# NUMERICAL METHODS FOR REACTING GAS FLOW SIMULATIONS 

Sander van Veldhuizen*, Kees Vuik* and Chris R. Kleijn ${ }^{\dagger}$<br>*Delft University of Technology, Delft Institute of Applied Mathematics and J.M. Burgers Center, Mekelweg 4, 2628 CD, The Netherlands<br>e-mail: \{s.vanveldhuizen, c.vuik\}@tudelft.nl<br>${ }^{\dagger}$ Delft University of Technology, Department of Multi Scale Physics and J.M. Burgers Center, Prins Bernardlaan 6, 2628 BW Delft, The Netherlands, e-mail: c.r.kleijn@tudelft.nl

Key words: Chemical Vapor Deposition, Transport phenomena, Advection Diffusion Reaction Equations, Time Integration


#### Abstract

In this study various numerical schemes for simulating 2D laminar reacting gas flows, as typically found in Chemical Vapor Deposition (CVD) reactors, are proposed and compared. These CVD systems are generally modeled by means of many stiffly coupled elementary gas phase reactions between a large number of reactants and intermediate species. The purpose of this study is to develop robust and efficient solvers for the stiff heat-reaction system, whereby the velocities are assumed to be given. For non-stationary $C V D$ simulation, an optimal combination in terms of efficiency and robustness between time integration, nonlinear solvers and linear solvers has to be found. Besides stability, which is important due to the stiffness of the problem, the preservation of non-negativity of the species is crucial. It appears that this extra condition on time integration methods is much more restrictive towards the time-step than stability. For a set of suitable time integration methods necessary conditions are represented. We conclude with a comparison of the workload between the selected time integration methods. This comparison has been done for a 2D test problem. The test problem does not represent a practical process, but represents only the computational problems.


## 1 Introduction

In Chemical Vapor Deposition (CVD) literature, and also other reactive flow literature, one is usually looking for the steady state solution of the species equations (1). The usual strategy to find this steady state solution is to perform a (damped/relaxed) Newton iteration with an (arbitrary) initial solution. Then, hopefully, the Newton iteration converges to the steady state. In the case of Newton divergence artificial time stepping is performed to find a better initial solution for the next Newton iteration.

In our research we are not looking for the steady state solution only, but we also want the transient solution. In order to simulate this transient, we have to use a time integration method that can handle stiff problems. In this paper we present suitable time integration methods for stiff problems. Furthermore, we compare these integration methods by their performance, in terms of efficiency.

## 2 Model for CVD Simulation

The mathematical model describing the CVD process consists of a set of PDEs with appropriate boundary and initial conditions, which describe the gas flow, the transport of energy, the transport of species and reactions in the reactor.

The gas mixture in the reactor is assumed to behave as a continuum. The gas flow in the reactor is assumed to be laminar. Since no large velocity gradients appear in CVD gas flows, viscous heating due to dissipation will be neglected. We also neglect the effects of pressure variations in the energy equation. The composition of the $N$ component gas mixture is described in terms of the dimensionless mass fractions $\omega_{i}=\frac{\rho_{i}}{\rho}, i=1, \ldots, N$, having the property $\sum_{i=1}^{N} \omega_{i}=1$. The transport of mass, momentum and heat are described respectively by the continuity equation, the Navier-Stokes equations and the transport equation for thermal energy expressed in terms of temperature $T$. See, for instance, Kleijn ${ }^{4}$ and Van Veldhuizen ${ }^{6}$.

We assume that in the gas-phase $K$ reversible reactions take place. For the $k^{\text {th }}$ reaction the net molar reaction rate is denoted as $R_{k}^{g}\left(\frac{\text { mole }}{m^{3} \cdot s}\right)$. For an explicit description of the net molar reaction rate, we refer to Kleijn ${ }^{4}$ and Van Veldhuizen ${ }^{6}$. The mass diffusion flux is decomposed into concentration diffusion and thermal diffusion. In this study we describe ordinary diffusion in terms of effective diffusion coefficients $\mathbb{D}_{i}^{\prime}$, such that we obtain

