**Inaugural-Dissertation**
zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen Gesamtfakultät
der
Ruprecht-Karls-Universität
Heidelberg

vorgelegt von
Dipl. Math. Daniel Gerecht
aus Friedberg(Hessen)

Tag der mündlichen Prüfung:

# Adaptive Finite Element Simulation of Coupled PDE/ODE Systems

## Modeling Intercellular Signaling

Betreuer: Prof. Dr. Dr. h.c. Rolf Rannacher

# Abstract

Intercellular signaling is a defining property of multicellular organisms, yet the spatio-temporal dynamics remain poorly understood. The subject of this work is the design of an efficient numerical algorithm for simulations of intercellular signaling in multicellular 3D environments modeled by coupled systems of partial differential equations (PDE) and ordinary differential equations (ODE).

The PDE part of these systems consists of reaction-diffusion equations and describes the concentration distribution of diffusible messengers, e.g. cytokines. Intracellular dynamics are described by a small number of ODEs per cell. Thus, every single iteration of a commonly used decoupling scheme has similar computational costs than solving the coupled PDE/ODE system at once. We therefore develop an efficient multilevel preconditioner for the coupled system. The computational cost of both coupled and decoupled solution methods are investigated for model problems of different coupling strength.

To keep the computational costs of the 3D simulations moderate, we use methods for adaptive mesh refinement. We discretize the system by different time meshes for the PDE and the ODE part to reduce the number of computationally expensive PDE time steps. Reliable a posteriori error estimations for coupled PDE/ODE systems are derived by means of the 'Dual Weighted Residual' (DWR) method. The discretization error is split into the contributions of the PDE and the ODE part. We compute local error indicators in space and time and set up an efficient adaptive mesh refinement method. The described methods are validated by numerical tests for several biologically motivated model problems.

We apply the developed numerical methods and simulate cytokine signaling between T cells in lymph nodes, which regulates the adaptive immune response in the human body. The numerical results show that, despite the high diffusivity of cytokines, highly localized cytokine concentrations with large gradients occur, which enables short-range cell-to-cell communication.

## Zusammenfassung

Die Signalübertragung zwischen Zellen ist von entscheidender Bedeutung für mehrzellige Lebewesen, jedoch ist über die Dynamiken in Raum und Zeit nur wenig bekannt. Der Fokus dieser Arbeit liegt auf der Erstellung eines effizienten numerischen Algorithmus zur 3D-Simulation von interzellulärer Signalübertragung, die durch gekoppelte Gleichungssysteme von partiellen Differentialgleichungen (PDE) und gewöhnlichen Differentialgleichungen (ODE) modelliert wird.

Der PDE-Teil dieser Systeme besteht aus Reaktions-Diffusionsgleichungen und beschreibt die Konzentrationsverteilung diffusibler Botenstoffe, z.B. Zytokine. Die intrazellulären Vorgänge werden durch wenige ODEs pro Zelle beschrieben. Die Berechnungskosten für ein gekoppeltes Verfahren für das gesamte PDE/ODE System sind daher ähnlich hoch, wie die Kosten für jede einzelne Iteration der gewöhnlich verwendeten entkoppelten Verfahren. Aus diesem Grund entwickeln wir einen effizienten Mehrgitter-Vorkonditionierer für das gekoppelte System. Die Effizienz sowohl gekoppelter als auch entkoppelter Lösungsverfahren wird für verschieden stark gekoppelte Modellprobleme untersucht.

Um die Berechnungskosten der 3D-Simulationen moderat zu halten, verwenden wir adaptive Gitterverfeinerungsstrategien. Wir diskretisieren die Gleichung mit unterschiedlichen Zeitgittern für den PDE- und den ODE-Teil, um die Anzahl der kostenintensiven PDE-Zeitschritte zu reduzieren. Auf Basis der dualgewichteten Residuenmethode (DWR) wird ein verlässlicher a posteriori Fehlerschätzer hergeleitet. Der Diskretisierungsfehler wird in die Beiträge des PDE- und ODE-Anteils aufgeteilt. Wir berechnen lokale Fehlerindikatoren in Raum und Zeit und verwenden diese für eine effiziente adaptive Gitterverfeinerungsstrategie. Die beschriebenen Methoden werden anhand einer Reihe biologisch motivierter Modellprobleme durch numerische Tests validiert.

Wir verwenden die entwickelten numerischen Methoden für die Simulation von Zytokinsignalen zwischen T-Zellen in Lymphknoten. Zytokinsignale regulieren die adaptive Immunantwort im menschlichen Körper. Die numerischen Ergebnisse zeigen, dass sich trotz hoher Diffusivität der Zytokine große Gradienten ausbilden, die eine zielgerichtete Kommunikation zwischen benachbarten Zellen ermöglichen.

# Contents

# 1 Introduction

In this work, we design an efficient numerical algorithm for simulating intercellular signaling in multicellular 3D environments.

Cellular signaling has been mathematically described by a variety of different models mostly relying on large systems of ordinary differential equations (ODE) [40]. These earlier models were extended by partial differential equations (PDE) to accurately describe concentration gradients [13, 19, 26, 44, 62]. In these receptor based models the concentration distributions of small signaling proteins are described by PDEs, i.e. reaction-diffusion equations, coupled on the cell surfaces with ODEs for the intracellular dynamics.

In this thesis, we focus on two parts of an effective method for solving coupled PDE/ODE systems, namely a coupled multilevel solver and adaptive mesh refinement. With the developed numerical methods, we gain insight into the competition of T helper cells for cytokines as part of the immune response.

### Fully Coupled Multilevel Solver

In literature, there are mainly two strategies for implicit solvers of nonlinear systems: the linearization based approaches considered in this thesis i.e. of Newton-type and nonlinear methods such as the nonlinear multigrid method [12, 29]. Newton-type methods provide a flexible and reliable framework and can be expanded by nonlinear methods such that for instance a nonlinear multigrid can be used as a preconditioner [47].

A decoupled solution scheme is often used when restrictions on accuracy can be relaxed in order to allow an easier numerical treatment of complicated problems. Such an approach makes it possible to reuse existing solvers and is widely used in numerical methods for coupled systems, see [23, 25, 33, 48, 49, 61]. In case of strongly coupled equations, this strategy can only be implemented at high computational cost through very small time steps or a higher number of iterations in the splitting scheme. Additionally, the size of the discretized PDE/ODE system is only slightly larger than the discretization of the PDE part alone, due to the small number of ODEs compared to the degrees of freedom for a discretization of the computational domain. We study the effect of splitting the linearization in a Newton-type solver and present a systematic comparison of coupled and decoupled solution schemes for PDE/ODE systems. A quantitative index based on a sensitivity analysis is used to compute the strength of the coupling.

It is well known [11, 29] that multilevel preconditioners are among the most efficient solving techniques for PDE. Therefore, we present a coupled solving scheme based on a Krylov-type solver preconditioned by a linear multigrid method. In particular, the transfer operators and smoother of the multigrid algorithm have been specifically adjusted to the formulation of the considered coupled problem.

### Adaptive Mesh Refinement

We investigate adaptive mesh refinement strategies to efficiently reduce the size of the discretized problem while preserving the accuracy of the approximation. A wide variety of adaptive methods exist, which are commonly used in the finite element context. The aim of these methods is mostly the reduction of local gradients or error reduction with respect to global norms. One is often interested in a specific quantity, e.g. the number of receptors at the final time point of the simulation for cellular signaling during the immune response. The Dual Weighted Residual (DWR) Method for a posteriori error estimation with respect to a given quantity of interest has been derived by Becker and Rannacher [8] and has been generalized for nonlinear parabolic problems by Schmich and Vexler [59]. Localized error indicators obtained by the DWR method allow to set up a versatile algorithm for adaptive mesh refinement and the equilibration of temporal and spatial discretization errors.

In this thesis, we extend the space-time Galerkin methods developed by Schmich and Vexler to coupled PDE/ODE systems in which a coarser temporal discretization is used for the PDE part of the system. This approach reduces the number of computationally expensive time steps of the PDE part and preserves the same accuracy by an increased number of cheaper ODE time steps. To cope with the instabilities triggered by inexact starting values, we derive the Galerkin finite element spaces corresponding to a temporal discretization by a damped Crank-Nicolson time stepping scheme, based on previous work by Goll et al. [28].

Space-time Galerkin methods have already been applied successfully to create adaptive grids, e.g. in the simulation of incompressible flows [7, 9, 37]. Coupled

PDE/ODE systems, emerging from cellular signaling, differ from previously investigated problems by the fast diffusivity of the PDE compared to the timescale of the activation process. We show in this thesis that the DWR method can be used as a reliable a posteriori error estimator for such systems and assess when, depending on the diffusivity, significant gains in accuracy by adaptive grids can be expected.

### Application: Cytokine signaling between T cells

Intercellular signaling is a defining property of multicellular organisms. In particular, the adaptive immune response relies on cell-to-cell communication of T helper cells by uptake and release of diffusible cytokines. Cytokines are small proteins, which play an important role in cell signaling. The regulation of the cytokine levels is key to type and strength of the immune response. Critically high values of cytokine can cause severe damage to the mammalian immune system, see Rochman et al. [55]. Quantitative understanding of such cytokine-driven cellular decisions is beginning to emerge [13, 24, 66], yet the underlying spatio-temporal cytokine dynamics remain poorly understood.

A combined mathematical and theoretical study by Busse et al. [13] has identified secretion and uptake of the cytokine interleukin-2 (IL-2) as a possible mechanism mediating immune suppression by 'regulatory T cells'. The model captures IL-2 signaling in the initial phase after antigen stimulation before the initiation of cell division (up to 30 hours). The proposed system of equations consists of a reaction diffusion equation describing the distribution of IL-2 in the intercellular area coupled with ODEs for the intracellular processes controlling IL-2 uptake and release.

A long-standing question in immunology has been how the spatial range of cytokine signaling is controlled in realistic multicellular geometries. In particular, the observation that diffusible messengers regulate the adaptive immune responses to infection has been puzzling because the long-range action of cytokines on neighboring cells would compromise the specificity of an immune response. It has therefore been suggested by Davis et al. [20] that cytokine is channeled through the tight synapse between two immune cells and that the synapse has an important function for IL-2 signaling. With the developed numerical methods we compute cytokine secretion and uptake within a population of immune cells in three-dimensional environments. We find by our numerical simulations that despite the high diffusivity of cytokines short-range spatial gradients occur. The rate of competitive cytokine uptake by cells causes highly localized cytokine concentrations with large gradients that result in short-range cell-to-cell communication.

### Outline:

The outline of this thesis is as follows:

#### Notation and Mathematical Model

In Chapter 2, we introduce basic notation, which will be used in this thesis, and give a brief introduction to coupled PDE/ODE systems for cellular signaling. Exemplarily, we present a biological model, which describes T cell interaction during the adaptive immune response. The existence and uniqueness of the solutions of the corresponding coupled PDE/ODE system is proven.

#### Discretization

In Chapter 3, the coupled equation is discretized in space and time by means of the Galerkin finite element method. We present a temporal discretization which corresponds to the damped Crank-Nicolson time marching scheme. Furthermore, we extend this discretization for different time meshes for the PDE and the ODE parts. The spatial discretization is performed by a standard finite element approach.

#### Coupled and Decoupled Linear Solver

In Chapter 4, we analyze the sensitivities of the coupling of PDE/ODE systems. We investigate both coupled and decoupled solving methods and interpret the results depending on the strength of the coupling. Existing multilevel preconditioners are extended for fully coupled solving schemes. The efficiency of these methods is shown numerically.

#### A Posteriori Error Estimation

In Chapter 5, we derive an a posteriori error estimator based on the DWR method for coupled PDE/ODE equations to estimate both the temporal and the spatial error with respect to a quantity of interest. Subsequently, we present an adaptive algorithm for successive mesh refinement of the spatial and temporal grids to equilibrate spatial and temporal errors from both the PDE and the ODE parts of the system. We motivate the investigation of a series of model problems in 2D and 3D for intercellular signaling and display the effectivity of the estimated error indicators.

### Application: Three-dimensional gradients of cytokine signaling between T cells

In Chapter 6, the developed numerical methods are applied to simulate large three-dimensional immune cell cluster and help to provide a better quantitative understanding of the molecular mechanism of the immune response. We conclude the chapter with an analysis of the signal range of T helper cells in large-scale simulations with over 2000 cells.

### Conclusion and Outlook

Finally, we summarize in Chapter 7 the presented results and give an outlook to extensions for future work.

# 2 Notation and Mathematical Model

In this chapter, we present the basic notation for this thesis and exemplarily present a PDE/ODE system modeling cellular signaling. We begin in Section 2.1 with the basic notation for this thesis. We give a brief introduction in Section 2.2 to systems of partial differential equations (PDE) coupled with ordinary differential equations (ODE) for intercellular signaling. A model for T cell interaction in the lymph node during the immune response is presented exemplarily in Section 2.3. We conclude the chapter in Section 2.4 with the proof of existence and uniqueness of the model.

## 2.1 Basic notations

For $d \in \{2, 3\}$, let $\Omega \subseteq \mathbb{R}^d$, be an open bounded domain with boundary $\partial\Omega$. We denote by $n$ the outer unit normal vector to $\partial\Omega$.

The standard Lebesgue space $L^p(\Omega), 1 \leq p < \infty$, consists of all measurable functions $u : \Omega \to \mathbb{R}$ which are Lebesgue-integrable to the $p$-th power. It is a Banach space with the norm

$$||u||_{L^p(\Omega)} = \left( \int\limits_{\Omega} |u(x)|^p \, dx \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty.$$

In the special case of $p = 2$ the space $L^2(\Omega)$ becomes a Hilbert space with the inner product

$$(u, v)_{L^2(\Omega)} := \int\limits_{\Omega} u(x)v(x) \, dx.$$

We denote the Sobolev spaces by $W^{m,p}(\Omega), m \in \mathbb{N}, 1 \leq p < \infty$, which are defined as the space of functions whose distributional derivatives of order up to $m$ are in $L^p(\Omega)$. They are Banach spaces with the norm

$$||u||_{W^{m,p}(\Omega)} := \left( \sum_{|\alpha| \leq m} ||\partial^\alpha u||^p_{L^p(\Omega)} \right)^{\frac{1}{p}}.$$

By $\alpha = (\alpha_1, ..., \alpha_d) \in \mathbb{N}^d$ we denote a multi-index. We use the following notation

$$|\alpha| := \sum_{j=1}^{d} \alpha_j, \quad \partial^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} ... \partial x_d^{\alpha_d}}.$$

In case of $p = 2$, the Sobolev space $H^m(\Omega) := W^{m,2}(\Omega)$ is a Hilbert space with the inner product

$$(u, v)_{H^m(\Omega)} := \sum_{|\alpha| \leq m} (\partial^\alpha u, \partial^\alpha v)_{L^2(\Omega)}.$$

We drop the subscripts in case $p = 2$ to shorten the notation

$$(u, v) := (u, v)_{L^2(\Omega)}, \quad ||u|| := ||u||_{L^2(\Omega)}.$$

The space $C^{k,\alpha}(I, X)$ with $k \in \mathbb{N}$ and $0 < \alpha \leq 1$ denotes the space of functions from $\bar{I}$ into any given space $X$, whose derivatives up to order $k$ are $\alpha$-Hölder continuous on $\Omega$. For a shorter notation, we set $C^k := C^{k,1}(\Omega, \mathbb{R})$.

Let $I := (0, T)$ with $0 < T < \infty$ be a bounded time interval. We denote by $L^p(I, X)$ the space of $L^p$-integrable functions $f$ from $I$ into $X$ for any Banach space $X$ and $1 \leq p < \infty$. It is itself a Banach space equipped with the norm

$$||f||_{L^p(I,X)} := \left( \int_I ||f(t)||^p_X \, dt \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty.$$

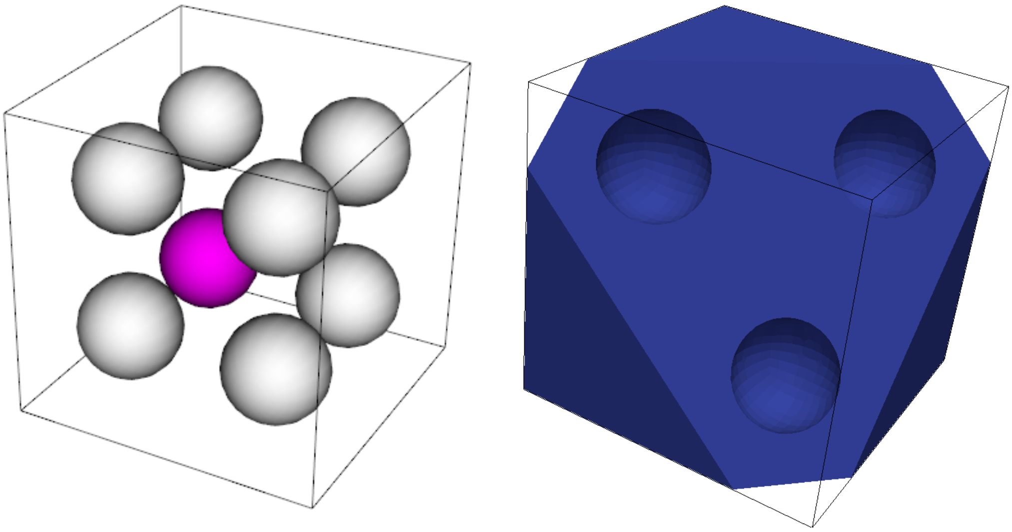
For a detailed derivation of these spaces, we refer to Wloka [69]. We use the shorter notation $((.,.))$ for the scalar product on $L^2(I \times \Omega)$

$$((u, v)) := (u, v)_{L^2(I \times \Omega)}.$$

## 2.2 Coupled PDE/ODE systems modeling cellular signaling

Cellular signaling has been modeled by a variety of different mathematical models, mostly relying on large systems of ordinary differential equations (ODE), see Kestler et al. [40]. In many cellular signaling pathways concentration gradients emerged, either because of the geometry of the cells [19], a competition between the cells for the signaling proteins [13] or biological reactions which enable pattern formation [44, 62]. To accurately model concentration gradients in the intercellular area, these earlier models were extended by partial differential equations (PDE). In these receptor based models, the concentration distribution of diffusible messengers are modeled by PDEs, i.e. reaction-diffusion equations, coupled on the cell surfaces with ODEs for the intracellular dynamics.

A model for intercellular signaling consists of PDEs for the interaction between the cells in the intercellular area $\Omega$ coupled with ODEs for intracellular processes. We denote by $N_c$ the number of cells in $\Omega$ and indicate by $\Gamma_i$ the boundary of each cell $i$ for $i = 1, \ldots, N_c$. The outer boundary of $\Omega$ is denoted by $\Gamma_{out}$. We indicate the union of all cell boundaries $\Gamma_i$ by $\Gamma$.



(a) 8 interacting cells with surfaces $\Gamma_i$

(b) Intercellular area $\Omega$ (sliced for visualization)

**Figure 2.1:** Model setup

Depending on the type of intercellular signaling, different nonlinear operators describe the dynamics in the intercellular area ($\mathcal{A}_\Omega$), e.g. degradation, the dynamics on the cell surfaces ($\mathcal{A}_{\Gamma_i}$) of each cell and the intracellular processes ($\mathcal{B}_i$). We denote the solution of the PDE part by $u$ and the vector of solutions of the ODE part by $v$.

$$\begin{aligned} \partial_t u - \mu\Delta u + \mathcal{A}_\Omega(u) &= 0 \quad \text{in } \Omega, \\ \mu\partial_n u + \mathcal{A}_{\Gamma_i}(u, v_i) &= 0 \quad \text{on } \Gamma_i, \quad \text{for all } 1 \leq i \leq N_c \\ \mu\partial_n u &= 0 \quad \text{on } \Gamma_{out}, \\ \partial_t v_i + \mathcal{B}_i(\widetilde{u}_i, v_i) &= 0, \qquad\qquad \text{for all } 1 \leq i \leq N_c \end{aligned} \tag{2.1}$$

with given initial values $u(0) = u^0$ and $v(0) = v^0$. We denote the average of $u$ on the surface of $\Gamma_i$ by $\widetilde{u}_i$ and the associated ODE values with this cell by $v_i$

$$\widetilde{u}_i(t) = \frac{\int_{\Gamma_i} u(t, s)\ \mathrm{ds}}{|\Gamma_i|}. \tag{2.2}$$

Similar systems arise in different applications, e.g. in closed-loop cardiovascular simulations, where the PDE part is coupled with the ODE part on the surfaces for in- and outflow, see Moghadam et al. [48].

### Variational Formulation and Notation

For a variational formulation, we introduce the Hilbert space $V^p = H^1(\Omega)$ for the PDE part of the equation and the vector space $V^o = \mathbb{R}^n$ for the ODE part, where $n$ denotes the number of ordinary differential equations in the system. Let $V^{p\prime}$ be the topological dual of $V^p$, then $V^p \hookrightarrow L^2(\Omega) \hookrightarrow V^{p\prime}$ constitutes a Gelfand triple. We combine the two vector spaces $V^p$ and $V^o$ to the product space $V$, such that

$$V := V^p \times V^o.$$

Additionally, we define the product space $L$ by

$$L := L^2(\Omega) \times V^o,$$

which contains e.g. the initial values of the system. We define the operators $\bar{a} : V \times V^p \to \mathbb{R}$ and $\bar{b} : V \times V^o \to \mathbb{R}$ by a multiplication of the equations (2.1) with test functions $\phi \in V^p$ and $\psi \in V^o$,

$$\begin{aligned}
\bar{a}\left( \begin{pmatrix} u \\ v \end{pmatrix} ; \phi \right) &:= (\mathcal{A}(u,v), \phi)_\Omega + \sum_{1 \le i \le N_c} (\mathcal{A}_{\Gamma_i}(u, v_i), \phi)_{\Gamma_i}, \\
\bar{b}\left( \begin{pmatrix} u \\ v \end{pmatrix} ; \psi \right) &:= \sum_{1 \le i \le N_c} (\mathcal{B}_i(\tilde{u_i}, v_i), \psi_i).
\end{aligned}$$

We introduce the following function spaces for time dependent functions

$$\begin{aligned}
W^p &:= \{u(t) \in L^2(I, V^p), \partial_t u(t) \in L^2(I, V^{p\prime}) \; ; \; t \in I\}, \\
W^o &:= \{v(t) \in L^2(I, V^o), \partial_t v(t) \in L^2(I, V^o) \; ; \; t \in I\}.
\end{aligned}$$

The product space $W$ is defined by

$$W := W^p \times W^o.$$

The product spaces $W^p$ and $W^o$ are embedded in the respective spaces of continuous functions. Thus, $W \subset C(I, V)$ holds true, see Wloka [69]. We define the operators $a : W \times W^p \to \mathbb{R}$ and $b : W \times W^o \to \mathbb{R}$ by integration over time of $\bar{a}$ and $\bar{b}$,

$$\begin{aligned}
a(w; \phi) &:= \int_0^T \bar{a}(w(t); \phi(t))\, dt, \\
b(w; \psi) &:= \int_0^T \bar{b}(w(t); \psi(t))\, dt.
\end{aligned}$$

We then write equation (2.1) for all $t > 0$ in the following weak formulation for a solution $w = \begin{pmatrix} u \\ v \end{pmatrix} \in W$, such that

$$\begin{aligned}
((\partial_t u, \phi)) + \mu\, ((\nabla u, \nabla \phi)) + a(w; \phi) + (u(0), \phi(0)) &= (u_0, \phi(0)), \quad \forall \phi \in W^p, \\
\int_0^T (\partial_t v \cdot \psi) + b(w; \psi)\, dt + (v(0), \psi(0)) &= (v_0, \psi(0)), \quad \forall \psi \in W^o,
\end{aligned} \tag{2.3}$$

where $\begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \in L$ denote the initial values of the system. With $\bar{w}, \bar{\varphi} \in V$ and $w, \varphi \in W$ and the notation

$$\bar{w} = \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix}, \quad w = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \bar{\varphi} = \begin{pmatrix} \bar{\phi} \\ \bar{\psi} \end{pmatrix}, \quad \varphi = \begin{pmatrix} \phi \\ \psi \end{pmatrix},$$

we can combine the two semi-linear forms to obtain the operators $\bar{\mathcal{F}} : V \times V \to \mathbb{R}$ and $\mathcal{F} : W \times W \to \mathbb{R}$, such that

$$\begin{aligned} \bar{\mathcal{F}}(\bar{w}; \bar{\varphi}) &:= \mu(\nabla \bar{u}, \nabla \bar{\phi}) + \bar{a}(\bar{w}; \bar{\phi}) + \bar{b}(\bar{w}; \bar{\psi}), \\ \mathcal{F}(w; \varphi) &:= \mu((\nabla u, \nabla \phi)) + a(w; \phi) + b(w; \psi). \end{aligned} \tag{2.4}$$

We restate equation (2.3) with initial value $w_0 := \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$ in the shorter notation:

*Given $w_0 \in L$, find $w \in W$, such that:*

$$((\partial_t w, \varphi)) + \mathcal{F}(w; \varphi) + (w(0), \varphi(0)) = (w_0, \varphi(0)), \quad \forall \varphi \in W. \tag{2.5}$$

Given appropriate parameters, the solution converges in time to a unique steady state. A solution scheme for the computation of the steady state has the advantage of being independent of a specific temporal discretization. In Chapter 4, we test some parts of the algorithm by a direct computation of such a steady state $\bar{w} \in V$, such that

$$\bar{\mathcal{F}}(\bar{w}; \varphi) = 0, \quad \forall \varphi \in V. \tag{2.6}$$

## 2.3 A model for intercellular signaling

Exemplarily, we focus in this thesis on a model for interleukin-2 expression by T helper cells, developed by Busse et al. [13]. Interleukin concentrations in the intercellular area regulate the type and strength of the immune response. The model consists of a reaction-diffusion equation, which describes the distribution of IL-2 between the T helper (Th) cells in the intercellular area $\Omega$, coupled with ODEs for the intracellular processes by a Robin boundary condition. In this section, we display the mathematical model and explain the used parameters. The biological role of Th cells as part of the adaptive immune system is explained in more detail in Section 6.1.

This system is prototypical for all receptor based model for intercellular communication by diffusing messengers. Other coupled PDE/ODE models for cellular signaling consist, e.g. of a larger PDE/ODE system, which describe the interaction with multiple diffusing proteins and more detailed intracellular processes [19, 26], or allow for receptor gradients on the cell surface [44, 62].

We use the following notation:

- $u(t,s) : I \times \Omega \to \mathbb{R}$ describes the concentration distribution of IL-2 in the intercellular area.
- $R_i, C_i$ and $E_i : I \to \mathbb{R}$ describe the number of IL-2 receptors (IL-2R), built IL-2/IL-2R receptor-complexes and internalized complexes for each of the simulated Th cells. The receptors are distributed homogeneously on the cell surfaces.

The mathematical model consists of a PDE

$$\begin{aligned}
\partial_t u(t,x) &= \mu \Delta u(t,x) - k_d u(t,x) && \text{for all } (t,x) \text{ in } (0,T] \times \Omega, \\
\mu \partial_n u(t,s) &= q_i(t,s) - k_{on} R_i(t) u(t,s) + k_{off} C_i(t) && \text{for all } (t,s) \text{ in } (0,T] \times \Gamma_i, \\
\partial_n u(t,s) &= 0 && \text{for all } (t,s) \text{ in } (0,T] \times \Gamma_{out},
\end{aligned} \tag{2.7a}$$

coupled with three ODEs for each Th cell

$$\begin{aligned}
\partial_t R_i(t) &= w_i^0 + w_i^1 \frac{C_i(t)^3}{K^3 + C_i(t)^3} - k_{on} R_i(t) \widetilde{u}_i(t) \\
&\qquad - k_{iR} R_i(t) + k_{off} C_i(t) + k_{rec} E_i(t) \quad \text{for all cells } i = 1, \dots, N_c, \\
\partial_t C_i(t) &= k_{on} R_i(t) \widetilde{u}_i(t) - (k_{off} + k_{iB}) C_i(t), \\
\partial_t E_i(t) &= k_{iB} C_i(t) - (k_{rec} + k_{deg}) E_i(t), \\
\widetilde{u}_i(t) &= \frac{\int_{\Gamma_i} u(t,s) \, \mathrm{ds}}{|\Gamma_i|},
\end{aligned} \tag{2.7b}$$

with the initial conditions

$$\begin{aligned}
u(0,x) &= 0 && \text{for all } x \in \Omega, \\
R_i(0) &= \frac{w_i^0}{k_{iR}} && \text{for all cells } i = 1, \dots, N_c, \\
C_i(0) &= 0, \\
E_i(0) &= 0.
\end{aligned} \tag{2.7c}$$

We consider two cell types which share the same receptor dynamics but differ in the IL-2 secretion rate:

- Secreting Th cells, which omit IL-2 with the secretion rate $q_i > 0$ and
- Responding Th cells with $q_i = 0$.

We consider fixed evenly distributed positions of the Th cells during the course of the simulations, as depicted in Figure 2.1(a).

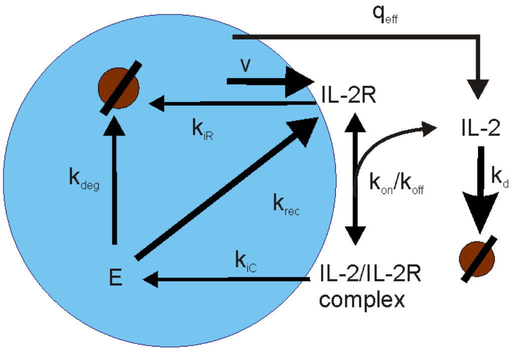


**Figure 2.2:** Intracellular receptor dynamics (picture taken from Thurley et al. [65])

The receptor dynamics depend on the IL-2 concentration in the surrounding medium, averaged over the cell surface $\Gamma_i$, because of the homogenous receptor expression. At the cell surface, IL-2 is bound to IL-2R molecules ($R_i$) forming IL-2/IL-2R complexes ($C_i$). Each Th cell has an uptake of IL-2 proportional to $k_{on}R_i$ on its surface. Receptor dynamics involve binding rates ($k_{on}$) of IL-2 by unbound receptors and unbinding rates by bound receptors ($k_{off}$), as well as internalization of bound and unbound receptors ($k_{iC}$, $k_{iR}$) and recycling and degradation of bound receptors ($k_{rec}$, $k_{deg}$).

The expression of receptors $R_i$ is regulated through a Hill Function, which constitutes a feedback term for the system. Because the feedback induced IL-2 receptor expression rate $w_i^1$ is much larger than the ordinary receptor expression rate $w_i^0$, a Th cell with enough receptor-complexes $C_i$ strongly up-regulates the number of free receptors $R_i$. If this 'activation' takes place, the cell captures more IL-2 than other Th cells with less free receptors. Internalized receptor-complexes enter an endosomal pool $E_i$, from which free receptors are recycled. Thus, receptor dynamics are governed by a set of nonlinear differential equations. The parameters of the system have been measured and estimated by Busse et al. [13], see Table 2.1. In

**Table 2.1:** Biological parameters, taken from Busse et al. [13]

| Symbol | Value | Parameter |
|---|---|---|
| $q_i$ | 0-22000 mol./$h$ | IL-2 secretion rate per cell |
| $\mu$ | 36000 $\mu m^2/h$ | Diffusion coefficient of IL-2 |
| $k_d$ | 0.1/$h$ | Extracellular IL-2 degradation |
| $w_i^0$ | 150 mol./cell/$h$ | Antigen stimulated IL-2 receptor expression rate |
| $w_i^1$ | 3000 mol./cell/$h$ | Feedback induced IL-2 receptor expression rate |
| $K$ | 1000 mol./cell | Half-saturation constant of feedback expression |
| $k_{on}$ | 111.6 /nM/$h$ | IL-2 association rate constant to IL-2 receptors |
| $k_{off}$ | 0.83/$h$ | IL-2 dissociation rate constant from IL-2 receptors |
| $k_{iR}$ | 0.64/$h$ | Internalization rate constant of IL-2 receptors |
| $k_{iC}$ | 1.7/$h$ | Internalization rate constant of receptor complexes |
| $k_{rec}$ | 9/$h$ | Recycling rate constant of IL-2 receptors |
| $k_{deg}$ | 5/$h$ | Endosomal degradation constant IL-2 receptors |
| $r$ | 5$\mu m$ | Cell radius |
| $d$ | 5$\mu m$ | Cell to cell distance |

the lymph node all T cells move slowly [56] and have the same neighbors during the course of hours. Thus, the model assumes fixed positions of the cells during the course of the simulations to reduce computational costs.

