of a stable traveling wave solution of a nonlinear reaction-diffusion equation [11]. The equation is the following:

$$\partial_t u - D \partial_x^2 u = k u^2(1 - u), \quad (4.9)$$

with homogeneous Neumann boundary conditions. We consider a 1D discretization with 5001 points on a  $[-70, 70]$  region for which we have negligible spatial discretization errors with respect to the ones coming from the numerical time integration.

The description of the dimensionless model and the structure of the exact solution can be found in [11] where the dimensionless analysis shows that in the case of  $D = 1$  and  $k = 1$ , the velocity of the self-similar traveling wave is  $c = 1/\sqrt{2}$  and the maximal gradient value reaches  $1/\sqrt{32}$ . The key point of this illustration is that the velocity of the traveling wave is proportional to  $(kD)^{1/2}$ , whereas the maximal gradient is proportional to  $(k/D)^{1/2}$ . Hence, we consider the case  $kD = 1$  for which one may obtain steeper gradients for the same speed of propagation.

Throughout all this paper, exact solution  $T^t u_0$  will be approximated by the resolution of the coupled reaction-diffusion problem performed by the Radau5 method [15] with fine tolerances,  $\eta_{Radau5} = 10^{-10}$ . This solution will be referred as the reference or *quasi-exact* solution. Strang approximations  $S^t u_0$  and  $S_\varepsilon^t u_0$  will be computed with a splitting technique recently introduced [8, 9], which considers Radau5 [15] to solve locally point by point the reaction term; and the ROCK4 method [1] for the diffusion problem. Radau5 [15] is a fifth order implicit Runge-Kutta method exhibiting  $A$ - and  $L$ -stability properties to efficiently solve stiff systems of ODEs, whereas ROCK4 [1] is formally a fourth order *stabilized* explicit Runge-Kutta method with extended stability domain along the negative real axis, well suited to numerically treat mildly stiff elliptic operators. Both methods implement adaptive time stepping techniques to guarantee computations within a prescribed accuracy tolerance.

14 *Descobes, Duarte, Dumont, Louwet, Massot*

In order to properly discriminate the previously estimated splitting errors from those coming from temporal integration of the substeps, we consider also fine tolerances,  $\eta_{\text{Radau5}} = \eta_{\text{ROCK4}} = 10^{-10}$ .

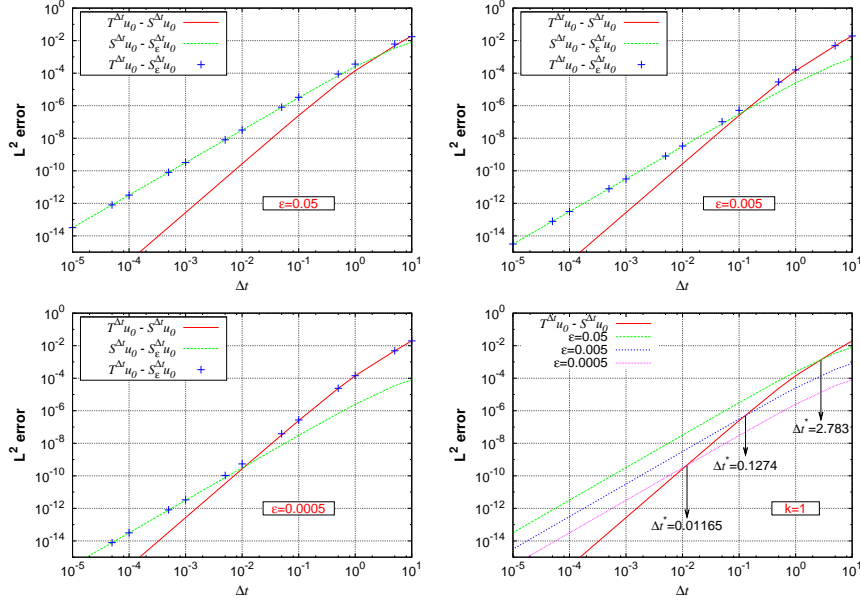


Fig. 1. KPP equation with  $k = 1$ . Local  $L^2$  errors for several splitting time steps  $\Delta t$  and  $\varepsilon = 0.05$  (top left), 0.005 (top right) and 0.0005 (bottom left). Bottom right: critical splitting time steps  $\Delta t^*$  obtained when  $\|T^{\Delta t}u_0 - S^{\Delta t}u_0\|_{L^2} \approx \|S^{\Delta t}u_0 - S_{\varepsilon}^{\Delta t}u_0\|_{L^2}$  in the numerical tests.

Figures 1 and 2 show  $L^2$  errors between  $T^t u_0$ ,  $S^t u_0$  and  $S_{\varepsilon}^t u_0$  solutions for  $k = 1$ ,  $k = 10$  and  $k = 100$  respectively, and several  $\varepsilon$ . Notice that estimates (3.5), (3.7) and (3.8) for all three errors in (3.9) are verified and in particular, for  $\Delta t$  larger than critical  $\Delta t^*$ , the estimated error  $err = \|S^{\Delta t}u_0 - S_{\varepsilon}^{\Delta t}u_0\|_{L^2}$  is no longer predicting the real local error given by  $T^t u_0 - S^t u_0$ .

With these results, we can also compare real  $\Delta t^*$ , obtained when  $\|T^{\Delta t}u_0 - S^{\Delta t}u_0\|_{L^2} \approx \|S^{\Delta t}u_0 - S_{\varepsilon}^{\Delta t}u_0\|_{L^2}$  in the numerical tests, with theoretically estimated  $\Delta t^*$  following (4.8). Table 1 summarizes these results where computation of estimated  $\Delta t^*$  in (4.8) is given by the computation of  $M_1$  and  $M_2$  with Maple<sup>©</sup> according to (4.3) and (4.4). A really good agreement can be observed even though theoretical results underestimate the real values. The loss of order depicted by the numerical results, is due to the influence of spatial gradients in the solution, as it was proven in [3]. This explains the error of the predicted critical  $\Delta t^*$  in (4.8) whenever one gets close to the order loss region.

Numerical results show also that  $\|S^{\Delta t}u_0 - S_{\varepsilon}^{\Delta t}u_0\|_{L^2} \propto \varepsilon$  according to (3.8) and consequently,  $\Delta t^* \propto \varepsilon$ ; therefore, the working region or domain of application

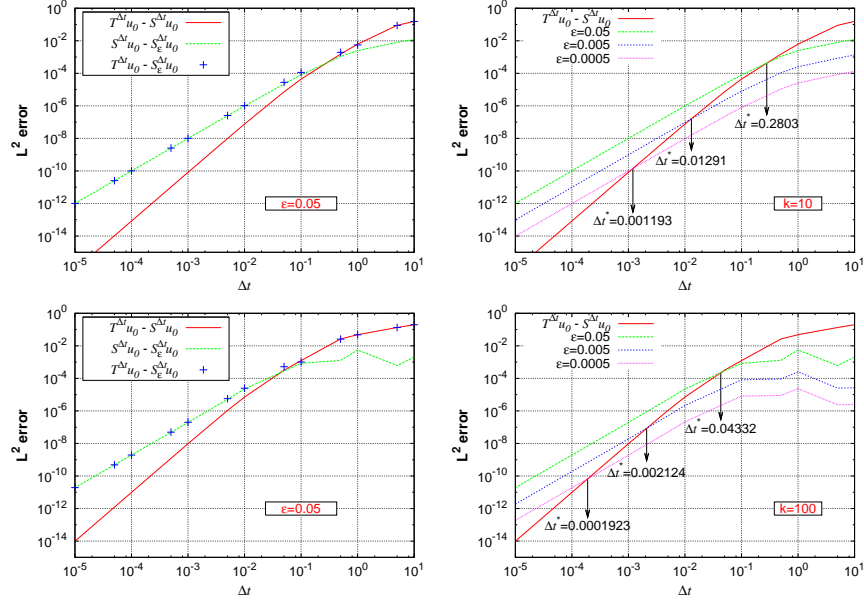


Fig. 2. KPP equation with  $k = 10$  (top) and  $k = 100$  (bottom). Local  $L^2$  errors for several splitting time steps  $\Delta t$  and  $\varepsilon = 0.05$  (left). Right: critical splitting time steps  $\Delta t^*$  obtained when  $\|T^{\Delta t} u_0 - S^{\Delta t} u_0\|_{L^2} \approx \|S^{\Delta t} u_0 - S_{\varepsilon}^{\Delta t} u_0\|_{L^2}$  in the numerical tests.

of the method,  $\Delta t < \Delta t^*$ , depends directly on the choice of  $\varepsilon$  as it can be seen in Table 1. Finally, in the context of traveling waves, these numerical experiments show that  $\Delta t^* \propto k^{-1} \propto 1/\|\partial u_0/\partial x\|_{\infty}$  according to Table 1; hence, application domains are reduced for stiffer configurations but numerical results show also that smaller time steps are required for the same level of accuracy. These conclusions are easily extrapolated to more general self-similar propagating waves.