$$
\begin{equation*}
\frac{\partial\left(\rho \omega_{i}\right)}{\partial t}=-\nabla \cdot\left(\rho \mathbf{v} \omega_{i}\right)+\nabla \cdot\left(\rho \mathbb{D}_{i}^{\prime} \nabla \omega_{i}\right)+\nabla \cdot\left(\mathbb{D}_{i}^{T} \nabla(\ln T)\right)+m_{i} \sum_{k=1}^{K} \nu_{i k} R_{k}^{g} \tag{1}
\end{equation*}
$$

where $\mathbb{D}_{i}^{T}$ the multi-component thermal diffusion coefficient for species $i$.
The main focus of our research is on efficient solvers for the above species equation(s) (1). Typically the time scales of the slow and fast reaction terms differ orders of magnitude from each other, and from the time scales of the diffusion and advection terms, leading to extremely stiff systems.

### 2.1 Simplified System

Since our research focuses on solving the species equations (1), we will only solve the coupled system of $N$ species equations, where $N$ denotes the number of gas-species in the reactor. The velocity field, temperature field, pressure field and density field are computed via another simulation package developed by Kleijn ${ }^{5}$. Other simplifications are the omission of surface reactions and thermal diffusion.

We consider a CVD process, which is a simplification of the CVD process considered in Kleijn ${ }^{5}$, that deposits silicon Si from silane $\mathrm{SiH}_{4}$. The gas-mixture consists of the carrier gas helium and 6 reactive species that satisfy the reaction mechanism

| G1: | $\mathrm{SiH}_{4}$ | $\rightleftarrows$ | $\mathrm{SiH}_{2}+\mathrm{H}_{2}$, |  |
| :--- | :---: | :--- | :---: | :--- |
| G2: | $\mathrm{Si}_{2} \mathrm{H}_{6}$ | $\rightleftarrows$ | $\mathrm{SiH}_{4}+\mathrm{SiH}_{2}$, |  |
| G3: | $\mathrm{Si}_{2} \mathrm{H}_{6}$ | $\rightleftarrows$ | $\mathrm{H}_{2} \mathrm{SiSiH}_{2}+\mathrm{H}_{2}$, |  |
| G4: | $\mathrm{SiH}_{2}+\mathrm{Si}_{2} \mathrm{H}_{6}$ | $\rightleftarrows$ | $\mathrm{Si}_{3} \mathrm{H}_{8}$, | and, |
| G5: | $2 \mathrm{SiH}_{2}$ | $\rightleftarrows$ | $\mathrm{H}_{2} \mathrm{SiSiH}_{2}$. |  |



Figure 1: Reactor geometry
The studied reactor configuration is illustrated in Figure 1. As computational domain we take, because of axisymmetry, one half of the $r-z$ plane. The pressure in the reactor is 1 atm . From the top a gas-mixture, consisting of silane, with mass fraction $f_{\mathrm{in}^{2}, \mathrm{SiH}_{4}}=0.001$, and helium (the rest), enters the reactor with a uniform temperature $T_{\text {in }}=300 \mathrm{~K}$ and a uniform velocity $u_{\mathrm{in}}$. At a distance of 10 cm . below the inlet a non rotating susceptor with temperature $T=1000 \mathrm{~K}$ and a diameter 30 cm . is placed.

We emphasize that this test-problem is not representing a practical process, but representing its computational problems. Further details on the test-problem can be found in Van Veldhuizen ${ }^{6}$.

## 3 Properties of Numerical Methods for Solving the Species Equations

As seen in the previous section the species eqns. (1) are PDEs of the advection-diffusion-reaction type. In order to have a unique solution appropriate boundary conditions and initial values have to be chosen.

To approximate the solution we use the Method of Lines (MOL), i.e., first discretize in space and then in time, resulting into the ODE system

$$
\begin{equation*}
w^{\prime}(t)=F(t, w(t)), \quad 0<t \leq T, \quad w(0) \text { given. } \tag{2}
\end{equation*}
$$

In (2) $F(t, w(t))$ is spatially discretized convection, diffusion and reactions. For spatial discretization a finite volume hybrid scheme is used, which uses a central scheme for

Peclet numbers less than 2 and locally first order upwinding whenever the Peclet number is larger than 2. The next step is to integrate the ODE system (2) with an appropriate time integration method. We note that the stiff reaction terms in CVD motivates to integrate parts of $F(t, w(t))$ implicitly. In general, due to the nonlinearities in the reaction term, (huge) nonlinear systems have to be solved.