**Remark 2.1.** *To study the dynamical process and validate the model, we compute the entire trajectory. Nevertheless, the simulations converge to a stable steady state during the course of the simulations. Therefore, we consider direct computations of the steady state* (2.6) *in Section 4.3.2 as well.*

## 2.4 Existence and uniqueness

This proof follows closely the proof of Marciniak-Czochra and Ptashnyk [45] for a similar system. We assume, as usual for biological systems, positive values for all constants. The following notation is used throughout the proof:

$$v := \begin{pmatrix} v_1 \\ ... \\ v_{N_c} \end{pmatrix} \quad \text{and} \quad v_i := \begin{pmatrix} R_i \\ C_i \\ E_i \end{pmatrix} \quad \text{for all } 1 \leq i \leq N_c.$$

**Theorem 2.1.** *Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with piecewise smooth boundaries. There exists a unique solution $u \in W^p \cap C^{0,\beta/2}(\bar{I}, C^{0,\beta}(\bar{\Omega}, \mathbb{R}))$ where $\beta \in (0, \alpha]$ and $v_i \in W^o \cap C^1(\bar{I}, V^o)$ to the system (2.7), if the starting values $u_0$ and $v_0$ have positive values and satisfy $u_0 \in H^1(\Omega) \cap C^{0,\alpha}(\bar{\Omega}, \mathbb{R})$.*

*Proof. Existence:* The existence of a solution of the system (2.7) in $V$ can be shown using the Schauder fixed point theorem, as detailed by Marciniak-Czochra and Ptashnyk [45]. We obtain the positivity of solutions $u(t) \geq 0, v(t) \geq 0$ and their boundedness, such that $u(t) \leq c$ and $|v(t)| \leq c$ for all $t \in I$ .

*Uniqueness:* We suppose that there exist two solutions to the problem $w^1$ and $w^2$, with $w^1 := \begin{pmatrix} u^1 \\ v^1 \end{pmatrix}$ and $w^2 := \begin{pmatrix} u^2 \\ v^2 \end{pmatrix}$. We denote the components of the vectors $v^k$ and the average concentrations of $u^k$ on the cell surface by

$$v^k = \begin{pmatrix} v_1^k \\ \dots \\ v_{N_c}^k \end{pmatrix}, \quad v_i^k = \begin{pmatrix} R_i^k \\ C_i^k \\ E_i^k \end{pmatrix}, \quad \widetilde{u}_i^k := \frac{\int_{\Gamma_i} u^k \, ds}{|\Gamma_i|}, \qquad \text{for } k \in \{1, 2\}.$$

We use the weak formulation of system (2.7) and test with $u := u^1 - u^2$. The following holds for the PDE part of the system, after integrating over time from 0 to any $t \in \bar{I}$.

$$\int_0^t \int_\Omega (\frac{1}{2}\partial_t |u|^2 + \mu |\nabla u|^2 + k_d |u|^2) \, dx \, dt$$

$$= \int_0^t \sum_{1 \leq i \leq N_c} \int_{\Gamma_i} ((k_{off} C_i^1 - k_{on} R_i^1 u^1) - (k_{off} C_i^2 - k_{on} R_i^2 u^2), u) \, ds \, dt.$$

We obtain for the ODE part of the system, with the notation $R_i := R_i^1 - R_i^2$, $C_i := C_i^1 - C_i^2$ and $E_i := E_i^1 - E_i^2$ and $\widetilde{u}_i = \widetilde{u}_i^1 - \widetilde{u}_i^2$, that

$$\begin{aligned}
\partial_t R_i &= w_i^1 \frac{C_i^1(t)^3}{K^3 + C_i^1(t)^3} - w_i^1 \frac{C_i^2(t)^3}{K^3 + C_i^2(t)^3} - (k_{on} R_i^1 \widetilde{u}_i^1 - k_{on} R_i^2 \widetilde{u}_i^2) \\
&\quad - k_{iR} R_i + k_{off} C_i + k_{rec} E_i \\
\partial_t C_i &= (k_{on} R_i^1 \widetilde{u}_i^1 - k_{on} R_i^2 \widetilde{u}_i^2) - (k_{off} + k_{iB}) C_i \\
\partial_t E_i &= k_{iB} C_i - (k_{rec} + k_{deg}) E_i.
\end{aligned}$$

We integrate over time and add the three equations for each cells, such that

$$\begin{aligned}|R_i| + |C_i| + |E_i| \le \int\limits_0^\tau c_0 w_i^1\,|C_i| + 2k_{on}\left(\max_{[0.T]}\left|\widetilde{u}_i^1\right|\cdot|R_i| + \max_{[0,T]}\left|R_i^2\right|\cdot|\widetilde{u}_i|\right)\,dt\\ +\int\limits_0^\tau k_{iR}\,|R_i| + (2k_{off}+2k_{iB})\,|C_i| + (2k_{rec}+k_{deg})\,|E_i|\;dt,\end{aligned}$$

where $c_0$ denotes the Lipschitz-constant of the Hill function. The Gronwall lemma leads us to

$$|R_i| + |C_i| + |E_i| \le \tilde{c}\int\limits_0^\tau |\widetilde{u}_i|\;dt. \tag{2.8}$$

Using the above estimate and $k_{on} \ge 0$, we obtain for positive $\delta_1$ and $\delta_2$

$$\begin{aligned}\int\limits_0^t\int\limits_\Omega \left(\frac{1}{2}\partial_t|u|^2 + \mu|\nabla u|^2 + k_d|u|^2\right)dx\,dt \le k_{off}\tilde{c}\frac{1}{2\delta_1}\sum_{1\le i\le N_c}\int\limits_0^t\int\limits_0^\tau |\widetilde{u}_i|^2\,\mathrm{d}\tau\,dt\\ +\frac{\delta_1}{2}\int\limits_0^t\int\limits_\Gamma |u|^2\,ds\,dt + \underbrace{2\tilde{c}k_{on}\max_{[0.T]\times\Gamma}\left|u^2\right|}_{:=c_1}\frac{1}{2\delta_2}\sum_{1\le i\le N_c}\int\limits_0^t\int\limits_0^\tau |\widetilde{u}_i|^2\,\mathrm{d}\tau\,dt\\ +\frac{\delta_2}{2}\int\limits_0^t\int\limits_\Gamma |u|^2\;ds\,dt - k_{on}\sum_{1\le i\le N_c}\min_{[0,T]}R_i^2\int\limits_0^t\int\limits_{\Gamma_i}|u|^2\,ds\,dt.\end{aligned}$$

We use a boundary trace embedding theorem, see e.g. Adams and Fournier [1]

$$\int\limits_\Gamma |u|^2\,ds \le c_s(\Omega)\left(\int\limits_\Omega |u|^2\,dx + \int\limits_\Omega |\nabla u|^2\,dx\right) \tag{2.9}$$

and obtain the boundedness of $\widetilde{u}_i$ by the $H^1$-norm of $u$ by (2.9) and the continuous embedding of $L^2$ in $L^1$ with a constant $c_e$

$$|\widetilde{u}_i|^2 \le \frac{1}{|\Gamma_i|^2}\left(\int\limits_{\Gamma_i}|u|\;ds\right)^2 \le \frac{c_e}{|\Gamma_i|^2}\int\limits_{\Gamma_i}|u|^2\;ds \le \underbrace{\frac{c_eC_s}{|\Gamma_i|^2}}_{:=c'}\int\limits_\Omega |u|^2 + |\nabla u|^2\,dx. \tag{2.10}$$

We obtain by using (2.9) and (2.10) that

$$\begin{aligned}\frac{1}{2}\int\limits_{\Omega}|u|^2\,dx &+ \underbrace{(\mu - c_s\frac{\delta_1}{2} - c'c_1\frac{\delta_2}{2})}_{:=c_2}\int\limits_0^t\int\limits_{\Omega}|\nabla u|^2\,dx\,dt\\ &+ \underbrace{(k_d - c_s\frac{\delta_1}{2} - c'c_1\frac{\delta_2}{2})}_{:=c_3}\int\limits_0^t\int\limits_{\Omega}|u|^2\,dx\,dt\\ \leq \underbrace{(k_{off}\tilde{c}c'\frac{N_c}{2\delta_1} + c_1c_s'\frac{N_c}{2\delta_2})}_{:=c_4}&\int\limits_0^t\int\limits_0^\tau\int\limits_{\Omega}|u|^2 + |\nabla u|^2\,dx\,\mathrm{d}\tau\,dt.\end{aligned}$$

We choose $\delta_1$ and $\delta_2$ small enough, such that $c_2$ and $c_3$ take positive values. Now we take the supremum over $t \in \bar{I}$ to obtain with $c_5 := \max(c_4, \frac{c_4}{c_2}, \frac{c_4}{c_3})$

$$\begin{aligned}&\frac{1}{2}\int\limits_{\Omega}|u|^2\,dx + c_2\int\limits_0^T\int\limits_{\Omega}|\nabla u|^2\,dx\,dt + c_3\int\limits_0^T\int\limits_{\Omega}|u|^2\,dx\,dt\\ \leq c_5\int\limits_0^\tau&\left(\frac{1}{2}\int\limits_{\Omega}|u|^2\,dx + c_2\int\limits_0^T\int\limits_{\Omega}|\nabla u|^2\,dx\,dt + c_3\int\limits_0^T\int\limits_{\Omega}|u|^2\,dx\,dt\right)\mathrm{d}\tau.\end{aligned}$$

We use Gronwall again to conclude with:

$$\frac{1}{2}\int\limits_{\Omega}|u|^2\,dx + c_2\int\limits_0^T\int\limits_{\Omega}|\nabla u|^2\,dx\,dt + c_3\int\limits_0^T\int\limits_{\Omega}|u|^2\,dx\,dt \leq 0.$$

Therefore, $u^1 = u^2$ in $[0,T]\times\Omega$ holds true. Equation (2.8) implies that $R_i^1 = R_i^2$, $C_i^1 = C_i^2$ and $E_i^1 = E_i^2$. □

# 3 Discretization

In this chapter, we describe the discretization of the weak formulation of coupled PDE/ODE-systems. The temporal and spatial discretizations are derived by means of the Galerkin finite element method.

Firstly, we present a semi-discretization in time in Section 3.1 by a Galerkin method, which corresponds to the damped Crank-Nicolson method. We deal with the spatial discretization of the semi-discrete problem in Section 3.2. In Section 3.3, we present a simplified time stepping scheme for the adjoint equation arising from a posteriori error analysis.

## 3.1 Temporal discretization

To introduce semi-discretizations in time, we subdivide the time interval $\bar{I} = [0, T]$ by time points $t_m$ for $0 \leq m \leq M$ such that

$$0 = t_0 < \ldots < t_m < \ldots t_M = T$$

with subintervals $I_m := (t_{m-1}, t_m]$ of length $k_m := t_m - t_{m-1}$. Thus, we obtain a partition $\mathcal{T}_k = \{0\} \cup \{I_m \mid 1 \leq m \leq M\}$ of the time interval $\bar{I}$. We introduce a shorter notation for the space-time $L_2$ scalar product on the interval $I_m$ on the product space $W$,

$$((w, z))_m := \int\limits_{I_m} w(t) \cdot z(t)\, dt.$$

To deal with the discontinuities in the solution $w \in W$, we introduce the standard notation for the limits from above $w^{m,+}$, the limits from below $w^{m,-}$ and the jumps $[w]_m$ of $w$ at time point $t_m$ by

$$w^{m,+} = \lim_{\epsilon \searrow 0} w(t_m + \epsilon), \quad w^{m,-} = \lim_{\epsilon \searrow 0} w(t_m - \epsilon), \quad [w]_m = w^{m,+} - w^{m,-}.$$

### 3.1.1 The damped Crank-Nicolson scheme

We first quote the well known Crank-Nicolson time marching scheme in which $w_k^m := w_k(t_m)$ denotes the approximated value of the semi-discretized solution $w_k$ at time $t_m$:

*Given $w_0 \in L$, find $w_k^m \in V$ for $m \leq M$, such that:*

$$
\begin{aligned}
(w_k^0 - w_0, \varphi) &= 0 \quad \forall \varphi \in L, && \text{(3.1a)}\\
(w_k^m - w_k^{m-1}, \varphi) + \frac{k_m}{2}\Big(\bar{\mathcal{F}}(w_k^m;\varphi) + \bar{\mathcal{F}}(w_k^{m-1};\varphi)\Big) &= 0 \quad \forall \varphi \in V. && \text{(3.1b)}
\end{aligned}
$$

We obtain from (3.1a) that a well posed semi-discrete problem needs initial data $w_0 \in L$ . Even for initial data in $V$, optimal convergence order is usually not obtained.

In the next chapters, we discretize and numerically solve PDE/ODE systems modeling intercellular signaling (2.5). The initial values $u_0$ of the PDE part of the model are set to a constant zero, see Section 2.3 with initial conditions (2.7c). The constant initial values model the biological fact that the simulation starts in the same moment in which the signaling cells start secreting diffusible messengers. A spatially constant $u_0$ does violate the boundary condition of robin type, which prescribes for secreting cells a non-zero normal derivative on the cell surfaces. Therefore, the problem is not well posed at the initial time point.

Due to the incompatibility of the initial values, we have to replace at least two of the Crank-Nicolson steps by an implicit Euler scheme to ensure the optimal convergence rate, see Luskin and Rannacher [42, 51]. Several implicit Euler steps may be necessary depending on the type of the irregularity. Let $J_0$ be the set of indices in which we use the implicit Euler scheme and $J_1$ the set of indices for the Crank Nicolson scheme, i.e. $J_0 \cup J_1 = \{1, ..., M\}$. The damped Crank-Nicolson scheme can then be formulated as:

*Given an initial condition $w_0 \in L$, find $w_k^m \in V$ such that the following holds for all $m \leq M$:*

$$
\begin{aligned}
(w_k^0, \varphi) &= (w_0, \varphi) && \forall \varphi \in L, \quad m = 0,\\
(w_k^m, \varphi) + k_m \bar{\mathcal{F}}(w_k^m;\varphi) &= (w_k^{m-1}, \varphi) && \forall \varphi \in V, \quad m \in J_0, \quad \text{(3.2)}\\
(w_k^m, \varphi) + \frac{k_m}{2}\Big(\bar{\mathcal{F}}(w_k^m;\varphi) + \bar{\mathcal{F}}(w_k^{m-1};\varphi)\Big) &= (w_k^{m-1}, \varphi) && \forall \varphi \in V, \quad m \in J_1
\end{aligned}
$$

We construct the temporal discretization in a patch-wise structure to easily evaluate the second order interpolations in Section 5.2. Therefore, we select the index sets $J_0$ and $J_1$ such that for an even index $m$ the adjacent index $m-1$ is part of the same index set: $m \in J_s \cap 2\mathbb{N} \Rightarrow m-1 \in J_s$ for $s \in \{0,1\}$. Necessarily, the total number of indices $M$ has to be chosen as an even number.

### 3.1.2 Galerkin version of the damped Crank-Nicolson scheme

The damped Crank-Nicolson method is interpreted as a Galerkin method to derive accurate error estimates. We follow the derivation by Goll et al. [28] and set up Galerkin spaces with different polynomial spaces on each of the time intervals as trial and test spaces of the Galerkin method. As in the previous section, we set up the two sets of indices $J_0$ and $J_1$ to indicate the use of the implicit Euler method for indices in $J_0$ and the use of the Crank-Nicolson method for indices in $J_1$.

We define the following semi-discrete spaces $\hat{\bar{X}}_k^{r-1}(Y)$ and $\bar{X}_k^r(Y)$ for a given function space $Y$ and $r \geq 1$ by making use of the sub intervals $I_m$

$$\begin{aligned}
\hat{\bar{X}}_k^{r-1}(Y) :=& \Big\{ \varphi_k : \bar{I} \to Y \;\Big|\; \varphi_k|_{I_m} \in P_{r-1}(I_m, Y), 1 \leq m \leq M, \varphi_k(0) \in L \Big\}, \\
\bar{X}_k^r(Y) :=& \Big\{ \varphi_k : \bar{I} \to Y \;\Big|\; \varphi_k|_{I_m} \in P_r(I_m, Y), [\varphi_k]_{m-1} = 0, \\
& \qquad \varphi_k|_{I_n} \in P_{r-1}(I_n, Y), m \in J_1, n \in J_0, \varphi_k(0) \in L \Big\}.
\end{aligned} \tag{3.3}$$

$P_r(I_m, Y)$ denotes the space of polynomials of degree lower or equal to $r$ on $I_m$ with values in $Y$. The semi-discrete space $\bar{X}_k^{r-1}(V)$ is used both as trial and test space for the discontinuous Galerkin method DG(r-1) of degree $r-1$. In contrast, the trial and test spaces of the continuous Galerkin method CG(r) of degree $r$, damped by the DG method, are chosen differently: the space $\bar{X}_k^{r-1}(V)$ is only used as the test space and $\hat{\bar{X}}_k^r(V)$ is used as the trial space.

We restate the damped Crank-Nicolson scheme in a Galerkin version:

*Given $w_0 \in L$, find $w_k \in \bar{X}_k^1(V)$ such that for all $\varphi_k \in \hat{\bar{X}}_k^0(V)$ the following holds*

$$((\partial_t w_k, \varphi_k)) + \bar{\mathcal{F}}(w_k; \varphi_k) + \sum_{m \in J_0} ([w_k]_{m-1}, \varphi_k^{m-1,+}) + (w_k(0) - w_0, \varphi_k^{0,-}) = 0. \tag{3.4}$$

We approximate the integrals on the time intervals either by the box rule or by the trapezoidal rule

$$\begin{aligned} &\int_{I_m} (w_k(t), \varphi_k(t))\,dt \approx k_m(w_k(t_m), \varphi_k(t_m)), && m \in J_0, \\ &\int_{I_m} (w_k(t), \varphi_k(t))\,dt \approx \frac{k_m}{2}\Big((w_k(t_m), \varphi_k(t_m)) + (w_k(t_{m-1}), \varphi_k(t_{m-1}))\Big) && m \in J_1, \end{aligned} \tag{3.5}$$

and obtain again the time stepping scheme (3.2) of the damped Crank-Nicolson scheme, with the notation $w_k^m := w_k(t_m)$.

### 3.1.3 Different time meshes for the PDE and the ODE part

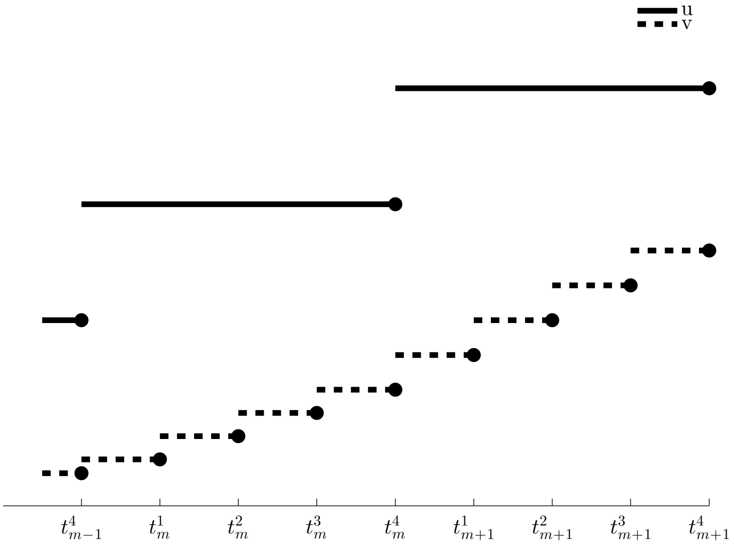
The main computational effort in solving coupled PDE/ODE systems comes from solving the PDE part. We show in Section 5.5.7 numerical results in which a finer time mesh of the ODE part can significantly reduce the discretization error. Balancing the errors between the ODE and the PDE part allows for a coarser time stepping in the PDE part of the system and therefore reduce the overall costs of solving the system to a certain accuracy. To the knowledge of the author, this thesis presents the first coupled solver approach which makes use of different temporal discretizations for the two parts of a coupled PDE/ODE system.

The strategy of an additionally refined temporal discretization of the ODE part is frequently used for segregated solving schemes, see e.g. Moghadam et al. [48]. In a segregated solving scheme the ODE part can be solved by higher order Runge-Kutta or backward differentiation formulas. Since we focus in this thesis on coupled solver approaches, we simplify the implementation by using the same time scheme but different time meshes for both parts of the system.

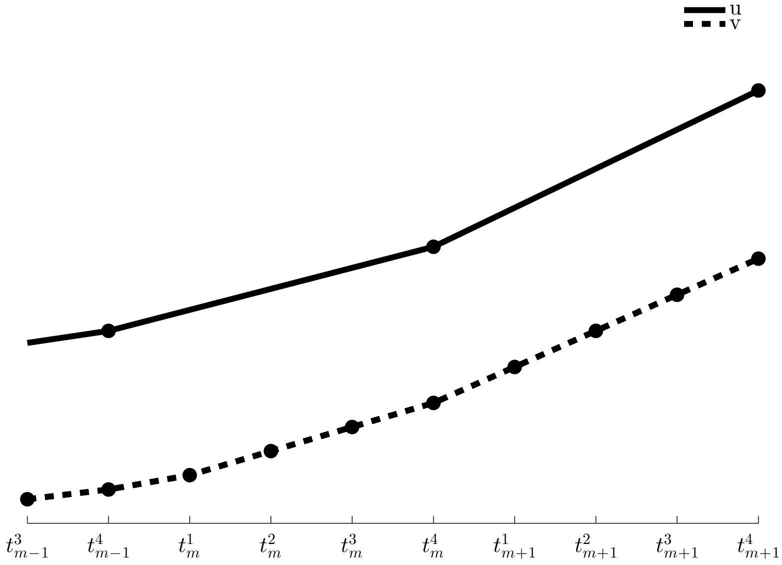
We subdivide each of the $M$ time intervals $I_m = (t_{m-1}, t_m]$ of the discretized PDE into $M_m^s$ new subintervals $I_m^j$. Therefore, we construct a partition of the time intervals $I_m$ for each $1 \leq m \leq M$ by the time points $t_m^j$ with $1 \leq j \leq M_m^s$ such that

$$t_{m-1} = t_m^0 < t_m^1 < ... < t_m^j < ... < t_m^{M_m^s} = t_m.$$

The interval $I_m$ is subdivided into subintervals $I_m^j := (t_m^{j-1}, t_m^j]$ of length $k_m^j = t_m^j - t_m^{j-1}$.



(a) DG(0)



(b) CG(1)

**Figure 3.1:** Different meshes in the temporal discretization for a coupled system with $M^s = 4$

We introduce the notation $M^p$ for the number of time steps of the PDE part and $M^o$ for the number of time steps of ODE part to distinguish the time stepping schemes of the PDE and the ODE part.

$$M^p := M, \qquad M^o := \sum_{m \leq M} M_m^s.$$

We denote the sums over all time intervals in a simplified notation by $\sum_{m \leq M}$. In this thesis, the sums over all time intervals start always with $m = 1$. Analogously, the sums over all subintervals of an interval $I_m$ are denoted by $\sum_{j \leq M_m^s}$. The sums over the subintervals of $I_m$ start with $j = 1$. We note that a patch-wise structure of the temporal discretization is needed for the second order interpolations in Section 5.2. Therefore, we choose even values for $M_m^s$. To shorten notation, we denote the beginning of each subintervals by $t_m^0 := t_{m-1}$ and the inner product on each time interval $I_m^j$ by

$$((w, z))_{m,j} = \int\limits_{I_m^j} w(t) \cdot z(t) \, dt.$$

Due to the different discretization, the point values of $w_k := \begin{pmatrix} u_k \\ v_k \end{pmatrix}$ at the intermediate time points $t_m^j$ are obtained differently for the PDE than for the ODE part. The linear interpolation on the time interval $I_m$ of the PDE part $u_k$ yield approximate intermediate values of $u_k$ at the time points $t_m^j$. The intermediate values of the ODE part on the other hand are directly available because of the finer temporal discretization

$$u_k^{m,j} := u_k(t_m^j) = \frac{t_{m+1} - t_m^j}{k^{m+1}} u_k^{m-1} + \frac{t_m^j - t_m}{k^{m+1}} u_k^m, \qquad v_k^{m,j} := v_k(t_m^j). \tag{3.6}$$

We formulate the damped Crank-Nicolson with different time meshes for the two parts as a Galerkin method. This enables us to derive error representations in Chapter 5. Therefore, we define the trial and test spaces $\hat{X}_k^{r-1}(Y)$ and $X_k^r(Y)$ for a given product space $Y = Y^p \times Y^o$, e.g. $Y := V$, by

$$\begin{aligned}
\hat{X}_k^{r-1}(Y) :=& \Big\{ \begin{pmatrix}\phi_k\\ \psi_k\end{pmatrix} : \bar{I} \to Y \,\Big|\, \phi_k|_{I_m} \in P_{r-1}(I_m, Y^p), \begin{pmatrix}\phi_k(0)\\ \psi_k(0)\end{pmatrix} \in L, \\
& \psi_k^j|_{I_m^j} \in P_{r-1}(I_m^j, Y^o), 1 \le m \le M, 1 \le j \le M_m^s \Big\}. \\
X_k^r(Y) :=& \Big\{ \begin{pmatrix}\phi_k\\ \psi_k\end{pmatrix} : \bar{I} \to Y \,\Big|\, [\phi_k]_{m-1} = 0, [\psi_k]_m^{j^m-1} = 0, \begin{pmatrix}\phi_k(0)\\ \psi_k(0)\end{pmatrix} \in L, \\
& \phi_k|_{I_m} \in P_r(I_m, Y^p), \psi_k|_{I_m^{j^m}} \in P_r(I_m^{j^m}, Y^o), 1 \le j^m \le M_m^s, m \in J_1, \\
& \phi_k|_{I_n} \in P_{r-1}(I_n, Y^p), \psi_k|_{I_n^{j^n}} \in P_{r-1}(I_n^{j^n}, Y^o), 1 \le j^n \le M_n^s, n \in J_0 \Big\}.
\end{aligned} \tag{3.7}$$

A notation for jumps of the solution at the intermediate time points is introduced by

$$v^{l,j,+} = \lim_{\epsilon \searrow 0} v(t_l^j + \epsilon), \quad v^{l,j,-} = \lim_{\epsilon \searrow 0} v(t_l^j - \epsilon), \quad [v]_l^j = v^{l,j,+} - v^{l,j,-}.$$

We discretize the system of equations (2.3) with the test space $X_k^1(V) \subset L^2(I,V)$ and the trial space $\hat{X}_k^0(V) \subset L^2(I,V)$ to obtain the Galerkin formulation:

*Given* $w_0 \in L$, *find* $w_k = \begin{pmatrix}u_k\\ v_k\end{pmatrix} \in X_k^1(V)$ *such that for all* $\varphi_k = \begin{pmatrix}\phi_k\\ \psi_k\end{pmatrix} \in \hat{X}_k^0(V)$ *the following holds*

$$\begin{aligned}
&((\partial_t w_k, \varphi_k)) + \bar{\mathcal{F}}(w_k; \phi_k) + \sum_{m \in J_0} ([u_k]_{m-1}, \phi_k^{m-1,+}) \\
&\quad + \sum_{\substack{m \in J_0, \\ j \le M_m^s}} ([v_k]_m^{j-1}, \psi_k^{m,j-1,+}) + (w_k(0) - w_0, \varphi_k^{0,-}) = 0.
\end{aligned} \tag{3.8}$$

We approximate the temporal integrals on the time intervals either by the box rule or by the trapezoidal rule.

$$\begin{aligned}
&\int\limits_{I_m^j} (w_k(t), \varphi_k(t))\, dt \approx k_m^j (w_k(t_m^j), \varphi_k(t_m^j)), && m \in J_0, \\
&\int\limits_{I_m^j} (w_k(t), \varphi_k(t))\, dt \approx \frac{k_m^j}{2} \Big( (w_k(t_m^j), \varphi_k(t_m^j)) + (w_k(t_m^{j-1}), \varphi_k(t_m^{j-1})) \Big), && m \in J_1.
\end{aligned}$$

In the following time marching scheme, we use the notation $w_k^{m,j}$ for the approximated value at time point $t_k^{m,j}$. The PDE part of the solution $u_k^{m,j}$ is obtained at the intermediate time points $t_k^{m,j}$ by the linear interpolation (3.6) of $u_k$. The time marching scheme resulting from the Galerkin formulation (3.8) reads as follows:

*Given initial values* $w_0 \in L$*, find* $w_k^m = \begin{pmatrix} u_k^m \\ v_k^m \end{pmatrix} \in V$ *for all* $1 \le m \le M$ *and* $1 \le j \le M_m^s$ *such that the following holds for all* $\phi \in V^p, \psi \in V^o$

$$
\begin{aligned}
(w_k^0, \varphi) &= (w_0, \varphi), && \text{for all } \varphi \in L \\
(u_k^m, \phi) + k_m \mu(\nabla u_k^m, \nabla \phi) & \\
+ \sum_{j \le M_m^s} k_m^j \bar{a}(w_k^{m,j}; \phi) &= (u_k^{m-1}, \phi), && m \in J_0, \\
(v_k^{m,j}, \psi) + k_m^j \bar{b}(w_k^{m,j}; \psi) &= (v_k^{m,j-1}, \psi) && m \in J_0, 1 \le j \le M_m^s \\
(u_k^m, \phi) + \frac{1}{2} k_m \mu(\nabla u_k^{m-1} + \nabla u_k^m, \nabla \phi) & \\
+ \frac{1}{2} \sum_{j \le M_m^s} k_m^j \Big( \bar{a}(w_k^{m,j}; \phi) + \bar{a}(w_k^{m,j-1}; \phi) \Big) &= (u_k^{m-1}, \phi), && m \in J_1 \\
(v_k^{m,j}, \psi) + \frac{1}{2} k_m^j \Big( \bar{b}(w_k^{m,j}; \psi) + \bar{b}(w_k^{m,j-1}; \psi) \Big) &= (v_k^{m,j-1}, \psi), && m \in J_1, 1 \le j \le M_m^s.
\end{aligned}
\tag{3.9}
$$

**Remark 3.1.** *We note that the PDE and ODE part are strongly coupled in the time marching algorithm for different time meshes for the two parts* (3.9)*. The discretized solution of the PDE part* $u_k^m$ *at time point* $t_m$ *is coupled to all of the intermediate values* $v_k^j$ *for* $j \le M_m^s$ *of the ODE part of the solution. The ODE values* $v_k^j$ *are coupled to* $u_k^m$ *as well. In a coupled solving approach, which we describe in Chapter 4, a finer time mesh of the ODE part increases the number of degrees of freedoms of the ODE part. Depending on the coupling, most degrees of freedoms of the ODE part are directly coupled to the PDE part. Hence, the discretization with a refined ODE time mesh results in a discrete PDE/ODE system, which needs to be solved as often as without the fine ODE time discretization. Only the number of ODEs, which are coupled in the discrete system to the PDE, is increased.*

## 3.2 Spatial discretization

We discretize the infinite dimensional function space $V$ by means of finite elements. To this end, we subdivide the domain $\Omega$ into disjoint open form-regular quadrilateral or hexahedron shape-regular cells $K$ (depending on the dimension of $\Omega$), following e.g. Ciarlet [18]. All cells together form the spatial mesh $\mathcal{T}_h = \{K\}$ of the domain, where $h$ is a parameter depending on the diameter of the mesh cells (e.g. the maximum of the diameters). The union of the closure of the mesh cells covers the computational domain $\Omega_h$ so that $\bar{\Omega}_h := \bigcup_{K \in \mathcal{T}_h} \bar{K}$. For $s \in \mathbb{N}, s \geq 1$, we construct the finite dimensional subspaces $V_h^s \subset V^p$

$$V_h^s := \Big\{ u \in C^0(\bar{\Omega}_h) \;\Big|\; u|_K \in Q^s(K), \text{ for all } K \in \mathcal{T}_h \Big\},$$

where $Q^s$ denotes the usual space of isoparametric finite elements. The space $Q^s$ is defined as the tensor product polynomial up to degree $s$ on a reference square or cube $\hat{K}$, mapped by a transformation of the same polynomial space onto the cell $K$.

We introduce hanging nodes between a refined cell and its neighbors, see Rheinboldt and Mesztenyi [53] for the adaptive mesh refinement of the triangulation $\mathcal{T}_h$, see Chapter 5. Only one hanging node is allowed per edge. No degrees of freedom are connected to the hanging nodes, instead they are eliminated by interpolation.

