ORDER–OPTIMAL PRECONDITIONERS FOR FULLY IMPLICIT RUNGE–KUTTA SCHEMES APPLIED TO THE BIDOMAIN EQUATIONS

TRYGVE K. NILSSEN *, GUNNAR A. STAFF †, AND KENT–ANDRE MARDAL ‡

Abstract. The PDE part of the bidomain equations is discretized in time with fully implicit Runge–Kutta methods, and the resulting block systems are preconditioned with a block diagonal preconditioner. By studying the time stepping operator in the proper Sobolev spaces we show that the preconditioned systems have bounded condition numbers given that the Runge–Kutta scheme is A–stable and irreducible with an invertible coefficient matrix. A new proof of order–optimality of the preconditioners for the one–leg discretization in time of the bidomain equations is also presented. The theoretical results are verified by numerical experiments. Additionally, the concept weakly positive definite matrices is introduced and analyzed.

1. Introduction. The electrical activity of the heart can be modelled using the bidomain equations [24]. Applications of the model include studies of arrhythmia, defibrillation and drug development. The bidomain model can be written as:

$$\frac{\partial \boldsymbol{s}}{\partial t} = \boldsymbol{F}(t, \boldsymbol{s}, v), \qquad \text{in } \Omega, \tag{1.1}$$

$$\frac{\partial v}{\partial t} = \nabla \cdot (\sigma_{\text{in}} \nabla v) + \nabla \cdot (\sigma_{\text{in}} \nabla u) - I(\boldsymbol{s}, v), \qquad \text{in } \Omega, \tag{1.2}$$

$$0 = \nabla \cdot (\sigma_{\text{in}} \nabla v) + \nabla \cdot ((\sigma_{\text{tot}}) \nabla u), \qquad \text{in } \Omega, \tag{1.3}$$

where the unknowns are the trans–membrane potential v, the extracellular potential u and the vector of state variables $\boldsymbol{s}$. The length of this vector varies from 1 in the simplest models and up to about 40 in the Winslow model, see e.g. [25]. The intra– and extracellular conductivity tensors are denoted σ_{in} and σ_{ex}, respectively, and $\sigma_{\text{tot}} = \sigma_{\text{in}} + \sigma_{\text{ex}}$. For notational convenience, the tensors are scaled by the membrane capacitance and the membrane surface area, see [21] for details. Depending on the membrane model, the rate function $\boldsymbol{F}$ might describe ionic fluxes, enzyme kinetics and possibly other entities. The function I is current density per membrane capacity. The computational domain denoted by $\Omega \subset \mathbb{R}^d$, $d \leq 3$ is a bounded connected polygonal domain approximating the geometry of the heart.

Usually these equations are solved by applying an operator splitting method, where the system is split into a non–linear system of ODEs and a linear system of PDEs, cf. [21]. Using complicated ion models like the Winslow model and with realistic 3D geometry, the CPU time for the PDE part is similar to the CPU time for the ODE part, cf. [22, 20] for simpler models the PDE part dominates. Realistic simulations of the electrical activity of the heart are time consuming; one heart beat typically requires thousands of time steps and millions of spatial degrees of freedom [17, 22, 20]. In this paper we will consider preconditioners for advanced time stepping schemes for the PDE part of the bidomain model, since this part is the bottleneck that prevent the use of more advanced and possibly efficient time stepping schemes for the fully

* University of Agder, Department of Economics and Business Administration, Service Box 422, 4604 Kristiansand, Norway

†SPT Group, 2027 Kjeller, Norway (gunnaran@simula.no)

‡Simula Research Laboratory, 1325 Lysaker, Norway, Department of Informatics, University of Oslo, Oslo, Norway, and Norwegian Research Council (grant 170650) (kent-and@simula.no)

coupled system. The PDE part has the following form:

$$\frac{\partial v}{\partial t} = \nabla \cdot (\sigma_{\text{in}} \nabla v) + \nabla \cdot (\sigma_{\text{in}} \nabla u) + f, \tag{1.4}$$

$$0 = \nabla \cdot (\sigma_{\text{in}} \nabla v) + \nabla \cdot (\sigma_{\text{tot}} \nabla u), \tag{1.5}$$

where $\sigma_{\text{tot}} = \sigma_{\text{in}} + \sigma_{\text{ex}}$. This is often referred to as an index 1 partial differential algebraic equation (PDAE[1]). Index 1 problems should be solved with stiffly accurate time stepping schemes suited for adaptive time step selection, cf. [10, Chapter 7]. Therefore, the RadauIIA and the LobattoIIIC schemes are desirable for solving (1.4)–(1.5).

A spatial semi–discretized version of these equations may be written as

$$I_h \frac{dv_h}{dt} = -M_{\text{in}} v_h - M_{\text{in}} u_h + f_h(t, x), \tag{1.6}$$

$$0 = -M_{\text{in}} v_h - M_{\text{tot}} u_h, \tag{1.7}$$

where I_h is the mass matrix, $M_{\text{in}} \in \mathbb{R}^{n \times n}$ is the spatial discretization matrix of $-\nabla \cdot (\sigma_{\text{in}} \nabla \cdot)$ (a stiffness matrix), $M_{\text{tot}} \in \mathbb{R}^{n \times n}$ is the discretization of $-\nabla \cdot (\sigma_{\text{tot}} \nabla \cdot)$ (a stiffness matrix) and f_h is the L^2 projection of f onto the discretization space. The variables v_h and u_h are the discrete approximations of v and u. The spatial discretization method will in our numerical experiments be a finite element method with Lagrange elements of order one to four. However, the theoretical results are valid for all conforming spatial discretizations.

In the PDE literature, efficient and accurate spatial discretization methods are well studied, while temporal discretization methods have been given less attention. In this paper we study the numerical solution of the index 1 DAE described in (1.6)–(1.7), which is the semi–discretized version of (1.4)–(1.5). We apply Runge–Kutta schemes on (1.6)–(1.7) and end up with block systems to be solved for each time step. The important point in this paper is that we are able to construct efficient preconditioners for these systems.

The method suggested in this paper is similar to the method the authors introduced in [14], where an order–optimal preconditioner for Runge–Kutta schemes applied to the linear parabolic equation

$$u_t = \Delta u + f$$

is presented. The preconditioner in [14] is block diagonal, and standard preconditioners are used for the diagonal blocks. Here, we extend this analysis to an index 1 DAE, namely the PDE part of the bidomain equations.

We use the common definition of an order–optimal preconditioner, which is that the preconditioner $\mathcal{B}_k$ is an order–optimal preconditioner for $\mathcal{A}_k$ with respect to the parameter k, given that the condition number $\kappa(\mathcal{B}_k \mathcal{A}_k) \leq c_0$, where c_0 is a constant independent of the parameter k. Additionally we also require that the evaluation and storage of $\mathcal{B}_k$ is similar to that of $\mathcal{A}_k$. In this paper the preconditioner will be order–optimal with respect to both the spatial and temporal discretization parameters h

[1]Strictly speaking PDAE is not a describing term for (1.4)–(1.5), since there is no algebraic equation involved in the equation system. The system is rather a time dependent PDE coupled with a stationary PDE. Still, (1.4)–(1.5) is often refereed to as an index 1 PDAE, since the semi–discretized version of it, (1.6)–(1.7), is an index 1 DAE (differential algebraic equation).

and δt. It will however depend on the Runge–Kutta scheme and its number of stages s.

The proof of order–optimality is done by showing that the continuous counterpart of the time stepping operator is an isomorphism in properly chosen Sobolev spaces. Then an "exact" preconditioner is defined as the Riesz identity mapping between the dual space and the chosen space[2]. Finally, we create an operator spectrally equivalent, in the meaning of independent of h and δt, to the Riesz identity mapping for creating an implementationally efficient algorithm. Other works that use this approach are [4, 11, 14, 15]. We also need the main result of [13], namely that an implicit Euler approximation of the bidomain equations can be preconditioned by order–optimal preconditioners based on the diagonal blocks of the time stepping operator. By Lemma 3.3 we get a new, more compact, and instructive proof of this result. We remark that this block preconditioner facilitate that the preconditioner for the higher–order schemes may reuse the preconditioner implemented for one–leg discretizations.