The topic of this research is to find the best combination of time integration, nonlinear and linear solvers in terms of efficiency. Note that if the computational cost of one time step is (very) expensive, then a time integration method that needs more, but computational cheaper, time steps is better in terms of efficiency.

Besides the efficiency criteria, also some other properties are desired for the numerical methods. As already mentioned in Section 2 the system of species equations is stiff. Following Hundsdorfer et al. ${ }^{2}$, we say that stiffness indicates a class of problems for which implicit methods perform (much) better than explicit methods. The eigenvalues of the Jacobian $\frac{\delta f}{\delta y}$ play certainly a role in this decision, but quantities such as the dimension of the system and the smoothness of the solution are also important.

### 3.1 Positivity

A natural property for mass fractions is that they are non-negative. As a consequence, it should also hold for the mathematical model, spatial and time integration of the process. While the first one is obvious, the latter two should not introduce any (small) negative components causing blow up of the solution. It appears that this extra condition on time integration methods is much more restrictive towards the time step than stability. We remark that positivity for spatial discretization with central scheme can be assured by locally first order upwinding.

An ODE system $w^{\prime}(t)=F(t, w(t)), t \geq 0$, is called positive if $w(0) \geq 0$ implies $w(t) \geq 0$ for all $t>0$. It is easy to prove that linear systems $w^{\prime}(t)=A w(t)$ are positive if and only if $a_{i j} \geq 0$ for $i \neq j$. See Van Veldhuizen ${ }^{7}$. For general nonlinear semi-discretizations $w^{\prime}(t)=F(t, w(t))$, it appears that unconditional positivity is a very restrictive requirement. Suppose that $F(t, w(t))$ satisfies the condition:

Condition 3.1 There is an $\alpha>0$, with $\alpha$ as large as possible, such that $\alpha \tau \leq 1$ and

$$
\begin{equation*}
v+\tau F(t, v) \geq 0 \quad \text { for all } \quad t \geq 0, v \geq 0 \tag{3}
\end{equation*}
$$

Application of Euler Forward to the nonlinear system $w^{\prime}(t)=F(t, w(t))$ gives

$$
\begin{equation*}
w_{n+1}=w_{n}+\tau F\left(t_{n}, w_{n}\right) \tag{4}
\end{equation*}
$$

Provided that $w_{n} \geq 0$, Condition 3.1 guarantees positivity for $w_{n+1}$ computed via Euler Forward (4). Furthermore, assume that $F(t, w(t))$ also satisfies :

Condition 3.2 For any $v \geq 0, t \geq 0$ and $\tau>0$ the equation

$$
\begin{equation*}
u=v+\tau F(t, u) \tag{5}
\end{equation*}
$$

has a unique solution that depends continuously on $\tau$ and $v$.
In Hundsdorfer et al. ${ }^{2}$ it has been proven that Condition 3.1 and 3.2 imply positivity for Euler Backward for any step-size $\tau$. We remark that Condition 3.2, unlike Condition 3.1, is not easy to be verified. As has been remarked in Hundsdorfer et al. ${ }^{2}$, it is sufficient to hold that $F(t, v)$ is continuously differentiable, and

$$
\begin{equation*}
\left\|I-\tau J_{F}(t, v)\right\| \leq C, \quad \text { for any } v \in \mathbb{R}^{n}, t \geq 0 \text { and } \tau>0 \tag{6}
\end{equation*}
$$

where $C$ is a positive constant and $J_{F}(t, v)$ the Jacobian matrix of derivatives of $F(t, v)$ with respect to $v$.

However, in practice the solutions of the resulting nonlinear systems have to be approximated, which might introduce small negative components. In the case that the negative components of the solution are the result of rounding errors, then it is justified to set them equal to zero. In the case one has negative components in the solution as consequence of the nonlinear (Newton) solver, then the most common method to avoid negative concentrations is clipping. This means that when a mass species is negative, it it set equal to zero. Clipping has the disadvantage that mass is no longer preserved. In practice it is better to avoid clipping.