Table 1. KPP equation. Comparison between real  $\Delta t_{\text{real}}^*$ , obtained when  $\|T^{\Delta t} u_0 - S^{\Delta t} u_0\|_{L^2} \approx \|S^{\Delta t} u_0 - S_{\varepsilon}^{\Delta t} u_0\|_{L^2}$  in the numerical tests, and theoretically estimated  $\Delta t_{\text{est}}^*$  following (4.8).

		$\varepsilon = 0.05$	$\varepsilon = 0.005$	$\varepsilon = 0.0005$
$k = 1$	$\Delta t_{\text{real}}^*$	2.783	0.1274	$1.17 \times 10^{-2}$
	$\Delta t_{\text{est}}^*$	1.107	0.1107	$1.11 \times 10^{-2}$
$k = 10$	$\Delta t_{\text{real}}^*$	0.2803	$1.29 \times 10^{-2}$	$1.19 \times 10^{-3}$
	$\Delta t_{\text{est}}^*$	0.1107	$1.11 \times 10^{-2}$	$1.11 \times 10^{-3}$
$k = 100$	$\Delta t_{\text{real}}^*$	$4.33 \times 10^{-2}$	$2.12 \times 10^{-3}$	$1.92 \times 10^{-4}$
	$\Delta t_{\text{est}}^*$	$1.11 \times 10^{-2}$	$1.11 \times 10^{-3}$	$1.11 \times 10^{-4}$



## 5. Construction of the Numerical Strategy

We have presented in § 2, a time adaptive numerical scheme fully based on theoretical error estimates developed in § 3. We have also studied the necessary general conditions in order to guarantee an effective error control based on local error estimates. In particular, this has been shown in the case of reaction traveling waves in § 4, for which theoretical studies give us some insight into the PDE solution. Nevertheless, this is not always possible and it is usually difficult to carry out such kind of analysis for more realistic models. Therefore, based on the theoretical analysis and previous illustrations on the influence of the various parameters of the scheme, a general numerical procedure that completes the adaptive scheme defined in § 2, is introduced in the following.

In a first part, we will settle the theoretical framework and the numerical procedure needed to estimate  $t^*$ , and to define the appropriate  $\varepsilon$ . This will be illustrated by numerical tests performed on a more complex model of time-space stiff propagating waves. These theoretical and numerical studies will allow to define, at the end, a final numerical strategy.

### 5.1. Numerical procedure to estimate critical $t^*$ and $\varepsilon$

Let us consider general system (2.2), based on theoretical estimates (3.7) and (3.8), we can write

$$S^{\Delta t}u_0 - T^{\Delta t}u_0 = C_0\Delta t^3, \quad (5.1)$$

where  $C_0 = C_1(u_0) + \mathcal{O}(\Delta t^4)$ , and

$$S^{\Delta t}u_0 - S_\varepsilon^{\Delta t}u_0 = \varepsilon C_\varepsilon \Delta t^2, \quad (5.2)$$

where  $C_\varepsilon = C_2(u_0) + \mathcal{O}(\varepsilon, \Delta t^3)$ ; the dependence of  $C_\varepsilon$  on  $\varepsilon$  is only given in the higher order terms and it is thus neglected.

For a given  $\varepsilon$ , in the same spirit as Corollary 4.2, we search for a critical  $\Delta t^*$  such that

$$\|S^{\Delta t}u_0 - T^{\Delta t}u_0\| \leq \|S^{\Delta t}u_0 - S_\varepsilon^{\Delta t}u_0\| \quad (5.3)$$

for all  $\Delta t \leq \Delta t^*$ . According to (5.1) and (5.2), we have then the following estimate:

$$\Delta t^* \approx \frac{\varepsilon C_\varepsilon}{C_0}. \quad (5.4)$$

For a given  $\varepsilon$ , this gives an upper bound for the time steps for which the local error estimate,  $err = \|S^{\Delta t}u_0 - S_\varepsilon^{\Delta t}u_0\|$ , is properly estimating the real Strang local error,  $\|S^{\Delta t}u_0 - T^{\Delta t}u_0\|$ , following (5.3).

In particular, when  $\Delta t \rightarrow \Delta t^*$ , we have that  $err \approx \|S^{\Delta t}u_0 - T^{\Delta t}u_0\|$ , and the local error estimate is predicting more accurately the real error of integration. The critical time step,  $\Delta t^*$ , is directly related to  $\varepsilon$  through (5.4) as we have already shown in the previous numerical results in § 4.2. Therefore, a suitable  $\varepsilon$  will define a critical  $\Delta t^*$  such that the estimated splitting time steps  $\Delta t$  for a given tolerance  $\eta$

will be close enough to critical  $\Delta t^*$ , in order to avoid an excessive overestimation of the Strang local error and thus, larger time steps can be chosen for a given accuracy tolerance  $\eta$ .

In order to compute  $\Delta t^*$  for a given  $\varepsilon$ , we must first estimate  $C_0$  in (5.4), since  $C_\varepsilon$  is computed out of the local error estimate,  $err$ , for known  $\Delta t$  and  $\varepsilon$  in (5.2). Estimating  $C_0$  amounts to directly estimate Strang local error through (5.1) and thus, the accuracy of the simulation might be controlled in this way without relying on a local error estimate as proposed in the embedded method strategy in § 2. Nevertheless, as we will see in the following, in order to estimate  $C_0$  and the Strang local error, we must define new local estimators and a numerical procedure that becomes rapidly very expensive if we want to implement such error control technique. Therefore, we must rely on a local error estimate given by a less expensive strategy for which the computation of  $C_0$  is only performed from time to time to guarantee the validity of local error estimates.

The next Lemma will be useful to define the numerical procedure to estimate  $C_0$ .

**Lemma 5.1.** *Let us consider system (2.2) and assume a local Lipschitz condition for  $f$ :*

$$\|f(u) - f(v)\| \leq \lambda \|u - v\|. \quad (5.5)$$

For a finite  $\Delta t$  the following holds

$$\|T^{\Delta t} u_0 - T^{\Delta t} v_0\| \leq \omega \|u_0 - v_0\|, \quad (5.6)$$

with  $\omega = 1 + \kappa \Delta t$  for small enough  $\Delta t$ .

**Proof.** Using Duhamel's formula for (2.2) yields

$$T^t u_0 - T^t v_0 = e^{t\partial_x^2} (u_0 - v_0) + \int_0^t e^{(t-s)\partial_x^2} (f(T^s u_0) - f(T^s v_0)) ds. \quad (5.7)$$

Taking norms and applying recursively (5.7),

$$\begin{aligned} \|T^t u_0 - T^t v_0\| &\leq \|u_0 - v_0\| + \lambda \int_0^t \|T^s u_0 - T^s v_0\| ds, \\ &\leq e^{\lambda t} \|u_0 - v_0\|, \end{aligned} \quad (5.8)$$

proves (5.6) for  $t = \Delta t$  finite.  $\square$

If we define a local estimator,  $e_1 = S^{a_1 \Delta t} u_0 - S^{b_1 \Delta t} (S^{c_1 \Delta t} u_0)$ , such that  $a_1 = b_1 + c_1$ , we obtain that

$$\begin{aligned} S^{b_1 \Delta t} (S^{c_1 \Delta t} u_0) - T^{a_1 \Delta t} u_0 &= S^{b_1 \Delta t} (S^{c_1 \Delta t} u_0) - T^{b_1 \Delta t} (S^{c_1 \Delta t} u_0) \\ &\quad + T^{b_1 \Delta t} (S^{c_1 \Delta t} u_0) - T^{b_1 \Delta t} (T^{c_1 \Delta t} u_0), \\ &= C_{S^{c_1 \Delta t} u_0} b_1^3 \Delta t^3 \\ &\quad + T^{b_1 \Delta t} (S^{c_1 \Delta t} u_0) - T^{b_1 \Delta t} (T^{c_1 \Delta t} u_0), \end{aligned} \quad (5.9)$$