Implicit Runge–Kutta temporal discretizations will typically neither result in positive nor symmetric time stepping operators, even if the spatial operators are both positive and symmetric. Therefore, we introduce a family of matrices which we refer to as weakly positive definite matrices. These matrices are designed to make the operators "weakly positive" such that the general existence theorem of Babuska can be applied. This approach was also taken when proving the order–optimality of the preconditioners presented in [14]. The papers [18, 19] discuss various ways to improve the preconditioners presented here. These preconditioners are often better in practice, but more technical to analyze.

A matrix is referred to as weakly positive definite if there exists a positive definite matrix such that the product of these two matrices also is positive definite [16]. We prove that a square matrix is weakly positive definite if and only if its real eigenvalues are positive. Implicit Runge–Kutta schemes which are A–stable, irreducible and have an invertible coefficient matrix have weakly positive coefficient matrices, as we will demonstrate. Weakly positive matrices can be seen as a generalization of the diagonally stable matrices discussed in e.g. [12].

The rest of this paper is organized as follows: In Section 2 we present the discretization of the bidomain model. Then in Section 3, we prove that the continuous version of the time stepping operator for the Runge–Kutta methods is an isomorphism bounded independently of δt. Then we present the order–optimal block preconditioners in Section 4. In Section 5 the weakly positive definite matrices are analyzed. Finally, we present some numerical results in Section 6.

2. Preliminaries and notation. In this paper the Kronecker product, $\otimes$, will be frequently used. For square matrices $A \in \mathbb{R}^{s\times s}$ and $B \in \mathbb{R}^{n\times n}$ the Kronecker product of $A \otimes B \in \mathbb{R}^{sn\times sn}$ is the block matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1s}B \\ \vdots & \ddots & \vdots \\ a_{s1}B & \dots & a_{ss}B \end{pmatrix}.$$

In the following B can also be a continuous operator.

The continuous weak form of (1.4)–(1.5) in the case of homogeneous Dirichlet boundary conditions can be expressed as:

[2] The Riesz identity mapping implied by the Riesz representation theorem.

Find $v, u \in L^2(0,T;H_0^1)$ with $\frac{\partial v}{\partial t} \in L^2(0,T;H^{-1})$ such that

$$\left(\frac{\partial v}{\partial t}, l\right) + (\sigma_{\text{in}}\nabla v, \nabla l) + (\sigma_{\text{in}}\nabla u, \nabla l) = (f, l), \quad \forall l \in H_0^1, \quad a.e.\ t \in [0,T],$$
$$(\sigma_{\text{in}}\nabla v, \nabla m) + (\sigma_{\text{tot}}\nabla u, \nabla m) = 0, \quad \forall m \in H_0^1, \quad a.e.\ t \in [0,T],$$

where $(\cdot,\cdot)$ denotes the L_2 inner product, but also the duality pairing between H_0^1 and H^{-1}. Furthermore, H^{-1} is the dual space of H_0^1, and we assume that $f \in L^2(0,T;H^{-1})$.

Similarly, the finite element formulation is defined by seeking an approximation $(v_h(t), u_h(t)) \in V_h \times U_h \subset H_0^1 \times H_0^1$ by:
Find $(v_h, u_h) \in L^2(0,T;V_h) \times L^2(0,T;U_h)$ with $\frac{\partial v_h}{\partial t} \in L^2(0,T;V_h)$ such that

$$\left(\frac{\partial v_h}{\partial t}, l\right) + (\sigma_{\text{in}}\nabla v_h, \nabla l) + (\sigma_{\text{in}}\nabla u_h, \nabla l) = (f, l), \quad \forall l \in V_h,$$
$$(\sigma_{\text{in}}\nabla v_h, \nabla m) + (\sigma_{\text{tot}}\nabla u_h, \nabla m) = 0, \quad \forall m \in U_h.$$

This is the variational form of (1.6)–(1.7).

The Runge–Kutta approximation of the bidomain equations (1.6)–(1.7) can be written as

$$I_h v_i = I_h v_h^{n-1} + \delta t \sum_{j=1}^{s} a_{ij}(-M_{\text{in}}v_j - M_{\text{in}}u_j + f_h(t_{n-1} + c_j\delta t)), \quad i = 1, \ldots, s, \tag{2.1}$$
$$0 = -M_{\text{in}}v_i - M_{\text{tot}}u_i, \quad i = 1 \ldots, s, \tag{2.2}$$

where v_i and u_i are the stage values for v_h and u_h, respectively, δt is the temporal discretization parameter, s is the number of quadrature nodes, a_{ij} are the Runge–Kutta coefficients, b_i are the quadrature weights and c_i are the quadrature points. The value of u_h and v_h at the next time step is then found by

$$v_h^{n+1} = v_s,$$
$$u_h^{n+1} = u_s,$$

assuming that the chosen Runge–Kutta scheme is stiffly accurate ($a_{si} = b_i$, $i = 1, \ldots, s$).

We remark that the "simplified" equation (2.2) requires that the Runge–Kutta matrix A is invertible. If A is singular, or we for some reason do not want to use the simplified formulation (2.2), the full formulation is

$$0 = \delta t \sum_{j=1}^{s} a_{ij}(-M_{\text{in}}v_j - M_{\text{tot}}u_j), \quad i = 1, \ldots, s. \tag{2.2b}$$

The focus of this paper is to solve the coupled system (2.1)–(2.2)/(2.2b).

2.1. Notation in the discrete case. We define two projection operators

$$P_1 = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}, \tag{2.3}$$

where I is the identity operator. However, in this subsection only, we let I denote the identity matrix in $\mathbb{R}^{n \times n}$, for notational convenience.

Equations (2.1) and (2.2)/(2.2b) can be written on the form

$$\mathcal{A}_h \mathbf{w}_h = \mathbf{b}_h, \tag{2.4}$$

where $\mathcal{A}_h$ is defined as

$$\mathcal{A}_h = I^s \otimes P_{1,h} + \delta t \, A \otimes \mathbb{M}_h \tag{2.5}$$

for (2.2b), and

$$\mathcal{A}_h = I^s \otimes P_{1,h} + \delta t \, \left(A \otimes P_1 \mathbb{M}_h + I^s \otimes P_2 \mathbb{M}_h \right), \tag{2.6}$$

for (2.2), where I^s is the identity matrix in $\mathbb{R}^{s \times s}$ and

$$P_{1,h} = \begin{pmatrix} I_h & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2n \times 2n}, \quad \mathbb{M}_h = \begin{pmatrix} M_{\text{in}} & M_{\text{in}} \\ M_{\text{in}} & M_{\text{tot}} \end{pmatrix} \in \mathbb{R}^{2n \times 2n},$$
$$P_1 \mathbb{M}_h = \begin{pmatrix} M_{\text{in}} & M_{\text{in}} \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2n \times 2n}, \quad P_2 \mathbb{M}_h = \begin{pmatrix} 0 & 0 \\ M_{\text{in}} & M_{\text{tot}} \end{pmatrix} \in \mathbb{R}^{2n \times 2n},$$
$$\mathbf{b}_h = \mathbb{1} \otimes \begin{pmatrix} v_h^{n-1} \\ 0 \end{pmatrix} + \delta t \, (A \otimes I^{2n}) \cdot \mathbf{f}_h \in \mathbb{R}^{2sn \times 1}, \quad \mathbb{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{s \times 1},$$
$$\mathbf{f}_h = \begin{pmatrix} f_h(t^{n-1} + \delta t \, c_1) \\ 0 \\ \vdots \\ f_h(t^{n-1} + \delta t \, c_s) \\ 0 \end{pmatrix} \in \mathbb{R}^{2sn \times 1}, \quad \mathbf{w}_h = \begin{pmatrix} v_1 \\ u_1 \\ \vdots \\ v_s \\ u_s \end{pmatrix} \in \mathbb{R}^{2sn \times 1}.$$

Note that (2.2) has been multiplied with δt to get (2.5). There is also a third option, where (2.2) is multiplied with $\delta t \, a_{ii}$, leading to

$$\mathcal{A}_h = I^s \otimes P_{1,h} + \delta t \, \left(A \otimes P_1 \mathbb{M}_h + \operatorname{diag} A \otimes P_2 \mathbb{M}_h \right) \tag{2.7}$$

where $\operatorname{diag} A$ is a matrix containing the diagonal elements from A. This is the operator used in the numerical experiments, since it gives slightly better results than the other formulations. In the following, we will refer to $\mathcal{A}_h$ either on the form (2.5), (2.6), or (2.7) as the discrete time stepping operator.