In our experience negative concentrations are avoided by implementing a variable step size algorithm. In that case clipping can be avoided. We briefly explain the variable time stepping algorithm as it is implemented in our code. Consider an attempted step from $t_{n}$ to $t_{n+1}=t_{n}+\tau_{n}$ with time step size $\tau_{n}$. Suppose an estimate $D_{n}$ of order $\hat{p}$ of the norm of the local error is available. Then, if $D_{n}<T o l$ this step $\tau_{n}$ is accepted, whereas if $D_{n}>T o l$ the step is rejected and redone with a smaller time step size $\tau_{n}$. The new step size is computed as

$$
\begin{equation*}
\tau_{\text {new }}=r \cdot \tau, \quad r=\left(\frac{T o l}{D_{n}}\right)^{\frac{1}{\bar{p}+1}} \tag{7}
\end{equation*}
$$

It is also possible to put bounds on the growth factor $r$ of the new step size. This is simply done by giving bounds on $r$. If the new time step introduces negative species or Newton divergence, we adjust it, by taking $\tau=\frac{1}{2} \tau$, such that positivity is assured.

We conclude this section with the claim that Euler Backward is the only time integration that is unconditionally positive. For a proof we refer to Bolley et al. ${ }^{1}$. This means that for any higher order (implicit) time integration method a time step criterion is needed to ensure preservation of non-negativity.

## 4 Suitable Time Integration Methods (TIM)

In this section we briefly present integration methods that are suitable, from a theoretic point of view, for the time integration of the species equations. More comprehensive descriptions are given in Hundsdorfer et al. ${ }^{2}$ and Van Veldhuizen ${ }^{6}$. At the end of this section we will also make some remarks on the linear and nonlinear solvers.

From the previous section it is clear that the Euler Backward method is a suitable method to perform time integration. It has the advantage of being unconditionally positive. Disadvantages for the transient simulations are the first order consistency and the probably high computational costs for one time step. The latter is due to the fact that the succeeding approximations are computed in an implicit way.

### 4.1 Time Integration Methods

We will discuss a selection of time integration methods that have good properties in both stability and positivity, or Total Variation Diminishing (TVD).

## Rosenbrock Methods

The two stage Rosenbrock method ROS2

$$
\begin{align*}
w_{n+1} & =w_{n}+b_{1} k_{1}+b_{2} k_{2} \\
k_{1} & =\tau F\left(w_{n}\right)+\gamma \tau \mathbf{A} k_{1} \\
k_{2} & =\tau F\left(w_{n}+\alpha_{21} k_{1}\right)+\gamma_{21} \tau \mathbf{A} k_{1}+\gamma \tau \mathbf{A} k_{2} \tag{8}
\end{align*}
$$

with $\mathbf{A}=F^{\prime}\left(w_{n}\right)$ is the Jacobian matrix of $F\left(w_{n}\right)$, and $b_{1}=1-b_{2}, \alpha_{21}=\frac{1}{2 b_{2}}$ and $\gamma_{21}=-\frac{\gamma}{b_{2}}$, is interesting. The method is second order accurate for arbitrary $\gamma$ as long as $b_{2} \neq 0$. The stability function is given as

$$
\begin{equation*}
R(z)=\frac{1+(1-2 \gamma) z+\left(\gamma^{2}-2 \gamma+\frac{1}{2}\right) z^{2}}{(1-\gamma z)^{2}} \tag{9}
\end{equation*}
$$

For $\gamma \geq \frac{1}{4}$ the method is unconditionally stable and for $\gamma_{+}=1+\frac{1}{2} \sqrt{2}$, we have the property that $R(z) \geq 0$, for all negative real $z$. For diffusion-reaction problems, which have negative real eigenvalues, this property ensures positivity of the solution. It appears that the second order Rosenbrock method performs quite well for advection diffusion reaction equations with respect to the positivity property, as has been shown in Verwer et al. ${ }^{8}$. In Verwer et al. ${ }^{8}$ it is conjectured that the property $R(z) \geq 0$ for all negative real $z$ plays a role.