We then define the spatially discretized spaces $X_{kh}^{r,s}$ and $\hat{X}_{kh}^{r-1,s}$ by

$$\begin{aligned} X_{kh}^{r,s} &:= X_k^r(V_h^s \times V^o), \\ \hat{X}_{kh}^{r-1,s} &:= \hat{X}_k^{r-1}(V_h^s \times V^o). \end{aligned} \tag{3.10}$$

We obtain the fully discretized version of problem (2.5) by:

*Given* $w_0 \in V_h^s \times V^o$, *find* $w_{kh} = \begin{pmatrix} u_{kh} \\ v_{kh} \end{pmatrix} \in X_{kh}^{1,s}$ *such that for all* $\varphi_{kh} = \begin{pmatrix} \phi_{kh} \\ \psi_k \end{pmatrix} \in \hat{X}_{kh}^{0,s}$ *the following holds*

$$\begin{aligned} ((\partial_t w_{kh}, \varphi_{kh})) + \bar{\mathcal{F}}(w_{kh}; \phi_{kh}) + \sum_{m \in J_0} ([u_{kh}]_{m-1}, \phi_{kh}^{m-1,+}) \\ + \sum_{\substack{m \in J_0, \\ j \leq M_m^s}} ([v_{kh}]_m^{j-1}, \psi_k^{m,j-1,+}) + (w_{kh}(0) - w_0, \varphi^{0,-}) = 0. \end{aligned} \tag{3.11}$$

We use the notation $w_{kh}^m := w_{kh}(t^m)$ and $w_{kh}^{m,j} := w_{kh}(t^{m,j})$, which are defined by linear interpolation analogously to (3.6). We obtain a fully discrete time marching scheme for a given $s \in \mathbb{N}$, a spatially discretized version of (3.9):

*Given initial values* $\begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \in V_h^s \times V^o$, *find* $w_{kh}^m = \begin{pmatrix} u_{kh}^m \\ v_{kh}^m \end{pmatrix} \in V_h^s \times V^o$ *for all*

*$1 \leq m \leq M$ such that the following holds for all $\phi_h \in V_h^s, \psi \in V^o$ and for all $1 \leq m \leq M$ , $1 \leq j \leq M_m^s$:*

$$
\begin{aligned}
(u_{kh}^0, \phi_h) &= (u_0, \phi_h, ), \\
(v_{kh}^0, \psi) &= (v_0, \psi), \\
(u_{kh}^m, \phi_h) + k_m \mu(\nabla u_{kh}^m, \nabla \phi_h) + \sum_{j \leq M_m^s} k_m^j \bar{a}(w_{kh}^{m,j}; \phi_h) &= (u_{kh}^{m-1}, \phi_h) \qquad m \in J_0 \\
(v_{kh}^{m,j}, \psi) + k_m^j \bar{a}(w_{kh}^{m,j}; \psi) &= (v_{kh}^{m,j-1}, \psi), \qquad m \in J_0, \\
(u_{kh}^m, \phi_h) + \frac{1}{2} k_m \mu(\nabla u_{kh}^{m-1} + \nabla u_{kh}^m, \nabla \phi_h) & \qquad (3.12) \\
+ \frac{1}{2} \sum_{j \leq M_m^s} k_m^j \Big( \bar{a}(w_{kh}^{m,j}; \phi_h) + \bar{a}(w_{kh}^{m,j-1}; \phi_h) \Big) &= (u_{kh}^{m-1}, \phi_h), \qquad m \in J_1, \\
(v_{kh}^{m,j}, \psi) + \frac{1}{2} k_m^j \Big( \bar{b}(w_{kh}^{m,j}; \psi) + \bar{b}(w_{kh}^{m,j-1}; \psi) \Big) &= (v_{kh}^{m,j-1}, \psi), \qquad m \in J_1.
\end{aligned}
$$

We denote the spatially discretized space of the PDE part $V^p$ by $V_h^p$. Since all computations in this thesis employ discretizations consisting of $Q^1$ elements, we select the spatially discretized space $V_h^p$ as $V_h^1$.

$$V_h^p := V_h^1$$

Additionally, a steady state of the system (2.5) is directly computed in Section 4.3.2. For this reason, we discretize the stationary equation (2.6) with finite elements and solve the following system for $w_h := \begin{pmatrix} u_h \\ v_h \end{pmatrix} \in V_h^p \times V^o$,

$$
\begin{aligned}
\mu(\nabla u_h, \nabla \phi_h) + \bar{a}(w_h; \phi_h) &= 0, \quad \text{for all } \phi_h \in V_h^s, \\
\bar{b}(w_h; \psi) &= 0, \quad \text{for all } \psi \in V^o.
\end{aligned} \tag{3.13}
$$

## 3.3 Dual time marching scheme

In Chapter 5, we locally estimate the error a posteriori and adapt the grid in time and space. Therefore, we consider the dual problem for a given functional $J \in W'$ (5.11) to the coupled system of equations (2.5):

*Find a dual solution* $z \in W$ *for a given linear functional* $J \in W'$ *and a solution of the primal problem* $w \in W$ *such that the following holds*

$$-((\varphi, \partial_t z)) + \mathcal{F}'(w; \varphi, z) + (\varphi(T), z(T)) = J(\varphi), \quad \forall \varphi \in V. \tag{3.14}$$

In contrast to the primal problem, the formulation of the Galerkin scheme of the dual problem yields a different time stepping scheme. The dual time marching scheme differs for adaptively refined time discretizations or when the time stepping scheme switches between the Crank-Nicolson scheme and the implicit Euler scheme, e.g. when $m \in J_1$ and $m+1 \in J_0$. Secondly, the coupling between the PDE and the ODE part in function $\bar{b}'_u$ results in an especially complicated formulation if different time meshes are used for the two parts of the equation.

In Goll et al. [28] and in Schmich and Vexlers work [59], the derived algorithm is completely implemented to ensure the adjoint consistency of the time stepping scheme, i.e. whether the calculated discrete adjoint corresponds to a discretization of the continuous adjoint. Adjoint inconsistencies can reduce the convergence of algorithms relying on the dual solution, e.g. in the context of optimization, see Sachs et al. [57].

Hartmann [31] derived the dual time stepping scheme of the Crank-Nicolson method for the adjoint of the discrete Galerkin formulation. The derived method is only first-order accurate in time on adaptively refined time discretizations. Further, he showed in [32] that for use in the dual weighted residual method mainly the approximation properties of the discrete dual solution are important: The dual solution itself is only needed to a certain accuracy.

We apply a dual time marching scheme, which we obtained by a discretization of the adjoint of the Galerkin formulation. This discretization corresponds to the time stepping scheme used for the primal problem in (3.12). In the numerical computation in Section 5.5, we were able to obtain reliable and precise error estimations with our discretized dual solutions. Therefore, we assume that for the PDE/ODE systems considered in this thesis the effect of adjoint inconsistencies can be neglected in the context of the Dual Weighted Residual method.

To compute the discretized dual solution $z_{kh} \in \hat{X}^{0,1}_{kh}$, we use the following dual time marching scheme:

*Find* $z^m = \begin{pmatrix} z_u^m \\ z_v^m \end{pmatrix} \in V_h^p \times V^o$ *for all* $m \leq M$ *with the starting values at the final* *time point $T$:*

$$(\phi_h, z_u^M) := J^{p,1}(\phi_h), \quad (\psi, z_v^M) := J^{o,1}(\psi),$$

*such that there holds for any $\phi_h \in V_h^p, \psi \in V^o$:*

- for all $m \in J_0$ and for all $j \leq M_m^s$:

$$\begin{aligned}(\phi_h, z_u^m) + k_m \mu(\nabla \phi_h, \nabla z_u^m) + k_m \bar{a}'_u(w^m; \phi_h, z_u^m)\\ + \sum_{j \leq M_m^s} k_m^j \bar{b}'_v(w^{m,j}; \phi_h, z_v^{m,j}) - J^{p,2}(\phi_h) = (\phi_h, z_u^{m+1}),\end{aligned} \tag{3.15a}$$

$$\begin{aligned}(\psi, z_v^{m,j-1}) + k_m^{j-1} \bar{a}'_v(w^m; \psi, z_u^m)\\ + k_m^{j-1} \bar{b}'_v(w^{m,j-1}; \psi, z_v^{m,j-1}) - J^{o,2}(\psi) = (\psi, z_v^{m,j}),\end{aligned} \tag{3.15b}$$

- for all $m \in J_1$ and for all $j \leq M_m^s$:

$$\begin{aligned}(\phi_h, z_u^m) + \frac{1}{2} k_m \mu(\nabla \phi_h, \nabla z_u^m) + \frac{1}{2} k_m \mu(\nabla \phi_h, \nabla z_u^{m+1})\\ + \frac{1}{2} k_m \bar{a}'_u(w^m; \phi_h, z_u^m) + \frac{1}{2} k_m \bar{a}'_u(w^{m+1}; \phi_h, z_u^{m+1})\\ + \frac{1}{2} k_m \bar{b}'_u(w^m; \phi_h, z_v^m) + \frac{1}{2} k_m \bar{b}'_u(w^{m+1}; \phi_h, z_v^{m+1}) - J^{p,2}(\phi_h) = (\phi_h, z_u^{m+1}),\end{aligned} \tag{3.15c}$$

$$\begin{aligned}(\psi, z_v^{m,j-1}) + \frac{1}{2} k_m^{j-1} \bar{a}'_v(w^{m,j-1}; \psi, z_u^m) + \frac{1}{2} k_m^{j-1} \bar{a}'_v(w^{m,j}; \psi, z_u^m)\\ + \frac{1}{2} k_m^{j-1} \bar{b}'_v(w^{m,j-1}; \psi, z_v^{m,j-1}) + \frac{1}{2} k_m^{j-1} \bar{b}'_v(w^{m,j}; \psi, z_v^{m,j}) - J^{o,2}(\psi) = (\psi, z_v^{m,j}).\end{aligned} \tag{3.15d}$$

- for $m = 0$ :

$$\begin{aligned}(\phi_h, z_u^0) = (\phi_h, z_u^1)\\ (\psi, z_v^0) = (\psi, z_v^1)\end{aligned} \tag{3.15e}$$

We point out that we use in equation (3.15c) a coarse approximation of the ODE part of the dual solution. A coarse trapezoidal rule, which uses only values of the dual solution at the start and end point of each interval $I_m$ is applied instead of taking the sum over all intermediate time points $t_m^j$.

We found that these coarser approximations became necessary in many of the computed examples in Section 5.5. When we used the finer approximations at all intermediate time points, we observed instability issues of the PDE values of the dual solution during the Crank-Nicolson time stepping ($m \in J_1$). These oscillations were never observed for the primal problem. We avoid the instabilities by using the trapezoidal rule for the dual problem (3.15c). Not all available information of the ODE part of the solution is used, but the time marching scheme still solves the system of equations (2.5) with an error of order $\mathcal{O}(k_m^2)$. Thus, the order of the method is not reduced. The ODE part of equation (3.15d) is solved on the finer time grid.

# 4 Coupled and Decoupled Linear Solvers

The results in this chapter have been submitted for publication and are available as a preprint [15].

In Chapter 3.1, we discretized the non-stationary coupled problem (2.5) in time with a damped Crank-Nicolson Scheme. In this chapter, we compare solution schemes for the arising nonlinear coupled discrete system (3.12) to be solved on each time interval $I_m$. Additionally, we investigate solution schemes for stationary computations of a steady state (2.6), which enables to measure the performance independently of a chosen temporal discretization.

We introduce a joint notation for both resulting discrete nonlinear systems by the operators $A_h : V_h^p \times V^o \to V_h^p$ and $B_h : V_h^p \times V^o \to V^o$ with corresponding right hand sides $f_h \in V^p$ and $g_h \in V^o$, such that

$$A_h(u, v) = f_h, \tag{4.1a}$$

$$B_h(u, v) = g_h. \tag{4.1b}$$

We use the index $h$ to indicate the dependence of the operator $B_h$ on the discretization through the coupling with the PDE part.

Often if a numerical scheme for solving coupled equations, e.g. multiphysics problems, is developed, decoupled methods are favored. A decoupled solution scheme allows solving both parts of the system with a solver at hand, possibly tuned to solve that specific part of the problem. The coupling can be implemented in an external (to the two solvers) framework allowing an easy implementation of the more complex coupled problem. This approach is widely used to solve coupled systems, see [23, 25, 33, 48, 49, 61]. The two solvers are used iteratively to solve the coupling of the system. These decoupled methods are much easier to implement but have serious drawbacks. Fixed point iterations tend to converge very slowly or not at all if the system is strongly coupled. The large iterative error in this strategy can only be reduced at high computational costs through very small time steps or a higher number of iterations between the two solvers. Other publications avoid decoupled methods for these reasons and instead develop a coupled solver scheme [34, 39]. Since the ODE part in systems for modeling the simulation of intercellular signaling does lead to a small discretization compared to the PDE part, it is not clear if a splitting method can outperform a coupled scheme. In case of a strong nonlinear coupling of the system, many iterations of the decoupled solution scheme can be avoided by using a fully coupled solver.

The effectivity of a coupled solution method depends on the impact of the coupling. Therefore, we first analyze the sensitivities of the coupled system in Section 4.1. We present coupled and decoupled solution schemes in Section 4.2 and discuss numerical results in Section 4.3.

## 4.1 Sensitivity analysis of the coupled system

Let $S_h : v_h \mapsto u_h$ and $T_h : u_h \mapsto v_h$ denote respectively the solution operator for the decoupled PDE part and ODE part of the discretized system of equations (4.1). With this notation, we can write the system (4.1) as

$$u = S_h(v), \tag{4.2a}$$

$$v = T_h(u). \tag{4.2b}$$

In a fixed point iteration, the PDE part of the equation (4.2) is solved for a given value of $v$, then the ODE part of the equation (4.2) is solved with the resulting value of $u$. The cycle is iterated until a given tolerance is reached. This process can also be written as a composition of the two operators:

$$u^{n+1} = S_h\big(T_h(u^n)\big). \tag{4.3}$$

A fixed point iteration to solve the system (4.1) has a slow convergence rate (typically only linear) and the number of fixed point iterations depends on the nature of the coupling and the model parameters. Nevertheless, a decoupled linear solver can be considered advantageous as part of a Newton scheme. Thus, instead of solving the update of the solution by using the Jacobian of the full system, one updates iteratively the two decoupled parts. We present these two solution schemes in the following sections, in which we compare two different strategies to solve the Newton update.

We continue by presenting a sensitivity approach to decide whether a fixed point iteration or the full system update should be used. As shown later, the choice depends on the actual model parameter and the method gives a quantitative index that can be used for practical implementations. The results shown here use the well known fact that a fixed point iteration can converge only if a specific condition on the iteration operator is fulfilled.

Considering the formulation (4.3), we write the Jacobian of the fixed point operator as

$$J = \frac{\partial S_h}{\partial v}\frac{\partial T_h}{\partial u}. \tag{4.4}$$

According to the Banach fixed point theorem, the following criterion has to be fulfilled for the convergence of the fixed point iteration

$$\|J\| < 1, \tag{4.5}$$

in some norm $\|\cdot\|$. A more convenient criterion is the substitution of the norm with the spectral radius of the matrix $J$,

$$|\lambda_{max}(J)| < 1. \tag{4.6}$$

If this condition is fulfilled a simple fixed point approach converges, therefore this criterion has been used, e.g. by Haftka et al. [30], to define whether the coupling of the system (4.1) is strong. Depending on the parameters, we show in Section 4.3.2 that the PDE/ODE systems for intercellular signaling are strong coupled and thus a full coupling is more effective than a decoupled method. Decoupled methods can still converge for strong coupled systems if embedded in a Newton's scheme, but require a large number of fixed point iterations, as we show in our numerical results.

We proceed with the calculation of the largest eigenvalue of the Jacobian $J$. We differentiate therefore the discretized operators $A_h$ and $B_h$ and obtain the sensitivity equations

$$A'_{h,u}(\hat{u},\hat{v})u_{\delta v} + A'_{h,v}(\hat{u},\hat{v})\delta v = 0, \quad \forall \delta v \in V^o, \tag{4.7a}$$

and

$$B'_{h,v}(\hat{u},\hat{v})v_{\delta u} + B'_{h,u}(\hat{u},\hat{v})\delta u = 0, \quad \forall \delta u \in V_h^p. \tag{4.7b}$$

We introduce the notation

$$u_{\delta v} := \frac{\partial u}{\partial v}(\delta v), \qquad v_{\delta u} = \frac{\partial v}{\partial u}(\delta u)$$

for the sensitivities. In the decoupled system, $u_{\delta v}$ indicates the variation of the PDE solution, which perturbs the solution of the ODE system, and equivalently $v_{\delta u}$ denotes the variation of the ODE system for a perturbation of the PDE system.

Since the sensitivities in the linear solver strongly depend on the used time stepping scheme, we consider only the sensitivities for a computation of the steady state. Then, the equations (4.7a) are stationary PDEs to be solved for each component of $\delta v$, while the ODE part (4.7b) consists of algebraic equations solved for each $\delta u$. Therefore, we compute the sensitivity matrices $\partial S_h/\partial v$ as a $N^o \times N^p$ matrix and $\partial T_h/\partial u$ as a $N^p \times N^o$ matrix, where $N^o$ denotes the number of ODE equations and $N^p$ the dimension of the PDE discretization.

**Remark 4.1.** *In the PDE/ODE system presented in Section 2.3, the coupling between the two parts appears only at the boundaries $\Gamma_i$ and only with the first two components of $v$. Thus, the product $\partial S_h/\partial v\ \partial T_h/\partial u$ decouples into a block diagonal matrix consisting of $2 \times 2$ matrices for each biological cell. In addition, we need to calculate the sensitivities* (4.7b) *only for the restriction of $\delta u$ on the boundaries $\Gamma_i$, which are nonetheless algebraic equations. Therefore, the major costs to calculate the sensitivities are given by the PDE part* (4.7a)*.*

For nonlinear systems of equations, the sensitivity analysis depends on a given point of linearization $(\hat{u}, \hat{v})$. We compute an approximate numerical solution of the system (4.1) for characteristic values of the parameters and choose the solution as point of linearization.

**Remark 4.2.** *We evaluated the PDE/ODE model modeling IL-2 signaling* (2.7) *with the presented sensitivity analysis. For different numbers, size and distribution of biological cells, as well as moderate secretion rates $q$ in the biological range, see Busse et al. [13], there exists a unique stationary state. Thus, it is possible to use a stationary solver to directly compute this steady state. We obtain maximal eigenvalues $\lambda$ of the sensitivity matrix such that*

$$5 < \lambda < 100.$$

*This analysis therefore shows quantitatively that the interaction between the PDE and the ODE part of the system of equations* (2.7) *is strong.*

## 4.2 Numerical schemes

In this section, we present different approaches for a solver of a strongly coupled PDE/ODE system. The different solvers are numerically compared to each other in Section 4.3.

### 4.2.1 Nonlinear solvers

There are mainly two strategies for implicit solvers of coupled systems:

- nonlinear methods, among them the nonlinear multigrid method also called 'full approximation scheme' (FAS), see Brandt or Hackbusch [12, 29],
- linearization based approaches, e.g. by Newton or Picard iterations, see a comparison by Paniconi and Putti [50].

These methods can be used in a combined approach, where for instance a Newton-type method can be used as smoother for a FAS and a linear or a nonlinear multigrid can be used as a preconditioner for a Newton-type method, see Cai and Keyes [14]. Nevertheless, the comparison and discussion of advantages and disadvantages of these strategies, that depend on many aspects like e.g. the accuracy of the Jacobian approximation, see Mavriplis [47], is not the focus of our work.

Newton-type methods provide a flexible and reliable framework for nonlinear problems by solving a series of linear equations with a quadratic convergence rate and thus much faster than Picard-methods. Since in the considered PDE/ODE system the linearization is not a critical point, we consider a Newton-type method preconditioned by a linear multigrid and study the effect of splitting the linearization in the following sections.

#### Fully Coupled Newton's Method

To apply Newton's method, we linearize the system and solve in each Newton step the system

$$\begin{pmatrix} A'_{h,u}(u^n, v^n) & A'_{h,v}(u^n, v^n) \\ B'_{h,u}(u^n, v^n) & B'_{h,v}(u^n, v^n) \end{pmatrix} \begin{pmatrix} \delta u^{n+1} \\ \delta v^{n+1} \end{pmatrix} = \begin{pmatrix} f_h - A_h(u^n, v^n) \\ g_h - B_h(u^n, v^n) \end{pmatrix}, \tag{4.8}$$

to obtain the Newton updates $\delta u^{n+1}$ and $\delta v^{n+1}$, with which we calculate the next iterates $u^{n+1} = u^n + \delta v^n$ and $v^{n+1} = v^n + \delta v^{n+1}$. We write $A'_{h,u}$ and $A'_{h,v}$ for the derivatives of $A_h$ with respect to $u$ and $v$ and analogously $B'_{h,u}$ and $B'_{h,v}$ for the derivatives of $B_h$.

### Decoupled Inexact Newton's Method

Secondly, we consider a decoupled solution scheme for the linear systems defined in each Newton-step. In a typical decoupled scheme, the two systems are solved iteratively in separate solvers. The coupled system is solved in each Newton step by this decoupled scheme until the residual of the system fulfills a certain accuracy or a maximum of iterations is reached.

A standard algorithm for a decoupled Newton's method is shown in Algorithm 4.1, which iterates until the approximated solution fulfills a prescribed accuracy ($TOL_{newton}$). In this scheme, we solve in each time step $m$ with a Newton type method the solution for the next time step $(u^{n+1}, v^{n+1})$ by calculating a few iterations of the decoupled subsystems. The decoupled system is solved for each Newton step $n$ by the following fixed point iteration

$$\begin{pmatrix} A'_{h,u}(u^n, v^n) & A'_{h,v}(u^n, v^n) \\ 0 & B'_{h,v}(u^n, v^n) \end{pmatrix} \begin{pmatrix} \delta u^{i+1} \\ \delta v^{i+1} \end{pmatrix} = \begin{pmatrix} f_h - A_h(u^n, v^n) \\ g_h - B_h(u^n, v^n) - B'_{h,v}(u^n, v^n)\delta u^i \end{pmatrix}, \tag{4.9}$$

until the Newton updates $(\delta u^{i+1}, \delta v^{i+1})$ fulfill the linear residual of the system (4.8) to an accuracy ($TOL_{fixpoint}$).

**Algorithm 4.1:** Decoupled algorithm: inexact Newton scheme

```
$n = 0$
repeat
    $i = 0$
    repeat
        compute Newton updates $(\delta u^{i+1}, \delta v^{i+1})$ by solving (4.9)
        evaluate the residual $res_{iter}$ of the linear system (4.8)
        $i = i + 1$
    until $res_{iter} < TOL_{fixpoint}$ or $i = MAX_{fixpoint}$
    update the iterate $u^n$ and $v^n$ by $\delta u^{n+1}$ and $\delta v^{n+1}$
    evaluate the residual $res_{newton}$ of the nonlinear system (4.1)
    $n = n + 1$
until $res_{newton} < TOL_{newton}$
```

A common approach to accelerate such a solution process is a quasi-Newton iteration in which the Jacobian matrix is only approximated. In this way the costs per Newton iteration are reduced, while the number of Newton iterations increase. A trade-off between accuracy and total costs can reduce the computing time with respect to a full Newton method. Such a quasi-Newton scheme is obtained if a low accuracy ($TOL_{fixpoint}$) or a small maximum number of fixed point iterations ($MAX_{fixpoint}$) is chosen. This decoupled method is compared for different parameters in numerical tests of Section 4.3.2 to the fully coupled Newton method.

### 4.2.2 Multigrid Schemes

In this section, we introduce a multilevel preconditioner which can cope with the strong coupling between PDE and ODEs. Such a coupling arises in the solver of the linear subsystems if the fully coupled Newton method is used instead of a splitting scheme. Coupled problems are commonly preconditioned by block preconditioning approaches, e.g. by simple block diagonal methods or a preconditioning of the Schur complement, see Mandel [43]. We do not use a block preconditioning approach because of the much smaller discretization of the ODE part. Instead, we set up a coupled preconditioner based on the linear multigrid method.

Since the condition number of the discrete problem depends on the refinement level, it is well known [11, 29] that the most efficient preconditioner for the PDE block is a multilevel preconditioner. We compare this approach to the previously described decoupled solution scheme in Figure 4.1. We consider a hierarchy of meshes $\{\mathbb{T}_l\}_{0\leq l\leq L}$, where the index 0 denotes the root mesh, i.e. the coarsest mesh from which all other meshes are derived by global or local refinement. In this section, we use the following notation for the system matrix of (4.8)

$$\mathcal{K}_l := \begin{pmatrix} A'^{,l}_{h,u} & A'^{,l}_{h,v} \\ B'_{h,u} & B'_{h,v} \end{pmatrix}, \tag{4.10}$$

where the index $l$ indicates the grid refinement level. Note, that we do not use the notation with superscript $l$ in the blocks of the ODE part. In fact, $B'_{h,v}$ does not depend on the mesh level, while $B'_{h,u}$ does depend on the mesh level through the coupling term $\widetilde{u}_h$ on the cell boundary. To reduce the computational costs, and simplify the implementation, we use the following approximation: the coupling ODE/PDE block is calculated at each level with the term $\widetilde{u}_h$ computed at the finest level. In this way, the whole ODE part does not depend on the refinement level $l$. We have observed that this modification does not influence considerably the performance of the multilevel algorithm. The aim of the algorithm is to solve the system $\mathcal{K}_L$ corresponding to the finest mesh level $L$

$$\mathcal{K}_L w_L = f_L. \tag{4.11}$$

Instead of system (4.11), one solves the following preconditioned system (4.12) with a Krylov space method like the conjugate gradient method (CG) or the generalized minimal residual (GMRES) method. We use a GMRES solver because of the asymmetry of the system matrix, but a different Krylov method as, e.g., the BiCG or BiCGStab would also be appropriate for our purpose. In the preconditioned system

$$\mathcal{P}^{-1}\mathcal{K}_L w_L = \mathcal{P}^{-1} f_L, \tag{4.12}$$

the operator $\mathcal{P}^{-1}$ is chosen as a possible cheap approximation of the inverse of $\mathcal{A}_L$. In our case $\mathcal{P}^{-1}$ is given by a multigrid algorithm.

We apply the well known V-cycle multigrid algorithm described by Algorithm 4.2. The convergence of the V-cycle multigrid algorithm has been shown, see Bramble [11]. We show numerically in Section 4.3.2 that the efficiency of the preconditioner developed in this thesis is independent of the mesh size. The principal components of the algorithm are the transfer operators and the smoothing operator.

### Transfer Operators

We use the following notation for the transfer operators

$$R_l^{l-1} : V_l \to V_{l-1} \text{ (restriction)}, \quad P_l^{l-1} : V_{l-1} \to V_l \text{ (prolongation)}. \tag{4.13}$$

The restriction and prolongation operators act only on the PDE part, i.e. the finite element discretization, while the ODE part is transferred by the identity in both directions. The restriction and prolongation for the PDE part are implemented as intergrid transfers induced by the natural embedding of hierarchical meshes, see Janssen and Kanschat [38]. The restriction of the whole residual is given in matrix notation by the application of the operator

$$\begin{pmatrix} R_l^{l-1} & 0 \\ 0 & I \end{pmatrix}, \tag{4.14}$$

**Algorithm 4.2:** Multilevel algorithm: V-cycle

```
Function MG($l, y_l$)
    if $l = 0$ then
        Solve on the coarsest grid
        $x_0 := \mathcal{K}_0^{-1} f_0$
        return $x_0$
    else
        Presmoothing
        $\bar{y}_l := S_l(y_l)$
        Calculation of defect
        $d_l := f_l - \mathcal{K}_l \bar{y}_l$
        Restriction of defect
        $d_{l-1} := R_l^{l-1} r_l$
        Defect equation
        $w_{l-1} :=$MG$(l-1, d_{l-1})$
        Prolongation of defect
        $w_l := P_{l-1}^l w_{l-1}$
        Correction
        $\bar{x}_l := \bar{y}_l + w_l$
        Postsmoothing
        $x_l := S_l(\bar{x}_l)$
    return $x_l$
```

where $I \in \mathbb{R}^{n \times n}$ denotes the identity matrix of the ODE part. The prolongation of the whole residual is defined analogously.

### Smoothers

In case of strong coupled problems, a common strategy for the smoothing process is to consider the full coupling only at the coarsest level and to smooth the two parts separately (decoupled) on the finer levels. Since in our case the ODE part is small compared to the PDE one, we expect that the marginally more expensive smoothing of the whole coupled system on all mesh levels would be efficient, given the strong coupling. Therefore, we compare smoothing the whole system to smoothing only the PDE part. For this comparison every efficient smoother would be appropriate, we choose the incomplete LU factorization (ILU). The two smoother are denoted $S_1$ and $S_2$:

$S_1$: incomplete factorization (ILU) of the whole matrix $\mathcal{A}_l$,

$S_2$: incomplete factorization (ILU) of the PDE block as part of a Block Gauss Seidel scheme.

We apply three smoothing iterations in each pre- and postsmoothing step of the V-cycle, see Algorithm 4.2.

## 4.3 Numerical results

In this section, we make a comparison of the different numerical schemes presented in the previous section. The following computations were performed by using the C++ library deal.II, developed by Bangerth et al. [6], and applying the UMFPACK library [21] as direct solver on the coarsest mesh level .

We consider the following test problem: we approximate the IL-2 signaling of eight evenly positioned T cells in a 3D environment modeled by system (2.7). The configuration is displayed in Figure 2.1(a). The responding T cells are marked in white and the secreting T cell in purple ($q = 2,500$ molecules/$h$). There exists a unique stationary state for this test problem. We obtain a maximal eigenvalue of the sensitivity matrix of $\lambda = 8.88$ with the methods of Section 4.1, which indicates a strong coupling between the PDE and the ODE part.

### 4.3.1 Multigrid preconditioners

We compare the different smoothers $S_1$ and $S_2$ in a series of numerical tests for a stationary solver of system (3.13), to avoid an influence on the comparison by the time stepping scheme. We compute the number of GMRES steps over all Newton steps ($\Sigma n$) and the average reduction rate ($r$) of the residual in each GMRES step. The Newton scheme is iterated until an accuracy of $10^{-6}$ is reached. The number of Newton steps is independent of the grid refinement, since it depends only on the coupling, the nonlinearity of the equation and the accuracy of the solver. Each of the seven resulting Newton steps is solved for an accuracy of $10^{-11}$.

We refine the grid globally until the finest grid with 885673 degrees of freedom, compared to 367 on the coarse grid for the PDE part. Additionally 24 degrees of freedom of the ODE part are coupled to the PDE part. Next, we apply the two smoothers $S_1$ and $S_2$ in the multilevel scheme. We compare two main approaches: smoothing only the PDE part or smoothing the whole matrix. Since the coupling

**Table 4.1:** Reduction rates of different preconditioners

| | MG-$S_1$ (ILU) | | MG-$S_2$ (Bl.-ILU) | |
|---|---|---|---|---|
| $L$ | $\log_{10} r$ | $\Sigma s$ | $log_{10}r$ | $\Sigma s$ |
| 2 | 2.00 | 46 | 1.41 | 69 |
| 3 | 1.92 | 51 | 1.27 | 77 |
| 4 | 1.85 | 54 | 1.21 | 81 |
| 5 | 1.81 | 54 | 1.15 | 84 |

*Notation:* $\Sigma s$ total GMRES iterations
$r$ average reduction rate
$L$ refinement level

between the PDE and the ODE part in the system is strong, we expect that an effective smoother for the coupled system yields better results. In fact, it can be observed that the coupled ILU smoother $S_1$ is 35% more effective than $S_2$ with much higher reduction rates. Therefore, we apply the smoother $S_1$ in the following sections.