2.2. Notation in the continuous case. The spatial continuous counterparts of variables defined in the previous subsection is analogous to the discrete operator above, but without the subscript h. The continuous time stepping operator related to (2.5) is defined as

$$\mathcal{A} = I^s \otimes P_1 + \delta t \, A \otimes \mathbb{M}, \tag{2.8}$$

where

$$\mathbb{M} = - \begin{pmatrix} \nabla \cdot (\sigma_{\text{in}} \nabla) & \nabla \cdot (\sigma_{\text{in}} \nabla) \\ \nabla \cdot (\sigma_{\text{in}} \nabla) & \nabla \cdot (\sigma_{\text{tot}} \nabla) \end{pmatrix}.$$

Similarly, the continuous operator for (2.6) and (2.7) is defined as

$$\mathcal{A} = I^s \otimes P_1 + \delta t \, \left(A \otimes P_1 \mathbb{M} + I^s \otimes P_2 \mathbb{M} \right), \tag{2.9}$$
$$\mathcal{A} = I^s \otimes P_1 + \delta t \, \left(A \otimes P_1 \mathbb{M} + \operatorname{diag} A \otimes P_2 \mathbb{M} \right). \tag{2.10}$$

A generic way of writing these different formulations is

$$\mathcal{A} = I^s \otimes P_1 + \delta t \left(A \otimes P_1 \mathbb{M} + \tilde{A} \otimes P_2 \mathbb{M} \right). \tag{2.11}$$

If $\tilde{A} = A$, we have (2.8), if $\tilde{A} = I$ we have (2.9) and if $\tilde{A} = \operatorname{diag} A$ we have (2.10). We remark that in exact arithmetic the solutions of (2.4) are identical for the three variants.

In this paper v is used for the transmembrane potential, u is used for the extracellular potential and w is used for vectors containing both v and u. Boldface notation is used for vectors including all stage variables and non-boldface for vectors including only one stage variable in the following way:

$$w = \begin{pmatrix} v \\ u \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} v_1 \\ u_1 \\ \vdots \\ v_s \\ u_s \end{pmatrix}.$$

In the proof of order–optimality we sometimes need two vectors of the same kind as $\mathbf{w}$, and we then write

$$\mathbf{w}^1 = \begin{pmatrix} v_1^1 \\ u_1^1 \\ \vdots \\ v_s^1 \\ u_s^1 \end{pmatrix}, \quad \mathbf{w}^2 = \begin{pmatrix} v_1^2 \\ u_1^2 \\ \vdots \\ v_s^2 \\ u_s^2 \end{pmatrix},$$

where v_i^1, u_i^1, v_i^2, u_i^2 are the stage values for stage i. We will also use

$$\mathbf{v}^k = \begin{pmatrix} v_1^k \\ \vdots \\ v_s^k \end{pmatrix}, \quad \mathbf{u}^k = \begin{pmatrix} u_1^k \\ \vdots \\ u_s^k \end{pmatrix}, \quad w_i^k = \begin{pmatrix} v_i^k \\ u_i^k \end{pmatrix} \quad k = 1, 2.$$

The block conductivity tensor is

$$\Sigma = \begin{pmatrix} \sigma_{\text{in}} & \sigma_{\text{in}} \\ \sigma_{\text{in}} & \sigma_{\text{tot}} \end{pmatrix} \in \mathbb{R}^{2d \times 2d}, \tag{2.12}$$

where σ_{in} and σ_{ex} are the positive definite conductivity tensors and $\sigma_{\text{tot}} = \sigma_{\text{in}} + \sigma_{\text{ex}}$. Then, with the above defined notation, the operator $\mathbb{M}$ can be written as

$$\mathbb{M}w = -\nabla \cdot (\Sigma \nabla w), \tag{2.13}$$

where the ∇ operator is applied to a vector it should be understood as it is applied to each component, e.g.,

$$\nabla w = \left(\frac{\partial v}{\partial x_1}, \ldots, \frac{\partial v}{\partial x_d}, \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_d} \right)^T \in (L^2(\Omega))^{2d},$$

and similar for larger vectors, e.g., $\nabla \mathbf{w} \in (L^2(\Omega))^{2sd}$. We define the divergence operator $\nabla \cdot$ weakly through the relation to the gradient operator

$$-(\nabla \cdot (\nabla \mathbf{w}^1), \mathbf{w}^2) = (\nabla \mathbf{w}^1, \nabla \mathbf{w}^2), \quad \forall \mathbf{w}^1, \mathbf{w}^2.$$

Further, we will need the intersection and the sum of two Hilbert spaces, cf. [6, Chapter 2]. If X and Y are Hilbert spaces, both continuously contained in some larger Hilbert space, then the sum $X+Y$ and the intersection $X \cap Y$ are Hilbert spaces, and the norms are defined as

$$\begin{aligned} \|z\|_{X\cap Y} &= (\|z\|_X^2 + \|z\|_Y^2)^{1/2}, \\ \|z\|_{X+Y} &= \inf_{\substack{z=x+y \\ x\in X, y\in Y}} (\|x\|_X^2 + \|y\|_Y^2)^{1/2}. \end{aligned}$$

The natural Sobolev space for the bidomain operator is

$$H = (L^2 \cap \sqrt{\delta t}\, H_0^1) \times \sqrt{\delta t}\, H_0^1,$$

where the norms are

$$\|u\|^2_{L^2\cap\sqrt{\delta t}\, H_0^1} = \int_\Omega u^2 + \delta t\, |\nabla u|^2 \, dx \quad \text{and} \quad \|u\|^2_{\sqrt{\delta t}\, H_0^1} = \int_\Omega \delta t\, |\nabla u|^2 \, dx.$$

and

$$\|w\|_H^2 = \|v\|_{L^2}^2 + \|w\|^2_{(\sqrt{\delta t}\, H_0^1)^2} = \|v\|_{L^2}^2 + \|v\|^2_{\sqrt{\delta t}\, H_0^1} + \|u\|^2_{\sqrt{\delta t}\, H_0^1}.$$

The corresponding dual space is

$$H^* = (L^2 + \frac{1}{\sqrt{\delta t}} H^{-1}) \times \frac{1}{\sqrt{\delta t}} H^{-1}.$$

The Sobolev space for $\mathbf{w}$ is

$$\boldsymbol{H} = H^s$$

and the corresponding norm is defined by

$$\|\mathbf{w}\|^2_{\boldsymbol{H}} = \|\mathbf{v}\|^2_{(L^2)^s} + \|\mathbf{w}\|^2_{(\sqrt{\delta t}\, H_0^1)^{2s}}.$$

and the dual space $\boldsymbol{H}^*$ of $\boldsymbol{H}$ is

$$\boldsymbol{H}^* = (H^*)^s.$$

Throughout the paper $\|\cdot\|$ denotes the L_2 norm and $(\cdot,\cdot)$ denotes both the L_2 inner product and the duality paring between a Hilbert space and its dual with respect to the L_2 inner product.

3. Theoretical study of the continuous time stepping operator. We will in the following show that $\mathcal{A} \in \mathcal{L}(\boldsymbol{H}, \boldsymbol{H}^*)$ is an isomorphism, where $\boldsymbol{H}^*$ is the dual space of $\boldsymbol{H}$.