We conclude this section with a remark on the implementation of the ROS2 scheme (10). We implemented the ROS2 scheme with the parameters $b_{1}=b_{2}=\frac{1}{2}$ and $\gamma=\gamma_{+}$. In order to avoid the matrix vector multiplication in the second stage the equivalent form

$$
\begin{align*}
w_{n+1} & =w_{n}+\frac{3}{2} \tilde{k}_{1}+\frac{1}{2} \tilde{k}_{2} \\
\tilde{k}_{1} & =\tau F\left(w_{n}\right)+\gamma \tau \mathbf{A} \tilde{k}_{1} \\
\tilde{k}_{2} & =\tau F\left(w_{n}+\tilde{k}_{1}\right)-2 \tilde{k}_{1}+\gamma \tau \mathbf{A} \tilde{k}_{2} \tag{10}
\end{align*}
$$

where $\tilde{k}_{1}=k_{1}$ and $\tilde{k}_{2}=k_{2}-k_{1}$, is implemented.

## Backward Differentiation Formulas (BDF)

The $k$-step BDF methods are implicit, of order $k$ and defined as

$$
\begin{equation*}
\sum_{j=0}^{k} \alpha_{j} w_{n+j}=\tau F\left(t_{n+k}, w_{n+k}\right), \quad n=0,1, \ldots \tag{11}
\end{equation*}
$$

which uses the $k$ past values $w_{n}, \ldots, w_{n+k-1}$ to compute $w_{n+k}$. Remark that the most advanced level is $t_{n+k}$ instead of $t_{n+1}$. The 1 -step BDF method is Backward Euler. The 2-step method is

$$
\begin{equation*}
\frac{3}{2} w_{n+2}-2 w_{n+1}+\frac{1}{2} w_{n}=\tau F\left(t_{n+2}, w_{n+2}\right) \tag{12}
\end{equation*}
$$

and the three step BDF is given by

$$
\begin{equation*}
\frac{11}{6} w_{n+3}-3 w_{n+2}+\frac{3}{2} w_{n+1}-\frac{1}{3} w_{n}=\tau F\left(t_{n+3}, w_{n+3}\right) \tag{13}
\end{equation*}
$$

Remark 4.1 A disadvantage of linear multi-step methods is that the first $k-1$ approximations cannot be computed with the linear $k$-step scheme. To compute the first $(k-1)$ approximations, one could use for the first step a BDF 1-step method, for the second approximation a BDF 2-step method, ... and for the $(k-1)^{\text {st }}$ approximation a BDF $(k-1)$-step scheme.

For the 2-step BDF method we obtain positivity, under Conditions 3.1 and 3.2, of $w^{\prime}(t)=$ $F(t, w(t))$ whenever $\alpha \tau \leq \frac{1}{2}$, provided that $w_{1}$ is computed from $w_{0}$ by a suitable starting procedure, i.e., $w_{1}$ has been computed such that $w_{1} \geq 0$ holds. For a derivation we refer to Hundsdorfer et al. ${ }^{2}$ and Van Veldhuizen ${ }^{6}$.

## IMEX Runge-Kutta Chebyshev Methods

The second order Runge-Kutta Chebyshev method is given as

$$
\begin{align*}
w_{n 0}= & w_{n} \\
w_{n 1}= & w_{n}+\tilde{\mu}_{1} \tau F\left(t_{n}+c_{0} \tau, w_{n 0}\right) \\
w_{n j}= & \left(1-\mu_{j}-\nu_{j}\right) w_{n}+\mu_{j} w_{n, j-1}+\nu_{j} w_{n, j-2}+j=1, \ldots, s \\
& +\tilde{\mu}_{1} \tau F\left(t_{n}+c_{j-1} \tau, w_{n, j-1}\right)+\tilde{\gamma}_{j} \tau F\left(t_{n}+c_{0} \tau, w_{n 0}\right),  \tag{14}\\
w_{n+1}= & w_{n s} .
\end{align*}
$$