**Remark 4.3.** *We note that the higher effectivity of a coupled smoother, compared to to a smoother which acts only for the PDE part, is connected to the type of smoothing. The ILU iterations, for which we showed results in this section, yield a good approximation of the inverse of the system matrix $\mathcal{K}_l$. Smoother with less computational cost, e.g. Jacobi iterations, can give effective approximations for the inverse of the Laplace operator but are not well suited for a coupled PDE/ODE system. A numerical investigation showed that a solver with a coupled Jacobi smoother needs* 8% *more iterations than a simple Jacobi smoother which acts only on the PDE part. Thus, the large gain in efficiency for a coupled smoother, as presented in this section, is only observed when effective smoothing techniques for PDE/ODE systems are applied.*

### 4.3.2 Comparison of coupled and decoupled schemes

In this section, we compare the described coupled and decoupled scheme with each other in different test cases. Hence, the two approaches are:

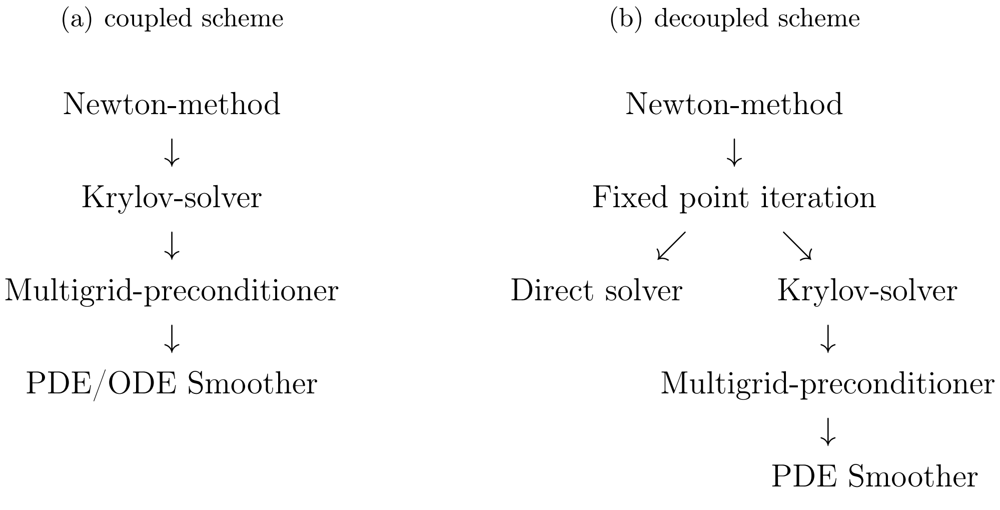
- a Newton-type method in which the linearized **coupled** system (4.8) is solved by a GMRES solver preconditioned by the multigrid method described in Section 4.2.2 with smoother $S_1$,

- a Newton-type method in which the linearized system is solved in a **decoupled** manner by a fixed point iteration defined by the system (4.9). The PDE block is solved by a GMRES solver preconditioned by a multigrid method following the implementation of Janssen and Kanschat [38]. The solver of choice for the symmetric part would be the conjugate gradient method (CG), but we use for a direct comparison instead the GMRES method. We have observed nevertheless that, in combination with the preconditioner, both solvers have similar performance.

We compare the two approaches in Figure 4.1.

To make the schemes comparable, we use a Newton-type method with the same accuracy of the GMRES solver of $TOL_{fixpoint} = 10^{-11}$ for both linearized systems. Thus, the number of Newton steps to solve the nonlinear problem is independent of the approach and we can compare the total number of GMRES steps to solve for an accuracy of $TOL_{newton} = 10^{-6}$. We compare the two solution schemes first in a stationary solver independent of a time stepping scheme and then in a non-stationary solver.

**Figure 4.1:** Comparison of multigrid schemes for nonlinear PDE/ODE systems


(a) coupled scheme

(b) decoupled scheme

### Stationary solvers

We compare the two solution schemes in a stationary computation of the test problem (2.6) in a setup of eight Th cells. The tests consist of simulations with biological parameters that correspond to strong coupling and of simulations with artificial parameters which correspond to a weakly coupled system.

- In the simulations with biological parameters, the maximal value of the eigenvalues of the sensitivity matrix is $\lambda = 8.88$. We expect thus a strong PDE/ODE coupling and hence that the decoupled approach is far less effective than the coupled one.
- A weakly coupled test case is created artificially by increasing the parameter $k_d$ from 0.1 to 1000. The consequent increment of the degradation of $u$ diminishes the influence of the uptake of the cells which depends on the components of $v$. Thus, the PDE part is 'decoupled' from the ODE part: the sensitivity analysis yields a maximal eigenvalue of $\lambda = 0.01$, which indicates that the coupling is very weak.

In Table 4.2, we compare the number of Newton steps ($n$) and the number of total GMRES iterations ($\Sigma s$) needed to obtain a solution of accuracy ($TOL_{newton}$). In each Newton step, the decoupled scheme described in Algorithm 4.1 is iterated until a residual $res_{iter} < TOL_{fixpoint}$ is reached without a given maximum for the number of fixed point iterations $MAX_{fixpoint}$. The average GMRES iterations per Newton step is denoted by $\bar{s}$ and the sum over all GMRES steps by $\Sigma s$. We globally refine the coarse grid three times up to a number of 114929 degrees of freedom.

**Table 4.2:** Coupled vs decoupled solver

| | biological problem $\lambda = 8.88$ | | | | | | modified problem $\lambda = 0.01$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | decoupled | | | coupled | | | decoupled | | | coupled | | |
| $L$ | $s$ | $\bar{n}$ | $\Sigma n$ | $s$ | $\bar{n}$ | $\Sigma n$ | $s$ | $\bar{n}$ | $\Sigma n$ | $s$ | $\bar{n}$ | $\Sigma n$ |
| 2 | 7 | 748 | 5236 | 7 | 6.6 | 46 | 3 | 11 | 33 | 3 | 7 | 21 |
| 3 | 7 | 921 | 6444 | 7 | 7.3 | 51 | 3 | 11 | 33 | 3 | 7 | 21 |
| 4 | 7 | 957 | 6699 | 7 | 7.7 | 54 | 3 | 11.7 | 35 | 3 | 7 | 21 |

*Notation:* $\Sigma n$ GMRES iterations in all Newton steps
$s$ Newton steps
$\bar{n}$ average GMRES iterations per Newton step
$L$ refinement level
$\lambda$ largest eigenvalue of the sensitivity matrices

The results show that the number of Newton steps is independent of the used solution scheme due to the high accuracy $TOL_{fixpoint}$ demanded of the decoupled solver. This accuracy comes at great cost for the decoupled solver, the sum of

GMRES iterations is significantly higher (by a factor of more than 100) for the strong coupled biological problem. The coupled solution scheme is very efficient both for the strong and weak coupled problem. The multigrid preconditioner of the coupled scheme reduces the number of GMRES iterations even in the strong coupled problem to around seven. The decoupled approach is more competitive for the weak coupled problem, though the coupled solver is still 40% faster.

### Non-stationary solvers

We solve in this section the time-dependent system (2.7) for 20 hours until a near stationary solution is reached. In each time step, a coupled nonlinear system (3.12) has to be solved. In contrast to the stationary case, the strength of the coupling is reduced when using small time steps.

We choose the maximum number of fixed point iterations in each Newton step ($MAX_{fixpoint}$) between one and four and compare the fully coupled Newton method to the decoupled quasi-Newton scheme.

**Table 4.3:** Decoupled and coupled schemes of the non-stationary problem

| | | $\Delta t = 0.1$ | | $\Delta t = 0.01$ | |
|---|---|---|---|---|---|
| | $MAX_{fixpoint}$ | $\Sigma s$ | $\Sigma n$ | $\Sigma s$ | $\Sigma n$ |
| decoupled | 1 | 1393 | 3375 | 5566 | 14016 |
| | 2 | 754 | 3792 | 3395 | 15217 |
| | 3 | 547 | 4363 | 2880 | 21153 |
| | 4 | 448 | 4775 | 2871 | 23658 |
| coupled | | 356 | 1799 | 2868 | 11299 |

*Notation:* $\Sigma s$ Newton steps in all time steps
$\Sigma n$ GMRES iterations in all Newton steps
$MAX_{fixpoint}$ max. iterations per Newton step

In Table 4.3 we report the number of computed newton steps ($\Sigma n$) and the number of computed GMRES steps ($\Sigma s$) over all time steps. The results are listed for computations on a once refined spatial grid (2189 degrees of freedom) with 200 ($\Delta t = 0.1$) or 2000 ($\Delta t = 0.01$) time steps.

A higher maximal number of fixed point iterations per Newton step increases the accuracy of the linear solver and thus reduces the number of Newton steps. The decoupled solution scheme with a maximum of four fixed point iterations result in near the same number of Newton steps compared to the coupled solution scheme but with more than twice the number of computed GMRES steps. Nevertheless, it can be observed that the number of total GMRES steps decreases with a reduced number of allowed fixed point iterations per Newton step. Thus, more than one fixed point iteration per Newton step should be avoided if the quasi-Newton method is still converging.

As already remarked, the effectiveness of the decoupled solver depends on the strength of the coupling and thus on the size of the time step. In fact for time steps $\Delta t = 0.1$, the coupled solver needs around half of the iterations of the decoupled solver. For smaller time steps ($\Delta t = 0.01$), the iterations of the coupled solver are reduced by 20% compared to the decoupled solver with $MAX_{fixpoint} = 1$.

**Remark 4.4.** *The use of a higher order time scheme, e.g. the Crank-Nicolson scheme, allows for larger time steps, and hence leads to a stronger coupling during the time integration. Since the coupled solution method is more effective (even for small time steps) in the implicit Euler scheme, it is even better in a higher order scheme.*

# 5 A Posteriori Error Estimation and Mesh Adaptation

The error estimator developed in this chapter is an extension of the error estimator for nonlinear parabolic PDE, developed by Schmich and Vexler [59], for coupled PDE/ODE equations. It is based on the Dual Weighted Residual (DWR) method developed by Becker and Rannacher [8]. We apply the DWR method to obtain accurate and reliable a posteriori error estimations and extend it for different time meshes of the PDE and the ODE part.

We estimate the discretization error with respect to a functional $J \in W'$ of a computed solution $w_{kh} \in X_{kh}^{1,1}$ compared to an exact solution $w \in W$ of a coupled PDE/ODE system. Additionally, we separate the error of the spatial discretization $\eta_h$ from the error of the time discretization $\eta_k$. We distinguish between the error of the time discretization of the PDE part $\eta_k^p$ and the ODE part $\eta_k^o$.

$$J(w) - J(w_{kh}) \approx \eta_k^p + \eta_k^o + \eta_h$$

We begin this chapter by presenting in Section 5.1 a general a posteriori error representation based on the DWR method. We evaluate the error estimator in Section 5.2 and localize the element-wise contributions for local mesh refinement in Section 5.3. In Section 5.4 we present an adaptive space-time mesh refinement strategy, which makes use of different time meshes for the PDE and the ODE part of the system. We conclude the chapter with Section 5.5 and present numerical results for adaptive mesh refinements in a series of 2D and 3D model problems.

## 5.1 A posteriori error representation

In this section, we recall an a posteriori error representation by Schmich an Vexler [59], which is based on the DWR method developed by Becker and Rannacher [8].

To this aim, we define the two operators $B : W \times W \to \mathbb{R}$ and $\tilde{B} : W \cup X_k^1(V) \times W \cup \hat{X}_k^0(V) \to \mathbb{R}$ with the notations (2.4) by

$$\begin{aligned}B(w;\varphi) :=&((\partial_t w,\varphi)) + \mathcal{F}(w;\varphi) + (w(0),\varphi(0)),\\ \tilde{B}(w_k;\varphi_k) :=&((\partial_t w_k,\varphi_k)) + \mathcal{F}(w_k;\varphi_k) + \sum_{m\in J_0}([u_k]_{m-1},\phi_k^{m,+})+\\ &\sum_{\substack{m\in J_0,\\ j\leq M_m^s}}([v_k]_m^{j-1},\psi_k^{m,j-1,+}) + (w_k(0),\varphi_k^{0,-}).\end{aligned}$$

with $w_k := \begin{pmatrix}u_k\\ v_k\end{pmatrix}$ and $\varphi_k = \begin{pmatrix}\phi_k\\ \psi_k\end{pmatrix}$. The non-discretized solution $w \in W$ of (2.5), the semi-discretized solution $w_k \in X_k^1(V)$ of (3.8), and the fully discretized solution $w_{kh} \in X_{kh}^{1,1}$ of (3.11) are given in a more compact formulation by

$$\begin{aligned}B(w;\varphi) &= (w_0,\varphi(0)), \quad \forall\varphi \in W,\\ \tilde{B}(w_k;\varphi_k) &= (w_0,\varphi_k^{0,-}), \quad \forall\varphi_k \in \hat{X}_k^0(V),\\ \tilde{B}(w_{kh};\varphi_{kh}) &= (w_0,\varphi_{kh}^{0,-}), \quad \forall\varphi_{kh} \in \hat{X}_k^{0,1}.\end{aligned} \tag{5.1}$$

This enables us to define the corresponding Lagrange functionals $\mathcal{L} : W \times W \to \mathbb{R}$ and $\tilde{\mathcal{L}} : X_k^1(V) \cup W \times \hat{X}_k^0(V) \cup W \to \mathbb{R}$ by

$$\begin{aligned}\mathcal{L}(w;z) &:= J(w) - B(w;z) + (w_0,z(0)),\\ \tilde{\mathcal{L}}(w_k;z_k) &:= J(w_k) - \tilde{B}(w_k;z_k) + (w_0,z_k^{0,-}),\\ \tilde{\mathcal{L}}(w_{kh};z_{kh}) &:= J(w_{kh}) - \tilde{B}(w_{kh};z) + (w_0,z_{kh}^{0,-}).\end{aligned} \tag{5.2}$$

The stationary points of the Lagrange functionals are the primal solutions $w, w_k$ and $w_{kh}$ together with the dual solutions $z, z_k$ and $z_{kh}$

$$\begin{aligned}\mathcal{L}'(w;z)(\delta w,\delta z) &= 0, \quad \forall(\delta w,\delta z) \in W \times W,\\ \tilde{\mathcal{L}}'(w_k;z_k)(\delta w_k,\delta z_k) &= 0 \quad \forall(\delta w,\delta z) \in X_k^1(V) \times \hat{X}_k^0(V),\\ \tilde{\mathcal{L}}'(w_{kh};z_{kh})(\delta w_{kh},\delta z_{kh}) &= 0, \quad \forall(\delta w_{kh},\delta z_{kh}) \in X_{kh}^{1,1}(V) \times \hat{X}_{kh}^{0,1}(V).\end{aligned} \tag{5.3}$$

The dual solutions $z \in W, z_k \in \hat{X}_k^0(V)$ and $z_{kh} \in \hat{X}_{kh}^{0,1}$ can be computed by solving the adjoint problems

$$\begin{aligned} B'_w(w;\varphi,z) &= J(\varphi), && \forall\varphi \in W,\\ \tilde{B}'_w(w_k;\varphi_k,z_k) &= J(\varphi_k), && \forall\varphi_k \in X_k^1(V),\\ \tilde{B}'_w(w_{kh};\varphi_{kh},z_{kh}) &= J(\varphi_{kh}), && \forall\varphi_{kh} \in X_{kh}^{1,1}, \end{aligned} \tag{5.4}$$

where $B'_w$ denotes the derivative of $B$ in the direction of $w$. The corresponding residuals operators $\rho : X_k^1(V) \times \hat{X}_k^0(V) \cup W \to \mathbb{R}$ and $\rho^* : X_k^1(V) \times \hat{X}_k^0(V) \cup W \times X_k^1(V) \to \mathbb{R}$ of the primal and dual equations are defined by

$$\begin{aligned} \rho(w)(\delta z) &:= \tilde{\mathcal{L}}'_z(w;z,\delta z) = -\tilde{B}(w;\delta z) + (w_0, \delta z(0)),\\ \rho^*(w,z)(\delta w) &:= \tilde{\mathcal{L}}'_w(w;z,\delta w) = -\tilde{B}'_w(w;z,\delta w) + J(\delta w). \end{aligned}$$

We denote the respective PDE and ODE part of the vectors $w, z, \delta w$ and $\delta z$ and the residuals $\rho$ and $\rho^*$ by

$$\begin{aligned} \rho &= \begin{pmatrix} \rho^{(u)} \\ \rho^{(v)} \end{pmatrix}, & w &= \begin{pmatrix} u \\ v \end{pmatrix}, & \delta w &= \begin{pmatrix} \delta u \\ \delta v \end{pmatrix},\\ \rho^* &= \begin{pmatrix} \rho^{*,(u)} \\ \rho^{*,(v)} \end{pmatrix}, & z &= \begin{pmatrix} z_u \\ z_v \end{pmatrix}, & \delta z &= \begin{pmatrix} \delta z_u \\ \delta z_v \end{pmatrix}. \end{aligned} \tag{5.5}$$

With this notation, we write the two parts of the residual operators $\rho$ and $\rho^*$ by

$$\begin{aligned} \rho(w)(\delta z) &= \rho^{(u)}(w)(\delta z_u) + \rho^{(v)}(w)(\delta z_v),\\ \rho^*(w,z)(\delta w) &= \rho^{*,(u)}(w,z)(\delta z_u) + \rho^{*,(v)}(w,z)(\delta z_v). \end{aligned} \tag{5.6}$$

We separate the influence of the temporal and spatial discretization on the goal functional $J$ by

$$J(w,z) - J(w_{kh},z_{kh}) = (J(w,z) - J(w_k,z_k)) + (J(w_k,z_k) - J(w_{kh},z_{kh})). \tag{5.7}$$

We recall the abstract error representation of Becker and Rannacher [8] and the extension for time-dependent equations by Schmich and Vexler [59] and apply it to the Lagrange functionals (5.2).

**Theorem 5.1.** *Let* $(w,z) \in W \times W$, $(w_k,z_k) \in X_k^1(V) \times \hat{X}_k^0(V)$ *and* $(w_{kh},z_{kh}) \in X_{kh}^{1,1} \times \hat{X}_{kh}^{0,1}$ *be solutions of* (5.3) *with a three times differentiable Lagrange functional* $\tilde{\mathcal{L}}$. *Then the following error representations hold for the temporal and spatial* *discretization error*

$$J(w) - J(w_k) = \tfrac{1}{2}\left(\rho(w_h)(z - \varphi_k) + \rho^*(w_k, z_k)(w - \psi_k)\right) + \mathcal{R}_1 \tag{5.8a}$$

$$J(w_k) - J(w_{kh}) = \tfrac{1}{2}\left(\rho(w_{kh})(z_k - \varphi_{kh}) + \rho^*(u_{kh}, z_{kh})(w_k - \psi_{kh})\right) + \mathcal{R}_2, \tag{5.8b}$$

*for arbitrarily chosen* $(\varphi_k, \psi_k) \in X_k^1(V) \times \hat{X}_k^0(V)$ *and* $(\varphi_{kh}, \psi_{kh}) \in X_{kh}^{1,1} \times \hat{X}_{kh}^{0,1}$. *The remainder terms* $\mathcal{R}_1$ *and* $\mathcal{R}_2$ *are defined by*

$$\mathcal{R}_1 := \int\limits_0^1 \tilde{\mathcal{L}}'''(x_k + se_k; e_k, e_k, e_k)s(s-1)\,ds$$

$$\mathcal{R}_2 := \int\limits_0^1 \tilde{\mathcal{L}}'''(x_{kh} + se_{kh}; e_{kh}, e_{kh}, e_{kh})s(s-1)\,ds$$

*with the notation*

$$x := \begin{pmatrix} w \\ z \end{pmatrix}, \quad x_k := \begin{pmatrix} w_k \\ z_k \end{pmatrix}, \quad x_{kh} := \begin{pmatrix} w_{kh} \\ z_{kh} \end{pmatrix},$$

*and the errors*

$$e_k := x - x_k, \qquad e_{kh} := x_k - x_{kh}.$$

*Proof.* The additional jump terms in $\tilde{\mathcal{L}}$ compared to $\mathcal{L}$ vanish for a continuous solution $w \in W$, such that

$$J(w) = \mathcal{L}(w, z) = \tilde{\mathcal{L}}(w, z). \tag{5.9}$$

It follows with (5.1) and identity (5.9) that

$$J(w) - J(w_k) = \tilde{\mathcal{L}}(x) - \tilde{\mathcal{L}}(x_k),$$
$$J(w_k) - J(w_{kh}) = \tilde{\mathcal{L}}(x_k) - \tilde{\mathcal{L}}(x_{kh}).$$

We obtain from (5.3) that $\tilde{\mathcal{L}}'(x; x) = 0$. Since $X_k^1(V) \not\subseteq W$, we can not follow directly that $\tilde{\mathcal{L}}'(x; x_k) = 0$. Instead, the validity of this equation has to be shown by a density argument, for details see Schmich [58]. Thus, we use of $\tilde{\mathcal{L}}'(x; e_k) = 0$

and obtain with the main theorem of calculus

$$\tilde{\mathcal{L}}(x) - \tilde{\mathcal{L}}(x_k) = \int_0^1 \tilde{\mathcal{L}}'(x_k + se_k)(e_k)\, ds + \tfrac{1}{2}\tilde{\mathcal{L}}'(x_k; e_k) - \tfrac{1}{2}\tilde{\mathcal{L}}'(x_k; e_k) - \tfrac{1}{2}\tilde{\mathcal{L}}'(x; e_k).$$

We use the fact that the last two terms are an approximation of the first one by the trapezoidal rule

$$\int_0^1 f(s)\, ds = \tfrac{1}{2} f(0) + f(1) + \tfrac{1}{2}\int_0^1 f''(s) \cdot s \cdot (s-1)\, ds,$$

and obtain that for any $y_k \in X_k^1(V) \times \hat{X}_k^0(V)$ holds

$$\tilde{\mathcal{L}}(x) - \tilde{\mathcal{L}}(x_k) = \tfrac{1}{2}\tilde{\mathcal{L}}'(x_k; x - y_k) + \int_0^1 \tilde{\mathcal{L}}'''(x_k + se_k; e_k, e_k, e_k) s(s-1)\, ds.$$

The error representation (5.8a) is then obtained with the definition of the residuals $\rho$ and $\rho^*$. The derivation of the error representation (5.8b) follows analogously. □

**Remark 5.1.** *Theorem 5.1 assumes that the domain $\Omega$ coincides with the computational domain $\Omega_h$. A small additional discretization error arises, due to the approximation of the curved boundaries $\Gamma_i$ by isoparametric finite elements. This error can be approximated by additional terms, derived by Bangerth and Rannacher [5]. Since this error is small in our application, compared to the spatial and temporal discretization errors, we neglect this additional error.*

## 5.2 Evaluation of the error estimator

Theorem 5.1 is not yet useful for error estimation because it contains several terms which need to be approximated. As proposed by Schmich and Vexler [59], we approximate the error representation (5.8) by neglecting the error term of higher order and replace the unknown semi-discrete solutions $(w_k, z_k)$ in the residual by

$(w_{kh}, z_{kh})$

$$J(w) - J(w_k) \approx \frac{1}{2}\Big(\rho(w_{kh})(z-\varphi_k) + \rho^*(w_{kh}, z_{kh})(w-\psi_k)\Big), \tag{5.10}$$
$$J(w_k) - J(w_{kh}) = \frac{1}{2}\Big(\rho(w_{kh})(z_k-\varphi_{kh}) + \rho^*(w_{kh}, z_{kh})(w_k-\psi_{kh})\Big).$$

There are several possible methods for approximating the interpolation error in time $z - \varphi_k$ and space $z_k - \varphi_{kh}$, see Becker and Rannacher [8]. We choose here to use local post-processing of the computed solution through patch-wise higher order interpolation, which circumvents the need for higher order Galerkin spaces for the solution of the adjoint problem. Therefore, we define the approximate the primal and dual interpolation errors denoted by $\Pi_k^{(w)}$ and $\Pi_k^{(z)}$ in the following way:

$$\Pi_k^{(w)} w_{kh} \approx w - \psi_k, \qquad \Pi_k^{(z)} z_{kh} \approx z - \varphi_k,$$
$$\Pi_h^{(w)} w_{kh} \approx w_k - \psi_{kh}, \qquad \Pi_h^{(z)} z_{kh} \approx z_k - \varphi_{kh}.$$

We define the temporal approximations

$$\Pi_k^{(w)} : X_{kh}^{1,1} \to X_{kh}^{2,1} \cap C^0(I,V), \qquad \Pi_k^{(w)} := \begin{cases} i_k^{(1)}\psi_{kh} - id, & \text{if } t \in I_m, m \in J^0 \\ i_{2k}^{(2)}\psi_{kh} - id, & \text{if } t \in I_m, m \in J^1, \end{cases}$$
$$\Pi_k^{(z)} : \hat{X}_{kh}^{0,1} \to \hat{X}_{kh}^{1,1}, \qquad \Pi_k^{(z)} := i_k^{(1)} - id,$$

and the spatial approximations by

$$\Pi_h^{(w)} : X_{kh}^{1,1} \to X_{kh}^{1,2}, \qquad \Pi_h^{(w)} := i_{2h}^{(2)} - id,$$
$$\Pi_h^{(z)} : \hat{X}_{kh}^{0,1} \to \hat{X}_{kh}^{0,2}, \qquad \Pi_h^{(z)} := i_{2h}^{(2)} - id,$$

where $id$ denotes the identity on the corresponding space. The PDE and ODE part of the interpolation operators $\Pi_k^{(w)}$ and $\Pi_k^{(z)}$ are denoted by

$$\Pi_k^{(w)} = \begin{pmatrix} \Pi_k^{(w),p} \\ \Pi_k^{(w),o} \end{pmatrix}, \quad \Pi_k^{(z)} = \begin{pmatrix} \Pi_k^{(z),p} \\ \Pi_k^{(z),o} \end{pmatrix}.$$

The linear interpolations in time $i_k^{(1)}$ are defined separately on the PDE part as $i_k^{(1),p}$ and on the ODE part as $i_k^{(1),o}$. They are obtained on each time interval $I_m$ or

$I_m^j$ by

$$\begin{aligned} i_k^{(1),p} u_{kh}(t) &:= \frac{t_m - t}{k_m} u^{m-1} + \frac{t - t_{m-1}}{k_m} u^m, \quad \text{for all } t \in I_m, \\ i_k^{(1),o} v_{kh}(t) &:= \frac{t_m^j - t}{k_m^j} v^{m,j-1} + \frac{t - t_m^{j-1}}{k_m^j} v^{m,j}, \quad \text{for all } t \in I_m^j. \end{aligned}$$

We separate the second order interpolations $i_{2k}^{(2)}$ into an interpolation on the PDE part $i_{2k}^{(2),p}$ and an interpolation on the ODE part $i_{2k}^{(2),o}$ analogously. The patch-wise structure of the time grid is used to define

$$i_{2k}^{(2),p} \phi_{kh}(t) := \begin{cases} \frac{(t_m - t)(t_{m+1})}{k_m(k_m + k_{m+1})} \phi_{kh}^{m-1} + \frac{(t - t_{m-1})(t_{m+1} - t)}{k_m k_{m+1}} \phi_{kh}^m \\ \qquad + \frac{(t - t_m)(t - t_{m-1})}{k_{m+1}(k_m + k_{m+1})} \phi_{kh}^{m+1} & \text{if } m - 1 \in 2\mathbb{N} \\ \frac{(t_{m-1} - t)(t_m)}{k_m(k_m + k_{m+1})} \phi_{kh}^{m-2} + \frac{(t - t_{m-2})(t_m - t)}{k_{m-1} k_m} \phi_{kh}^{m-1} \\ \qquad + \frac{(t - t_{m-1})(t - t_m)}{k_m(k_{m-1} + k_m)} \phi_{kh}^m, & \text{if } m \in 2\mathbb{N}, \end{cases}$$

$$i_{2k}^{(2),o} \psi_{kh}(t) := \begin{cases} \frac{(t_m^j - t)(t_m^{j+1})}{k_m^j(k_m^j + k_m^{j+1})} \psi_{kh}^{m,j-1} + \frac{(t - t_m^{j-1})(t_m^{j+1} - t)}{k_m^j k_m^{j+1}} \psi_{kh}^{m,j} \\ \qquad + \frac{(t - t_m^j)(t - t_m^{j-1})}{k_m^{j+1}(k_m^j + k_m^{j+1})} \psi_{kh}^{m,j+1}, & \text{if } j - 1 \in 2\mathbb{N} \\ \frac{(t_m^{j-1} - t)(t_m^j)}{k_m^{j-1}(k_m^{j-1} + k_m^j)} \psi_{kh}^{m,j-2} + \frac{(t - t_m^{j-2})(t_m^j - t)}{k_m^{j-1} k_m^j} \psi_{kh}^{m,j-1} \\ \qquad + \frac{(t - t_m^{j-1})(t - t_m^{j-2})}{k_m^j(k_m^{j-1} + k_m^j)} \psi_{kh}^{m,j}, & \text{if } j \in 2\mathbb{N}. \end{cases}$$

We assume that a patch structure of the spatial mesh exists as well, i.e. eight adjacent cells in 3D can be combined to one macro cell. Thus, a quadratic interpolation operator $i_{2h}^{(2)}$ is applied piecewise on each patch of cells.

At last, we define a computable error estimation by $\eta = \eta_k^p + \eta_k^o + \eta_h$, which consists of an estimation $\eta_h$ of the spatial discretization error and estimations $\eta_k^p$ and $\eta_k^o$ for the PDE and the ODE temporal discretization errors. The error estimations are evaluated by

$$\begin{aligned} \eta_k^p &:= \tfrac{1}{2}(\rho^p(w_{kh})(\Pi_k^{(z),p} z_{kh}^u) + \rho^{*,(p)}(w_{kh}, z_{kh}^u)(\Pi_k^{(w),p} u_{kh}), \\ \eta_k^o &:= \tfrac{1}{2}(\rho^o(w_{kh})(\Pi_k^{(z),o} z_{kh}^v) + \rho^{*,(o)}(w_{kh}, z_{kh}^v)(\Pi_k^{(w),o} v_{kh}), \\ \eta_h &:= \tfrac{1}{2}(\rho(w_{kh})(\Pi_h^{(z)} z_{kh}) + \rho^*(w_{kh}, z_{kh})(\Pi_h^{(w)} w_{kh}). \end{aligned}$$

## 5.3 Localization of the error estimations

The aim of this section is to localize error estimations into elementwise contributions for local mesh refinement. Therefore, we split the error estimations $\eta_k^p, \eta_k^o$ and $\eta_h$ into their contributions on each time interval $I_m$. The localized error indicators are denoted by $\eta_{k,m}^p, \eta_{k,m}^o$ and $\eta_{h,m}$ on each time interval $I_m$, such that

$$\eta_k^p = \sum_{m=1}^{M} \eta_{k,m}^p, \quad \eta_k^o = \sum_{m=1}^{M} \eta_{k,m}^o, \quad \eta_h = \sum_{m=1}^{M} \eta_{h,m}.$$

Before we define the localized error indicators in (5.12), we introduce the following notations. We use the abbreviations $a_m$ and $b_m$ for the time integrals of the operators $\bar{a}$ and $\bar{b}$ over each interval $I_m$:

$$\begin{aligned} a_m(w;z) &:= \int\limits_{I_m} \bar{a}(z(t);z(t))\,dt, \\ b_m^j(w;z) &:= \int\limits_{I_m^j} \bar{b}(w(t);z(t))\,dt. \end{aligned}$$

The respective PDE and ODE part of the local residuals $\rho_m$ and $\rho_m^*$ and the vectors $w_{kh}, z_{kh}, \delta w$ and $\delta z$ are denoted by

$$\begin{aligned} \rho_m &:= \begin{pmatrix} \rho_m^p \\ \rho_m^o \end{pmatrix}, \quad w_{kh} := \begin{pmatrix} u_{kh} \\ v_{kh} \end{pmatrix}, \quad \delta w := \begin{pmatrix} \delta u \\ \delta v \end{pmatrix}, \\ \rho_m^* &:= \begin{pmatrix} \rho_m^{*,p} \\ \rho_m^{*,o} \end{pmatrix}, \quad z_{kh} := \begin{pmatrix} z_{kh}^u \\ z_{kh}^v \end{pmatrix}, \quad \delta z := \begin{pmatrix} \delta z_u \\ \delta z_v \end{pmatrix}. \end{aligned}$$

We split the local primal residuals $\rho_m(w_{kh})(\delta z)$ and dual residuals $\rho_m^*(w_{kh}, z_{kh})(\delta w)$ into a PDE and an ODE part

$$\begin{aligned} \rho_m(w_{kh})(\delta z) &= \rho_m^p(w_{kh})(\delta z_u) + \rho_m^o(w_{kh}, z_{kh})(\delta z_v), \\ \rho_m^*(w_{kh}, z_{kh})(\delta w) &= \rho_m^{*,p}(w_{kh}, z_{kh})(\delta u) + \rho_m^{*;o}(w_{kh}, z_{kh})(\delta v). \end{aligned}$$

The local primal residuals $\rho_m$ for all $0 \leq m \leq M$ are defined as follows:

- for $m \in J_0, m = 0$:

$$\begin{aligned}\rho_0^p(w_{kh})(\delta z_u) &:= -\left(u_0 - u_{kh}(0), \delta z_u^{0,-}\right),\\ \rho_0^o(w_{kh})(\delta z_v) &:= -\left(v_0 - v_{kh}(0), \delta z_v^{0,-}\right),\end{aligned}$$

- for $m \in J_0, m > 0$:

$$\begin{aligned}\rho_m^p(w_{kh})(\delta z_u) :=& -((\partial_t u_{kh}, \delta z_u))_m - \mu((\nabla u_{kh}, \nabla \delta z_u))_m - a_m(w_{kh}; \delta z_u)\\ &- ([u_{kh}]_{m-1}, \delta z_u^{m-1,+}),\\ \rho_m^o(w_{kh})(\delta z_v) :=& -\sum_{j \le M_m^s} \left( ((\partial_t v_{kh}, \delta z_v))_{m,j} - b_m^j(w_{kh}; \delta z_v) - ([v_{kh}]_m^{j-1}, \delta z_v^{m,j-1,+}) \right).\end{aligned}$$

- for $m \in J_1, m > 0$:

$$\begin{aligned}\rho_m^p(w_{kh})(\delta z_u) &:= -((\partial_t u_{kh}, \delta z_u))_m - \mu((\nabla u_{kh}, \nabla \delta z_u))_m - a_m(w_{kh}; \delta z_u),\\ \rho_m^o(w_{kh})(\delta z_v) &:= -\sum_{j \le M_m^s} \left( ((\partial_t v_{kh}, \delta z_v))_{m,j} - b_m^j(w_{kh}; \delta z_v) \right).\end{aligned}$$

We assume that the functional $J(w)$ is linear and of the form

$$\begin{aligned}J(w) &:= J^p(u) + J^o(v),\\ J^p(u) &:= J^{p,1}(u(T)) + \int_I J^{p,2}(u)\,dt,\\ J^o(v) &:= J^{o,1}(v(T)) + \int_I J^{o,2}(v)\,dt,\end{aligned} \tag{5.11}$$

with functionals $J^{p,1} \in V^{p\prime}$ and $J^{o,1} \in V^{o\prime}$, which depend only on the value at the final time point $T$ and time dependent functionals $J^{p,2} \in W^{p\prime}$ and $J^{o,2} \in W^{o\prime}$.