18 *Descombes, Duarte, Dumont, Louwet, Massot*

where  $C_{S^{c_1\Delta t}u_0} = C_1(S^{c_1\Delta t}u_0) + \mathcal{O}(\Delta t^4)$ . Therefore, assuming that  $C_{S^{c_1\Delta t}u_0} \approx C_0$  and considering Lemma 5.1, it follows from the difference between (5.1) at  $a_1\Delta t$  and (5.9):

$$\begin{aligned} \|e_1 - (a_1^3 - b_1^3)C_0\Delta t^3\| &\leq \omega \|T^{c_1\Delta t}u_0 - S^{c_1\Delta t}u_0\|, \\ &\leq \omega C_0 c_1^3 \Delta t^3. \end{aligned} \quad (5.10)$$

Hence, defining a second local estimator,  $e_2 = S^{a_2\Delta t}u_0 - S^{b_2\Delta t}(S^{c_2\Delta t}u_0)$ , such that  $a_2 = b_2 + c_2$ , we obtain a second expression similar to (5.10) with  $e_2$  and  $(a_2, b_2, c_2)$ , and we can estimate  $C_0$  and  $\omega$ . In particular, we notice that  $b_1$  should be close to  $b_2$  in order to better approximate  $\omega$  into (5.6) and (5.10), and that  $c_1$  and  $c_2$  should also be small enough to guarantee  $C_{S^{c_1\Delta t}u_0} \approx C_0$  and  $C_{S^{c_2\Delta t}u_0} \approx C_0$ . On the other hand, to optimize the required number of extra computations from a practical point of view, we can use the estimator  $e_2$  to compute estimator  $e_1$  by setting  $a_2 = c_1$ , and we can also fix  $a_1 = 1$  so we can use  $S^{a_1\Delta t}u_0$  for the time integration of the problem. In this way, the extra computations needed to compute local estimators  $e_1$  and  $e_2$  will be given by  $S^{c_2\Delta t}u_0$ ,  $S^{b_2\Delta t}(S^{c_2\Delta t}u_0)$ ,  $S^{c_1\Delta t}u_0$  and  $S^{b_1\Delta t}(S^{c_1\Delta t}u_0)$  within a time step  $\Delta t$ . Then, we will be able to compute  $\omega$  and  $C_0$ , by solving two expressions of type (5.10). The next numerical example illustrates the validity of this numerical procedure.

## 5.2. Numerical example of evaluation of critical $t^*$ : BZ equation

We are concerned with the numerical approximation of a model of the Belousov-Zhabotinski reaction, a catalyzed oxidation of an organic species by acid bromated ion (for more details and illustrations, see [10]). We thus consider the model introduced in [11] and coming from the classic work of Field, Koros and Noyes (FKN) (1972), which takes into account three species:  $\text{HBrO}_2$  (hypobromous acid), bromide ions  $\text{Br}^-$  and cerium(IV). Denoting by  $a = [\text{Ce(IV)}]$ ,  $b = [\text{HBrO}_2]$  and  $c = [\text{Br}^-]$ , we obtain a very stiff system of three partial differential equations:

$$\left. \begin{aligned} \partial_t a - D_a \partial_x^2 a &= \frac{1}{\mu} (-qa - ab + fc), \\ \partial_t b - D_b \partial_x^2 b &= \frac{1}{\epsilon} (qa - ab + b(1 - b)), \\ \partial_t c - D_c \partial_x^2 c &= b - c, \end{aligned} \right\} \quad (5.11)$$

with diffusion coefficients  $D_a$ ,  $D_b$  and  $D_c$ , and some real positive parameters  $f$ , small  $q$ , and small  $\epsilon$ ,  $\mu$ , such that  $\mu \ll \epsilon$ .

The dynamical system associated with this system models reactive excitable media with a large time scale spectrum (see [11] for more details). Moreover, the spatial configuration with addition of diffusion generates propagating wavefronts with steep spatial gradients. Hence, this model presents all the difficulties associated with a stiff time-space multi-scale configuration. The advantages of applying a splitting strategy to these models have already been studied and presented in [4].

We consider the 1D application of problem (5.11) with homogeneous Neumann boundary conditions in a space region of  $[0, 80]$  with a spatial discretization of 4001 points, good enough to prevent important spatial discretization errors, and the following parameters, taken from [11]:  $\epsilon = 10^{-2}$ ,  $\mu = 10^{-5}$ ,  $f = 3$  and  $q = 2 \times 10^{-4}$ , with diffusion coefficients  $D_a = 1$ ,  $D_b = 1$  and  $D_c = 0.6$ . Reference solution and Strang approximations are defined in the same way as in the KPP application with the same tolerances for the time integration solvers.

First of all, we validate theoretical order estimates (3.5), (3.7) and (3.8) and verify relation (3.9). Figure 3 shows  $L^2$  errors between  $T^t u_0$ ,  $S^t u_0$  and  $S_\epsilon^t u_0$  solutions for several  $\epsilon$  and the real  $\Delta t^*$  such that  $\|T^{\Delta t} u_0 - S^{\Delta t} u_0\|_{L^2} \approx \|S^{\Delta t} u_0 - S_\epsilon^{\Delta t} u_0\|_{L^2}$ , obtained after treating the numerical results. Maximum  $L^2$  error considers the maximum value between normalized local errors for  $a$ ,  $b$  and  $c$  variables; in these numerical tests, it corresponds usually to variable  $b$ .

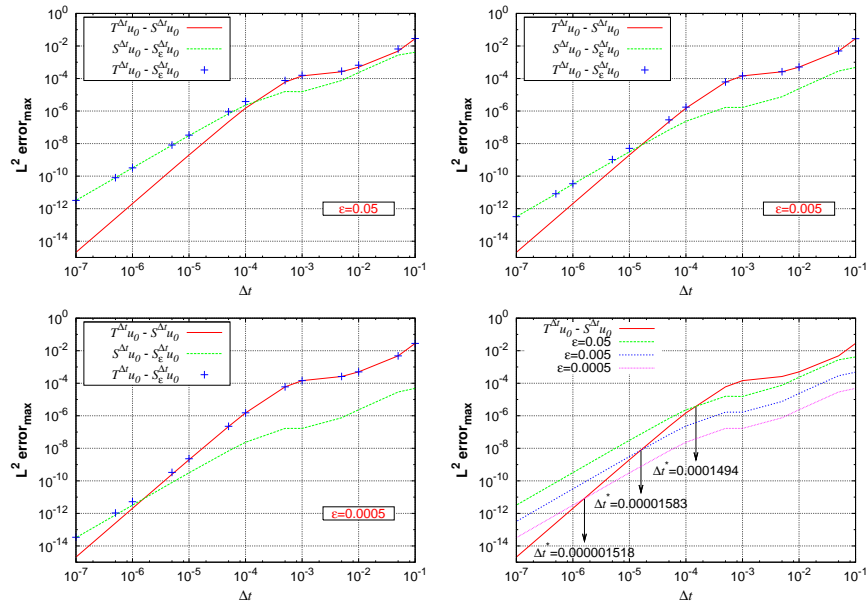


Fig. 3. BZ equation. Maximum local  $L^2$  errors for several splitting time steps  $\Delta t$  and  $\epsilon = 0.05$  (top left), 0.005 (top right) and 0.0005 (bottom left). Bottom right: critical splitting time steps  $\Delta t^*$  obtained when  $\|T^{\Delta t} u_0 - S^{\Delta t} u_0\|_{L^2} \approx \|S^{\Delta t} u_0 - S_\epsilon^{\Delta t} u_0\|_{L^2}$  in the numerical tests.

Let us now define the two sets  $(a_1, b_1, c_1)$  and  $(a_2, b_2, c_2)$ , and compute local estimators  $e_1$  and  $e_2$  in order to obtain  $C_0$  according to (5.10) with  $\Delta t = \Delta t_0 = 10^{-5}$ ; that is a time step for which there is no order loss yet, as seen in Figure 3. As it was previously detailed, we consider  $a_1 = 1$  and  $a_2 = c_1$  to avoid some extra computations. Furthermore,  $b_2$  should be close to  $b_1$ , and  $c_1$  and  $c_2$  small enough. Setting  $b_1$  larger than  $1/2$  would yield more different  $b_2$  since  $c_1 = a_2$ . On

20 *Descombes, Duarte, Dumont, Louwet, Massot*

the other hand, for  $b_1$  smaller than  $1/2$  we can even set  $b_2 = b_1$  but in this case  $c_1$  will be larger than  $1/2$ . Therefore, we reach a compromise by setting  $b_1 = 1/2$  that yields  $c_1 = a_2 = 1/2$ , so we can choose for instance  $b_2 = 2/5$  close to  $b_1$ , and thus,  $c_2 = 1/10$ .