The coefficients $\tilde{\mu}_{1}, \mu_{j}, \nu_{j}, \tilde{\mu}_{j}$ and $\tilde{\gamma}_{j}$ are available in analytical form for $s \geq 2$ :

$$
\begin{gather*}
\tilde{\mu}_{1}=b_{1} \omega_{1} \quad \text { and for } \quad j=2, \ldots, s  \tag{15}\\
\mu_{j}=\frac{2 b_{j} \omega_{0}}{b_{j-1}}, \quad \nu_{j}=\frac{-b_{j}}{b_{j-2}}, \quad \tilde{\mu}_{j}=\frac{2 b_{j} \omega_{1}}{b_{j-1}}, \quad \tilde{\gamma}_{j}=-a_{j-1} \tilde{\mu}_{j} \tag{16}
\end{gather*}
$$

where

$$
\begin{equation*}
b_{0}=b_{2}, \quad b 1=\frac{1}{\omega_{0}}, \quad b_{j}=\frac{T_{j}^{\prime \prime}\left(\omega_{0}\right)}{\left(T_{j}^{\prime}\left(\omega_{0}\right)\right)^{2}}, \quad j=2, \ldots, s \tag{17}
\end{equation*}
$$

with

$$
\begin{equation*}
\omega_{0}=1+\frac{\varepsilon}{s^{2}}, \quad \text { and } \quad \omega_{1}=\frac{T_{s}^{\prime}\left(\omega_{0}\right)}{T_{s}^{\prime \prime}\left(\omega_{0}\right)} \tag{18}
\end{equation*}
$$

Furthermore,

$$
\begin{equation*}
c_{0}=0, c_{1}=c_{2}, c_{j}=\frac{T_{s}^{\prime}\left(\omega_{0}\right)}{T_{s}^{\prime \prime}\left(\omega_{0}\right)} \frac{T_{j}^{\prime \prime}\left(\omega_{0}\right)}{T_{j}^{\prime}\left(\omega_{0}\right)}, c_{s}=1 \tag{19}
\end{equation*}
$$

and,

$$
\begin{equation*}
a_{j}=1-b_{j} T_{j}\left(\omega_{0}\right) . \tag{20}
\end{equation*}
$$

In (17) - (20) $T_{j}(x)$ are the Chebyshev polynomials of the first kind satisfying the recursion

$$
\begin{equation*}
T_{j}(x)=2 x T_{j-1}(x)-T_{j-2}(x), \quad j=2, \ldots, s \tag{21}
\end{equation*}
$$

with $T_{0}(x)=1$ and $T_{1}(x)=x$. Furthermore, in (18) $\varepsilon$ is a free parameter. In Figure 2 its stability region is given. The parameter $\beta(s)$ moves to $-\infty$ when the number of stages $s$ increases. For $\varepsilon$ small, by which we mean $\frac{\varepsilon}{s^{2}} \ll 1$, the stability bound $\beta(s)$ satisfies

$$
\begin{equation*}
\beta(s) \approx \frac{2}{3}\left(s^{2}-1\right)\left(1-\frac{2}{15} \varepsilon\right) . \tag{22}
\end{equation*}
$$



Figure 2: Stability region of (14) for $s=5$.
The IMEX extension of (14) is as follows. Suppose we have an ODE system $w^{\prime}(t)=$ $F(t, w(t))$, where $F(t, w(t))$ can be split as

$$
\begin{equation*}
F(t, w(t))=F_{E}(t, w(t))+F_{I}(t, w(t)) \tag{23}
\end{equation*}
$$

In (23) the term $F_{I}(t, w(t))$ is the part of $F$ which is (supposed to be) too stiff to be integrated by an explicit Runge-Kutta Chebyshev method. Obviously, the term $F_{E}(t, w(t))$ is the moderate stiff part of $F$ that can be integrated in an explicit manner using RKC methods. The first stage of (14) becomes in the IMEX-RKC scheme

$$
\begin{equation*}
w_{n 1}=w_{n}+\tilde{\mu}_{1} \tau F_{E}\left(t_{n}+c_{0} \tau, w_{n 0}\right)+\tilde{\mu}_{1} \tau F_{I}\left(t_{n}+c_{1} \tau, w_{n 1}\right) \tag{24}
\end{equation*}
$$

with $\tilde{\mu}_{1}$ as defined before. Note that the highly stiff part of $F$ is treated implicitly. The other $(s-1)$ subsequent stages of (14) will be modified in a similar way, such that in each of the remaining $s-1$ stages the solution of a system of nonlinear algebraic equations

$$
\begin{equation*}
w_{n j}-\tilde{\mu}_{1} \tau F_{I}\left(t_{n}+c_{j} \tau, w_{n j}\right)=v_{j} \tag{25}
\end{equation*}
$$

with $v_{j}$ a given vector, is required.
With respect to stability of this IMEX extension of (14) we remark that the implicit part is unconditionally stable, whereas the stability condition for the explicit part remains unchanged. For more background we refer to Verwer et al. ${ }^{9,10}$.