The local dual residuals $\rho_m^*$ for all $0 \le m \le M$ are defined by:

- for $m = M$:

$$\begin{aligned}
\rho_M^{*,p}(w_{kh}, z_{kh})(\delta u) :=& J^{p,1}(\delta u(t_M)) - \mu((\nabla u_{kh}, \nabla \delta u))_m - a'_{m,u}(w_{kh}; z_{kh}, \delta u) \\
&- b'_{M,u}(w_{kh}; z_{kh}, \delta u) - (\delta u(t_M), z_{kh}^{u;M,-}), \\
\rho_M^{*,o}(w_{kh}, z_{kh})(\delta v) :=& J^{o,1}(\delta v(t_M)) - \sum_{j \leq M_m^s} \left( a^{j\,\prime}_{m,v}(w_{kh}; z_{kh}, \delta v) + b^{j\,\prime}_{m,v}(w_{kh}; z_{kh}, \delta v) \right) \\
&- \sum_{j \leq M_m^s - 1} (\delta v(t_M^j), [z_{kh}^v]_M^j) - (\delta v(t_M), z_{kh}^{v;M,-}).
\end{aligned}$$

- for $0 < m < M$:

$$\begin{aligned}
\rho_m^{*,p}(w_{kh}, z_{kh})(\delta u) :=& J^{p,2}(\delta u) - \mu((\nabla u_{kh}, \nabla \delta u))_m - a'_{m,u}(w_{kh}; z_{kh}, \delta u) \\
&- b'_{m,u}(w_{kh}; z_{kh}, \delta u) - (\delta u(t_m), [z_{kh}^u]_m), \\
\rho_m^{*,o}(w_{kh}, z_{kh})(\delta v) :=& J^{o,2}(\delta v) - \sum_{j \leq M_m^s} \left( a^{j\,\prime}_{m,v}(w_{kh}; z_{kh}, \delta v) + b^{j\,\prime}_{m,v}(w_{kh}; z_{kh}, \delta v) \right) \\
&- (\delta v(t_m^j), [z_{kh}^v]_m^j).
\end{aligned}$$

- for $m = 0$:

$$\begin{aligned}
\rho_0^{*,p}(w_{kh}, z_{kh})(\delta u) :=& -(\delta u^{0,+}, [z_{kh}^u]_0), \\
\rho_0^{*,o}(w_{kh}, z_{kh})(\delta v) :=& -(\delta v^{0,+}, [z_{kh}^v]_0).
\end{aligned}$$

We can then define the local error indicators for $m \leq M$ by

$$\begin{aligned}
\eta_{k,m}^p :=& \frac{1}{2} \left( \rho_m^p(w_{kh})(\Pi_k^{(z),p} z_{kh}^u) + \rho_m^{*,p}(w_{kh}, z_{kh})(\Pi_k^{(w),p} u_{kh}) \right), \\
\eta_{k,m}^o :=& \frac{1}{2} \left( \rho_m^o(w_{kh})(\Pi_k^{(z),o} z_{kh}^v) + \rho_m^{*,o}(w_{kh}, z_{kh})(\Pi_k^{(w),o} v_{kh}) \right), \\
\eta_{h,m} :=& \frac{1}{2} \left( \rho_m(w_{kh})(\Pi_h^{(z)} z_{kh}) + \rho_m^{*}(w_{kh}, z_{kh})(\Pi_h^{(w)} w_{kh}) \right).
\end{aligned} \tag{5.12}$$

To localize the error indicators $\eta_{h,m}$ on each cell of the mesh, we use the notation $\eta_{h,K,m}$ for the cell wise contributions at time point $t_m$ for a mesh cell $K \in \mathcal{T}_h$ with

$$\eta_{h,m} = \sum_{K \in \mathcal{T}_h} \eta_{h,K,m}.$$

As usual in the DWR method, the residual is integrated by parts to exploit oscillatory behavior of the residuals. Therefore, we replace $((\nabla u, \nabla \varphi))_m$ in the local and dual residual such that the cell wise contributions of the residuals are obtained by evaluating the corresponding cell and the surrounding edges. The jump of the normal derivative of $u$ between to neighboring cells $K$ and $K'$ with common edge $\Gamma$ is denoted by $[n \cdot \nabla u] := n \cdot (\nabla u_{h|K \cap \Gamma} - \nabla u_{h|K' \cap \Gamma})$.

$$((\nabla u, \nabla \varphi))_m = \sum_{K \in \mathcal{T}_h} \int_{I_m} (-\Delta u, \varphi)_K + \frac{1}{2}([n \cdot \nabla u], \varphi)_{\partial K \backslash \partial \Omega_h} + (n \cdot \nabla u, \varphi)_{\partial K \cap \partial \Omega_h} \, dt$$

We do not use different spatial grids for different time points in this thesis. Thus, we are not interested in the time dependence of the spatial error. Therefore, we sum over the contributions of all time intervals to obtain

$$\eta_{h,K} = \sum_{m \leq M} \eta_{h,K,m}. \tag{5.13}$$

The computed local error indicators $\eta_{h,K}$ enable us to set up a strategy for an adaptive refinement of the spatial mesh .

## 5.4 Refinement strategies

In this section, we present a mesh refinement strategy in Algorithm 5.1 for space-time adaptivity based on the above a posteriori error estimates. We aim to compute the functional value $J(w)$ to a given accuracy TOL, such that

$$|\eta| = |\eta_h + \eta_k^p + \eta_k^o| \leq \text{TOL}.$$

We refine the discretization locally until the estimated error is reduced below the given accuracy. In each cycle the subsequent refinement is chosen with the aim of equilibrating the discretization errors in space and time. We set the equilibration parameter to $\kappa_1 = 4$ for our computations such that

$$\frac{1}{\kappa_1} \leq \frac{\eta_h}{\eta_k} \leq \kappa_1,$$

is ensured during each step of the mesh refinement process. Additionally, we have the option to balance the estimated error indicators $\eta_k^p$ and $\eta_k^o$ due to the different temporal discretization between the PDE and the ODE part. There exist a multitude of different options how to achieve a balancing between the two error indicators. Therefore, we set a second equilibration parameter $\kappa_2$, e.g. $\kappa_2 = 1$. We select the additional refinement of the ODE time mesh, such that

$$\eta_k^o \leq \kappa_2 \eta_k^p.$$

In the case of a large estimated error in the PDE part, such that $\eta_k^p \gg \eta_k^o$, we use the same time mesh for the PDE and the ODE part. A coarser refinement in the ODE time mesh than in the PDE time mesh would only slightly reduce the computational effort, while eventually reducing the accuracy of the PDE part due to the coupling.

There are several techniques available for the refinement of both the spatial and the temporal discretization. E.g. Richter [54] proposed an optimal mesh refinement strategy based upon the minimization of the product of expected error and the computational effort for obtaining a solution on the new mesh. The drawback of the strategy is, that if in some cells the error is largely overestimated this approach refines too few cells. This could lead to a multitude of mesh refinement cycles. Therefore we avoid Richter's minimization strategy and choose a simpler strategy, which refines both in the spatial grid $\mathcal{T}_h$ and in the temporal grid $\mathcal{T}_k$ a subset consisting of a fixed percentage of the total number of cells. Additionally to the cells with the largest localized error indicators we refine some of their neighbors as well to ensure the patch-wise structure of the grid. All cells are refined simultaneously in each patch, which consist of $2^d$ cells in the spatial grid and of 2 time intervals in the temporal grid. To maintain the regularity of the spatial grid and to minimize the number of hanging nodes, we make sure that two neighboring cells do not differ from each other by more than one refinement level.

The different temporal discretization between the PDE part and the ODE part is described by the number of time steps $M_m^s$ of the ODE part during one time step of the PDE part. In principal $M_m^s$ could be chosen on each interval $I_m$ separately by evaluating the ratio between the localized error indicators $\eta_{k,m}^p$ and $\eta_{k,m}^o$. Coupled solution schemes on a fine ODE time mesh have small computational costs compared to a refinement of the PDE time mesh. Therefore, we refine all time intervals $I_m$ with the same number of ODE time steps ($M^s$). We set the parameter $M^s$, approximately, as the fraction of the error indicators:

$$M^s \approx \frac{\eta_k^o}{\eta_k^p}. \tag{5.14}$$

We note that in Algorithm 5.1 the ODE time mesh is refined even if the estimated error of the temporal discretization is small compared to the error of the spatial discretization. As we already emphasized a computation on refined ODE time meshes comes with few additional costs. Therefore, the algorithm delays a refinement of the PDE time mesh up to the point until the temporal discretization error exceeds the spatial discretization error.

**Algorithm 5.1:** Adaptive mesh refinement algorithm

```
Select a initial temporal and spatial discretization
$n = 0$
repeat
    compute the primal solution $w_{k_n,h_n}$ by solving (3.12)
    compute the dual solution $z_{k_n,h_n}$ by solving (3.15)
    evaluate the error indicators $\eta^p_{k_n,m}, \eta^o_{k_n,m}$ for each time step $m$ via (5.12)
    evaluate the error indicators $\eta_{h,K}$ for each cell $K \in \mathcal{T}_{h_n}$ via (5.13)
    if $|\eta| \leq TOL$ then
        return
    end
    if $|\eta_h| \geq \kappa_1 |\eta^p_{k_n} + \eta^o_{k_n}|$ then
        Adapt the spatial discretization
    else if $|\eta^p_{k_n} + \eta^o_{k_n}| \geq \kappa_1 |\eta_h|$ then
        Adapt the temporal discretization
    else
        Adapt both the temporal and the spatial discretization
    end
    if $\kappa_2 |\eta^p_{k_n}| \geq |\eta^o_{k_n}|$ then
        Adapt the time discretization of the ODE part of the equation
    end
    $n = n + 1$
```

Depending on the solving method and the nonlinearity of the equation, there are limits to the number of ODE time steps per PDE time step $M^s$. We were able to obtain fast and reliable solver for parameters $M^s$ of up to 1000 by using a coupled solution approach, see Chapter 4. Commonly used PDE solvers, such as preconditioned CG or GMRES-methods, are not well suited to solve equations resulting from coupled PDE/ODE equations. A few lines of coupled ODEs do not significantly change the properties of the system and a reasonable performance of the solver is still observed. On the other hand, the number of ODEs is increased by a large parameter $M^s$, see Remark 3.1. When such a system is solved on a coarse spatial mesh $\mathcal{T}_h$, the size of the PDE part can be equal or even smaller than the number of coupled ODE equations. Specialized PDE solvers are not suitable in such cases. We note that for highly nonlinear problems, even independently of the used linear solver, a large parameter $M^s$ can significantly increase the number of Newton steps for large time steps of the PDE part. Hence, we advise to set a maximum for the parameter $M^s$ and achieve further reduction of the temporal error by refinement of both the PDE and the ODE time mesh.

## 5.5 Numerical results

In this section, we present numerical results for adaptive mesh refinement based on the previously derived error indicators. We begin this section by discussing implementation details in Section 5.5.1 and heuristic error indicators in Section 5.5.2. We motivate in Section 5.5.3 the investigation of a series of model problems in 2D and 3D based on the nonlinear PDE/ODE system presented in Section 2.3. We display the effectivity of adaptive temporal grids in Section 5.5.4 and show that the advantage of spatial adaptive grids relates closely to the diffusion rate of the investigated PDE/ODE system in Section 5.5.5. In Section 5.5.6, we test the proposed mesh adaptation strategy on 2D and 3D applications and discuss the effectivity of adaptive grids compared to grids obtained by heuristic oder global refinement strategies. The effectivity of different time discretizations between the PDE and the ODE part is investigated in Section 5.5.7 and possible applications are discussed.

To evaluate the quality of the error estimator, we define the effectivity index of an estimated error $\eta$ by the ratio between the estimated and the actual error:

$$I_{eff} = \frac{\eta}{J(w) - J(w_{kh})}.$$

Since we do not know the exact solution $w$ of the considered model problems, we have to approximate it by computations on fine grids and extrapolation. Therefore, the 'actual' error $J(w) - J(w_{kh})$ can only be roughly estimated, especially on fine grids.

### 5.5.1 Implementation aspects

The domain is discretized by piecewise semi linear isoparametric finite elements ($\mathcal{Q}^1$). We compute a primal solution $w_{kh} \in X_{kh}^{1,1}$ and a dual solution $z \in \hat{X}_{kh}^{0,1}$. In each time step of both the primal and the dual solving process, a nonlinear coupled problem needs to be solved in terms of the definitions in (3.10). We apply the coupled solution techniques detailed in Chapter 4. These coupled solution schemes enable us to neglect iterations errors, which would inevitably occur in decoupled schemes.

Depending on the choice of the quantity of interest $J$, both the time discretization of the primal problem and the dual problem have to cope with irregular initial data. Since the dual time stepping is solved on the temporal discretization of the primal problem, but backwards in time, damping steps are necessary both at the beginning and at the end of the time interval. Hence, we choose a symmetric time stepping such that for $m_d$ damping steps the set of indices $J_0$ is given by

$$J_0 := \{1, ..., m_d\} \cup \{M + 1 - m_d, ..., M\}.$$

We keep the number of implicit Euler steps $m_d$ constant for each mesh refinement cycle.

Rannacher [51] showed that, depending on the severity of the irregularity of the initial conditions, two or more damping steps are necessary. We made the observation in our computations that the starting conditions (especially for the adjoint problem) could not be sufficiently smoothed with two damping steps. Therefore, we use in our computations six damping steps to avoid a loss of accuracy by the lack of strong A-stability of the Crank-Nicolson method.

### 5.5.2 Heuristic error indicators

We compare in the next sections adaptively refined discretizations with discretizations obtained by uniform refinement. The a posteriori error estimation based on the DWR method provides reliable error indicators for a given quantity of interest. A mesh refinement algorithm based on these dually weighted error indicators enables a fast error reduction.

Heuristic error indicators allow for adaptive mesh refinement without solving an additional dual problem. They do rely on smoothness properties of the measured solution and therefore have the drawback that they cannot detect the influence of the discretization error on the evaluation of a given functional. Especially in the case of localized quantities of interest, e.g. point evaluations, heuristic indicators do not produce suitable discretizations.

Sharp upper bounds for the error in the $L^2(I, H^1(\Omega))$-norm for the parabolic heat equation (5.15) have been derived by Verfürth [68].

$$\begin{aligned} \partial_t u - \Delta u &= f && \text{in } \Omega \times (0,T], \\ u &= 0 && \text{on } \partial\Omega \times (0,T], \\ u &= u_0 && \text{in } \Omega. \end{aligned} \tag{5.15}$$

Verfürth showed that the error of the solution approximated in time by the Crank-Nicolson scheme and in space by finite elements is bounded in $L^2(I, H^1)$-norm from above by a temporal error indicator $\tilde{\eta}_k$ and a spatial error indicator $\tilde{\eta}_h$, such that

$$\int\limits_0^T ||u - u_{kh}||^2_{H^1(\Omega)}\, dt \le c(\tilde{\eta}_h^2 + \tilde{\eta}_k^2). \tag{5.16}$$

The constant $c$ in (5.16) depends only on the geometry of the finite element approximations. The error indicators consist of local error indicators $\tilde{\eta}_{h,K}$ for each cell $K$ of the spatial mesh, indicators $\tilde{\eta}_{h,\partial K}$ for the edges of each cell $K$ of the spatial mesh and indicators $\tilde{\eta}_{k,m}$ for each time interval $I_m$,

$$\begin{aligned} \tilde{\eta}_h^2 &:= \sum_{K \in \mathcal{T}} \left( \tilde{\eta}_{h,K}^2 + \tilde{\eta}_{h,\partial K}^2 \right), \\ \tilde{\eta}_k^2 &:= \sum_{m \le M} \tilde{\eta}_{k,m}^2 \end{aligned}$$

with the local indicators defined by

$$\tilde{\eta}_{h,K}^2 := \sum_{m \le M} k_m h_K^2 ||f - \frac{u_{kh}^m - u_{kh}^{m-1}}{k_m} + \frac{1}{2}(\Delta u_{kh}^m + \Delta u_{kh}^{m-1})||^2_{L^2(K)}, \tag{5.17}$$

$$\tilde{\eta}_{h,\partial K}^2 := \sum_{m \le M} k_m h_K || \frac{1}{2}[n \cdot \nabla u_{kh}^m + n \cdot \nabla u_{kh}^{m-1}]_{\partial K \setminus \partial\Omega} ||^2_{L^2(E)}, \tag{5.18}$$

$$\tilde{\eta}_{k,m}^2 := k_m ||\nabla (u_{kh}^m - u_{kh}^{m-1})||^2_{L^2(\Omega)}. \tag{5.19}$$

The jump along the face for each cell $K \in \mathcal{T}_h$ is denoted by $[\cdot]_{\partial K}$. The a posteriori analysis of Verfürth [67] further showed that similar bounds hold true for nonlinear systems. The same edge residuals $\tilde{\eta}_{h,\partial K}$ and temporal error indicators

$\tilde{\eta}_{k,m}$ are obtained for nonlinear systems without convective terms. It is known that edge residuals dominate the cell residuals $\tilde{\eta}_{h,K}$ for lower order finite element discretizations, see Carstensen and Verfürth [16]. Therefore, we neglect the cell residuals for the heuristic error indicators used in this thesis.

In the following sections, we compute the heuristic error indicators for spatial mesh refinement by $\tilde{\eta}_{h,\partial K}$. Since the ODE part of the solution is not spatially resolved, the spatial gradients in the heuristic error indicators $\tilde{\eta}_{k,m}$ can only be evaluated for the PDE part. Heuristic error bounds for the time discretization error in $L^2$-norm are obtained by measuring the variation of the solution on the time mesh, see e.g. Hartmann [8, 31]. Thus, we apply the following error indicators for the temporal discretization error of the ODE part by

$$\tilde{\eta}_{k,o}^2 := \sum_{m \leq M} k_m ||(v_{kh}^m - v_{kh}^{m-1})||_{L^2}^2.$$

The balance of the heuristic error contributions of ODE part with the PDE part depends on the given quantity of interest. In this thesis, we simply add over both ODE and PDE contributions to obtain local error indicators for the adaptive refinement of the time mesh. We compute the efficiency of adaptively refined discretizations obtained by the heuristic indicators and by dually weighted error indicators in the following sections and compare the results.

### 5.5.3 Discussion of model problems

Space-time Galerkin methods have already been applied successfully to create adaptive grids in the simulation of incompressible flows by Bänsch [7], Hoffman [37], and Besier and Rannacher [9]. Other applications include the acoustic wave equation by Bangerth and Rannacher [4], a propagating flame by Schmich and Vexler [59], and the Black-Scholes equation by Goll et al. [28]. In all publications it was shown that adaptive discretizations yield significant gains in accuracy compared to heuristic or global mesh refinement strategies. Both the control of the discretization error and the increased accuracy compensated for the additional computational work of solving the adjoint equation and evaluating the local error indicators.

The large class of problems resulting from coupled PDE/ODE systems has, to the knowledge of the author, never been adaptively discretized with error indicators obtained by the DWR method. The aim for the next sections is firstly to show that the DWR method can be used as a reliable a posteriori error estimator for coupled PDE/ODE systems and secondly to assess in which cases a significant gain in accuracy by adaptive grids can be expected.

The coupled PDE/ODE models for cellular signaling differ from the previously investigated problems mainly by the enormously fast diffusivity of the PDE compared to the time scales of the processes governed by the ODE part. While the receptors (ODE part) increase during the course of multiple hours, the proteins (PDE part) diffuse in the intercellular area with a diffusion coefficient of 36,000 $\mu m^2/h$. The diffusion coefficient ranges between 360,000 and 150 $\mu m^2/h$ on similar spatial scales in PDE/ODE models for cellular signaling, see [19, 22, 26, 64]. The intracellular processes governed by the ODEs take place on similar time scales, i.e. multiple hours of ligand-binding. Furthermore many applications result in similar PDE/ODE systems. E.g. Moghadam et al. [48] simulated cardiovascular networks in which the Navier-Stokes equation (PDE) is coupled with an ODE system modeling blood in- and outflow.

The high diffusivity of proteins renders spatial simulations dispensable in many intracellular pathways. A study by Claus et al. [19] investigated receptor-ligand interactions in the intracellular SMAD signaling pathway both by a full spatial modeling of the cytoplasm and an ODE model. The ODE model simplified the former PDE part with homogeneously distributed ligand concentrations by assuming infinitely fast diffusion. The study showed, for a diffusion coefficient of 54,000 $\mu m^2/h$ and regularly shaped cells, that the simplified ODE model produces similar results compared to the more elaborate PDE/ODE system. On the other hand, the same comparison for models with a lower diffusivity (3,600 $\mu m^2/h$) showed significantly different results for the PDE/ODE system. Therefore, the necessity of local spatial mesh refinement clearly depends on the given diffusion rate compared to the reaction speed.

To show the efficiency over a large variety of different models, we investigate the efficiency of adaptively refined discretizations with respect to the diffusion rate in two model problems. Firstly, we simulate the model presented in Section 2.3 with a diffusion rate of 36,000 $\mu m^2/h$. We solve the same system in a second model problem but with a diffusion rate of only 1 $\mu m^2/h$. This emulates the simulation of a receptor-ligand system with low diffusion rate, e.g. the NF$\kappa$B intracellular pathway simulated by Terry and Chaplain [64].

We consider both two and three-dimensional models, because 2D models allow us a more detailed study of local spatial mesh refinement due to the reduced computational costs. We compose the setup for the 2D model problems of four cells including one secreting cell. The setup of the 3D model problems are composed

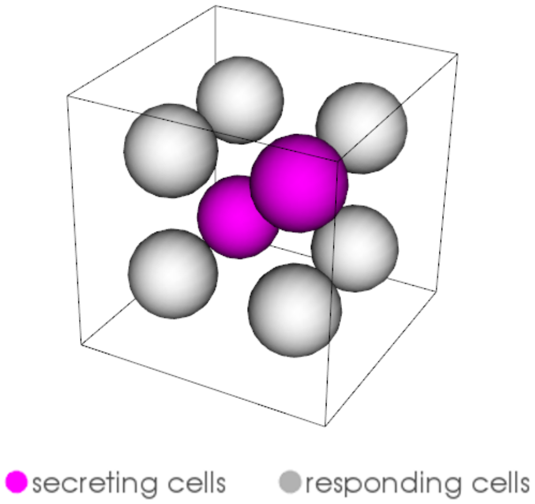


**Figure 5.1:** Three dimensional model problem

of eight cells including two secreting cells. The cells are arranged as depicted in Figure 5.1. We choose a secretion rate of $q = 3,500$ molecules/$h$.

We display in Figure 5.3 the distribution of IL-2 concentration in the extracellular area for both a diffusion coefficient of 36,000 and 1 $\mu m^2/h$, visualized by the VisIt software [17]. One clearly sees the higher cytokine-levels around the secreting cell in the lower left corner. The reduced diffusion coefficient gives rise to much sharper gradients between the T cells. We compare the average Interleukin concentrations on the cell surface and the receptor levels during the activation process for 15 hours in Figure 5.4. The figure shows the higher receptor levels of the activated cell compared the three non-activated cells. The results are displayed for a diffusion coefficient of 36,000 $\mu m^2/h$. We observe a similar time course in the case of slower diffusivity, though with slightly slower receptor increase and therefore a delayed process of activation.

The DWR method gives us the opportunity to approximate the error of a given quantity of interest. In the following sections, we consider two different goal functionals for a given solution $w \in W$. Firstly, we define the functional $J_R$ as the sum over all receptors at the final time point $T = 15\ h$ of the simulations:

$$J_R(w) := \sum_{i \leq N_c} R_i(T).$$

The number of ligand free receptors is closely connected to the number of activated cells in general and can be approximately measured in biological experiments, see Busse et al. [13]. Therefore, solving the IL-2 pathway with respect to the functional $J_R$ precisely computes the number of activated cells after the activation process. Secondly, we define the functional $J_I$:

$$J_I(w) := \int_A u(T),$$

where $A \subset \Omega$ denotes a square of two $\mu m$ width in the lower right corner of the extracellular area $\Omega$. The concentration level of cytokines is usually measured globally after the end of the experiment for biological 'in vitro simulations' in petri dishes. The functional is of lesser biological relevance but emulates the structure of more localized PDE/ODE applications, e.g. the pressure at a specific point of a cardiovascular network.

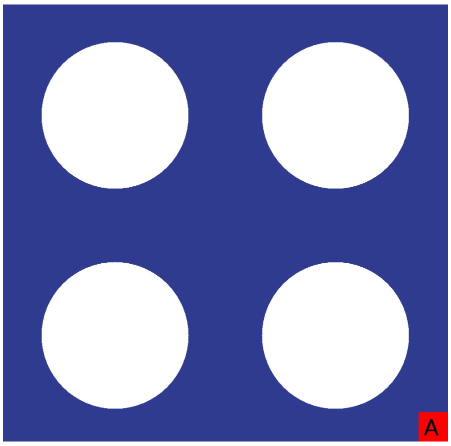
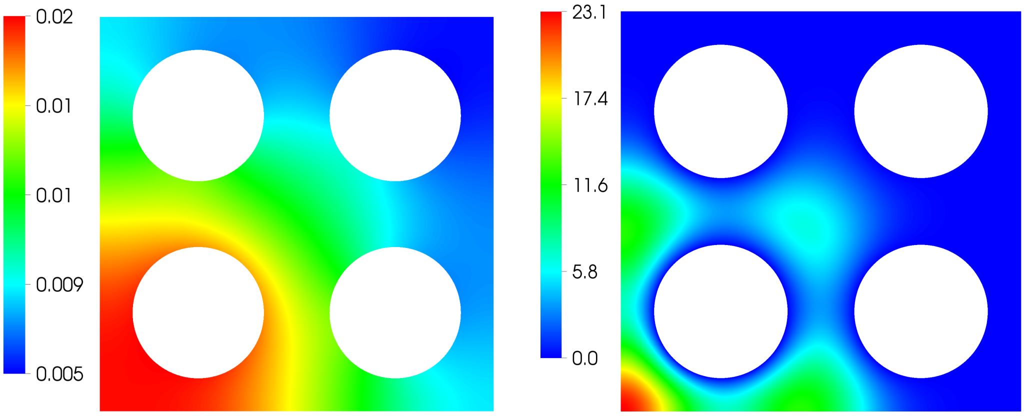


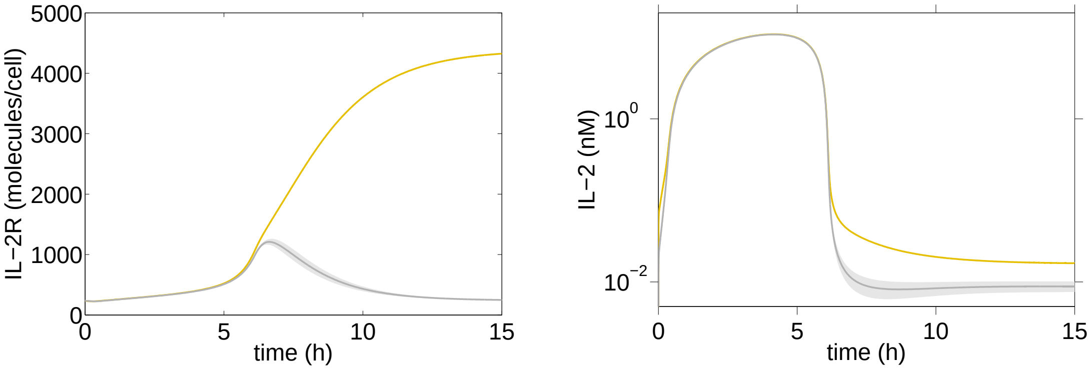
**Figure 5.2:** The highlighted area $A \subset \Omega$ denotes the lower right corner of the extracellular area (width 2 $\mu m$)

To assess the influence of residual errors onto errors of the quantity of interest $J_R$, we display the corresponding dual solutions in Figure 5.5. The dual solution can be interpreted as the sensitivity of the goal functional to perturbations of the solution. Even though the functional $J_R$ measures the number of receptors only at the final time point, especially the ODE part of the equation acts a great influence during the first hours of the activation process. On the other hand, the PDE part of the equation is not part of the quantity of interest. Thus, the dual solution shows that the PDE part only influences the functional during the final hours significantly.



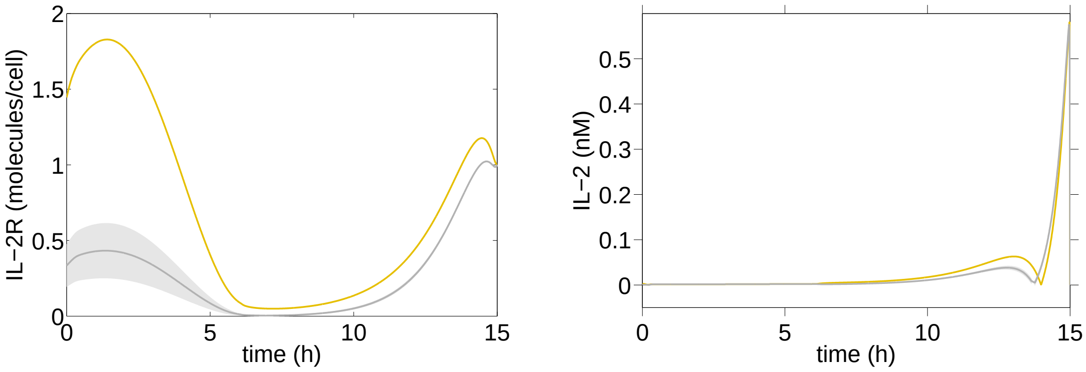
(a) biological diffusivity, $\mu = 36{,}000\ \mu m^2/h$ (b) decreased diffusivity, $\mu = 1\ \mu m^2/h$.

**Figure 5.3:** Distribution of IL-2 concentration at T=15$h$



(a) Receptors (b) Interleukin

**Figure 5.4:** Time course of the receptor number and average IL-2 concentration at the surface for the activated cell (marked gold) and the non-activated cells (marked grey) of the model problem in 2D for $\mu = 36{,}000\ \mu m^2/h$. Solid lines indicate averages, blurred regions standard deviations.



(a) Receptors (b) Interleukin

**Figure 5.5:** Dual solution of the model problem in 2D for $\mu = 36,000\mu m^2/h$ and the functional $J_R$: time course of the receptor number and average IL-2 concentration at the surface for the activated cell (marked gold) and the non-activated cells (marked grey). Solid lines indicate averages, blurred regions standard deviations.

### 5.5.4 Adaptive temporal refinement

To check for independent error estimations $\eta_k$ and $\eta_h$, we observe the behavior of the spatial error estimate $\eta_h$ under adaptive temporal mesh refinement. We note that the results of different time discretizations for the ODE part of system are found in Section 5.5.7.