With the local error estimate,  $err = \|S^{\Delta t}u_0 - S_\varepsilon^{\Delta t}u_0\|_{L^2}$ , for the various time steps and several  $\varepsilon$  shown in Figure 3, Figure 4 presents the estimated critical  $\Delta t^*$  calculated with (5.4) from the estimated  $C_0(\Delta t_0)$  and  $err$ . These critical time steps,  $\Delta t^*$ , estimated with (5.4) are in good agreement with numerically measured  $\Delta t^*$  in Figure 3, and depend on the value of  $\varepsilon$ . Hence, the domain of application or working region of the method,  $\Delta t \leq \Delta t^*$ , might be settled depending on the desired level of accuracy by means of an appropriate choice of  $\varepsilon$ . For instance, if we consider the case  $\varepsilon = 0.05$  in Figure 3, for  $\Delta t = 10^{-6}$ , the local error estimate is given by  $err \approx 10^{-10}$  whereas the real Strang local error is  $\sim 10^{-12}$ . This overestimation of the local error will certainly imply an underestimation in the required size of the time steps for a given tolerance. Therefore, for a given tolerance  $\eta$  a more suitable configuration should consider an  $\varepsilon$  such that  $\Delta t \approx \Delta t^*$  in order to reduce excessive overestimations of local errors.

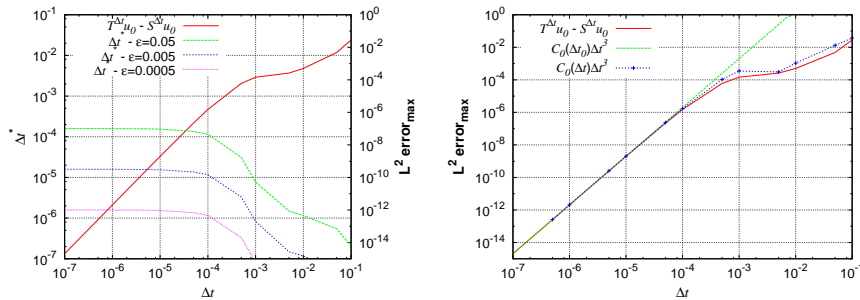


Fig. 4. BZ equation. Working region of the method  $\Delta t \leq \Delta t^*$  with  $\Delta t^*$  calculated with  $C_0$  estimated at  $\Delta t_0 = 10^{-5}$  and  $err$  obtained for several splitting time steps  $\Delta t$  and  $\varepsilon$  (left). Right: predicted Strang error calculated with  $C_0$  estimated at  $\Delta t_0 = 10^{-5}$  and locally at several splitting time steps  $\Delta t$ .

In the illustration shown in Figure 4,  $C_0$  was estimated in the third order region of the method and therefore, all values are well approximated as long as  $\Delta t$  remains in this region. In particular, critical  $\Delta t^*$  will be progressively underestimated for larger  $\varepsilon$  and consequently, it will impose smaller time steps for a given tolerance; this is already the case for  $\varepsilon = 0.05$ , for which  $\Delta t^*$  is in the transition zone towards the lower order region. Even though the computation of  $C_0$  with small time steps will be less expensive, a much more accurate procedure considers current time step as shown in Figure 4. In particular, by estimating locally  $C_0$ , we are estimating real Strang error and thus,  $\Delta t \leq \Delta t^*$  guarantees prescribed accuracy even if asymptotic order estimates are no longer verified. This allows to properly extend the domain

of application over the whole range of possible time steps for a given accuracy; an extremely important issue for real applications for which splitting time steps may go far beyond asymptotic behavior including the potential order reduction region associated with the stiffness of the problem.

### 5.3. Numerical strategy

Previous studies conducted in § 5.1 and § 5.2 allow to properly complete the adaptive splitting strategy introduced in § 2. In this part we conduct the final description of the numerical strategy.

Let us consider general problem (2.1) for  $\mathbf{u} \in \mathbb{R}^m$ , for which we use  $S_2^t$  in (2.6) as resolution scheme. Depending on the problem, the adaptive method will be applied considering time evolution of  $l \leq m$  variables:  $\tilde{\mathbf{u}} \in \mathbb{R}^l$ . Let us denote  $\Omega_l$  the set of indices of these variables. In order to consider only  $l < m$  variables, the formers must be decoupled of the remaining  $m - l$  variables in the reactive term  $\mathbf{f}(\tilde{\mathbf{u}})$  in (2.1). To simplify the presentation, we will only consider  $\varepsilon \in (0, \varepsilon_{max})$ ,  $\varepsilon_{max} < 1/2$ .

We set the accuracy tolerance  $\eta$ , an initial time step  $\Delta t^0$  and initial  $\varepsilon_0$ , and perform the time integration of (2.1) with the Strang scheme  $S_2^t$  and the embedded shifted one  $S_{2,\varepsilon}^t$  given by (2.7). We compute local error estimate *err* and new time step  $\Delta t^{new}$  according to (2.11). If *err* is smaller than  $\eta$ , current time step solution is accepted and simulation time evolves; otherwise, current solution is rejected and the time integration is recomputed with  $\Delta t^{new}$ . In particular, it is better to choose rather small  $\Delta t^0$  to avoid initial rejections.

In order to guarantee an effective error control, we define the working region  $\Delta t \leq \Delta t^*$  by estimating the corresponding  $\Delta t^*$  for current  $\varepsilon$ . This is done for the first time step  $\Delta t^0$  and then periodically after  $N$  accepted time steps depending on the problem, based on the numerical procedure introduced in § 5.1. Computation of critical  $\Delta t^*$  is also performed with  $\tilde{\mathbf{u}}$ , and a rather large initial  $\varepsilon_0$  is suitable to initially guarantee  $\Delta t \leq \Delta t^*$ .

We define then a suitable working region  $\Delta t \in [\beta \Delta t^*, \gamma \Delta t^*]$  with  $0 < \beta < \gamma \leq 1$ , for which splitting time steps are close to  $\Delta t^*$ . A new  $\varepsilon$  is then computed if  $\Delta t$  is much lower than  $\Delta t^*$  ( $\Delta t < \beta \Delta t^*$ ) in order to avoid unnecessary small time steps; or if  $\Delta t$  is very close or possibly larger than  $\Delta t^*$  ( $\Delta t > \gamma \Delta t^*$ ) with  $\gamma$  close to one, in order to increase upper bound of the domain of application. This guarantees that  $\varepsilon$  is dynamically computed and properly adapted to the dynamics of the phenomenon.

Finally, the numerical resolution strategy can be summarized as follows, where  $\mathbf{U} \in \mathbb{R}^{m \times n}$  stands for the spatial discretization of  $\mathbf{u}$  over  $n$  points,  $\mathbf{U} := (u^{(j,k)})$  such that  $j \in [1, m]$  and  $k \in [1, n]$ .

- **Input parameters.** Define accuracy tolerance  $\eta$ , time domain of study  $[t_0, T]$ , initial time step  $\Delta t^0$ , initial  $\varepsilon_0$ , and period of computation of  $\Delta t^*$ :  $N$ .
- **Initialization.** Set iteration counter  $i = 0$  and  $t = t_0$ ,  $\mathbf{U} = \mathbf{U}_0$ ,  $\Delta t = \Delta t^0$ ,  $\varepsilon = \varepsilon_0$ . We define a flag *estimate* initialized as `.false..` Throughout the

whole computation, we need to store  $\mathbf{U}$ , **an array of size  $m \times n$ .**

- **Time evolution.** If  $t < T$ :

- (1) Only if  $\frac{i}{N} = \left\lfloor \frac{i}{N} \right\rfloor$  or *estimate* is `.true.`:

**Computation of critical  $\Delta t^*$  I:** For the sets  $(a_1, b_1, c_1)$  and  $(a_2, b_2, c_2)$  with  $a_1 = 1$  and  $a_2 = c_1$ , we compute successively:

- $\tilde{\mathbf{U}}_1 = S^{c_2 \Delta t} \tilde{\mathbf{U}}_0$ , where  $\tilde{\mathbf{U}}_0$  is built out of  $\mathbf{U}$ ,  $\tilde{\mathbf{U}}_0 = (u^{(j,\cdot)})_{j \in \Omega_l}$ ;
- $\tilde{\mathbf{U}}_1 = S^{b_2 \Delta t} \tilde{\mathbf{U}}_1$ ;
- $\tilde{\mathbf{U}}_2 = S^{c_1 \Delta t} \tilde{\mathbf{U}}_0$ ;
- $e_1 = \max_{j \in \Omega_l} \|\tilde{u}_2^{(j,\cdot)} - \tilde{u}_1^{(j,\cdot)}\|$ ;
- $\tilde{\mathbf{U}}_2 = S^{b_1 \Delta t} \tilde{\mathbf{U}}_2$ ;
- *estimate* is set to `.true.`.