### 4.2 Nonlinear and Linear Solvers

In all TIM from Section 4.1, except for the ROS2 scheme, nonlinear systems $F(x)=0$, $x \in \mathbb{R}^{n}$ have to be solved. The Newton iteration is, with its second order convergence, an obvious choice. The disadvantage of having local convergence will disappear if one uses a line-search algorithm, such that the succeeding iterates are norm reducing, i.e.,

$$
\begin{equation*}
\left\|F\left(x^{k+1}\right)\right\| \leq\left\|F\left(x^{k}\right)\right\| \quad k=0,1,2, \ldots \tag{26}
\end{equation*}
$$

for some norm in $\mathbb{R}^{n}$. More background information can be found in Kelley ${ }^{3}$ and Van Veldhuizen ${ }^{6}$.

In the Newton iteration linear systems have to be solved. In most 2D applications direct solvers like LU factorization are still applicable. To reduce the amount of work one usually reorders the unknowns, in order to reduce the bandwidth of the matrix. Also in our case it is possible to reduce the bandwidth of the Jacobian considerably. The way to do this is described in Van Veldhuizen ${ }^{7}$.

For 3D problems direct solvers (LU factorization) are no longer applicable. To approximate the solution of the linear systems one has to switch to iterative linear solvers like, for instance, Krylov Subspace methods. On the last topic research is still in progress.

## 5 Numerical Results

In this section we compare the performance of the TIM of Section 4.1 for solving the species equations. If necessary, nonlinear systems will be solved by the full Newton iteration. Linear systems will be solved using the LU factorization with rearranging, as mentioned in Section 4.2. As mentioned before, the velocity, temperature, density and pressure field are computed by Kleijn ${ }^{5}$. The experiments are done in FORTRAN. The computations are done on a serial Pentium $4(2.8 \mathrm{GHz})$ computer with 1 Gb memory capacity. Moreover, the code is compiled with FORTRAN g 77 on LINUX.

At $t=0$ we start with the zero concentration profile for all species, except the carrier gas, and let the reactive specie silane $\mathrm{SiH}_{4}$ enter the reactor at the inflow boundary. Then, we stop the simulation at steady state, which is reached when the relative change of the
solution vector is less than $10^{-6}$, i.e.,

$$
\begin{equation*}
\frac{\left\|u_{n+1}-u_{n}\right\|}{\left\|u_{n}\right\|} \leq 10^{-6} \tag{27}
\end{equation*}
$$

For a comparison between the workloads of the various TIM, we look to the amount of CPU time, the number of time steps and the total number of Newton iterations (if needed) it takes to reach steady state.

The solutions computed by different TIM also have been compared with a reference solution $u_{r e f}$, which has been computed with high accuracy. It appeared that the solutions of the different TIM, denoted by $u_{T I M}$, all had the same quality, by which we mean that

$$
\begin{equation*}
\frac{\left\|u_{T I M}-u_{r e f}\right\|}{\left\|u_{r e f}\right\|}=\mathcal{O}\left(10^{-7}\right) . \tag{28}
\end{equation*}
$$

For all TIM a variable step size controller is implemented as described in Section 3.1. Moreover, for the IMEX-RKC scheme an algorithm is implemented which determines the number of stages $s$ based on the estimated time step size given by the time step size controller. For a comprehensive description of this algorithm we refer to Van Veldhuizen ${ }^{7}$ and Verwer et al. ${ }^{10}$.