THEOREM 3.1. *$\mathcal{A} \in \mathcal{L}(\boldsymbol{H}, \boldsymbol{H}^*)$ is an isomorphism from $\boldsymbol{H}$ to its dual space $\boldsymbol{H}^*$.*

Proof. Using the Theorem from Babuska and Aziz (cf. [5]), the theorem can be proved by showing the following three properties:
There exists a c_1 independent of δt such that (boundedness)

$$|(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2)| \le c_1 \|\mathbf{w}^1\|_{\boldsymbol{H}} \|\mathbf{w}^2\|_{\boldsymbol{H}}, \quad \forall \mathbf{w}^1, \mathbf{w}^2 \in \boldsymbol{H}. \tag{3.1}$$

There exists a c_2 independent of δt such that (inf–sup)

$$\sup_{\mathbf{w}^2 \in \boldsymbol{H}} \frac{(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2)}{\|\mathbf{w}^2\|_{\boldsymbol{H}}} \geq \frac{1}{c_2}\|\mathbf{w}^1\|_{\boldsymbol{H}}, \quad \forall \mathbf{w}^1 \in \boldsymbol{H}. \tag{3.2}$$

For any $\mathbf{w}^2 \in \boldsymbol{H}$ there exists a $\mathbf{w}^1 \in \boldsymbol{H}$ such that

$$(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2) \neq 0. \tag{3.3}$$

$\square$

Before we start to prove Theorem 3.1, we will prove two preliminary results. The first result is that the operator $\mathbb{M}$ is elliptic, and the next is that the Runge–Kutta coefficient matrices have a property which we refer to as weakly positive definite.

Lemma 3.2. *The operator* $\mathbb{M}$ *is strictly elliptic.*

Proof. The main step of this proof is the observation that $\Sigma \in \mathbb{R}^{2d\times 2d}$ is a pointwise symmetric positive definite (SPD) matrix, whenever the block matrices σ_{in} and σ_{ex} are so. To see this, let $x, y \in \mathbb{R}^d$

$$\begin{aligned}
\left(x^T, y^T\right) \begin{pmatrix} \sigma_{\text{in}} & \sigma_{\text{in}} \\ \sigma_{\text{in}} & \sigma_{\text{tot}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= x^T\sigma_{\text{in}}x + x^T\sigma_{\text{in}}y + y^T\sigma_{\text{in}}x + y^T\sigma_{\text{in}}y + y^T\sigma_{\text{ex}}y \\
&= \|{\sigma_{\text{in}}}^{1/2}(x+y)\|^2 + \|{\sigma_{\text{ex}}}^{1/2}y\|^2 \qquad (3.4)\\
&> 0, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \neq 0, \qquad (3.5)
\end{aligned}$$

where $\sigma_{\text{tot}} = \sigma_{\text{in}} + \sigma_{\text{ex}}$. To see that (3.4) leads to (3.5) we observe that the last term in (3.4) is positive if $y \neq 0$ and the first term is positive if $y = 0$. Since Σ is symmetric we can therefore conclude that it is SPD.

We further have that any positive definite and bounded matrix is pointwise spectrally equivalent with the identity matrix, i.e. $\Sigma \sim I$ (this follows from the fact that all finite dimensional norms are equivalent). Thus we can conclude that the operator $\mathbb{M}$ is strictly elliptic and there exist two constants γ_1 and γ_2 such that

$$\gamma_1(\nabla w, \nabla w) \leq (\Sigma\nabla w, \nabla w) \leq \gamma_2(\nabla w, \nabla w), \quad \forall w \in H. \tag{3.6}$$

Note that γ_1 is bounded from below by the smallest eigenvalue of Σ pointwise in Ω. Similarly γ_2 is bounded from above by the largest eigenvalue of Σ pointwise in Ω. $\square$

Remark 3.1. *Notice that typical variations in the conductivity tensors are mild. The ratio* γ_2/γ_1 *is typically around 10 for a human heart.*

As a small digression, we shall now see how Lemma 3.2 can help us to analyze the implicit Euler discretization of (1.6)–(1.7). The time stepping operator for the implicit Euler scheme can be written (when $\tilde{A} = A$): $\mathcal{A}_E = P_1 + \delta t\,\mathbb{M}$.

Lemma 3.3. *The operator* $\mathcal{A}_E = P_1 + \delta t\,\mathbb{M}$ *is an isomorphism mapping* H *to* H^*

Proof.

$$c_0^E\|w\|_H \leq (\mathcal{A}_E w, w) \leq c_1^E\|w\|_H, \quad \forall w \in H,$$

where $c_0^E = \min(1, \delta t\,\gamma_1)$ and $c_1^E = \max(1, \delta t\,\gamma_2)$. $\square$

Remark 3.2. *Lemma 3.3 gives a simple alternative proof of the main result in [13], namely that block diagonal preconditioners for the one–leg temporal discretizations of the bidomain equations* (1.4)–(1.5) *are order optimal given that the preconditioner is an isomorphism mapping* H^* *to* H*. Further discussions on the construction of such preconditioners are postponed to Section 4.*

To prove the inf–sup condition in Theorem 3.1 we will use a family of matrices which we refer to as weakly positive definite c.f. [16]. These matrices are defined as follows.

DEFINITION 3.4. *A matrix $A \in \mathbb{R}^{s\times s}$ is weakly positive definite if there exists a positive definite $C \in \mathbb{R}^{s\times s}$ such that the product CA is positive definite, i.e.*

$$x^T Cx > 0, \quad \forall x \in \mathbb{R}^s,\, x \neq 0 \tag{3.7}$$

$$x^T CAx > 0, \quad \forall x \in \mathbb{R}^s,\, x \neq 0. \tag{3.8}$$

Note that the neither A, C nor CA needs to be symmetric in Definition 3.4.

LEMMA 3.5. *A real square matrix is weakly positive definite if and only if the real eigenvalues are positive.*

Note that the eigenvalues of a weakly positive definite matrix can lie anywhere in the complex plane, except for at zero and along the negative real axis. The proof of Lemma 3.5 is postponed to Section 5.

The important observation now is that an A–stable irreducible Runge–Kutta scheme with an invertible A–matrix[3] will have a weakly positive definite A–matrix. In fact a stronger result is known, namely that the real part of the eigenvalues are positive. To see this we study the stability function of a general Runge–Kutta scheme applied to the Dahlquist test–equation $y' = \lambda y$,

$$R(z) = \frac{\det(I - zA + z\mathbf{1}b^T)}{\det(I - zA)}, \quad z = \delta t\, \lambda.$$

Assume now that there exists an eigenvalue μ of A with $Re(\mu) < 0$. Consequently the stability function would have a pole in $z = 1/\mu$, which lies in the left half plane. But then A–stability is impossible since it requires that $|R(z)| \leq 1\ \forall z \in \mathbb{C}^-$. Since we have assumed that the Runge–Kutta scheme is irreducible, this leads to a contradiction. Additionally, since we have claimed that the A–matrix is invertible, a zero–eigenvalue is not possible. Therefore we conclude that the eigenvalues of the A–matrix have positive real part, under the given assumptions.

We have now established that $\mathbb{M}$ is elliptic and that the A–matrix is weakly positive definite, and are now ready to prove Theorem 3.1. It remains to prove (3.1)–(3.3), which is proved in three lemmas. Lemma 3.6 proves (3.1), Lemma 3.8 proves (3.2) and Lemma 3.9 proves (3.3).

LEMMA 3.6. *There exists a $c_1 > 0$ such that*

$$\left(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2\right) \leq c_1 \|\mathbf{w}^1\|_{\boldsymbol{H}} \|\mathbf{w}^2\|_{\boldsymbol{H}}. \tag{3.9}$$

That is,

$$\|\mathcal{A}\|_{\mathcal{L}(\boldsymbol{H},\boldsymbol{H}^*)} \leq c_1. \tag{3.10}$$

[3] A-stable schemes produce stable numerical solutions of linear problems of the form $y' = \lambda y,\ \lambda < 0$, independent of δt (the exact solution is also asymptotically stable). This is an advantage for stiff problems, characterized by $|\lambda| \gg 1$, which would otherwise get very strict requirements on δt. All Runge–Kutta schemes discussed in this paper and all the standard implicit Runge–Kutta schemes are A–stable, irreducible and has invertible A–matrix (see [9]).