These operations needs to store  $\tilde{\mathbf{U}}_1$  and  $\tilde{\mathbf{U}}_2$ , **two arrays of size  $l \times n$ .**

- (2) **Time integration over  $\Delta t$ :** We compute successively:

- for each  $k \in [1, n]$ ,  $u_{new}^{(\cdot,k)} = Y^{\Delta t/2} u^{(\cdot,k)}$ ;
- for each  $k \in [1, n]$ ,  $\tilde{u}_1^{(\cdot,k)} = Y^{\varepsilon \Delta t} u_{new}^{(j,k)} \Big|_{j \in \Omega_l}$ ;
- $\mathbf{U}_* = X^{\Delta t} \mathbf{U}_*$ , with  $\mathbf{U}_* = {}^t(\mathbf{U}_{new}, \tilde{\mathbf{U}}_1)$ ;
- for each  $k \in [1, n]$ ,  $u_*^{(\cdot,k)} = Y^{(1/2-\varepsilon)\Delta t} u_*^{(\cdot,k)}$ ;
- for each  $k \in [1, n]$ ,  $u_{new}^{(\cdot,k)} = Y^{\varepsilon \Delta t} u_{new}^{(\cdot,k)}$ ;
- $err = \max_{j \in \Omega_l} \|\tilde{u}_{new}^{(j,\cdot)} - \tilde{u}_1^{(j,\cdot)}\|$ .

We need to store  $\mathbf{U}_{new}$ , **an array of size  $m \times n$ .**

- (3) Only if *estimate* is `.true.`:

**Computation of critical  $\Delta t^*$  II:** We compute successively:

- $e_2 = \max_{j \in \Omega_l} \|\tilde{u}_{new}^{(j,\cdot)} - \tilde{u}_2^{(j,\cdot)}\|$ ;
- $C_0$  using (5.10) with  $e_1$  and  $e_2$ ;
- estimate  $\Delta t^*$  out of (5.4) and set  $\Delta t^* = \zeta \Delta t^*$  with security factor  $0 < \zeta \leq 1$  close to one;
- *estimate* is set to `.false.`.
- If  $\Delta t \notin [\beta \Delta t^*, \gamma \Delta t^*]$  with  $0 < \beta < \gamma \leq 1$ : *estimate* is set to `.true.`.

- (4) Only if *estimate* is `.true.` and  $i > 0$ :

**Computation of  $\varepsilon$ :** According to (5.4) with  $err$ ,  $C_0$  and  $\Delta t^* = \Delta t$ :

- $\varepsilon = \min\{\theta \varepsilon, \varepsilon_{max}\}$  with  $\theta \geq 1$  as security factor;
- computation of  $\Delta t^*$  with new  $\varepsilon$ ;
- *estimate* is set to `.false.`.

- (5) **Computation  $\Delta t^{new}$ :** According to (2.11) with security factor  $0 < v \leq 1$  close to one.

- If  $\Delta t > \Delta t^*$ : set  $err = tol + C$  with  $C > 1$ . Used to potentially reject initial  $\Delta t = \Delta t^0$ .
- If  $\Delta t^{new} > \Delta t^*$  and  $\varepsilon \neq \varepsilon_{max}$ : *estimate* is set to `.true.`.

- $\Delta t = \min\{\Delta t^{new}, \Delta t^*\}$ .
- If  $err \leq tol$ :  $t = t + \Delta t$ ,  $i = i + 1$ ,  $\Delta t = \min\{\Delta t, T - t\}$  and  $\mathbf{U} = \mathbf{U}_{new}$ .

In this strategy, reaction is always integrated point by point if the reactive term is modeled by a system of ODEs without spatial coupling. This integration can be performed completely in parallel [9, 7]. On the other hand, for linear diffusion problems, another alternative considers a variable by variable resolution, for each  $j \in [1, m] \cup \Omega_l$ :

$$u_{\star}^{(j,\cdot)} = X^{\Delta t} u_{\star}^{(j,\cdot)}, \quad (5.12)$$

that can also be performed in parallel [9].

Depending on the problem, either the computation of critical  $\Delta t^*$  (steps (1), (3) and (4)), or the computation of  $\varepsilon$  (step (4)) can be potentially removed if one considers large enough  $\varepsilon_0$  and fine enough  $\eta$ . Finally, the whole strategy with all steps needs to store at worst two arrays of size  $l \times n$  and other two of size  $m \times n$ , beyond memory requirements of diffusion and reaction solvers.

## 6. Final Numerical Evaluation of the Method

In this last part, we evaluate the performance of the method in terms of accuracy of the simulation, and show that an effective control of the simulation error is performed in the context of two different problems. First, we will consider a propagating wave featuring time-space multi-scale character. Then, the potential of the method is fully exploited for a more complex configuration of repetitive gas discharges generated by high frequency pulsed applied electric fields followed by long time scale relaxation, for which a precise description of discharge and post-discharge phases is achieved.

### 6.1. BZ equation revisited

Coming back to BZ model, we perform a time integration of (5.11) with several accuracy tolerances  $\eta$ . First of all, we consider the numerical strategy detailed in § 5.3 without taking into account steps (1), (3) and (4), that is without computation of neither critical  $\Delta t^*$  nor  $\varepsilon$ . We set  $\Delta t^0 = 10^{-7}$  and  $\varepsilon_0 = 0.05$  in all cases, with  $t \in [0, 2]$ . In this example, a rather small initial splitting time step is chosen to avoid initial rejections even though this initial rejection phase usually does not take many steps as it will shown in the next example. On the other hand, we have chosen a intermediary value for  $\varepsilon$  in order to clearly distinguish the different behaviors of the strategy in terms of prediction of the local errors depending on the proposed tolerance.

Figure 5 shows time evolution of accepted splitting time steps  $\Delta t$ . In this case, BZ equation models a propagating self-similar wave, so splitting time step stabilizes once the overall phenomenon is solved within the prescribed tolerance  $\eta$ . Local error



estimates  $err$  are also shown, which naturally verify prescribed accuracy, since we impose time steps for which  $err$  is limited by  $\eta$  through (2.11).

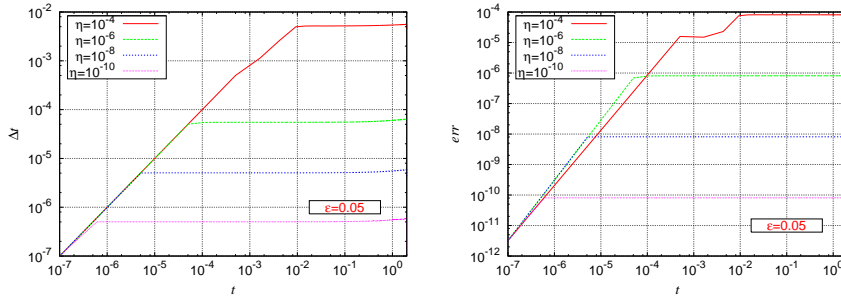


Fig. 5. BZ equation. Time evolution of accepted splitting time steps  $\Delta t$  (left) and local  $L^2$  error estimates  $err = \|S^{\Delta t} u_0 - S_\varepsilon^{\Delta t} u_0\|_{L^2}$  (right), for several tolerances  $\eta$  and  $\varepsilon = 0.05$ .