In Figure 3 the residual $\|F(w(t))\|_{2}$ versus the time step, for different TIM except the IMEX-RKC scheme, is given. In Figure 4 the time step size versus time step is given for the same TIM. Results for the IMEX-RKC scheme are given in Figure 5. Recall that $F(w(t))$ is the right hand side of the semi-discretization (2). We conclude with the contour plots of the steady state solution in Figure 6.

| TIM | CPU time | \# time steps | \# Newton iterations |
| :---: | :---: | :---: | :---: |
| Euler Backward | 1061 CPU sec | 120 | 236 |
| ROS2 | 579 CPU sec | 190 | - |
| BDF-2 | 689 CPU sec | 99 | 182 |
| IMEX-RKC | 13000 CPU sec | 1127 | 9075 |

Table 1: Workloads of various TIM

## 6 Conclusions

Based on Table 1, we conclude that for this 2D test-problem Rosenbrock is the cheapest TIM to solve the system of specie equations (1). The 'bad' performance of the IMEX RKC scheme is due to that per time step $s$ nonlinear systems have to be solved. Although these nonlinear systems are cheaper to solve, it did not pay off in this 2D test case. However, this property can become interesting in 3D simulations.


Figure 3: Residual $\|F(w(t))\|_{2}$ versus time step for Euler Backward (EB), second order Rosenbrock (ROS2) and the BDF-2 method (BDF)


Figure 4: Time step size versus time step for Euler Backward (EB), second order Rosenbrock (ROS2) and the BDF-2 method (BDF)


Figure 5: Residual $\|F(w(t))\|_{2}$ and time step size for the IMEX Runge-Kutta-Chebyshev scheme versus time step


Figure 6: Contour plots of the mass fractions of $\mathrm{SiH}_{4}$ (a), $\mathrm{SiH}_{2}$ (b), $\mathrm{H}_{2} \mathrm{SiSiH}_{2}$ (c), $\mathrm{Si}_{2} \mathrm{H}_{6}$ (d), $\mathrm{Si}_{3} \mathrm{H}_{8}$ (e) and $\mathrm{H}_{2}(\mathrm{f})$. The outflow boundary is situated on $r=15.0 \mathrm{~cm}$. to $r=17.5 \mathrm{~cm}$.

## REFERENCES

[1] C. Bolley and M. Crouzeix, Conservation de la Positivité Lors de la Discrétisation des Problèmes d'Évolution Paraboliques, RAIRO Anal. Numer. 12, pp. 237-245, (1973)
[2] W. Hundsdorfer and J.G. Verwer, Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations, Springer Series in Computational Mathematics, 33, Springer, Berlin, (2003)
[3] C.T. Kelley, Solving Nonlinear Equations with Newton's Method, Fundamentals of Algorithms, SIAM, Philadelphia, (2003)
[4] C.R. Kleisn, Transport Phenomena in Chemical Vapor Deposition Reactors, PhD thesis, Delft University of Technology, Delft, (1991)
[5] C.R. Kleijn, Computational Modeling of Transport Phenomena and Detailed Chemistry in Chemical Vapor Deposition- A Benchmark Solution, Thin Solid Films, 365, pp. 294-306, (2000)
[6] S. van Veldhuizen, Efficient Solution Methods for Stiff Systems of Advection-Diffusion-Reaction Equations, Literature Study, Technical Report at the Delft University of Technology, TWA-05-05, Delft, (2005)
[7] S. van Veldhuizen, Numerical Methods for CVD Simulation, Technical Report at the Delft University of Technology, TWA-06-07, Delft, (2006)
[8] J.G. Verwer, E.J. Spee, J.G. Blom and W. Hundsdorfer, A Second-Order Rosenbrock Method Applied to Photochemical Dipersion Problems, SIAM Journal on Sci. Comp., 20, pp.1456-1480, (1999)
[9] J.G. Verwer and B.P. Sommeijer, An Implicit-Explicit Runge-KuttaChebyshev Scheme for Diffusion-Reaction Equations, SIAM Journal on Sci. Comp., 25, pp.1824-1835, (2004)
[10] J.G. Verwer, B.P. Sommeijer and W. Hundsdorfer, RKC Time-Stepping for Advection-Diffusion-Reaction Problems, Journal of Comp. Physics, 201, pp. 61-79, (2004)