We denote the number of time steps by $M$ and the number of degrees of freedom of the PDE part by $N^p$ and of the ODE part by $N^o$. The total number of degrees of freedom, which has to be solved by the non-stationary solver, is obtained by $M \cdot (N^p + N^o)$. We consider a fine spatial discretization of the 2D model problem with $N^p = 33405$ to ensure that the temporal error is dominant. The intracellular dynamics of the four cells are described by $N^o = 12$ degrees of freedom.

We display the results for a diffusion coefficient of $\mu$ = 36,000 $\mu m^2/h$ and the quantity of interest $J = J_R$. The results are presented in Table 5.1. We observe that the spatial error remains unchanged, while the temporal error is reduced in each step of the adaptive mesh refinement cycle. The spatial error dominates the temporal error after four temporal mesh refinement cycles. To check the effectivity of the error estimates, we approximate the exact solution of the goal functional by $J(w) \approx 3013.0032$ by computation on a highly refined discretization. We observe that both the dominant temporal error in the first cycles and the dominant spatial error in the later cycles are effectively estimated by comparison of the estimated error to the exact error $|J(w) - J(w_{kh})|$.

**Table 5.1:** Invariance of $\eta_h$ under adaptive temporal mesh refinement and effectivity of the temporal error estimator $\eta_k$ for the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

| $M$ | $N^p$ | $\eta_k$ | $\eta_h$ | $J(w) - J(w_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|
| 150 | 33405 | $4.52 \cdot 10^{-1}$ | $6.40 \cdot 10^{-3}$ | $5.69 \cdot 10^{-1}$ | 0.80 |
| 372 | 33405 | $2.22 \cdot 10^{-2}$ | $7.06 \cdot 10^{-3}$ | $3.52 \cdot 10^{-2}$ | 0.83 |
| 768 | 33405 | $2.14 \cdot 10^{-3}$ | $7.07 \cdot 10^{-3}$ | $6.75 \cdot 10^{-3}$ | 1.36 |
| 1602 | 33405 | $-2.88 \cdot 10^{-4}$ | $7.07 \cdot 10^{-3}$ | $5.87 \cdot 10^{-3}$ | 1.16 |

To assess the efficiency of adaptively refined temporal meshes, we compare the mesh refinement strategy based on error indicators obtained by the DWR method to heuristic error indicators (see Section 5.5.1) and uniform mesh refinement. A comparison of the error reduction by mesh refinement is shown in Figure 5.6. Adaptive temporal mesh refinement by the DWR method is clearly more efficient than uniform mesh refinement. Using heuristic error indicators does not lead to an efficient refinement of the temporal discretization. We compare the temporal discretizations in Figure 5.7 and observe that the heuristic error indicators lead to a coarse temporal grid near the final time point $T = 15h$. Since the ODE part of the solution is approaching a steady state, the heuristic methods indicate no mesh refinement near the final time point. In contrast, the error indicators based on the DWR method sense effectively the dependence of the goal functional $J_R$ on the final time point. Hence, the temporal discretization in last hours is not coarsened even though it is approaching a steady state.

### 5.5.5 Adaptive spatial refinement

In this section, we observe the behavior of the temporal error estimate $\eta_k$ under spatial mesh refinement. We evaluate the effectivity of the error estimates obtained by the DWR method and the efficiency of the adaptive mesh refinement strategy for the model problems presented in Section 5.5.3. We compare this mesh refinement strategy with heuristic and uniform mesh refinement schemes.

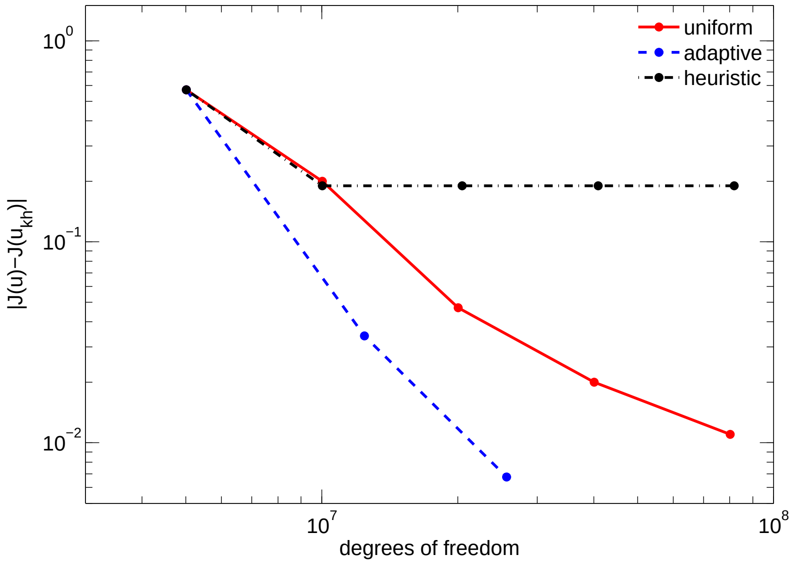
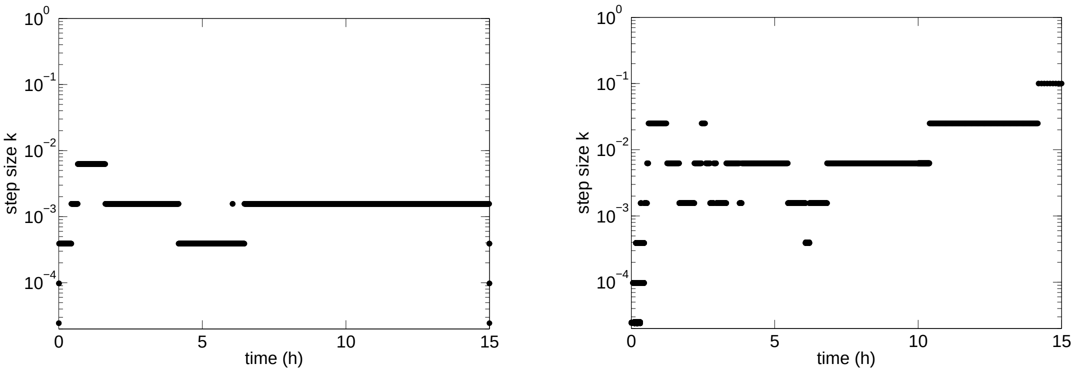


**Figure 5.6:** Comparison of the error $|J(w) - J(w_{kh})|$ in the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$ for different temporal mesh refinement strategies



(a) Adaptive

(b) Heuristic

**Figure 5.7:** Temporal discretizations after six refinement cycles used for the computations of the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

### Global quantity of interest $J = J_R$

We begin the evaluation of the error estimator of the 2D model problem with the diffusion coefficient $\mu = 36{,}000\ \mu m^2/h$ and the sum over the receptors of all biological cells as quantity of interest $J = J_R$. We check for independent temporal and spatial error estimates by observing the behavior of the temporal error estimate $\eta_k$ under adaptive spatial mesh refinement. The results are presented in Table 5.2.

**Table 5.2:** Invariance of $\eta_k$ under adaptive temporal mesh refinement and effectivity of the temporal error estimator $\eta_h$ for the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

| $M$ | $N^p$ | $\eta_k$ | $\eta_h$ | $J(w) - J(w_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|
| 1500 | 165 | $9.50 \cdot 10^{-3}$ | $1.86 \cdot 10^{0}$ | $1.92 \cdot 10^{0}$ | 0.97 |
| 1500 | 336 | $9.45 \cdot 10^{-3}$ | $7.79 \cdot 10^{-1}$ | $7.71 \cdot 10^{-1}$ | 1.02 |
| 1500 | 740 | $9.40 \cdot 10^{-3}$ | $4.11 \cdot 10^{-1}$ | $4.05 \cdot 10^{-1}$ | 1.04 |
| 1500 | 1749 | $9.38 \cdot 10^{-3}$ | $2.08 \cdot 10^{-1}$ | $1.88 \cdot 10^{-1}$ | 1.15 |
| 1500 | 4050 | $9.37 \cdot 10^{-3}$ | $8.83 \cdot 10^{-2}$ | $8.50 \cdot 10^{-2}$ | 1.15 |
| 1500 | 9711 | $9.37 \cdot 10^{-3}$ | $3.93 \cdot 10^{-2}$ | $4.29 \cdot 10^{-2}$ | 1.13 |
| 1500 | 23597 | $9.37 \cdot 10^{-3}$ | $1.65 \cdot 10^{-2}$ | $2.43 \cdot 10^{-2}$ | 1.06 |

We consider a fine temporal discretization with $M = 15000$ to ensure that the spatial error is dominant. While the spatial error is reduced in each step of the adaptive mesh refinement cycle, we observe that the temporal error converges to a fixed value. The spatial error indicators estimate the 'exact' error closely. Since the exact solution $w$ is not known and was only numerically approximated, the effectivity of the error estimator is only approximately evaluated.

Next, we compare the efficiency of adaptively refined spatial grids to heuristic error estimates (see Section 5.5.1) and uniform mesh refinement. We show the spatial discretizations after six refinement cycles based on dually weighted error estimates and simple heuristic methods in Figure 5.8. The goal functional $J_R$ causes a spatial mesh refinement near the surface of the cells. Heuristic error indicators give rise to a similar discretization since the solution expresses large gradients in the same area.

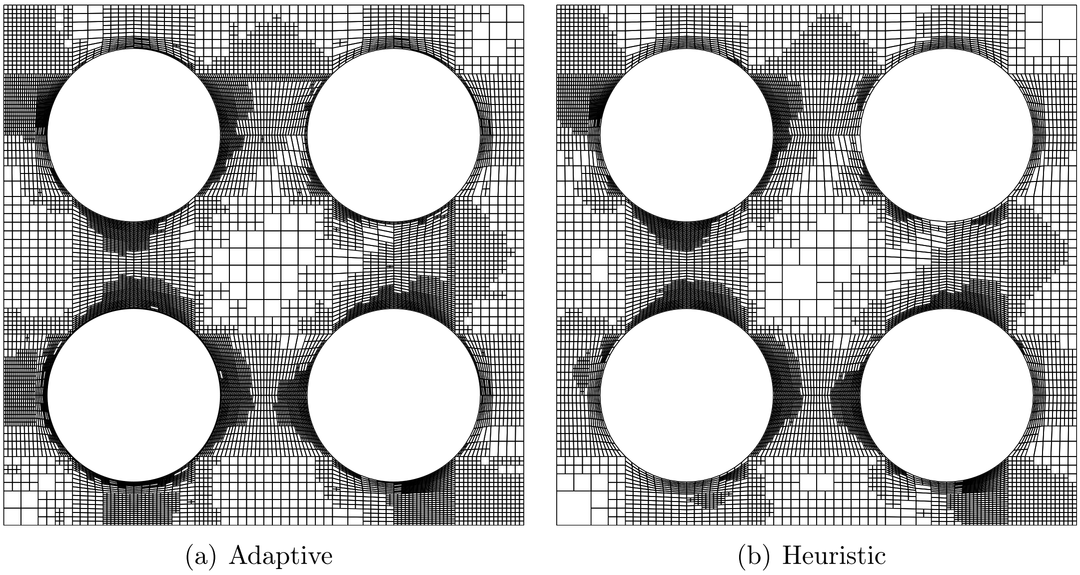

(a) Adaptive (b) Heuristic

**Figure 5.8:** Spatial discretizations after seven refinement cycles of the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

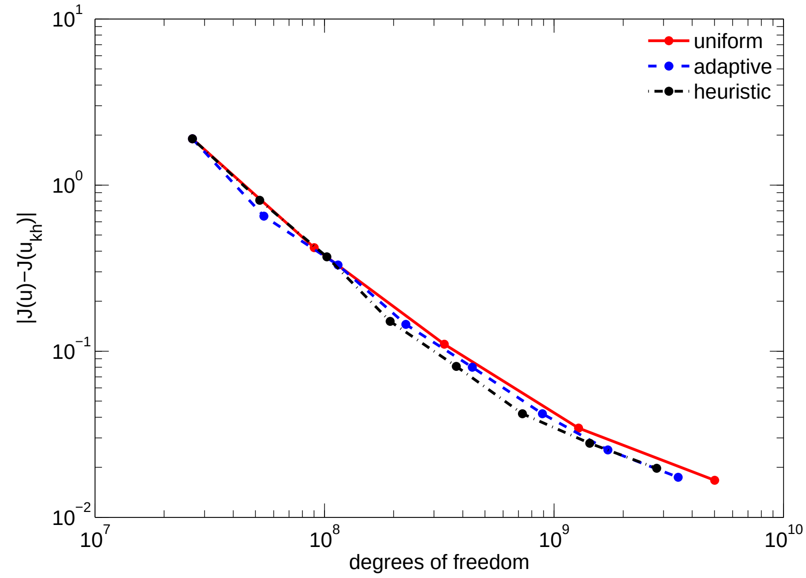


**Figure 5.9:** Comparison of the error $|J(w) - J(w_{kh})|$ in the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$ for different spatial mesh refinement strategies

A comparison of the error reduction by mesh refinement is shown in Figure 5.9. We see that adaptive temporal refinement by the DWR method, heuristic refinements and the uniform refinement all lead to similar error reductions. We conclude that adaptive discretizations yield no significant advantage over globally refined discretizations for coupled PDE/ODE systems with a large diffusion coefficient and a non spatially localized quantity of interest.

**Remark 5.2.** *The DWR method senses the high sensitivities of the goal functional* $J_R$ *to perturbations of the solution on the whole spatial domain for model problem with high diffusivity. These strong sensitivities result in a mesh refinement, which depends mainly on the local residuals and not on the dual weights. Thus, the mesh obtained by dually weighted error indicators near coincides with the mesh obtained by heuristic error indicators. Since most parts of the mesh are refined by both indicators, a uniform mesh refinement shows no significant drawback.*

### Localized quantity of interest $J = J_I$

In this section, we investigate the efficiency of adaptive spatial discretizations derived by the DWR method for the localized quantity of interest $J_I$. We consider the 2D model problem with both a diffusion coefficient of $\mu = 36{,}000\ \mu m^2/h$ and $\mu = 1\ \mu m^2/h$ to evaluate the spatial error estimator for a range of diffusivity.

We approximately compute the exact solution of the goal functional by $J(w) \approx 0.03286$ for $\mu = 36{,}000\ \mu m^2/h$ and $J(u) \approx 0.03466$ for $\mu = 1 \mu m^2/h$ to compare the computed error indicators to the 'exact' error. To measure mainly the spatial error indicators, we refine the temporal discretization sufficiently by $M = 1500$. We display the results for $\mu = 1\ \mu m^2/h$ in Table 5.3 and observe reliable error indicators after one refinement cycle and efficiency indicators $I_{eff}$ which converge to one.

We compare the spatial discretizations for a high diffusion coefficient of $\mu = 36{,}000\ \mu m^2/h$ after six refinement cycles in Figure 5.10. The heuristic error indicators do not result in a refinement of the spatial discretization near area $A \subset \Omega$, in contrast to the error indicators obtained by the DWR method. Even though the dually weighted error indicators lead to a high refinement of the spatial mesh of area $A \subset \Omega$, a comparison of the error reduction rates in Figure 5.12 yields that all mesh refinement strategies result in a similar error reductions. Thus, uniform spatial mesh refinement is for high diffusivity the most efficient refinement strategy, even for the localized goal functional $J_I$.

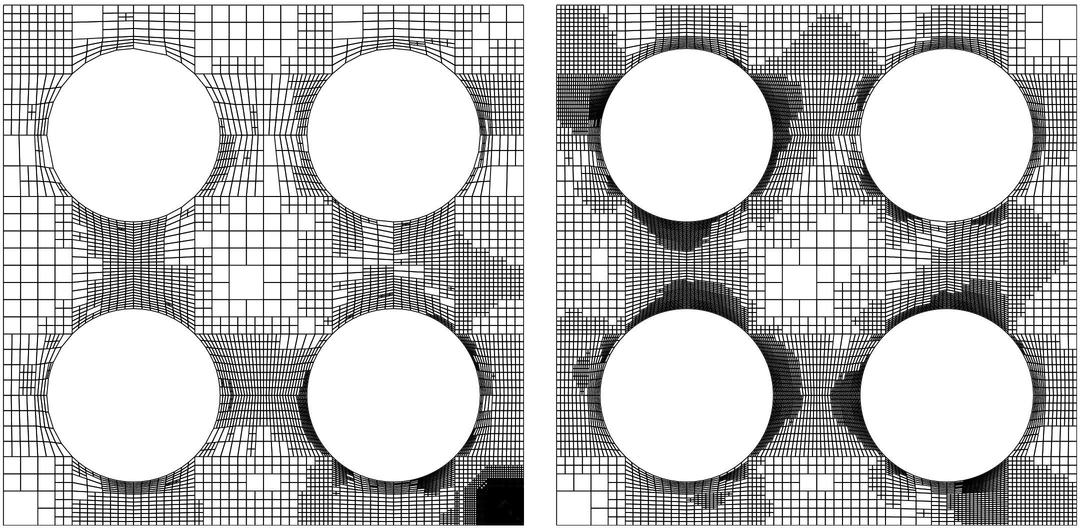
**Table 5.3:** Effectivity of the temporal error estimator $\eta_h$ for the 2D model problem with $\mu = 1\ \mu m^2/h$ and $J = J_I$

| $M$ | $N^p$ | $\eta_k$ | $\eta_h$ | $J(u) - J(u_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|
| 1500 | 165 | $1.57 \cdot 10^{-5}$ | $4.60 \cdot 10^{-1}$ | $3.98 \cdot 10^{-2}$ | 11.55 |
| 1500 | 322 | $6.18 \cdot 10^{-6}$ | $2.82 \cdot 10^{-2}$ | $1.79 \cdot 10^{-2}$ | 1.58 |
| 1500 | 620 | $1.37 \cdot 10^{-5}$ | $6.25 \cdot 10^{-3}$ | $4.57 \cdot 10^{-3}$ | 1.37 |
| 1500 | 1228 | $1.67 \cdot 10^{-5}$ | $1.69 \cdot 10^{-3}$ | $1.26 \cdot 10^{-3}$ | 1.36 |
| 1500 | 2406 | $1.75 \cdot 10^{-5}$ | $4.77 \cdot 10^{-4}$ | $3.97 \cdot 10^{-4}$ | 1.25 |
| 1500 | 4744 | $1.83 \cdot 10^{-5}$ | $2.50 \cdot 10^{-4}$ | $2.20 \cdot 10^{-4}$ | 1.22 |
| 1500 | 9553 | $1.95 \cdot 10^{-5}$ | $1.24 \cdot 10^{-4}$ | $1.27 \cdot 10^{-4}$ | 1.13 |

Spatial discretizations for a low diffusion coefficient of $\mu = 1\ \mu m^2/h$ show a more localized refinement pattern, see Figure 5.11. Only the error indicators obtained by the DWR method refine the spatial mesh of area $A \subset \Omega$ sufficiently. A comparison of the error reduction rates of the spatial mesh refinement strategies in Figure 5.13 displays a strong advantage of the DWR method over heuristic error indicators. The lack of refinement of the spatial mesh near area $A$ by the heuristic error indicators reduces the error reduction rate of the corresponding strategy greatly. Even a uniform mesh refinement is more favorable than the heuristic scheme.

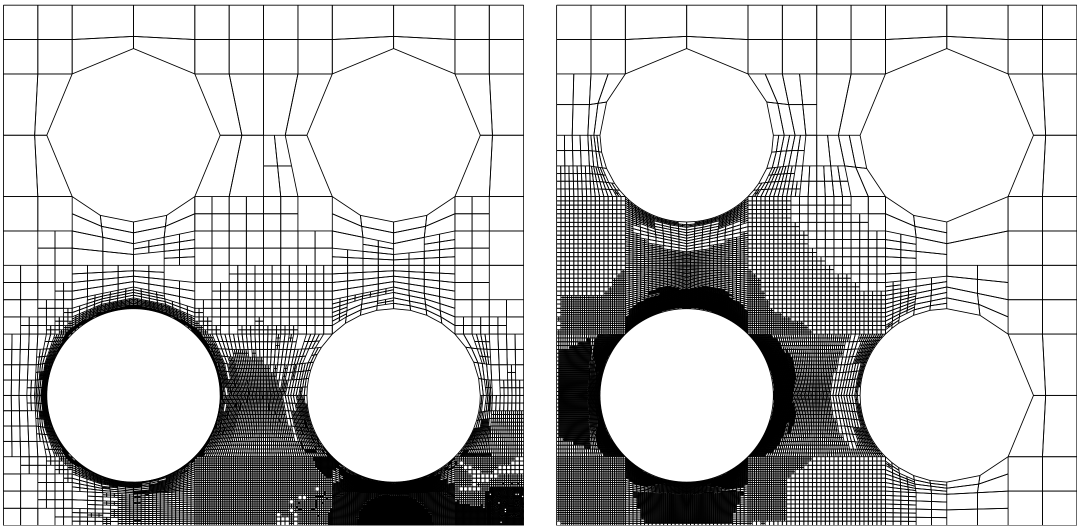
High and low diffusion coefficients distinctively influence the effectivity of spatial discretizations by the DWR method. The extra effort for computing the dual solution can be justified only for low diffusion coefficients by a significantly increased error reduction rate. In contrast, problems with high diffusivity are effectively solved by uniform mesh refinement. Even spatially localized goal functionals demand a fine spatial discretization of the whole domain.

**Remark 5.3.** *The results in this section display the difference of the sensitivities of the goal functional $J_I$ for low and high diffusivity of the 2D model problem. While the heuristic error indicator indicate mesh refinement near the gradients of the solution, the DWR method senses the high sensitivities of the localized goal functional to perturbations of the solution in area $A$. The solution in area $A$, for the model problem with high diffusivity, is dependent on a high accuracy of the whole intercellular area $\Omega$. Thus, the computational domain is near uniformly refined. In contrast, the sensitivities of the solution for the model problem with low diffusivity in area $A$ are highly localized. Heuristic methods instead indicate refinement near the gradients of the solution. Hence, the heuristic indicators are*



(a) Adaptive (b) Heuristic

**Figure 5.10:** Spatial discretizations after seven refinement cycles of the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_I$



(a) Adaptive (b) Heuristic

**Figure 5.11:** Spatial discretizations after seven refinement cycles of the 2D model problem with $\mu = 1\ \mu m^2/h$ and $J = J_I$

*not able to sense the high dependence of the goal functional $J_I$ on the area between A and the secreting cell in the lower left corner. Therefore, the DWR method provides local error indicators for an effective local mesh refinement.*

### 5.5.6 Space-time error balancing

In this section, we make use of both temporal and the spatial error indicators to obtain efficient space-time discretizations. We balance the temporal and spatial discretization errors, as described in Algorithm 5.1. Both two and three dimensional model problems are considered. We do not use different time meshes for the ODE part of the system and refer to the corresponding investigation in Section 5.5.7.

#### Two dimensional model problems

We focus on the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$ and compare the following mesh refinement strategies:

- 'uniform': global mesh refinement on the whole domain.
- 'adaptive in time': uniform spatial mesh refinement and adaptive refinement of the time mesh based on the temporal error indicators.
- 'adaptive in space': uniform temporal mesh refinement and adaptive refinement of the spatial mesh based on the spatial error indicators.
- 'fully adaptive': balancing of the spatial and temporal errors by Algorithm 5.1.

The error reduction rate of the different mesh refinement strategies are displayed in Figure 5.14. As already indicated by Section 5.5.5, uniform spatial mesh refinement is a reasonable strategy for the model problem with high diffusivity. A faster error reduction rate is obtained by adaptively chosen time discretizations.

We display the error estimations of the fully adaptive algorithm in Table 5.4. The spatial discretization is refined in each refinement cycle, due to the large spatial error indicators, while the temporal discretization is refined only twice. The error indicators for the adaptive discretizations both in space and in time closely approximate the exact errors obtained by a comparison of the computed functional value $J(w_{kh})$ to an approximated functional value $J(w) \approx 3013.0032$.

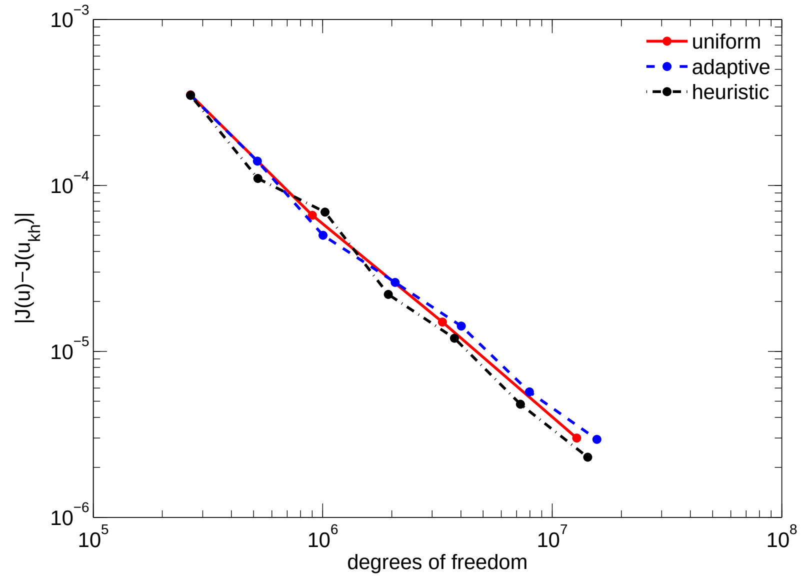


**Figure 5.12:** Comparison of different spatial mesh refinement strategies for the reduction of the error $|J(w) - J(w_{kh})|$ in the 2D model problem for a high diffusion coefficient $\mu = 36,000\mu m^2/h$ and localized quantity of interest $J = J_I$

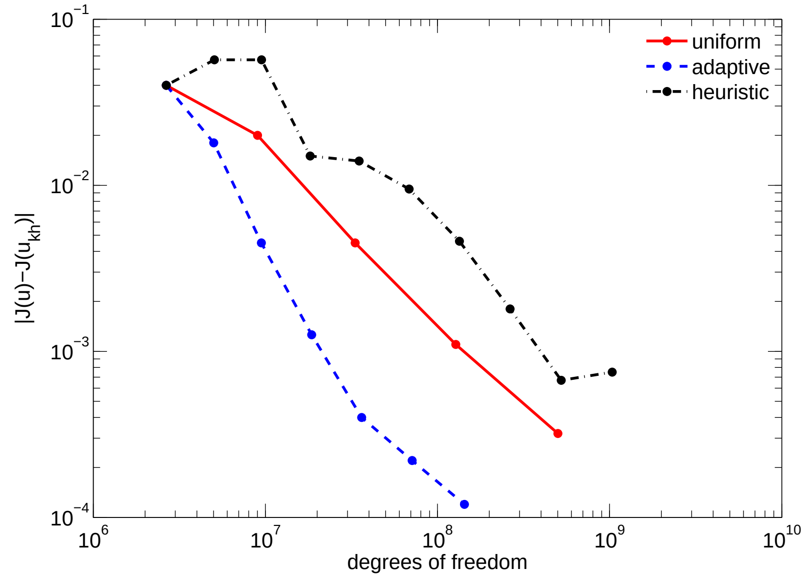


**Figure 5.13:** Comparison of different spatial mesh refinement strategies for the reduction of the error $|J(w) - J(w_{kh})|$ in the 2D model problem for a low diffusion coefficient $\mu = 1\mu m^2/h$ and localized quantity of interest $J = J_I$

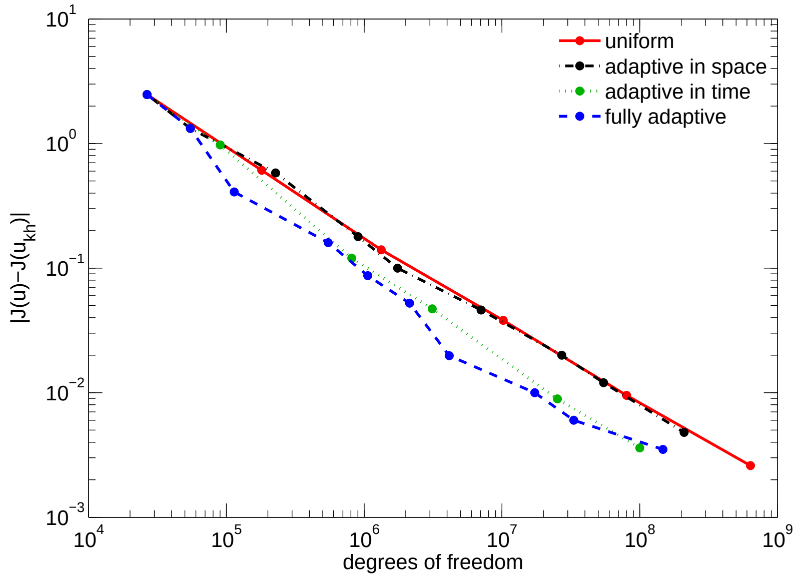


**Figure 5.14:** Comparison of different space-time mesh refinement strategies for the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

**Table 5.4:** Adaptive mesh refinement for the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

| $M$ | $N^p$ | $\eta_k$ | $\eta_h$ | $J(w) - J(w_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|
| 150 | 165 | $4.50 \cdot 10^{-1}$ | $1.78 \cdot 10^{0}$ | $2.47 \cdot 10^{0}$ | 0.90 |
| 150 | 353 | $4.56 \cdot 10^{-1}$ | $6.68 \cdot 10^{-1}$ | $1.32 \cdot 10^{0}$ | 0.85 |
| 366 | 747 | $1.99 \cdot 10^{-2}$ | $3.44 \cdot 10^{-1}$ | $4.10 \cdot 10^{-1}$ | 0.89 |
| 366 | 1485 | $1.97 \cdot 10^{-2}$ | $1.53 \cdot 10^{-1}$ | $1.61 \cdot 10^{-1}$ | 1.07 |
| 366 | 2890 | $1.93 \cdot 10^{-2}$ | $7.37 \cdot 10^{-2}$ | $8.80 \cdot 10^{-2}$ | 1.06 |
| 366 | 5829 | $1.81 \cdot 10^{-2}$ | $3.34 \cdot 10^{-2}$ | $5.34 \cdot 10^{-2}$ | 0.96 |
| 762 | 11310 | $2.27 \cdot 10^{-3}$ | $1.72 \cdot 10^{-2}$ | $1.50 \cdot 10^{-2}$ | 1.30 |
| 762 | 22674 | $2.18 \cdot 10^{-3}$ | $8.18 \cdot 10^{-3}$ | $7.47 \cdot 10^{-3}$ | 1.39 |
| 762 | 43549 | $2.00 \cdot 10^{-3}$ | $4.01 \cdot 10^{-3}$ | $4.00 \cdot 10^{-3}$ | 1.50 |

### Three dimensional model problems

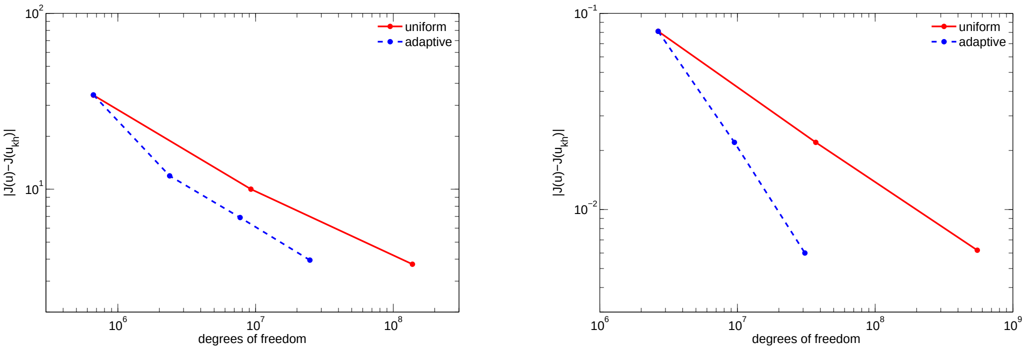
After testing several two dimensional problems, we focus now on the more realistic three dimensional model problems. Exemplarily, we display in Table 5.5 the effective error indicators for the model problem with diffusion coefficient $\mu = 36{,}000$ $\mu m^2/h$ and $J = J_R$. The functional value $J(w)$ of the exact solution has been approximated on a highly refined discretization by $J(w) = 4900.9$. Only the spatial discretization is refined, in all three refinement cycles, due to the small temporal discretization errors.