Table 2 summarizes global  $L^2$  errors between splitting and reference solutions at the end of the time domain of study,  $t = 2$ . For a fine enough  $\eta$  and consequently, small enough time steps, a precise error control is achieved by the local error control strategy as we could have expected from previous results in Figure 3 for  $\varepsilon = 0.05$ . Nevertheless, for  $\eta = 10^{-4}$  we can see rather high global errors even if this configuration considers naturally less time integration steps and thus, less accumulation of local approximation errors. If we take a look at Figure 3, we note that for  $\varepsilon = 0.05$  and local errors of about  $10^{-4}$ , the local error estimate,  $err$ , is not predicting properly real Strang errors, as it was previously discussed, since  $\Delta t > \Delta t^*$ . Therefore, a strategy that introduces a more precise description of errors for a larger range of time steps must be considered, whenever the required accuracy casts the method away from its asymptotic behavior. This is an under covered difficulty of any time adaptive technique based on a lower order embedded method, and to our knowledge, an open problem that has not been studied much, and that this work tries to overcome.

Table 2. BZ equation.  $L^2$  errors at final time  $t = 2$  for  $a, b, c$  variables and several tolerances  $\eta$ .

$\eta$	$L^2$ error $a$	$L^2$ error $b$	$L^2$ error $c$
$10^{-4}$	$7.97 \times 10^{-3}$	$1.07 \times 10^{-2}$	$4.72 \times 10^{-3}$
$10^{-6}$	$1.71 \times 10^{-6}$	$1.83 \times 10^{-6}$	$7.98 \times 10^{-7}$
$10^{-8}$	$1.45 \times 10^{-8}$	$1.54 \times 10^{-8}$	$6.78 \times 10^{-9}$
$10^{-10}$	$1.74 \times 10^{-10}$	$1.75 \times 10^{-10}$	$1.08 \times 10^{-10}$

Let us now consider the entire strategy with all steps for several tolerances with  $\Delta t^0 = 5 \times 10^{-7}$  and  $\varepsilon_0 = 0.05$ . In the following illustrations we have considered the following parameters:  $\varepsilon_{max} = 0.999$ ;  $a_1 = 1$ ,  $b_1 = c_1 = a_2 = 1/2$ ,  $b_2 = 2/5$  and  $c_2 = 1/10$  for intermediary time steps evaluations;  $\zeta = 0.9$  as security factor of critical  $\Delta t^*$  estimate;  $\beta = 0.1$  and  $\gamma = 0.95$  to define the working region  $\Delta t \in [\beta\Delta t^*, \gamma\Delta t^*]$ ;  $\theta = 10$  as security factor of  $\varepsilon$  estimate;  $C = 10$  to potentially reject initial time step  $\Delta t^0$ ; and  $v = 0.9$  as security factor of  $\Delta t^{new}$  estimate. All local estimators,  $err$ ,  $e_1$  and  $e_2$ , are computed with normalized  $L^2$  norms.

Considering the propagating phenomenon, we set  $N = 10$ , but we estimate  $\Delta t^*$  only twice for  $i = 0$  and  $i = N$ . Figure 6 shows time evolution of splitting time steps; there are different scenarios depending on the required accuracy. In all cases for  $\varepsilon_0 = 0.05$ , we estimate initially  $\Delta t^* \approx 1.4 \times 10^{-4}$ . For  $\eta = 10^{-4}$ , this limitation implies smaller time steps than what is required for the prescribed tolerance. Thus,  $\Delta t$  increases until  $\Delta t^{new} > \Delta t^*$  and a new  $\varepsilon$  is estimated:  $\varepsilon \approx 0.43$ . No substantial changes are made when  $i = N$ , since  $\Delta t \in [\beta\Delta t^*, \gamma\Delta t^*]$  for the current  $\eta$ .

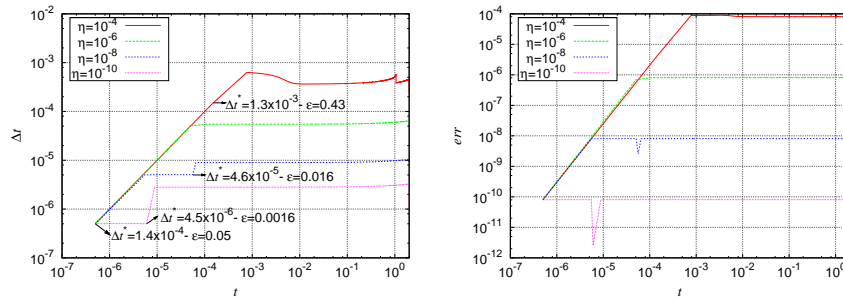


Fig. 6. BZ equation. Time evolution of accepted splitting time steps  $\Delta t$  (left) and local  $L^2$  error estimates  $err = \|S^{\Delta t} u_0 - S_\varepsilon^{\Delta t} u_0\|_{L^2}$  (right), for several tolerances  $\eta$ , considering critical  $\Delta t^*$  and computation of  $\varepsilon$ .

For  $\eta = 10^{-6}$ , we keep initial  $\Delta t^*$  and  $\varepsilon_0$  since  $\Delta t \in [\beta\Delta t^*, \gamma\Delta t^*]$  as we can see in Figure 3. Finally, for  $\eta = 10^{-8}$  and  $\eta = 10^{-10}$ ,  $\Delta t < \beta\Delta t^*$  and thus,  $\varepsilon$  is recomputed, giving respectively  $\varepsilon \approx 0.016$  and  $0.0016$ . In particular, we consider larger splitting time steps for which Strang local errors are better predicted. Table 3 shows that error control is this time guaranteed for all values of tolerance  $\eta$ , and thus, for a larger range of time steps. Compared with previous results in Table 2, we correct completely the errors in the prediction of local error estimates, which yields more accurate resolutions for the largest tolerances; whereas slightly less accurate results are obtained for the smallest tolerances since larger splitting time steps are considered.

Table 3. BZ equation.  $L^2$  errors at final time  $t = 2$  for  $a, b, c$  variables and several tolerances  $\eta$ , considering critical  $\Delta t^*$  and computation of  $\varepsilon$ .

$\eta$	$L^2$ error $a$	$L^2$ error $b$	$L^2$ error $c$
$10^{-4}$	$6.85 \times 10^{-5}$	$9.04 \times 10^{-5}$	$4.06 \times 10^{-5}$
$10^{-6}$	$1.71 \times 10^{-6}$	$1.83 \times 10^{-6}$	$7.98 \times 10^{-7}$
$10^{-8}$	$4.53 \times 10^{-8}$	$4.84 \times 10^{-8}$	$2.12 \times 10^{-8}$
$10^{-10}$	$4.48 \times 10^{-9}$	$4.77 \times 10^{-9}$	$2.15 \times 10^{-9}$

## 6.2. Simulation of multi-pulsed gas discharges

In this section, we consider a simplified model of plasma discharges at atmospheric pressure for which we analyze the performance of the proposed numerical strategy in a configuration of nanosecond repetitively pulsed discharges. This kind of phenomenon is studied for plasma assisted combustion or flow control, for which the enhancement of the gas flow chemistry or momentum transfer during typical time scales of the flow of  $10^{-4} - 10^{-3}$ s, is due to consecutive discharges generated by high frequency (in the kHz range) sinusoidal or pulsed applied voltages [23]. As a consequence, during the post-discharge phases of the order of tens of microseconds, not only time scales are very different from those during discharges of a few tens of nanoseconds, but a complete different physics is taking place. Then, to the rapid multi-scale configuration during discharges, we have to add other rather slower multi-scale phenomena in the post-discharge, such as recombination of charged species, heavy-species chemistry, diffusion, gas heating and convection. Therefore, it is very challenging to efficiently simulate this kind of highly multi-scale problems and to accurately describe the physics of the plasma/flow interaction between consecutive discharge/post-discharge phases.