Proof. We have

$$\begin{aligned}
\big(\mathcal{A}\mathbf{w}^1,\mathbf{w}^2\big) &= \Big((I^s\otimes P_1+\delta t\,(A\otimes P_1\mathbb{M}+\tilde{A}\otimes P_2\mathbb{M}))\mathbf{w}^1,\mathbf{w}^2\Big)\\
&= \big((I^s\otimes P_1)\mathbf{w}^1,\mathbf{w}^2\big)+\delta t\,\Big((A\otimes P_1\mathbb{M}+\tilde{A}\otimes P_2\mathbb{M})\mathbf{w}^1,\mathbf{w}^2\Big)\\
&= \big(\mathbf{v}^1,\mathbf{v}^2\big)+\delta t\,\big(((A\otimes\sigma_{\text{in}})\nabla\mathbf{v}^1,\nabla\mathbf{v}^2)+((A\otimes\sigma_{\text{in}})\nabla\mathbf{u}^1,\nabla\mathbf{v}^2)\\
&\qquad\qquad +((\tilde{A}\otimes\sigma_{\text{in}})\nabla\mathbf{v}^1,\nabla\mathbf{u}^2)+((\tilde{A}\otimes\sigma_{\text{tot}})\nabla\mathbf{u}^1,\nabla\mathbf{u}^2)\big)\\
&\le \gamma_2 a_{max}\left(\|\mathbf{v}^1\|\|\mathbf{v}^2\|+\delta t\sum_{ij}\|\nabla w_i^1\|\|\nabla w_j^2\|\right) &(3.11)\\
&\le c_1\left(\|\mathbf{v}^1\|\|\mathbf{v}^2\|+\delta t\,\|\nabla\mathbf{w}^1\|\|\nabla\mathbf{w}^2\|\right) &(3.12)\\
&\le c_1\left(\|\mathbf{v}^1\|^2+\delta t\,\|\nabla\mathbf{w}^1\|^2\right)^{\frac{1}{2}}\left(\|\mathbf{v}^2\|^2+\delta t\,\|\nabla\mathbf{w}^2\|^2\right)^{\frac{1}{2}}\\
&= c_1\|\mathbf{w}^1\|_{\boldsymbol{H}}\|\mathbf{w}^2\|_{\boldsymbol{H}},
\end{aligned}$$

where $\sigma_{\text{tot}}=\sigma_{\text{in}}+\sigma_{\text{ex}}$, (3.11) comes from

$$\big(\mathbb{M}w_i^1,w_j^2\big)\le\gamma_2\|\nabla w_i^1\|\|\nabla w_j^2\|, \tag{3.13}$$

the Cauchy–Schwarz inequality, and the definition

$$a_{max}=\max\left(\max_{ij}|a_{ij}|,\max_{ij}|\tilde{a}_{ij}|,1\right).$$

Furthermore, (3.12) comes from the fact that

$$\begin{aligned}
\sum_{ij}\|\nabla w_i^1\|\|\nabla w_j^2\| &= \sum_i\|\nabla w_i^1\|\sum_j\|\nabla w_j^2\|\\
&\le s\|\nabla w^1\|\|\nabla w^2\|. &(3.14)
\end{aligned}$$

Finally, (3.14) follows from the equivalence of the two finite dimensional spaces ℓ^1 and ℓ^2. ◻

In order to prove (3.2), we need the following intermediate result. Notice that the following lemma can be seen as a weaker form of the Lax–Milgram theorem. The Runge–Kutta coefficient matrix A is typically nonsymmetric and indefinite and therefore $A\otimes\mathbb{M}$ is typically not elliptic. In our analysis, we employ an extra matrix C to make the operator $CA\otimes\mathbb{M}$ elliptic,

Lemma 3.7. *There exists a constant α, such that*

$$((CA\otimes\mathbb{M})\mathbf{w},\mathbf{w})\ge\alpha\gamma_1\|\nabla\mathbf{w}\|,\quad\forall\mathbf{w}\in\boldsymbol{H}, \tag{3.15}$$

where γ_1 comes from Lemma 3.2, and C comes from Definition 3.4.

Proof. From Definition 3.4 and Lemma 3.5 and since $\|x\|=1$ gives us a compact set, we know that there exists a positive α such that

$$\alpha=\min\left(\min_{\|x\|=1}x^TCx,\ \min_{\|x\|=1}x^TCAx\right)>0. \tag{3.16}$$

We will also need the symmetric part of CA, which will be denoted $S(CA)=\frac{1}{2}((CA)^T+CA)$.

By simple linear algebra calculation we know that

$$
\begin{aligned}
((CA \otimes \mathbb{M})\mathbf{w}, \mathbf{w}) &= \sum_{ij}^{s} (CA)_{ij} (\nabla \cdot (\Sigma \nabla w_j), w_i) \\
&= \sum_{ij}^{s} (CA)_{ij} (\Sigma \nabla w_j, \nabla w_i) \\
&= \sum_{ij}^{s} (S(CA))_{ij} (\Sigma \nabla w_j, \nabla w_i) && (3.17) \\
&= ((S(CA) \otimes \Sigma)\nabla \mathbf{w}, \nabla \mathbf{w}) && (3.18) \\
&\geq \alpha\gamma_1 \|\nabla \mathbf{w}\|^2. && (3.19)
\end{aligned}
$$

We know that $(\Sigma\nabla\cdot, \nabla\cdot)$ is symmetric, and therefore we can substitute CA with the symmetric part $S(CA)$, which justifies (3.17). In (3.18), $S(CA)$ is SPD since CA is positive definite. This is also true for Σ, and consequently true for $S(CA) \otimes \Sigma$. The last line (3.19) is given by (3.16), the lower bound of (3.6) and the fact that the eigenvalues of the tensor product of two SPD matrices equals the product of the eigenvalues of the same matrices (see e.g. [8]). □

LEMMA 3.8. *There exists a constant $c_2 > 0$ such that*

$$
\sup_{\mathbf{w}^2 \in \boldsymbol{H}} \frac{(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2)}{\|\mathbf{w}^1\|_{\boldsymbol{H}} \|\mathbf{w}^2\|_{\boldsymbol{H}}} > \frac{1}{c_2}, \quad \forall \mathbf{w}^1 \in \boldsymbol{H}. \tag{3.20}
$$

Proof. Given $\mathbf{w}^1 \in H$, let $\mathbf{w}^2 = (C^T \otimes P_1 + D^T \otimes P_2)\mathbf{w}^1$, where $D = CA\tilde{A}^{-1}$. The matrix D is constructed such that $CA \otimes \mathbb{M} = CA \otimes P_1\mathbb{M} + D\tilde{A} \otimes P_2\mathbb{M}$, where $\tilde{A}$ is the matrix from the generic form (2.11). We can further calculate:

$$
\begin{aligned}
\sup_{\mathbf{w}^2 \in \boldsymbol{H}} \frac{(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2)}{\|\mathbf{w}^1\|_{\boldsymbol{H}} \|\mathbf{w}^2\|_{\boldsymbol{H}}} &\geq \frac{(\mathcal{A}\mathbf{w}^1, (C^T \otimes P_1 + D^T \otimes P_2)\mathbf{w}^1)}{\|\mathbf{w}^1\|_{\boldsymbol{H}} \|(C^T \otimes P_1 + D^T \otimes P_2)\mathbf{w}^1\|_{\boldsymbol{H}}} \\
&= \frac{((C \otimes P_1 + D \otimes P_2)\mathcal{A}\mathbf{w}^1, \mathbf{w}^1)}{\|\mathbf{w}^1\|_{\boldsymbol{H}} \|(C^T \otimes P_1 + D^T \otimes P_2)\mathbf{w}^1\|_{\boldsymbol{H}}} \\
&\geq \frac{1}{\beta} \frac{1}{\|\mathbf{w}^1\|^2_{\boldsymbol{H}}} \big(((C \otimes P_1 + D \otimes P_2)(I^s \otimes P_1)\mathbf{w}^1, \mathbf{w}^1) \\
&\qquad + \delta t \left((C \otimes P_1 + D \otimes P_2)(A \otimes P_1\mathbb{M} + \tilde{A} \otimes P_2\mathbb{M})\mathbf{w}^1, \mathbf{w}^1 \right) \big) \\
&= \frac{1}{\beta} \frac{((C \otimes P_1)\mathbf{w}^1, \mathbf{w}^1) + \delta t \left((CA \otimes \mathbb{M})\mathbf{w}^1, \mathbf{w}^1 \right)}{\|\mathbf{w}^1\|^2_{\boldsymbol{H}}} && (3.21) \\
&\geq \frac{\gamma_1\alpha}{\beta} \frac{\|\mathbf{v}^1\|^2 + \delta t \, \|\nabla \mathbf{w}^1\|^2}{\|\mathbf{w}^1\|^2_{\boldsymbol{H}}} && (3.22) \\
&= \frac{\gamma_1\alpha}{\beta} > 0,
\end{aligned}
$$

where $\beta = \max(\|C\|, \|D\|)$. The step (3.21) is just insertion of $D = CA\tilde{A}^{-1}$, while (3.22) is justified by Lemma 3.7.□

We remark that the proof of Lemma 3.8 is similar to the proof of the inf–sup condition in [14]. The main differences are the following. The laplace operator Δ in [14] is switched to the elliptic operator $\mathbb{M}$ in this paper. The equation in [14] is not a PDAE and therefore the projection operators P_1 and P_2 are unnecessary, which simplifies the proof. Furthermore, the corresponding construction of $\mathbf{w}^1$ is simpler in the parabolic case because we do not have to handle the $\tilde{A}$–matrix. Finally, Lemma 3.7 is trivially true in [14].

LEMMA 3.9. *For any* $\mathbf{w}^1 \in H$ *there exists a* $\mathbf{w}^2 \in \boldsymbol{H}$ *such that*

$$(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2) \neq 0. \tag{3.3}$$

Proof. First we notice that A and A^T share the same set of eigenvalues, and therefore A^T is also weakly positive definite. By Lemma 3.5, we know that there exist a C_2 such that C_2 and C_2A^T are positive definite. Further, let $\mathbf{w}^2$ be given, and let $\mathbf{w}^1 = (C_2^T \otimes P_1 + D_2^T \otimes P_2)\mathbf{w}^2$, where $D_2^T = \tilde{A}^{-1}AC_2^T$. Then we have

$$\begin{aligned}
(\mathcal{A}\mathbf{w}^1, \mathbf{w}^2) &= (\mathcal{A}(C_2^T \otimes P_1 + D_2^T P_2)\mathbf{w}^2, \mathbf{w}^2) \\
&= \Big((I^s \otimes P_1 + \delta t\,(A \otimes P_1\mathbb{M} + \tilde{A} \otimes P_2\mathbb{M})(C_2^T \otimes P_1 + D_2^T \otimes P_2)\mathbf{w}^2, \mathbf{w}^2\Big) \\
&= (\mathbf{v}^2, C_2\mathbf{v}^2) + ((AC_2^T \otimes P_1\mathbb{M} + \tilde{A}D_2^T \otimes P_2\mathbb{M})\mathbf{w}^2, \mathbf{w}^2) \\
&\geq (\mathbf{v}^2, C_2\mathbf{v}^2) + (\nabla\mathbf{w}^2, (C_2A^T \otimes I^{2d})\nabla\mathbf{w}^2) > 0.
\end{aligned}$$

□

4. The preconditioner.

In the following we construct the preconditioner for the continuous operator $\mathcal{A}$ based on the proper Sobolev spaces. The discrete preconditioner can be viewed as an operator acting on the discrete subspaces. This approach was also taken in [4, 11, 14, 15].

4.1. Block preconditioner for the Runge–Kutta discretization.

In previous works on block preconditioners for Runge–Kutta discretizations of the parabolic PDEs [14, 19], a block Jacobi and a block Gauss–Seidel preconditioner are presented. Here, we have extended this work to the Runge–Kutta discretization of the bidomain equations. We prove that the block Jacobi preconditioner is order–optimal and demonstrate this with numerical experiments. Additional numerical experiments for both the block Jacobi and the block Gauss–Seidel preconditioner for Runge–Kutta discretizations of the bidomain equations can be found in [18].

Above, we showed that the continuous operator $\mathcal{A}$ was an isomorphism mapping $\boldsymbol{H}$ to $\boldsymbol{H}^*$. Therefore, let $\mathcal{B} \in \mathcal{L}(\boldsymbol{H}^*, \boldsymbol{H})$ be an isomorphism i.e.,

$$\|\mathcal{B}\|_{\mathcal{L}(\boldsymbol{H}^*,\boldsymbol{H})} \leq d_1, \quad \text{and} \quad \|\mathcal{B}^{-1}\|_{\mathcal{L}(\boldsymbol{H},\boldsymbol{H}^*)} \leq d_2, \tag{4.1}$$

where d_1 and d_2 are independent of δt. Then, by applying the preconditioner to the continuous operator, we have that

$$\mathcal{B}\mathcal{A} : \boldsymbol{H} \to \boldsymbol{H}$$

and

$$\|\mathcal{B}\mathcal{A}\|_{\mathcal{L}(\boldsymbol{H},\boldsymbol{H})} \leq c_1 d_1, \quad \text{and} \quad \|(\mathcal{B}\mathcal{A})^{-1}\|_{\mathcal{L}(\boldsymbol{H},\boldsymbol{H})} \leq c_2 d_2.$$

Consequently the condition number of the continuous preconditioned system is bounded by

$$\kappa(\mathcal{B}\mathcal{A}) = \|\mathcal{B}\mathcal{A}\|_{\mathcal{L}(\boldsymbol{H},\boldsymbol{H})}\|(\mathcal{B}\mathcal{A})^{-1}\|_{\mathcal{L}(\boldsymbol{H},\boldsymbol{H})} \leq c_1 d_1 c_2 d_2. \tag{4.2}$$

We now suggest a preconditioner for (2.11), which is a block Jacobi preconditioner, given by

$$\mathcal{B}^{-1} = I^s \otimes P_1 + \delta t \left(\operatorname{diag} A \otimes P_1 \mathbb{M} + \operatorname{diag} \tilde{A} \otimes P_2 \mathbb{M} \right). \tag{4.3}$$

The block Jacobi preconditioner (4.3) is basically a Riesz identity mapping from the dual space $\boldsymbol{H}^*$ to $\boldsymbol{H}$. However, $\mathcal{B}$ is the exact preconditioner which will not be used in practice. Instead we use an approximation of these operators which is fast to evaluate. We use an algebraic multigrid preconditioner. Another choice could have been a domain decomposition approximation of the exact operator. Or if higher–order elements are used in space, the preconditioner could be approximated using lower–order elements, yielding an efficient preconditioner.