**Table 5.5:** Adaptive refinement for the 3D model problem with $\mu = 36{,}000$ $\mu m^2/h$ and $J = J_R$

| $M$ | $N^p$ | $\eta_k$ | $\eta_h$ | $J(w) - J(w_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|
| 300 | 2189 | 0.60 | 39.5 | 34.3 | 1.17 |
| 300 | 8497 | 0.58 | 15.3 | 11.9 | 1.33 |
| 300 | 31111 | 0.59 | 6.68 | 6.92 | 1.05 |
| 300 | 120555 | 0.58 | 3.32 | 3.95 | 0.99 |

We display the error reduction rates of the adaptive mesh refinement strategy in Figure 5.15(a). The study of two dimensional model problems already showed that for high diffusion coefficients, adaptive discretization do not improve the accuracy compared to uniform mesh refinement. The advantage of the adaptive mesh refinement strategy relies here on accurately estimating the small temporal discretization error. Therefore, it indicates correctly that precise computations are possible on a coarse time mesh.

These results are contrasted in Figure 5.15(b) by the error reduction rates of the model problem with low diffusion coefficient $\mu = 1$ $\mu m^2/h$ and localized quantity of interest $J = J_I$. The 'true' functional value $J(w) \approx 0.0427$ has been approximately computed on a fine discretization. To obtain an error reduction of factor 10, the use of uniformly refined grids increases the number of degrees of freedom 25-fold, compared to discretizations created by two adaptive refinement cycles. We conclude, for models with slow diffusion, that the adaptive mesh refinement strategy based on the DWR method is even more effective in three dimensions than in the two dimensional results in Section 5.5.5.



(a) $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

(b) $\mu = 1\ \mu m^2/h$ and $J = J_I$

**Figure 5.15:** Comparison of the error $|J(w) - J(w_{kh})|$ for two 3D model problems and different mesh refinement strategies

## 5.5.7 Different time meshes for the PDE and the ODE part

In this section we present the effectivity of error estimations for different time discretizations in the ODE and the PDE part. Moreover, we display the efficiency of a refined temporal ODE discretization compared to other approaches and discuss possible applications.

### Effectivity of the error indicators

To assess the effectivity of the error indicators obtained by the DWR method for different time meshes of the PDE and the ODE part, we devise a simple 2D model problem. We choose the same setup and parameters as in the last sections for $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$. To reduce the coupling between the PDE and the ODE part, we change the following parameters: $k_{on} = 1\ /\mathrm{nM}/h$, $T = 1\ h$ and $q = 1$ mol.$/h$, and thereby obtain dominating errors in the ODE part of the PDE/ODE-system. The 'exact' quantity of interest for the evaluation of the effectivity indices was computed as $J(w) \approx 442.96256$ on a fine discretization. We display the results in Table 5.6.

Firstly, we note the near exact error indicators $\eta_k^o$ of the ODE part of the equation. The parameters have been chosen such that the coupling of the ODE part to the PDE part is weak and the nonlinearities are reasonable. Thus, the resulting system poses a rather easy task for the error estimator $\eta_k^o$ of the dominating ODE time discretization error. We observe that $\eta_k^o$ is reduced iteratively by increasing the

**Table 5.6:** Effectivity of the temporal error estimator $\eta_k^o$ for the 2D model problem with $\mu = 36{,}000$ $\mu m^2/h$, $k_{on} = 1$ /nM/$h$, $T = 1$ $h$ and $q = 1$ mol./$h$ and $J = J_R$ for a maximum of 10 ODE time steps per PDE time step

| $M^p$ | $M^o$ | $N^p$ | $\eta_k^p$ | $\eta_k^o$ | $\eta_h$ | $J(w) - J(w_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|---|---|
| 10 | 10 | 165 | $5.03 \cdot 10^{-3}$ | $1.00 \cdot 10^{1}$ | $-9.28 \cdot 10^{-7}$ | $9.80 \cdot 10^{0}$ | 1.02 |
| 10 | 100 | 165 | $4.17 \cdot 10^{-3}$ | $1.01 \cdot 10^{0}$ | $-7.51 \cdot 10^{-7}$ | $1.01 \cdot 10^{0}$ | 1.00 |
| 22 | 220 | 165 | $3.45 \cdot 10^{-4}$ | $6.44 \cdot 10^{-1}$ | $-9.03 \cdot 10^{-7}$ | $6.43 \cdot 10^{-1}$ | 1.00 |
| 46 | 460 | 165 | $-2.60 \cdot 10^{-5}$ | $4.02 \cdot 10^{-2}$ | $-1.01 \cdot 10^{-6}$ | $4.02 \cdot 10^{-2}$ | 1.00 |
| 94 | 940 | 165 | $-4.14 \cdot 10^{-6}$ | $2.47 \cdot 10^{-3}$ | $-1.03 \cdot 10^{-6}$ | $2.46 \cdot 10^{-3}$ | 1.00 |

**Table 5.7:** Effectivity of the temporal error estimator $\eta_k^o$ for the 2D model problem with $\mu = 36{,}000$ $\mu m^2/h$, $k_{on} = 1$ /nM/$h$, $T = 1$ $h$ and $q = 1$ mol./$h$ and $J = J_R$ for a maximum of 100 ODE time steps per PDE time step

| $M^p$ | $M^o$ | $N^p$ | $\eta_k^p$ | $\eta_k^o$ | $\eta_h$ | $J(w) - J(w_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|---|---|
| 10 | 10 | 165 | $5.03 \cdot 10^{-3}$ | $1.00 \cdot 10^{1}$ | $-9.28 \cdot 10^{-7}$ | $9.80 \cdot 10^{0}$ | 1.02 |
| 10 | 1000 | 165 | $4.10 \cdot 10^{-3}$ | $1.04 \cdot 10^{-1}$ | $-8.12 \cdot 10^{-7}$ | $1.05 \cdot 10^{-1}$ | 1.01 |
| 22 | 2200 | 165 | $3.19 \cdot 10^{-4}$ | $6.46 \cdot 10^{-2}$ | $-9.88 \cdot 10^{-7}$ | $6.49 \cdot 10^{-2}$ | 1.00 |
| 46 | 4600 | 165 | $-2.57 \cdot 10^{-5}$ | $4.03 \cdot 10^{-3}$ | $-1.10 \cdot 10^{-6}$ | $4.00 \cdot 10^{-3}$ | 1.00 |
| 94 | 9400 | 165 | $-4.97 \cdot 10^{-6}$ | $2.47 \cdot 10^{-4}$ | $-1.13 \cdot 10^{-6}$ | $2.46 \cdot 10^{-4}$ | 1.00 |

number of ODE time steps $M^o$. A temporal discretization of ten ODE time steps for each PDE time step reduces the error by the same factor. The maximum number of ODE time steps for each PDE time step has been selected as $M^s = 10$ on all time intervals $I_m$. We call to mind the discussion in Section 5.4 in case of a coupled solving scheme: Time discretizations with a much larger refinement level in the ODE part make a carefully chosen linear solver necessary. To demonstrate the reliable performance of the error indicators even for time discretizations of widely different refinement levels, we display the efficiencies for a maximum number of 100 ODE refinement steps per PDE time step in Table 5.7.

### Efficiency of time meshes for the PDE and the ODE part

Next, we evaluate the efficiency of time meshes for the PDE and the ODE part for the biologically motivated 2D-problem with $\mu = 36{,}000\ \mu m^2/h$. We apply the full adaptive mesh refinement, see Algorithm 5.1, and check in each refinement cycle if the ratio between PDE and ODE steps needs to be adapted. The results of the refinement algorithm are displayed in Table 5.8. Steeper error reduction rates are observed in Table 5.4 than for the previously used refinement algorithm without different time meshes for the PDE and the ODE part.

**Table 5.8:** Adaptive algorithm for the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$

| $M^p$ | $M^o$ | $N^p$ | $\eta_k^p$ | $\eta_k^o$ | $\eta_h$ | $J(u) - J(u_{kh})$ | $I_{eff}$ |
|---|---|---|---|---|---|---|---|
| 150 | 150 | 165 | $-4.78 \cdot 10^{-2}$ | $4.98 \cdot 10^{-1}$ | $1.78 \cdot 10^{0}$ | $2.47 \cdot 10^{0}$ | 0.90 |
| 150 | 1500 | 165 | $-6.03 \cdot 10^{-2}$ | $2.12 \cdot 10^{-2}$ | $1.74 \cdot 10^{0}$ | $1.98 \cdot 10^{0}$ | 0.86 |
| 150 | 1500 | 351 | $-5.56 \cdot 10^{-2}$ | $2.09 \cdot 10^{-2}$ | $6.01 \cdot 10^{-1}$ | $8.02 \cdot 10^{-1}$ | 0.70 |
| 150 | 1500 | 746 | $-5.39 \cdot 10^{-2}$ | $2.08 \cdot 10^{-2}$ | $3.00 \cdot 10^{-1}$ | $3.93 \cdot 10^{-1}$ | 0.67 |
| 150 | 1500 | 1478 | $-5.30 \cdot 10^{-2}$ | $2.08 \cdot 10^{-2}$ | $1.26 \cdot 10^{-1}$ | $2.00 \cdot 10^{-1}$ | 0.46 |
| 150 | 1500 | 2919 | $-5.24 \cdot 10^{-2}$ | $2.08 \cdot 10^{-2}$ | $6.74 \cdot 10^{-2}$ | $1.37 \cdot 10^{-1}$ | 0.26 |
| 366 | 3660 | 5851 | $-1.21 \cdot 10^{-2}$ | $4.47 \cdot 10^{-3}$ | $3.32 \cdot 10^{-2}$ | $1.68 \cdot 10^{-2}$ | 1.52 |
| 366 | 3660 | 11287 | $-1.20 \cdot 10^{-2}$ | $4.47 \cdot 10^{-3}$ | $1.64 \cdot 10^{-2}$ | $2.31 \cdot 10^{-3}$ | 3.82 |

We observe that the temporal error of the PDE part of the equation $\eta_k^p$ is by a factor ten smaller than the error of the ODE part of the equation $\eta_k^o$. Therefore, the adaptive algorithm refines the temporal discretization such that ten ODE time steps are computed for each PDE time step. The temporal error is even reduced by a factor of more than 20, which diminishes the temporal error for the following iterations. Consequently only the spatial error has to be reduced by spatial mesh refinement and a coarser temporal discretization suffices. In many iteration cycles, the error indicators for the temporal and the spatial discretization nearly cancel each other out. Thus, the 'true' error is underestimated by the sum of the error indicators, even though a fast error reduction is achieved.

We compare the error reduction rates of this approach to the earlier results in Section 5.5.6 without different time meshes for the PDE and the ODE part in Figure 5.16. We note that the same error reduction is achieved in both algorithms, in the first refinement cycles, because the spatial discretization error dominates the temporal error. In the later cycles, when the temporal discretization becomes relevant, the customized temporal discretization for the ODE and the PDE part enhances the error reduction rate.

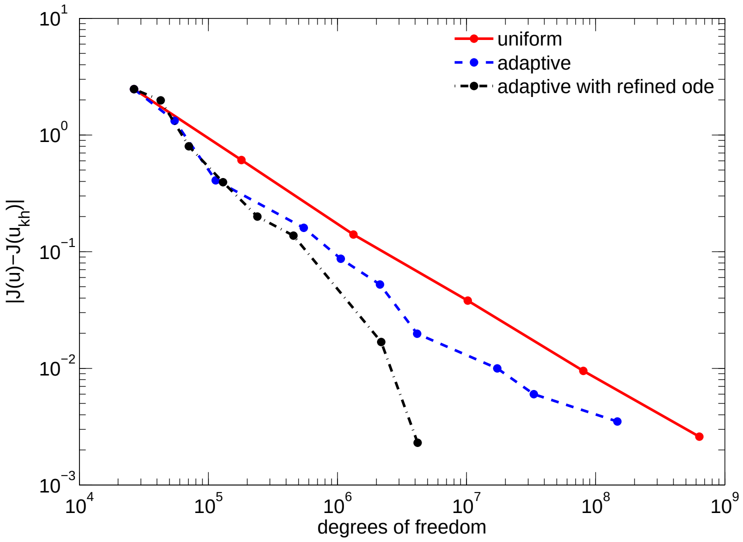


**Figure 5.16:** Comparison of the error $|J(w) - J(w_{kh})|$ in the 2D model problem with $\mu = 36{,}000\ \mu m^2/h$ and $J = J_R$ for mesh refinement strategies with and without different time meshes for the PDE and the ODE part

We conclude that different time discretizations for the two parts of a system of equations can greatly reduce the computational effort. The error indicators obtained by the DWR method prove to be reliable even in nonlinear strongly coupled biological applications and can be put to use in an adaptive mesh refinement strategy. A large error in the less computational expensive part of the system is efficiently reduced by only refining the necessary part of the temporal discretization.

**Remark 5.4.** *We remark that the error reduction by the different time meshes of the PDE and the ODE part comes with little additional costs. The refined time mesh of the ODE part adds additional degrees of freedom to the ODE part of the solution vector in a coupled solution scheme, see Section 3.1.3. Hence, the total number of degrees of freedom is only slightly enlarged. Different time meshes for the two parts of a coupled system reduce the error significantly if the computationally less expensive part of the system is responsible for a larger part of the time discretization error. Specifically tailored time meshes reduce the error* *most efficiently for systems in which the outcome of the goal functional is dependent on one part of the system and the two parts are only weakly coupled to each other.*

# 6 Application: Three-dimensional Gradients of Cytokine Signaling between T cells

In this chapter, we apply the developed numerical methods and simulate large clusters of immune cells. The numerical results provide a better quantitative understanding of the molecular mechanism of the immune response.

We explain in Section 6.1 the role of T cells as part of the adaptive immune response. In Section 6.2 we include the immunological synapse, which has an important function for cytokine signaling, in the presented model of Section 2.3. We conclude the chapter with numerical results in Section 6.3. We investigate the dependence of cytokine signaling between T helper cells on the presence of regulatory T cells. Further, we analyze of the signal range of T helper cells in large-scale simulations with over 2000 cells.

The results in this chapter have been accepted for publication [65].

## 6.1 IL-2 signaling of T cells as part of the adaptive immune response

The molecular mechanism of the immune response is the subject of intense research in medicine and biology. The adaptive immune response relies on selective expansion of clones of T and B cells. Cell-to-cell communication is a defining property of multicellular organisms. In particular, the release, sensing and uptake of cytokines, small signaling proteins, by cells is essential for the regulation of the mammalian immune system [55]. High levels of cytokines, so-called 'cytokine storms', can cause severe damage to the human body. Adaptive immune responses must be rapid and effective in the case of strong infection, but also carefully controlled to avoid autoimmune diseases. Cytokine concentrations regulate critical processes such as the type and strength of the immune response. Quantitative understanding of such cytokine-driven cellular decisions is beginning to emerge [13, 24, 27, 60, 66], yet the underlying spatio-temporal cytokine dynamics remain poorly understood. Cytokines act in heterogeneous environments, typically with high cell-densities. It is not known how they diffuse under such conditions and in turn regulate immune responses.

An important cytokine with corresponding receptor expression is interleukin(IL)-2. It is secreted by T helper (Th) cells early after antigenic stimulation and taken up by high-affinity IL-2 receptors (IL-2R) on Th cells and regulatory T (Treg) cells. Proliferation and differentiation of Th cells are tightly regulated by the interplay of regulatory T cells and the Th cells itself. Experimentally tested models for the IL-2R dynamics are available [13, 24, 66]. Th cells recognize an antigen of an antigen-presenting cell (APC) and then activate responding T cells as part of the immune response. Regulatory T cells instead suppress the activation of the responding T cells. Understanding the role of regulatory T cells is important for the treatment of auto immune diseases and cancer. The success of organ transplantations and cancer immunotherapy is directly linked to the suppressing activity of regulatory T cells.

A combined experimental and theoretical study [13] has identified secretion and uptake of IL-2 as a possible mechanism mediating immune suppression by regulatory T cells. The model captures IL-2 signaling in the initial phase after antigen stimulation where the cells are primed for proliferation but have not yet entered initiated cell division (up to 30 hours). The coupled PDE/ODE system is displayed in Section 2.3. Secreting Th cells represent cells which have recognized the antigen of an APC before the start of the simulations. APC recognition is not part of the model, instead the spatial positions of the secreting Th cells are chosen randomly. APCs themselves do not express IL-2 receptors but constitute simply ‘excluded volumes’ with respect to the IL-2 dynamics. We do not consider the APC in the extracellular area to focus on the role of IL-2 uptake by T cells by a simplified spatial discretization. Regulatory T cells are distinguished from responding T cells by higher receptor expression rates $w_i^0$ and $w_i^1$ (Table 6.1).

**Table 6.1:** Biological parameters for regulatory T cells, see [13]

| Symbol | Value | Parameter |
|---|---|---|
| $w_i^0$ | 1000 mol./cell/$h$ | Antigen stimulated IL-2 receptor expression rate |
| $w_i^1$ | 8000 mol./cell/$h$ | Feedback induced IL-2 receptor expression rate |
| $r$ | $5\mu m$ | Cell radius |

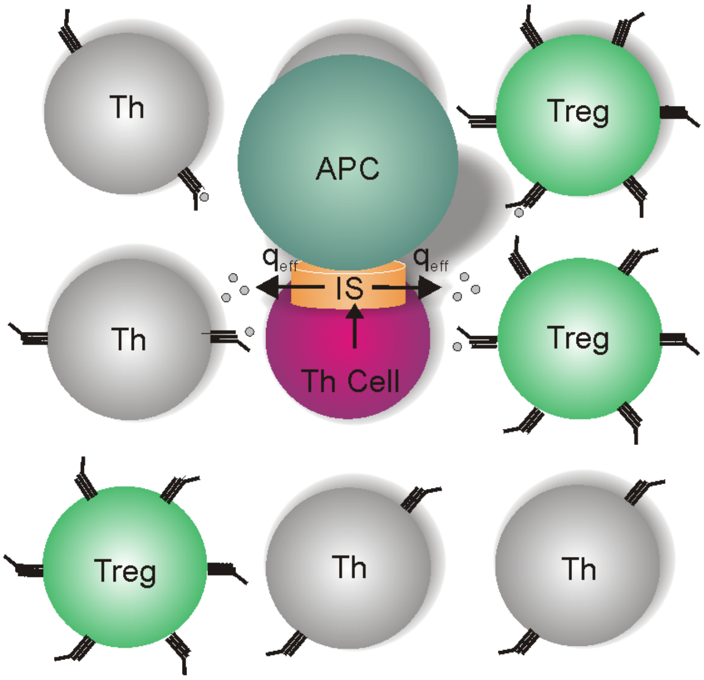


**Figure 6.1:** Model scheme: activated Th cells release cytokine molecules into the immunological synapse (IS).

Upon receiving an antigen stimulus, only about one quarter of a Th cell population releases IL-2 molecules. The activation of the other T cells is decided by their competition for IL-2 with the regulatory T cells, which thus inhibit activation by absorbing IL-2. Th cells release IL-2 into the extracellular medium and regulatory T cells express high affinity IL-2 receptors. The responding Th cells express high levels of IL-2 receptors only if activated. As index of activation, we therefore consider the number of receptors $R_i + C_i$ after the time necessary to reach a stationary condition. The competition between T cells is significantly influenced by their spatial position relative to IL-2 sources. Thus, the secreting Th cells have a great advantage regarding activation compared to responding Th cells, since they emit IL-2 with a secretion rate $q_i$.

## 6.2 Immunological synapse and model adaptation

The model for T cell interaction published by Busse et al. [13] (see Section 2.3) assumed homogeneous secretion of the cytokine over the cell surface. However, T

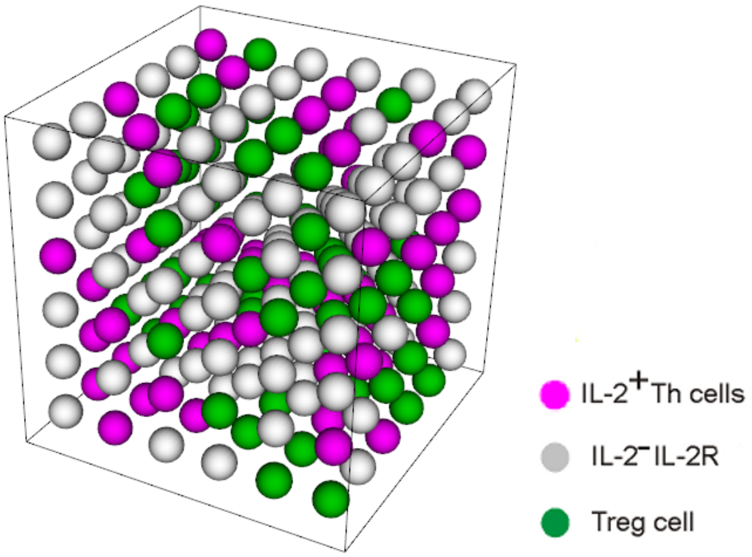
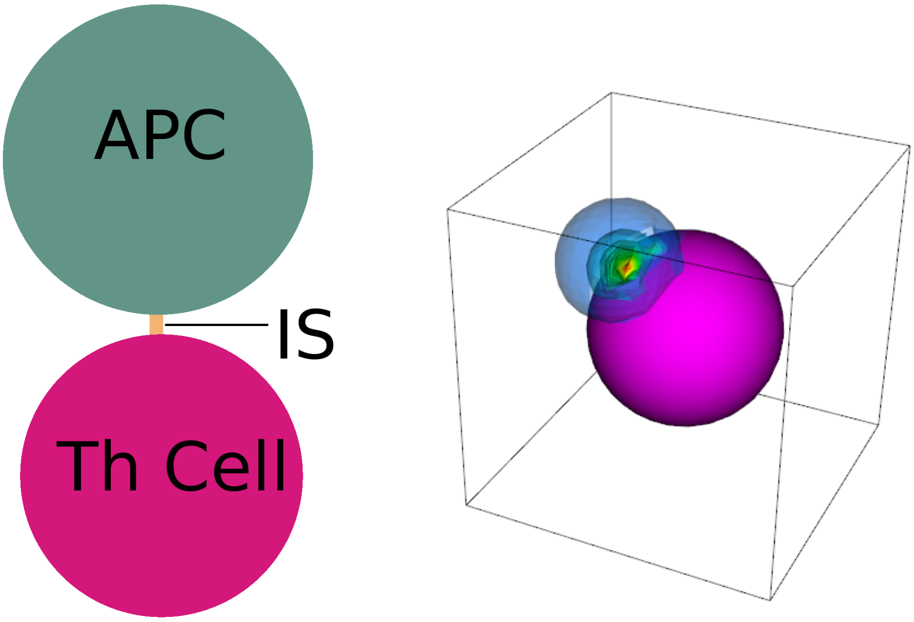


**Figure 6.2:** Setup of the 3D-simulation of a T cell population

cells release IL-2 and other cytokines in a polarized fashion into the immunological synapse. Immunological synapses are formed between immune cells by surface proteins after antigen recognition and have been observed between T cells and APC (Figure 6.1). Many cytokines are secreted preferentially into the immunological synapse and a range of cytokine receptors have been found to be specifically located in the immunological synapse, too. Therefore, it is likely that the synapse has an important function for cytokine signaling, beyond its role for T cell receptor signaling on which theoretical studies have focused, see Davis et al. [20].

We analyze a model of cytokine secretion and uptake in the immunological synapse, represented by a small cylindrical region between a Th cell and an opposed APC or second Th cell (Figure 6.3). The distance between cells is in the range of 10 to 40 $nm$. Not all cytokine molecules escape into the extracellular space ($q_{eff}$), potentially reaching other nearby cells. Cytokine molecules may also induce signals to the secreting cell by binding to receptors at the cell itself. With a synapse length of 20 $nm$ and a cell diameter of $10,000$ $nm$, it is not possible to accurately include the synapse in the 3D-grid. Instead, we implement the synapse by polarized secretion on a single grid point on the cell surface of a secreting cell which is randomly chosen, see Figure 6.3(b). Periodic boundary conditions are considered on the outer boundary $\Gamma_{out}$ instead of free boundary conditions. By this addition, we avoid artificial behaviors in the corners of the intercellular area $\Omega$ to have a more biological depiction of large cell clusters.



(a) Model (b) Discretization

**Figure 6.3:** Cytokine molecules are released into the immunological synapse (IS) between a Th cell and an opposed antigen presenting cell (APC). The synapse is discretized by polarized secretion, the opposed cell is omitted in the discretization.

In a further addition to the model described in Section 2.3, only the activation of non-secreting Th cells and regulatory T cells is considered. Therefore the IL-2 secretion rate ($q$) of the Th cells upon antigen recognition is replaced by a smaller effective secretion rate ($q_{eff}$). The rest of the secreted interleukin is used for the activation of the cell itself (autocrine signaling). The intracellular processes of secreting Th cells are neglected, only activation of other cells is considered (paracrine signaling). In recent experiments by Höfer et al. [36], it was determined that one quarter of T cells release IL-2 molecules by antigen stimulation with a secretion rate of $q = 36,000$ molecules$/h$. Simulations of the immunological synapse, see Thurley et al. [65], estimated an effective secretion rate of $q_{eff} = 3,500$ molecules$/h$.

## 6.3 Numerical results

Earlier numerical computations of the competition of T cells for IL-2 were published by Busse et al. [13] in 2010. They performed two-dimensional simulations to analyze the dynamics and pattern formation of the intercellular signaling between up to 170 cells. We use the developed numerical methods in this thesis, consisting of the coupled solver developed in Chapter 4 and the adaptive refinement strategies of Chapter 5, for an efficient computation of the signaling between up to 2000 cells in realistic 3D environments. We present in this section results that show the important influence of 3D effects in cell interactions and bring new insights into the study of the immune response.

Graphs were visualized by Matlab [46]. Visualizations of the three dimensional data were created by Marcus Schaber and Lisa Kolb of the Visualization and Numerical Geometry Group of the Heidelberg University, see [41].

### 6.3.1 IL-2 signaling of T helper cells

It is unknown under which conditions the cytokine molecules that escape the synapse can transmit a paracrine signal. To investigate the origins and consequences of spatially inhomogeneous dynamics of cytokine signaling, we perform extensive three-dimensional simulations of a T cell population. We simulate the IL-2 dynamics for a large number of T cells (216 cells in a volume of 1 nl). After antigenic stimulation in a fraction of T cells and subsequent IL-2 secretion, the IL-2 concentration increases rapidly and nearly homogeneously for several hours. Analysis of the time course (Figure 6.5) shows that IL-2R expression is upregulated in non-secreting T cells in response to the increased IL-2 concentration in the first hours after antigenic stimulation. High IL-2 receptor levels cause a fast IL-2 uptake from the medium and as a result concentration gradients occur. In large parts of the simulated region, the IL-2 concentration reaches a steady-state at around 10 pM while locally it is more than twice as large. This inhomogeneity in IL-2 concentration corresponds to receptor upregulation (activation) of non-secreting Th cells (IL-2$R^+$ cells with more than 4000 receptors). These cells are found near the regions with high IL-2 concentration.

However we see in Figure 6.5, as the IL-2R are being upregulated, IL-2 becomes increasingly depleted in the medium. As a result, only a fraction of the cells receive a sufficient IL-2 stimulus to sustain high IL-2R expression, whereas the remaining cells downregulate IL-2R expression (IL-2$R^-$ cells). Interestingly, the time courses

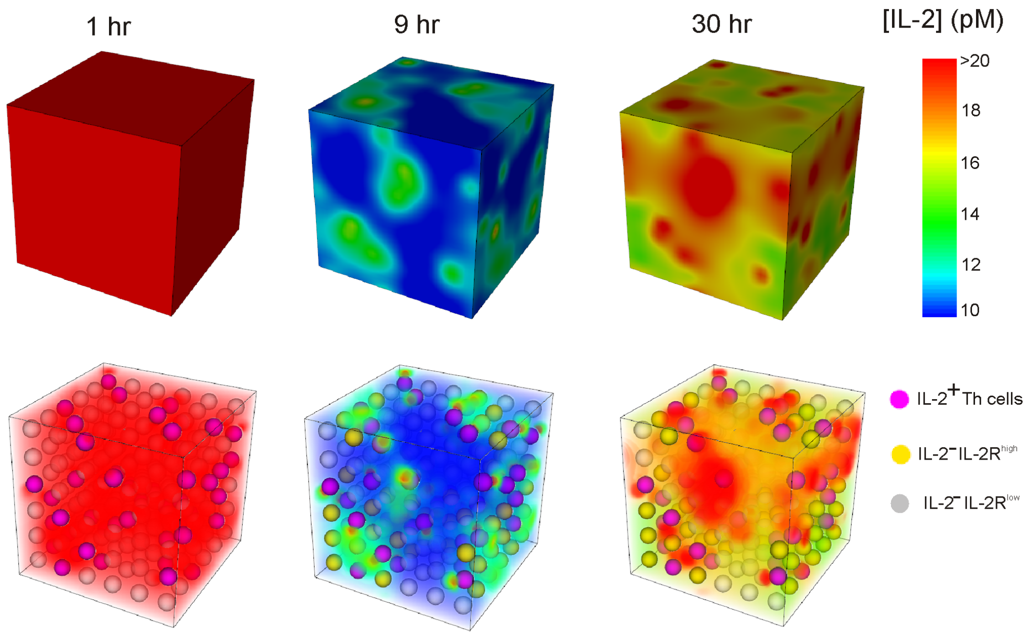


**Figure 6.4:** Spatiotemporal cytokine dynamics of an Th cell cluster showing the IL-2 concentration at indicated time points. 59 of the 162 non-secretory cells are activated after 30 hours.

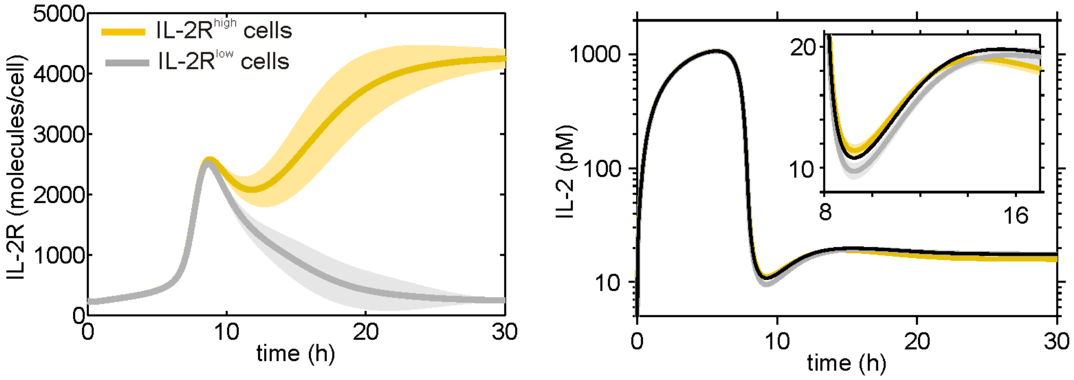


**Figure 6.5:** Time course of the receptor number and average IL-2 concentration at the surface. Solid lines indicate averages, blurred regions standard deviations. The black line indicates the average IL-2 concentration in the simulated region.

of IL-2 concentrations at the surfaces of the cells show only small differences between IL-2$R^+$ and IL-2$R^-$ cells: In the beginning, IL-2 equally rises near IL-2$R^+$ and IL-2$R^-$ cells, but as IL-2 depletion sets in, the cells that eventually become IL-2$R^-$ cells receive slightly less IL-2. Later at steady-state, the IL-2 concentration is somewhat higher in the microenvironment of IL-2$R^-$ cells, because they do not consume as many IL-2 molecules.

This form of local bistability, which occurs in the expression of IL-2R on Th cells, was observed already in Busse et al. [13]: Based on a quasi-stationary state assumption, they showed that in the model without regulatory T cells, the IL-2R expression rate responds to the increase of the secretion rate in a digital way and the cells are activated only after a certain threshold is exceeded. A small bistable region around the threshold is observed. These findings were supported by experimental data by Busse et al. Here we find that the extended model (see Section 6.2) qualitatively reproduces the predicted bistable system behavior in 3D simulations.

The simulations indicate the amount of IL-2 escaping from the immunological synapse can sustain paracrine signaling in at least a fraction of surrounding cells. However, competition for the cytokine can cause heterogeneity in the response of a cell population and result in IL-2 levels that are much lower than local concentration peaks. The paracrine cytokine signals are not only characterized by these cytokine gradients, but also by a rapid temporary cytokine boost occurring in the first hours after stimulation. Such a temporary cytokine signal has been observed by single-cell experiments by Sojka et al. [63] and recently by ELISA experiments [66], although with conflicting time-scales. Single cell experiments evoked a peak in the number of IL-2 secreting cells at one to six hours after antigen stimulation, while Tkach et al. [66] report a peak in the IL-2 concentration measured after 50 hours. Our simulations point to an IL-2 peak in the first 10 hours after stimulation, and thus support the earlier suggestion by Sojka et al. [63] that ELISA studies have limitations in reflecting the time-course of cytokine signals in the living organism.