General model to study gas discharge dynamics is based on the following drift-diffusion equations for electrons and ions, coupled with Poisson's equation [2, 20]:

$$\left. \begin{aligned} \partial_t n_e - \partial_{\mathbf{x}} \cdot n_e \mathbf{v}_e - \partial_{\mathbf{x}} \cdot (D_e \partial_{\mathbf{x}} n_e) &= n_e \alpha |\mathbf{v}_e| - n_e \eta |\mathbf{v}_e| + n_e n_p \beta_{ep} + n_n \gamma, \\ \partial_t n_p + \partial_{\mathbf{x}} \cdot n_p \mathbf{v}_p - \partial_{\mathbf{x}} \cdot (D_p \partial_{\mathbf{x}} n_p) &= n_e \alpha |\mathbf{v}_e| - n_e n_p \beta_{ep} + n_n n_p \beta_{np}, \\ \partial_t n_n - \partial_{\mathbf{x}} \cdot n_n \mathbf{v}_n - \partial_{\mathbf{x}} \cdot (D_n \partial_{\mathbf{x}} n_n) &= n_e \eta |\mathbf{v}_e| - n_n n_p \beta_{np} - n_n \gamma, \end{aligned} \right\} \quad (6.1)$$

$$\varepsilon_0 \partial_{\mathbf{x}}^2 V = -q_e (n_p - n_n - n_e), \quad (6.2)$$

where  $\mathbf{x} \in \mathbb{R}^d$ ,  $n_i$  is the density of species  $i$  (e: electrons, p: positive ions, n: negative ions),  $V$  is the electric potential,  $\mathbf{v}_i = \mu_i \mathbf{E}$  ( $\mathbf{E}$  being the electric field) is the drift velocity.  $D_i$  and  $\mu_i$ , are diffusion coefficient and absolute value of mobility of charged species  $i$ ,  $q_e$  is the absolute value of electron charge, and  $\varepsilon_0$  is permittivity of free space.  $\alpha$  is the impact ionization coefficient,  $\eta$  stands for electron attachment on neutral molecules,  $\beta_{ep}$  and  $\beta_{np}$  accounts respectively for electron-positive ion and

negative-positive ion recombination, and  $\gamma$  is the detachment coefficient. Electric field  $\mathbf{E}$  and potential  $V$  are related by

$$\mathbf{E} = -\partial_{\mathbf{x}}V. \quad (6.3)$$

Nevertheless, in this paper, we will consider a simplified reaction-diffusion 1D model based on (6.1):

$$\left. \begin{aligned} \partial_t n_e - D \partial_x^2 n_e &= n_e \alpha |\mathbf{v}_e| - n_e \eta |\mathbf{v}_e| + n_e n_p \beta_{ep}, \\ \partial_t n_p - D \partial_x^2 n_p &= n_e \alpha |\mathbf{v}_e| - n_e n_p \beta_{ep} + n_n n_p \beta_{np}, \\ \partial_t n_n - D \partial_x^2 n_n &= n_e \eta |\mathbf{v}_e| - n_n n_p \beta_{np}. \end{aligned} \right\} \quad (6.4)$$

As in (6.1), all the coefficients of the model are functions of the local reduced electric field  $E/N_{\text{gas}}$ , where  $E$  is the electric field magnitude and  $N_{\text{gas}}$  is the air neutral density. Transport parameters and reaction rates for air are taken from [21], with attachment coefficients taken from [19].

In this numerical illustration, we consider an air gap of 0.5 cm where we have a high initial distribution of electrons and ions over the region  $[0, 0.01]$  cm. A constant electric field of  $\sim 40$  kV/cm is then applied over this region during 10 ns with a pulse period of 1  $\mu$ s. All parameters in (6.4) are computed with the imposed field without solving neither (6.2) nor (6.3). Finally, we consider a constant diffusion coefficient:  $D = 50$  cm<sup>2</sup>/s and a spatial discretization of 1001 points. Figure 7 shows the spatial distribution of electron density just before and after each pulse. Globally, there are at least two completely different physical configurations given either by high reactive activity whenever the electric field is applied, or rather by the propagative nature of the post-discharge phase.

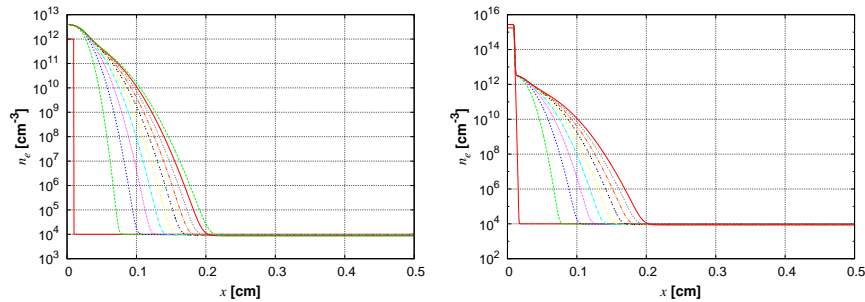


Fig. 7. Repetitive gas discharge model. Spatial distribution of electron density before (left) and after (right) each pulse, starting from initial distribution (left) and for a duration of ten pulses.

Considering the adaptive strategy described in § 5.3 with  $\Delta t^0 = 10^{-10}$ ,  $\varepsilon_0 = 0.05$  and the same parameters used for the previous BZ simulation, computation is initialized with a time step included in the pulse duration. Figure 8 shows the corresponding splitting time steps for a tolerance of  $\eta = 10^{-3}$ . Splitting time step features

28 *Descombes, Duarte, Dumont, Louwet, Massot*

a periodic behavior and succeed to consistently adapt itself to the discharge/post-discharge phenomena. This yields high varying time steps going from  $\sim 10^{-10}$  to  $\sim 10^{-7}$ . Therefore, after each post-discharge phase, since the new time step is computed based on the previous one according to (2.11), this new time step will surely skip the next pulse. In order to avoid this, each time we get into a new period, we initialize time step with the length of the pulse:  $\Delta t = 10$  ns; this time step is obviously rejected as seen in Figure 8, as well as the next ones, until we are able to retrieve the right dynamics of the phenomenon for the required accuracy tolerance. No other intervention is needed neither for modeling parameters nor for numerical solvers in order to automatically adapt time step to describe the several time scales of the phenomenon within a prescribed accuracy.

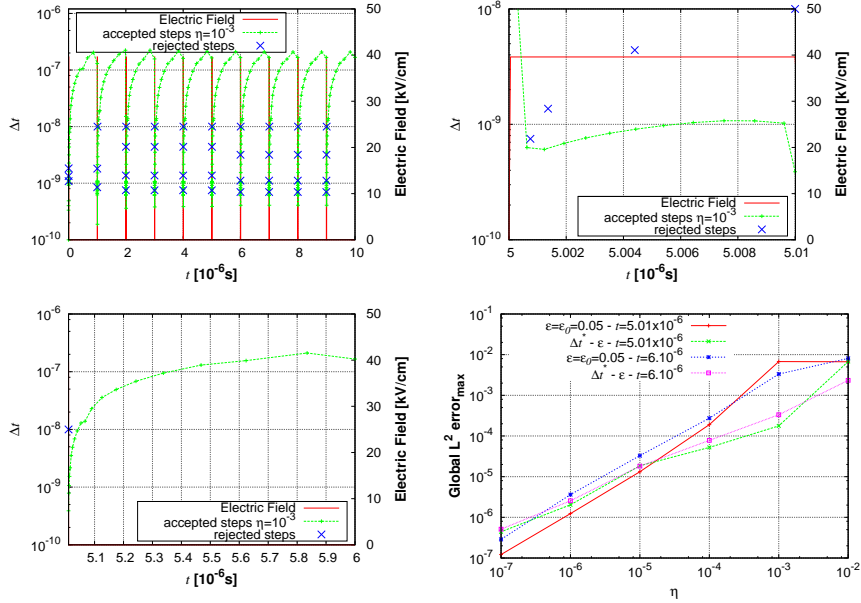


Fig. 8. Repetitive gas discharge model. Time evolution of accepted and rejected splitting time steps, and imposed electric field for  $t \in [0, 10] \mu\text{s}$  (top left), during pulse  $t \in [5, 5.01] \mu\text{s}$  (top right) and post-discharge  $t \in [5.01, 6] \mu\text{s}$  (bottom left). Bottom right: global  $L^2$  errors at the end of the pulse ( $t = 5.01 \mu\text{s}$ ) and the post-discharge phase ( $t = 6 \mu\text{s}$ ) with and without  $\Delta t^*$  and  $\varepsilon$  computation.

For this application, we compute critical  $\Delta t^*$  and possibly  $\varepsilon$ , for  $N = 10$  and  $N = 100$  in each period in order to perform these computations at least once during the discharge and post-discharge regimes. For example, for  $t \in [5, 6] \mu\text{s}$  as in Figure 8,  $\varepsilon = \varepsilon_{max}$  with  $\Delta t^* \approx 4.3 \times 10^{-9}$  during the pulse, and  $\varepsilon \approx 0.26$  with  $\Delta t^* \approx 1.6 \times 10^{-7}$  for the rest of the period. Similar values are found for the other periods. Notice that after each pulse,  $\Delta t^*$  is automatically updated because

$\Delta t$  increases and then  $\Delta t$  gets equal to  $\Delta t^*$ . In particular, the important difference between  $\Delta t^*$  for each region, comes naturally from the completely different modeling parameters and hence, physics description of each regime.