We will now relate these results to the discrete case. Let Ω_h be the discretized and meshed counterpart of Ω. We choose the discrete spaces as above, $\boldsymbol{V}_h \times \boldsymbol{U}_h \subset \boldsymbol{H}$. Since the discrete space $\boldsymbol{V}_h \times \boldsymbol{U}_h$ is a subset of $\boldsymbol{H}$, we get that

$$\kappa(\mathcal{B}_h \mathcal{A}_h) \leq \kappa(\mathcal{B}\mathcal{A}).$$

Finally we need to incorporate the effects from the approximation $\tilde{\mathcal{B}}_h$ of the exact preconditioner $\mathcal{B}_h$. By using a simple Cauchy–Schwarz like argument, we can estimate the preconditioning effect of the approximated preconditioner by

$$\kappa(\tilde{\mathcal{B}}_h \mathcal{A}_h) \leq \kappa(\tilde{\mathcal{B}}_h \mathcal{B}_h^{-1}) \kappa(\mathcal{B}_h \mathcal{A}_h). \tag{4.4}$$

We need to estimate the deterioration of the preconditioner due to the approximation of the exact preconditioner, $\kappa(\tilde{\mathcal{B}}_h \mathcal{B}_h^{-1})$. This can be done easily for the block Jacobi preconditioner (4.3). Let $\tilde{\mathcal{B}}_{h,i}$ be a cheap approximation of the block Jacobi preconditioner $\mathcal{B}_{h,i}$ for the bidomain equations cf. [13],

$$\mathcal{B}_{h,i}^{-1} = P_{1,h} + \delta t\, a_{ii} \mathbb{M}_h.$$

$$\mathcal{B}_h = \begin{pmatrix} \mathcal{B}_{h,1} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \mathcal{B}_{h,s} \end{pmatrix}.$$

Assume that $\tilde{\mathcal{B}}_{h,i}$ is constructed such that it is symmetric and spectrally equivalent to $\mathcal{B}_{h,i}$, i.e.

$$c_3(\tilde{\mathcal{B}}_{h,i} w_{h,i}, w_{h,i}) \leq (\mathcal{B}_{h,i} w_{h,i}, w_{h,i}) \leq c_4(\tilde{\mathcal{B}}_{h,i} w_{h,i}, w_{h,i}), \quad w_{h,i} \in V_h \times U_h \subset H, \tag{4.5}$$

where c_3 and c_4 is chosen such that (4.5) valid for $i = 1, \dots, s$.

We remark that the diagonal blocks $\mathcal{B}_{h,i}$ of the Jacobi preconditioner is similar to a block diagonal preconditioner for the Euler or Crank–Nicolson schemes (see [13] and Lemma 3.3). This is an implementational advantage, since the preconditioner can be developed for the lower–order temporal discretizations, and later be reused in the same form on the diagonal blocks for the higher–order Runge–Kutta preconditioner.

By combining (4.2), (4.5) and (4.4), we find that

$$\kappa(\tilde{\mathcal{B}}_h \mathcal{A}_h) \leq \frac{c_1 d_1 c_2 d_2 c_4}{c_3},$$

which states that the preconditioner is bounded independent of the spatial discretization parameter h, and the time step δt.

5. Weakly positive definite matrices. The purpose of this section is to prove Lemma 3.5. This will be done by first proving two lemmas, Lemma 5.1 and Lemma 5.2, and then proving Lemma 3.5.

LEMMA 5.1. *Any $A \in \mathbb{R}^{2\times 2}$ with nonreal eigenvalues is weakly positive definite.*

Proof. In this proof we construct a positive definite C, such that CA is positive definite. To do so, we study the angle between x and Ax, written

$$\alpha(x) \equiv \angle(x, Ax) : \mathbb{R}^2 \to (-\pi, \pi].$$

Notice that if there exists an x such that $\alpha(x) = 0$, then A will have a positive real eigenvalue, which is against the assumption. Similarly, if there exists an x with $\alpha(x) = \pi$, then there will be a negative eigenvalue of A. Thus $\alpha(x) \notin \{0, \pi\}$ for $x \neq 0$.

Notice further that $\alpha(x)$ is a continuous function of $x \neq 0$. This gives that for a given A we will either have

$$\alpha(x) \in (0, \pi) \quad \text{or} \quad \alpha(x) \in (-\pi, 0) \quad \forall x \neq 0.$$

Assume that

$$\alpha(x) \in (0, \pi), \quad \forall x \neq 0,$$

and let

$$\theta = \sup_{x \neq 0} \alpha(x).$$

In the following we show that $\theta < \pi$. We have that $\alpha(x) = \angle(x, Ax) = \angle(\frac{x}{\|x\|}, \frac{Ax}{\|x\|})$, because changing the length of vectors does not change the angle between them. Therefore

$$\sup_{x \neq 0} \alpha(x) = \sup_{\|x\|=1} \alpha(x).$$

Since the set defined by $\|x\| = 1$ is a compact set, and $\alpha(x)$ is a continuous function of x, we can conclude that the supremum value is attained and that it is less than π, i.e.

$$\theta = \max_{x \neq 0} \alpha(x) \in (0, \pi).$$

Then set $C = R_{-\theta/2}$ where $R_{-\theta/2}$ is the rotation matrix with an angle $-\frac{\theta}{2}$ i.e.,

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{5.1}$$

We have that

$$\begin{gathered} \angle(x, Ax) \in (0, \theta), \quad \forall x \neq 0, \\ \Downarrow \\ \angle(x, CAx) \in (-\frac{\theta}{2}, \frac{\theta}{2}) \subset (-\frac{\pi}{2}, \frac{\pi}{2}), \quad \forall x \neq 0. \end{gathered}$$

And further since $\angle(x, Cx) = -\frac{\theta}{2} \in (-\frac{\pi}{2}, 0)$, $\forall x \neq 0$, we get

$$\begin{aligned} x^T C x > 0, &\quad \forall x \neq 0 \\ x^T CAx > 0, &\quad \forall x \neq 0, \end{aligned}$$

which means that both C and CA are positive definite.

Finally, if

$$\alpha(x) \in (-\pi, 0), \quad \forall x \neq 0,$$

we define $\theta = \inf \alpha(x)$ and the result can be proved similarly. □

The next lemma proves that a special block matrix can be made positive definite. This result will be used as an induction step for extending the above 2×2–result to Lemma 3.5.

LEMMA 5.2. *Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$ be positive definite. For any $C \in \mathbb{R}^{n\times m}$ there exists an $\epsilon > 0$ such that the block matrix*

$$E = \begin{pmatrix} \epsilon A & \epsilon C \\ 0 & B \end{pmatrix}$$

is positive definite.

Proof. Since $S(A)$ is symmetric positive definite, $S(A)^\gamma$ is defined for $\gamma \in \mathbb{R}$, cf. [23]. Further let

$$\alpha = \min_{\|x\|=1} x^T A x,$$
$$\beta = \min_{\|y\|=1} y^T B y.$$

Both α and β exist and are positive because the objective functions are continuous and positive, plus that the admissible sets are compact.

Let $z^T = (x^T, y^T)$. We have

$$\begin{aligned} z^T E z &= \epsilon x^T S(A) x + \epsilon x^T C y + y^T B y \\ &= \|(\epsilon S(A))^{\frac{1}{2}} x + \frac{\epsilon}{2}(\epsilon S(A))^{-\frac{1}{2}} C y\|^2 - \frac{\epsilon}{4} y^T C^T (S(A))^{-1} C y + y^T B y \\ &\geq -\frac{\epsilon}{4} y^T C^T A^{-1} C y + y^T B y \\ &\geq \left(\beta - \frac{\epsilon}{4}\frac{\|C\|^2}{\alpha}\right)\|y\|^2 \qquad (5.2) \\ &> 0, \end{aligned}$$

when $\epsilon < \frac{4\alpha\beta}{\|C\|^2}$ and $y \neq 0$. Here (5.2) follows by the definitions of α and β. □

We are now ready to prove Lemma 3.5. In Lemma 5.2 m and n are arbitrary positive integers, but in the following proof only $n \leq 2$ is needed.

Proof. [Proof of Lemma 3.5] First, we prove that a weakly positive definite matrix can not have negative eigenvalues. To see this, assume that C is positive definite and that A has a negative eigenvalue, $Ax = -\lambda x$, $\lambda \in \mathbb{R}^+$ where x is an eigenvector. Then

$$x^T C A x = -\lambda x^T C x < 0,$$

and therefore CA is not positive definite.

Next, we show that if A has no real negative eigenvalues, then A is weakly positive definite. This is done by constructing a positive definite matrix C, such that CA is positive definite. The construction utilizes the Schur decomposition.