### 6.3.2 Competitive uptake by regulatory T cells

Regulatory T cells mediate immune tolerance and are critical for the prevention of autoimmune reactions. They express high levels of IL-2R but do not secrete IL-2. To study the effect of regulatory T cells on the IL-2 dynamics after activation of conventional Th cells, we simulate a T-cell population consisting of antigen-stimulated IL-2 secreting and non-secreting Th cells as well as 33% regulatory T cells in a random configuration shown in Figure 6.2. Compared to the situation in the absence of regulatory T cells, the IL-2 concentration attains a spatially inhomogeneous steady state more rapidly, with the overall IL-2 concentration being lower (Figure 6.6 and 6.7). Importantly, the non-secreting Th cells do not permanently upregulate IL-2R in the presence of regulatory T cells because the regulatory T cells bind IL-2 themselves and thus suppress the paracrine IL-2 signal. The comparison with the simulations without regulatory T cells imply that Th cells require for sustained IL-2 signaling both a temporarily strong and a stable weak IL-2 stimulus. The insight that Th cells can sustain IL-2 signaling at low cytokine concentration, but only after initial stimulation with high cytokine concentration, is a spatio-temporal phenomenon similar to hysteresis: Active cells express more cytokine receptors, which bind more cytokine molecules even at lower concentration and thus stabilize the active state once it is achieved. Regulatory T cells can suppress prolonged IL-2 signaling in Th cells by inhibiting the strong initial IL-2 signal and the resulting upregulation of the IL-2 receptors.

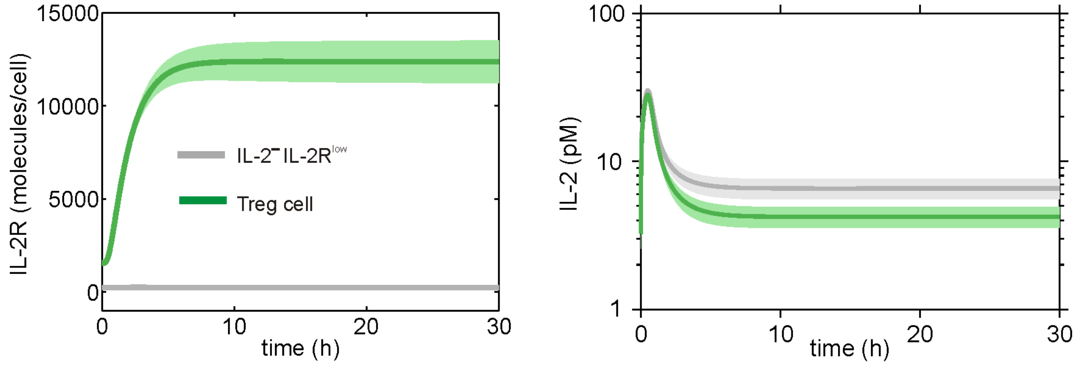


**Figure 6.6:** Time course of the receptor number and average IL-2 concentration at the surface, including regulatory T cells. Solid lines indicate averages, blurred regions standard deviations

Having established that an effective paracrine IL-2 signal is possible in our model, and that it can be suppressed by regulatory T cells, we analyze in Figure 6.8 to which extent IL-2 secretion rate, cell-to-cell distance and the fraction of IL-2 secreting cells shape the spatio-temporal dynamics. Without regulatory T cells, the number of activated Th cells increases linearly with the effective IL-2 secretion rate $q_{eff}$, until, eventually, all cells in the simulated region become active. By contrast, the presence of regulatory T cells creates a threshold at an effective secretion rate of $q_{eff} \approx 18,000$ molecules/h, below which there is no paracrine IL-2 signaling between Th cells. The same pattern is observed if we vary the fraction of cytokine secreting cells instead of the effective secretion rate. Hence the presence of regulatory T cells changes the paracrine IL-2 dynamics from a gradual to an all-or-none response: Either the paracrine signal is completely suppressed by competitive uptake or the suppression is overrun and all cells are activated. However, effective IL-2 secretion rates are around $3,500$ molecules/h, which is likely to be too small to overcome the suppression of the regulatory T cells in a physiological setting.

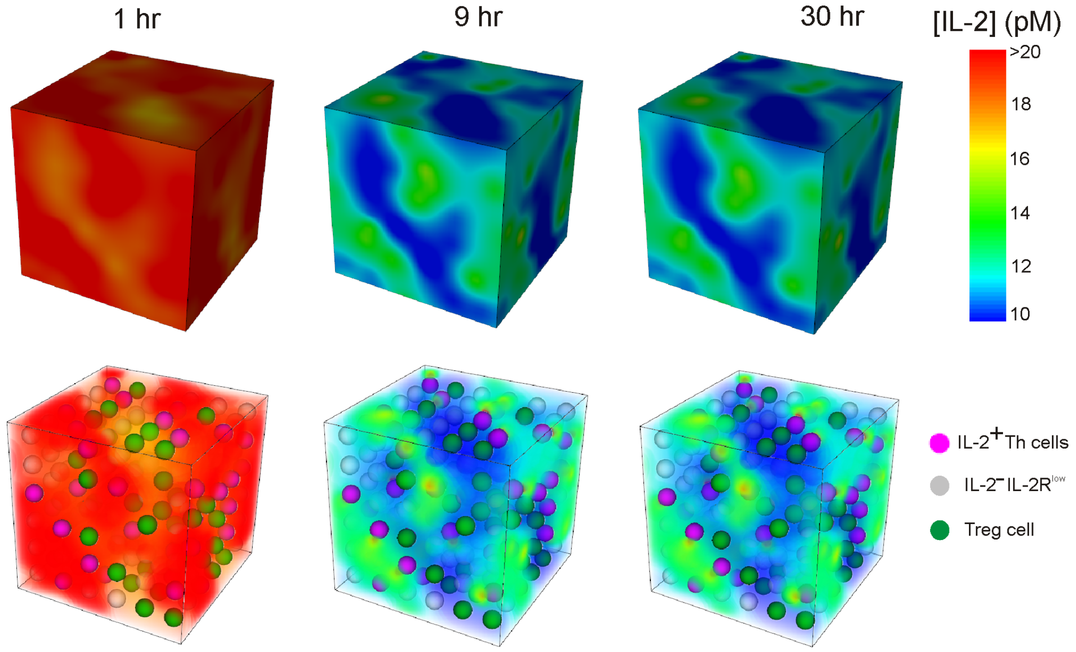


**Figure 6.7:** Spatiotemporal cytokine dynamics of a coculture of both T helper and regulatory T cells at indicated time points

Within the physiological range range from 2 to $20\mu m$, the cell-to-cell distance does not influence the number of Th cells that become activated by paracrine IL-2 signaling. This is the case because cytokine molecules can reach nearby cells rapidly by diffusion compared to the slower time scales of changes in IL-2R expression and IL-2 internalization. Therefore the exact cell-to-cell distance is unimportant for the number of activated Th cells, which is in agreement with recent experimental data [66].

Regulatory T cells efficiently suppress paracrine IL-2 signals, because they express

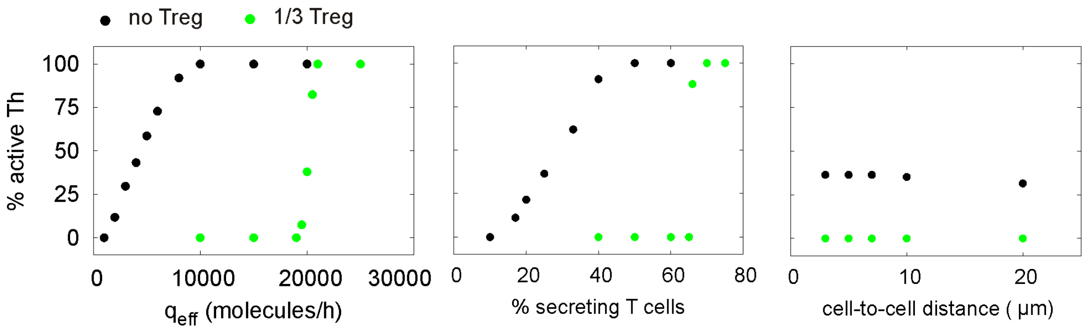


**Figure 6.8:** Fraction of activated Th cells (IL-2$R^+$ cells) after 30h simulation time with and without the suppression of regulatory T cells.

high levels of IL-2R, preventing a strong cytokine signal. The suppression of IL-2 signals is an important mechanism contributing to immune tolerance mediated by regulatory T cells. In line with earlier work by Feinerman et al. [24] and Busse et al. [13], this suggests that the IL-2 uptake by regulatory T cells avoids autoimmune diseases even in the case of strong infection.

### 6.3.3 Long-range signals of IL-2 secretor cells

Our simulations yielded global elevations in IL-2 concentration only temporarily before the target cells expressed high levels of IL-2R. Beyond this point, only short-range IL-2 gradients with local concentrations were observed. Generally, we expect that the balance between cytokine secretion, dilution through diffusion in the three-dimensional extracellular space and cellular consumption will determine the signaling range. To understand the interplay of these three factors, we simulated a single IL-2-secreting Th cell surrounded by non-secreting Th cell which all are potential responders to the IL-2 (Figure 6.9). Although we use the specific parameters for IL-2 here, this model is of more general interest to situations with few signaling cells and many responder cells (e.g., IL-4 secreting Th cells in a B cell population).

We simulated the signal range during Th cell activation for varying values of $q_{eff}$ with (Figure 6.10). We found that for the estimated effective secretion rates ($q_{eff} = 3,500$ molecules$/h$), high IL-2 concentrations are restricted to the microenvironment of the cytokine secreting cell. Remarkably, although secretion is polarized through the synapse, the cytokine concentration is higher along the entire surface of the secreting cell, including the pole opposite to the synapse, than at

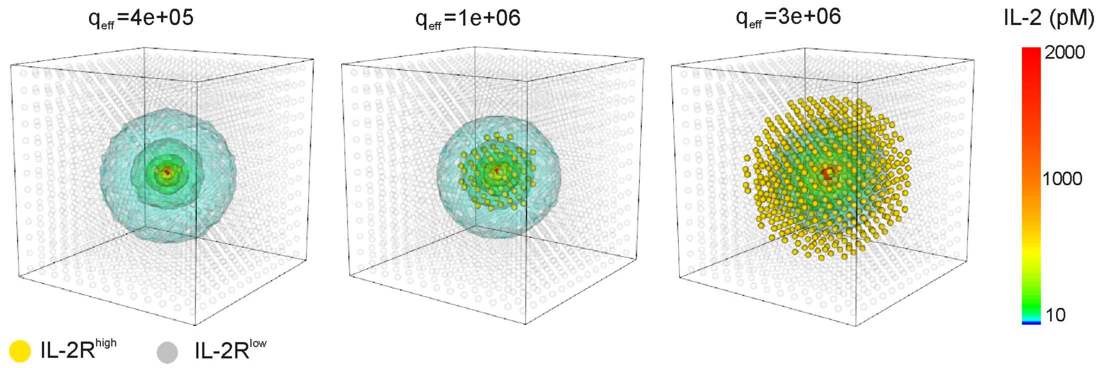


**Figure 6.9:** Large-scale simulation (2198 Th cells) with one IL-2 secreting Th cell placed in the center

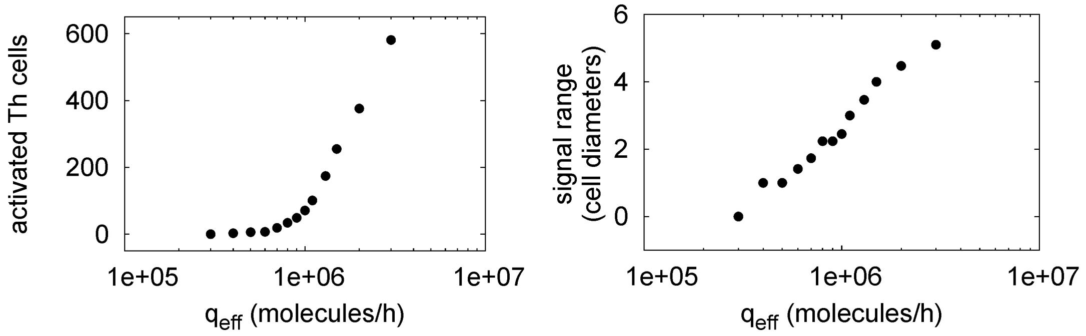


**Figure 6.10:** Simulations for varying values of $q_{eff}$ with the signal range determined by the distance from the center to the most distant activated cell, normalized by the cell-to-cell distance ($15\mu m$).

nearby cells. For larger secretion rates (of the order to $10^6$ molecules$/h$), the IL-2 signal reaches hundreds of cells. However, with the estimated IL-2 secretion rate, hundreds of secreting cells would be needed to realize such a high rate. Therefore, IL-2 from an individual producer will act locally whereas large clusters of activated cells could cause long-range signals. Very high secretion rates ($3 \cdot 10^6$ molecules$/h$) evoke an organ-wide cytokine signal. The spatial range scales linearly with the logarithm of the secretion rate. Thus, the cell population recognizes relative rather than absolute increases of the secreted cytokine molecules.

# 7 Conclusion and Outlook

In this thesis, we developed an adaptive numerical solver for accurate 3D-simulations of intercellular signaling modeled by coupled PDE/ODE systems. The proposed method uses Galerkin finite element discretizations in space and time. The temporal Galerkin space corresponds to the damped Crank-Nicolson method and was extended in this thesis for different time meshes of the PDE and the ODE part of the system.

Both decoupled and coupled solution schemes were considered for solving the linear subsystems arising in each iteration of Newton's method. We showed that the interaction between the PDE and the ODE part of the biological system is strong by a quantitative sensitivity analysis. A comparison with a decoupled scheme showed the better performance of the coupled scheme especially for systems with large eigenvalues of the sensitivity matrices, e.g. by large time steps. We compared the performance of different multigrid preconditioners for coupled solving schemes based on a Krylov-type solver. The number of necessary Krylov iterations was reduced by considering the coupling of the system in the smoothing iterations.

We derived an a posteriori error estimator by means of the DWR method for coupled PDE/ODE systems with different time meshes for the two parts of the system. We demonstrated in a numerical investigation of a series of two and three dimensional model problems that the estimated errors can be used as a reliable a posteriori error estimator. Significant gains in accuracy were obtained for systems with low diffusion coefficients by adaptively refined grids, based on error indicators derived by the DWR method, compared to uniformly refined grids. Adaptively determined temporal discretizations, differently chosen for PDE and ODE part, significantly reduced the computational effort.

The developed numerical methods were used for simulations of the competition between T cells for cytokines in realistic 3D environments. In our simulations, the spatial distribution of cytokine secretion and uptake within a population of immune cells had a huge impact on the cellular response. The numerical results brought new insights into rapid adaptive immune responses in the human body. We found that, despite the high diffusivity of cytokines, localized cytokine sinks and sources occurred. Cytokine signaling was specifically targeted to neighboring cells, which resulted in short-range cell-to-cell communication. However, collective secretion by many producing cells also allowed cytokines to signal over multicellular distances.

Based on the results achieved in this thesis, we present several ideas for further research:

### Parameter estimation by experimental data

Given recent developments in micro fluidic devices and cell culture on structured surfaces, the investigation of the spatio-temporal dynamics of cytokine signaling is being attempted now by Adutler-Lieber et al. [2]. Reliable data obtained by experiments in synthetic cellular micro-environments would enable the development of a more accurate model for T cell interaction by parameter estimation and model adaptation.

### Homogenization techniques

In the simulation of large cell clusters a geometry including all individual cells with their respective surface largely expands the spatial discretization even on coarse grids. There exist analytical approaches which derive a macroscopic model using homogenization techniques by Marciniak-Czochra and Ptashnyk [45]. These techniques allow for coarser discretizations. So far only analytical work exists for cell cluster of the same cell type. For applying homogenization techniques to cell clusters consisting of multiple cell types, e.g. as the application considered in this thesis, a new analytical framework which makes use of stochastic homogenization has to be developed.

### Application of the numerical methods for different PDE/ODE systems

Coupled PDE/ODE systems are found as well in many other applications like combustion, microelectronics, biology and medicine [3, 10, 35, 48]. In these applications the ODE part is coupled with the PDE part on the whole computational domain. A different solver should be considered to apply the adaptive mesh refinement methods derived in this thesis on systems with a much larger ODE part, compared to the systems considered in this thesis.

### A posteriori analysis of decoupled solution methods

In this thesis, we showed that error indicators obtained by the DWR method estimated the error of different time meshes for the PDE and the ODE part of the system closely. We applied the presented accurate coupled solution method as implicit solver and could thus neglect iteration errors. In many applications like coupled PDE systems, a decoupled solution scheme is more advantageous than a coupled scheme. Iteration errors arise in decoupled solution schemes and have to be controlled. Rannacher and Vihharev [52] presented an a posteriori error analysis which balances discretization and iteration errors. The investigation focused on stationary PDEs, but the methods could be extended to non-stationary applications. The control of the iteration error in each time step would extend the applicability of different time meshes for different parts of the system to decoupled solution schemes.

# Acknowledgments

I would like to express my gratitude to my supervisor Prof. Dr. Dr. h.c. Rolf Rannacher for giving me the opportunity to work on this interesting subject and for continuously supporting this work. I thank Prof. Dr. Thomas Höfer for many fruitful discussions on cellular signaling and for the strong support of this work.

I thank Dr. Elfriede Friedmann for continuously supporting this work and many helpful discussions. I would like to thank Dr. Kevin Thurley for a great and successful teamwork.

Further my gratitude goes to Dr. Thomas Carraro for countless discussions and suggestions on my thesis, I particularly thank him for enabling me the research stay at the University of Dundee. I thank Prof. Dr. Mark Chaplain and Dr. Dumitru Trucu for a warm welcome in Scotland and fruitful discussions on mathematical modeling of biological processes.

I thank Prof. Dr. Anna Marciniak-Czochra for many suggestions and enabling me the opportunity to present my work at the University of Warsaw and the Tohoku University in Sendai.

I thank Lisa Kolb, Markus Schaber and especially Dr. Susanne Krömker as head of the Visualization and Numerical Geometry Group of Heidelberg University for all their support in the visualization of three-dimensional data.

I wish to thank the whole Numerical Analysis Group of the Heidelberg University: it was a great time for me to be a part of this group. Last but not least I would like to thank my roommate Christian for many helpful ideas and discussions.

I gratefully acknowledge the funding received during my doctoral studies at the *Faculty of Mathematics and Computer Science at Heidelberg University* by the German Research Foundation (DFG) through ViroQuant and travel funds by the *Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences* and by the *Excellence Initiative of Heidelberg University.*

Darüberhinaus möchte ich mich besonders meiner Familie und meinen Freunden für die viele Unterstützung danken.

# Bibliography

[1] Robert A Adams and John JF Fournier. *Sobolev spaces.* Academic press, 2003.

[2] Shimrit Adutler-Lieber, Irina Zaretsky, Ilia Platzman, Janosch Deeg, Nir Friedman, Joachim P. Spatz, and Benjamin Geiger. Engineering of synthetic cellular microenvironments: Implications for immunity. *Journal of Autoimmunity*, 54(0):100 – 111, 2014.

[3] G. Ali, M. Culpo, and S. Micheletti. Domain decomposition techniques or microelectronic modeling. In Hans-Georg Bock et al., editors, *Progress in Industrial Mathematics at ECMI 2006*, pages 500–505. Springer Berlin Heidelberg, 2008.

[4] Wolfgang Bangerth and Rolf Rannacher. Finite element approximation of the acoustic wave equation: error control and mesh adaptation. *East West Journal of Numerical Mathematics*, 7(4):263–282, 1999.

[5] Wolfgang Bangerth and Rolf Rannacher. *Adaptive finite element methods for differential equations.* Springer Science & Business Media, 2003.

[6] Wolfgang Bangerth, Ralf Hartmann, and Guido Kanschat. deal.II – a general purpose object oriented finite element library. *ACM Trans. Math. Softw.*, 33 (4):24/1–24/27, 2007.

[7] Eberhard Bänsch. An adaptive finite-element strategy for the three-dimensional time-dependent Navier-Stokes equations. *Journal of Computational and Applied Mathematics*, 36(1):3 – 28, 1991.

[8] Roland Becker and Rolf Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.

[9] Michael Besier and Rolf Rannacher. Goal-oriented space–time adaptivity in the finite element Galerkin method for the computation of nonstationary incompressible flow. *International Journal for Numerical Methods in Fluids*, 70(9):1139–1166, 2012.

[10] Muriel Boulakia, Serge Cazeau, Miguel Fernández, Jean-Frédéric Gerbeau, and Nejib Zemzemi. Mathematical modeling of electrocardiograms: A numerical study. *Annals of Biomedical Engineering*, 38:1071–1097, 2010.

[11] J. H. Bramble. *Multigrid Methods*. Longman Scientific and Technical, London, 1993.

[12] Achi Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation*, 31(138):333–390, April 1977.

[13] Dorothea Busse, Maurus de la Rosa, Kirstin Hobiger, Kevin Thurley, Michael Flossdorf, Alexander Scheffold, and Thomas Höfer. Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments. *Proceedings of the National Academy of Sciences*, 2010.

[14] Xiao-Chuan Cai and David E Keyes. Nonlinearly preconditioned inexact Newton algorithms. *SIAM Journal on Scientific Computing*, 24(1):183–200, 2002.

[15] Thomas Carraro, Elfriede Friedmann, and Daniel Gerecht. Coupled vs decoupled approaches for PDE/ODE systems modeling intercellular signaling. *Preprint*, 2015. URL `http://numerik.iwr.uni-heidelberg.de`.

[16] Carsten Carstensen and Rüdiger Verfürth. Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM journal on numerical analysis*, 36(5):1571–1587, 1999.

[17] Hank Childs et al. VisIt: An end-user tool for visualizing and analyzing very large data. In *High Performance Visualization–Enabling Extreme-Scale Scientific Insight*, pages 357–372. Oct 2012.

[18] Philippe G Ciarlet. *The finite element method for elliptic problems*. Elsevier, 1978.

[19] Juliane Claus, Elfriede Friedmann, Ursula Klingmüller, Rolf Rannacher, and Tibor Szekeres. Spatial aspects in the SMAD signaling pathway. *Journal of mathematical biology*, 67(5):1171–1197, 2013.

[20] Daniel M Davis and Michael L Dustin. What is the importance of the immunological synapse? *Trends in immunology*, 25(6):323–327, 2004.

[21] Timothy A Davis. Algorithm 832: Umfpack v4. 3—an unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software (TOMS)*, 30(2):196–199, 2004.

[22] James S Economou and Hyun S Shin. Lymphocyte-activating factor I. generation and physicochemical characterization. *The Journal of Immunology*, 121 (4):1446–1452, 1978.

[23] Charbel Farhat and M Lesoinne. Fast staggered algorithms for the solution of three-dimensional nonlinear aeroelastic problems. *Numerical Unsteady Aerodynamics and Aeroelastic Simulation*, pages 7–1, 1998.

[24] Ofer Feinerman, Garrit Jentsch, Karen E Tkach, Jesse W Coward, Matthew M Hathorn, Michael W Sneddon, Thierry Emonet, Kendall A Smith, and Grégoire Altan-Bonnet. Single-cell quantification of IL-2 response by effector and regulatory T cells reveals critical plasticity in immune response. *Molecular systems biology*, 6(1), 2010.

[25] Carlos A Felippa, KC Park, and Charbel Farhat. Partitioned analysis of coupled mechanical systems. *Computer methods in applied mechanics and engineering*, 190(24):3247–3270, 2001.

[26] Elfriede Friedmann, Andrea C Pfeifer, Rebecca Neumann, Ursula Klingmüller, and Rolf Rannacher. Interaction between experiment, modeling and simulation of spatial aspects in the JAK2/STAT5 signaling pathway. In Hans Georg Bock, Thomas Carraro, Willi Jäger, Stefan Körkel, Rolf Rannacher, and Johannes P. Schlöder, editors, *Model Based Parameter Estimation*, pages 125–143. Springer Berlin Heidelberg, 2013.

[27] Amanda V Gett and Philip D Hodgkin. A cellular calculus for signal integration by T cells. *Nature immunology*, 1(3):239–244, 2000.

[28] Christian Goll, Rolf Rannacher, and Winnifried Wollner. On the adjoint to the damped Crank–Nicolson time marching scheme: Applications to goal-oriented mesh adaptation for the Black–Scholes equation. *J. Comput. Finance, in press*, 2013.

[29] Wolfgang Hackbusch. *Multi-grid methods and applications.* Springer Series in Computational Mathematics, 4. Berlin etc.: Springer- Verlag. XIV, 377, 1985.

[30] Raphael T. Haftka, Jaroslaw Sobieszczanski-Sobieski, and Sharon L. Padula. On options for interdisciplinary analysis and design optimization. *Structural optimization*, 4:65–74, 1992.

[31] Ralf Hartmann. A-posteriori Fehlerschätzung und adaptive Schrittweiten- und Ortsgittersteuerung bei Galerkin-Verfahren für die Wellengleichung. Diplomarbeit, University Heidelberg, 1998.

[32] Ralf Hartmann. Adaptive finite element methods for the compressible Euler equations. Phd thesis, University Heidelberg, 2002.

[33] Matthias Heil. Stokes flow in an elastic tube—a large-displacement fluid-structure interaction problem. *International journal for numerical methods in fluids*, 28(2):243–265, 1998.

[34] Matthias Heil. An efficient solver for the fully coupled solution of large-displacement fluid–structure interaction problems. *Computer Methods in Applied Mechanics and Engineering*, 193(1):1–23, 2004.

[35] J. D Hoernel. A non-stationary problem coupling PDEs and ODEs modelizing an automotive catalytic converter. *Applicable Analysis*, 84(6):617–630, 2005.

[36] Thomas Höfer, Oleg Krichevsky, and Grégoire Altan-Bonnet. Competition for IL-2 between regulatory and effector T cells to chisel immune responses. *Frontiers in immunology*, 3, 2012.

[37] J. Hoffman. Efficient computation of mean drag for the subcritical flow past a circular cylinder using general Galerkin G2. *International Journal for Numerical Methods in Fluids*, 59(11):1241–1258, 2009.

[38] Bärbel Janssen and Guido Kanschat. Adaptive multilevel methods with local smoothing for $h^1$- and $h^{\mathrm{curl}}$-conforming high order finite element methods. *SIAM J. Sci. Comput.*, 33:2095–2114, August 2011.

[39] Bärbel Janssen and Thomas Wick. Block preconditioning with schur complements for monolithic fluid-structure interactions. *ECCOMAS CFD*, 2010.

[40] Hans A Kestler, Christian Wawra, Barbara Kracher, and Michael Kühl. Network modeling of signal transduction: establishing the global view. *Bioessays*, 30(11-12):1110–1125, 2008.

[41] Lisa Kolb. Visualizing high-resolution numerical data with isosurfaces using topological methods. Bachelor's thesis, Heidelberg University, 2013.

[42] Mitchell Luskin, Rolf Rannacher, and Wolfgang Wendland. On the smoothing property of the Crank-Nicolson scheme. *Applicable Analysis*, 14(2):117–135, 1982.

[43] Jan Mandel. On block diagonal and schur complement preconditioning. *Numerische Mathematik*, 58(1):79–93, 1990.

[44] Anna Marciniak-Czochra. Receptor-based models with diffusion-driven instability for pattern formation in hydra. *Journal of Biological Systems*, 11(03): 293–324, 2003.

[45] Anna Marciniak-Czochra and Mariya Ptashnyk. Derivation of a macroscopic receptor-based model using homogenization techniques. *SIAM Journal on Mathematical Analysis*, 40(1):215–237, 2008.

[46] MATLAB. *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts, 2010.

[47] Dimitri J. Mavriplis. An assessment of linear versus nonlinear multigrid methods for unstructured mesh solvers. *Journal of Computational Physics*, 175(1): 302–325, 2002.

[48] Mahdi Esmaily Moghadam, Irene E. Vignon-Clementel, Richard Figliola, and Alison L. Marsden. A modular numerical method for implicit 0D/3D coupling in cardiovascular finite element simulations. *Journal of Computational Physics*, 244(0):63 – 79, 2013.

[49] DP Mok and WA Wall. Partitioned analysis schemes for the transient interaction of incompressible flows and nonlinear flexible structures. *Trends in computational structural mechanics*, pages 689–698, 2001.

[50] Claudio Paniconi and Mario Putti. A comparison of Picard and Newton iteration in the numerical solution of multidimensional variably saturated flow problems. *Water Resources Research*, 30(12):3357–3374, 1994.

[51] Rolf Rannacher. Finite element solution of diffusion problems with irregular data. *Numerische Mathematik*, 43(2):309–327, 1984.

[52] Rolf Rannacher and Jevgeni Vihharev. Balancing discretization and iteration error in finite element a posteriori error analysis. In *Numerical Methods for Differential Equations, Optimization, and Technological Problems*, pages 109–132. Springer, 2013.

[53] Werner C Rheinboldt and Charles K Mesztenyi. On a data structure for adaptive finite element mesh refinements. *ACM Transactions on Mathematical Software (TOMS)*, 6(2):166–187, 1980.

[54] Thomas Richter. Parallel multigrid method for adaptive finite elements with application to 3D flow problems. Phd thesis, University Heidelberg, 2005.

[55] Yrina Rochman, Rosanne Spolski, and Warren J Leonard. New insights into the regulation of T cells by $\gamma$c family cytokines. *Nature Reviews Immunology*, 9(7):480–490, 2009.

[56] Catherine Sabatos, Junsang Doh, Sumone Chakravarti, Rachel S. Friedman, Priya G. Pandurangi, Aaron J. Tooley, and Matthew F. Krummel. A synaptic basis for paracrine interleukin-2 signaling during homotypic T cell interaction. *Immunity*, 29(2):238 – 248, 2008.

[57] Ekkehard W Sachs and Matthias Schu. Gradient computation for model calibration with pointwise observations. In *Control and Optimization with PDE Constraints*, pages 117–136. Springer, 2013.

[58] Michael Schmich. Adaptive finite element methods for computing nonstationary incompressible flows. Phd thesis, University Heidelberg, 2009.

[59] Michael Schmich and Boris Vexler. Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations. *SIAM Journal on Scientific Computing*, 30(1):369–393, 2008.

[60] Edda G Schulz, Luca Mariani, Andreas Radbruch, and Thomas Höfer. Sequential polarization and imprinting of type 1 T helper lymphocytes by interferon-$\gamma$ and interleukin-12. *Immunity*, 30(5):673–683, 2009.

[61] G. Seemann, F.B. Sachse, M. Karl, D.L. Weiss, V. Heuveline, and O. Dössel. Framework for modular, flexible and efficient solving the cardiac bidomain equations using PETSc. In Alistair D. Fitt et al., editors, *Progress in Industrial Mathematics at ECMI 2008*, Mathematics in Industry, pages 363–369. Springer Berlin Heidelberg, 2010.

[62] JA Sherratt, PK Maini, W Jäger, and WA Muller. A receptor based model for pattern formation in hydra. *Forma*, 10(2):77–95, 1995.

[63] Dorothy K Sojka, Denis Bruniquel, Ronald H Schwartz, and Nevil J Singh. IL-2 secretion by CD4+ T cells in vivo is rapid, transient, and influenced by tcr-specific competition. *The Journal of Immunology*, 172(10):6136–6143, 2004.

[64] Alan J. Terry and Mark A.J. Chaplain. Spatio-temporal modelling of the NF-$\kappa$BNF-$\kappa$B intracellular signalling pathway: The roles of diffusion, active transport, and cell geometry. *Journal of Theoretical Biology*, 290(0):7 – 26, 2011.

[65] Kevin Thurley, Daniel Gerecht, Elfriede Friedmann, and Thomas Höfer. Three-dimensional gradients of cytokine signaling between T cells. *PLoS Computational Biology*, 2015.

[66] Karen E Tkach, Debashis Barik, Guillaume Voisinne, Nicole Malandro, Matthew M Hathorn, Jesse W Cotari, Robert Vogel, Taha Merghoub, Jedd Wolchok, Oleg Krichevsky, et al. T cells translate individual, quantal activation into collective, analog cytokine responses via time-integrated feedbacks. *eLife*, 3, 2014.

[67] R Verfürth. A posteriori error estimates for non-linear parabolic equations. *Preprint, Ruhr-Universität Bochum, Fakultät für Mathematik, Bochum, Germany*, 2004.

[68] R. Verfürth. A posteriori error estimates for finite element discretizations of the heat equation. *CALCOLO*, 40(3):195–212, 2003.

[69] Joseph Wloka. *Partielle Differentialgleichungen: Sobolevräume und Randwertaufgaben.* Teubner, 1982.