An effective error control is achieved for each part of the phenomenon, as we can deduce from the global error between splitting and reference solutions at the end of the pulse ( $t = 5.01 \mu\text{s}$ ) and at the end of the post-discharge phase ( $t = 6 \mu\text{s}$ ). If we compare these results with the ones obtained without estimating neither  $\Delta t^*$  nor  $\varepsilon$  with  $\varepsilon = \varepsilon_0 = 0.05$ , we can draw the same conclusions as in the BZ application. For less accurate resolutions with high tolerances, the proposed strategy corrects the error in the local error estimates made with  $\varepsilon = \varepsilon_0 = 0.05$ ; in particular, for  $\eta = 10^{-3}$  there is a ratio of about 10 between both solutions. For higher tolerances,  $\eta \geq 10^{-2}$ , both methods yield a time step equal to the pulse duration,  $\Delta t = 10 \text{ ns}$ . On the other hand, for the smallest tolerances, slightly more accurate solutions are obtained with a fixed  $\varepsilon = \varepsilon_0$  because smaller splitting time steps are used.

## 7. Conclusions

The present work proposes a new resolution strategy for stiff evolutionary PDEs based on an efficient splitting scheme previously developed [9, 8] that considers high order dedicated integration methods for each subproblem in order to properly solve the fastest time scales associated with each one of them, and in such a way that the main source of error is led by the operator splitting error. Then, to control the error of the resolution, it relies on an adaptive splitting time technique that allows to discriminate the global time scales related to the coupled phenomenon, given a required level of accuracy of computations. Compared with a standard procedure for which accuracy is guaranteed by considering time steps of the order of the fastest scale, the error control featured by our method implies an effective accurate resolution for problems modeling various physical scenarios, independent of the fastest physical time scale, and an important improvement of computational efficiency whenever highly unsteady phenomena is simulated. In particular, we have successfully applied the proposed strategy to a simplified model of plasma discharges that nevertheless exhibits a broad time scale spectrum coming from the modeling equations and also important and discontinuous variation of parameters in time and in space that notably increase the numerical complexity of the problem.

A numerical analysis of the method has been developed in order to settle a solid mathematical background, and a complementary numerical procedure was conceived in order to overcome classical restrictions of adaptive time stepping schemes whenever asymptotic estimates fail to predict the dynamics of the problem. A both mathematical and numerical detailed study of the method has thus led to a fully complete adaptive time stepping strategy that guarantees an effective control of the errors of integration for a large range of time steps; a key issue for problems for which splitting time steps can go beyond the fastest physical scales of the problem. The contribution of this paper is then mainly given by a dedicated adaptive time

splitting method for stiff PDEs, and by a complete study of the behavior of time stepping schemes based on lower order embedded methods, for the whole set of potential time steps. In this paper we have always considered fine enough spatial discretizations in order to perform an evaluation of the theoretical estimates introduced for the proposed time integration scheme. For higher dimensional problems, fine spatial discretization becomes a critical issue in terms of computational costs and a technique of local grid refinement might be a good solution to guarantee the theoretical behavior of the splitting schemes (see for instance [8]). Nevertheless, a mathematical study on the splitting errors with discretized operators will certainly be an useful tool to yet improve the performance of these techniques. This and other related theoretical aspects are particular topics of our current research.

### Acknowledgments

This research was supported by a fundamental project grant from ANR (French National Research Agency - ANR Blancs): *Séchelles* (project leader S. Descombes), and by a Ph.D. grant for M. Duarte from Mathematics (INSMI) and Engineering (INSIS) Institutes of CNRS and supported by INCA project (National Initiative for Advanced Combustion - CNRS - ONERA - SAFRAN).

### References

1. A. Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput.*, 23:2041–2054, 2002.
2. N. Y. Babaeva and G. V. Naidis. Two-dimensional modelling of positive streamer dynamics in non-uniform electric fields in air. *J. Phys. D: Appl. Phys.*, 29:2423–2431, 1996.
3. 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. of Computer Mathematics*, 84(6):749–765, 2007.
4. 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.
5. 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.
6. S. Descombes and M. Thalhammer. The Lie-Trotter splitting method for nonlinear evolutionary problems involving critical parameters. An exact local error representation and application to nonlinear Schrödinger equations in the semi-classical regime. *Preprint, available on HAL (<http://hal.archives-ouvertes.fr/hal-00557593>)*, 2010.
7. 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. *ESAIM Proc. (to app.)*, 2011.
8. 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, space adaptive multiresolution and dedicated high order implicit/explicit time in-

- tegrators. *Submitted to SIAM J. Sci. Comput.*, available on HAL (<http://hal.archives-ouvertes.fr/hal-00457731>), 2011.
9. T. Dumont, M. Duarte, S. Descombes, M.A. Dronne, M. Massot, and V. Louvet. Simulation of human ischemic stroke in realistic 3D geometry: A numerical strategy. *Submitted to Bulletin of Math. Biology*, available on HAL (<http://hal.archives-ouvertes.fr/hal-00546223>), 2011.
  10. I. R. Epstein and J. A. Pojman. *An Introduction to Nonlinear Chemical Dynamics*. Oxford University Press, 1998. Oscillations, Waves, Patterns and Chaos.
  11. P. Gray and S. K. Scott. *Chemical oscillations and instabilities*. Oxford University Press, 1994.
  12. W. Gröbner. *Die Liereihen und ihre Anwendungen*. VEB Deutscher Verlag der Wiss., Berlin 1960, 1967. Second Edition.
  13. E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer-Verlag, Berlin, second edition, 2006. Structure-Preserving Algorithms for Ordinary Differential Equations.
  14. E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations I*. Springer-Verlag, Berlin, second edition, 1993. Nonstiff problems.
  15. E. Hairer and G. Wanner. *Solving ordinary differential equations II*. Springer-Verlag, Berlin, second edition, 1996. Stiff and differential-algebraic problems.
  16. O. M. Knio, H. N. Najm, and P. S. Wyckoff. A semi-implicit numerical scheme for reacting flow. II. Stiff, operator-split formulation. *J. Comput. Phys.*, 154:482–467, 1999.
  17. O. Koch and M. Thalhammer. Embedded split-step formulae for the time integration of nonlinear evolution equations. *Preprint*, 2010.
  18. A. N. Kolmogoroff, I. G. Petrovsky, and N. S. Piscounoff. Etude de l'équation de la diffusion avec croissance de la quantité de matière et son application a un problème biologique. *Bulletin de l'Université d'état Moscou, Série Internationale Section A Mathématiques et Mécanique*, 1:1–25, 1937.
  19. I. A. Kossyi, A. Yu Kostinsky, A. A. Matveyev, and V. P. Silakov. Kinetic scheme of the non-equilibrium discharge in nitrogen-oxygen mixtures. *Plasma Sources Sci. Technol.*, 1(3):207–220, 1992.
  20. A. A. Kulikovskiy. Positive streamer between parallel plate electrodes in atmospheric pressure air. *J. Phys. D: Appl. Phys.*, 30:441–450, 1997.
  21. R. Morrow and J. J. Lowke. Streamer propagation in air. *J. Phys. D: Appl. Phys.*, 30:614–627, 1997.
  22. E. S. Oran and J. P. Boris. *Numerical simulation of reacting flows*. Cambridge University Press, 2001. Second Edition.
  23. G. Pilla, D. Galley, D. Lacoste, F. Lacas, D. Veynante, and C. O. Laux. Stabilization of a turbulent premixed flame using a nanosecond repetitively pulsed plasma. *IEEE Trans. Plasma Sci.*, 34(6, Part 1):2471–2477, 2006.
  24. M. A. Singer, S. B. Pope, and H. N. Najm. Modeling unsteady reacting flow with operator splitting and ISAT. *Combustion and Flame*, 147(1-2):150 – 162, 2006.
  25. G. Strang. Accurate partial difference methods. I. Linear Cauchy problems. *Arch. Ration. Mech. Anal.*, 12:392–402, 1963.
  26. G. Strang. On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.*, 5:506–517, 1968.