Let $A = QTQ^T$ be the real Schur decomposition of A, where Q is orthogonal and T is a block[4] upper triangular matrix with 1×1 or 2×2 blocks on the diagonal, see

[4]The standard Schur decomposition has an upper (not block) triangular matrix T, but is in general complex.

[23]. The real eigenvalues of A, which are positive, can be found on the 1×1 blocks on the diagonal of T, and the nonreal eigenvalues can be found as the eigenvalues of the 2×2 blocks on the diagonal of T. T may be written

$$T = \begin{pmatrix} T_1 & \multicolumn{3}{c}{\boxed{\quad t_1 \quad}} \\ 0 & T_2 & \multicolumn{2}{c}{\boxed{\quad t_2 \quad}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & T_k \end{pmatrix}$$

where t_i denote the nonzero offdiagonal (possibly nonsquare) blocks, i.e. t_i is the matrix consisting of all the entries to the right of T_i.

The diagonal blocks of T are called T_i. An important observation now is that each T_i is weakly positive definite. This is trivially true if T_i is a positive real number, and shown by Lemma 5.1 if T_i has truly complex eigenvalues.

We now construct

$$C = QDQ^T,$$

where D is block diagonal, and where the diagonal blocks are denoted D_i. Then each D_i should have the same dimension as T_i. We get that

$$CA = QDTQ^T.$$

In the following we show that the D_is can be constructed in such a way that both D and DT are positive definite. This is further equivalent to both C and CA being positive definite, since positive definite matrices are invariant under orthogonal change of basis. Note that D is positive definite if and only if all the D_is are.

By multiplying D and T, we see that

$$\begin{aligned} DT &= \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & D_k \end{pmatrix} \begin{pmatrix} T_1 & \multicolumn{3}{c}{\boxed{\quad t_1 \quad}} \\ 0 & T_2 & \multicolumn{2}{c}{\boxed{\quad t_2 \quad}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & T_k \end{pmatrix} \\ &= \begin{pmatrix} D_1T_1 & \multicolumn{3}{c}{\boxed{\quad D_1t_1 \quad}} \\ 0 & D_2T_2 & \multicolumn{2}{c}{\boxed{\quad D_2t_2 \quad}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & D_kT_k \end{pmatrix}. \end{aligned} \tag{5.3}$$

Therefore, to finish this proof we start in the lower right corner of DT and show by induction that positive definite D_i blocks can be constructed to make the block matrix (5.3) positive definite. Since T_k is weakly positive definite, there exists a positive definite D_k such that the lower right block D_kT_k is positive definite. Similarly, there exists a positive definite $\tilde{D}_{k-1}$ such that $\tilde{D}_{k-1}T_{k-1}$ is positive definite. Choosing $D_{k-1} = \epsilon\tilde{D}_{k-1}$, Lemma 5.2 gives us that the extended lower right block

$$\begin{pmatrix} \epsilon\tilde{D}_{k-1}T_{k-1} & \epsilon\tilde{D}_{k-1}t_{k-1} \\ 0 & D_kT_k \end{pmatrix} \tag{5.4}$$

is positive definite (for any t_{k-1}) if ϵ is small enough.

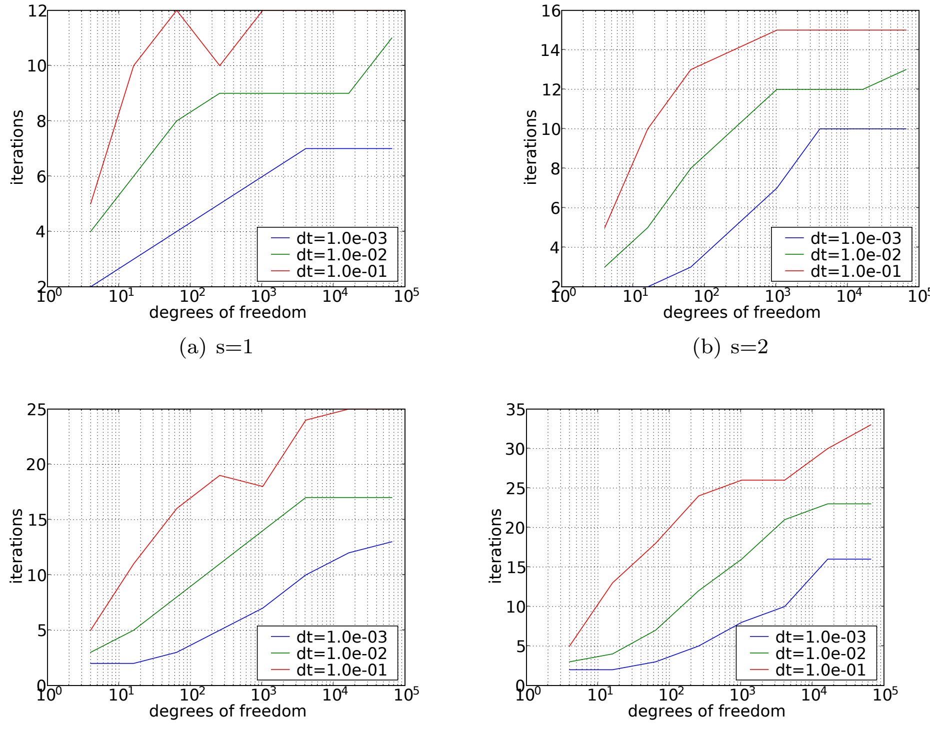
By Lemma 5.2 and induction we see that the system (5.4) can be extended to the full block system (5.3), and the proof is complete. □

6. Numerical Results. The purpose of this section is to show numerically the performance of the preconditioners. All implementation is done in the framework of PyCC [2], which is a Python library interfacing compiled packages for matrix–storage, preconditioners, element discretization, etc. It supports higher–order Lagrange elements [7], generated using SyFi [3]. We will use the ML [1] algebraic multigrid package to approximate the exact preconditioner, using one V–cycle with symmetric Gauss–Seidel as point smoother.

6.1. The order–optimality for the Runge–Kutta block preconditioners. We proved above that the block Jacobi preconditioner is order–optimal with respect to the spatial discretization parameter h and the time step δt, assuming that the single blocks of the preconditioner are order–optimal as in (4.5). Therefore, we test the efficiency of the preconditioner on one time step of (1.6)–(1.7) using (2.7) on a regular 2D grid with Lagrangian finite elements of various orders. We solve the linear problem using left preconditioned BiCGStab with a relative tolerance of 10^{-5} for the residual in L_2 norm. The start vector is an oscillatory random vector. The result can be seen in Figure 6.1. It seems clear that for each s the number of iterations to reach convergence is small and bounded. Furthermore, we demonstrate the same behavior with third and fifth order elements in space in Figure 6.2 and Figure 6.3, respectively. In all cases, the number of iterations is small. The number of iterations increases slightly as the number of degrees of freedom increases. However, notice that this increase is also present for $s = 1$. Hence, we attribute this behavior to the algebraic multigrid preconditioner which is only almost order–optimal. We have confirmed this behavior with similar experiments on the Poisson equation and one time step of a heat equation. It is also clear that the AMG preconditioner performance deteriorates as the order of the element increases. In fact, fifth order elements typically require about twice as many iterations as linear elements in our experiments. We remark that the number of iterations decreases as δt decreases. This is to be expected for low–order time discretizations of parabolic PDEs since the matrix becomes more diagonal dominant, but it appears to be valid also in this more complicated setting.

Here, we have only solved the discrete system for one time step. The efficiency of higher–order methods in space and time depends on the fact that larger time steps δt and characteristic mesh size h can be used when the solution is smooth in space and time. Furthermore, due to a higher–order polynomial approximation, the accuracy of the approximated solution is of higher order. Therefore, higher–order methods, given order–optimal preconditioners, are beneficial if the solution is smooth and the accuracy requirements are high. In [18] it is demonstrated a speed–up by a factor $\approx 10^4$ when comparing 4th order Lagrangian elements and RadauIIA with four stage variables and linear Lagrangian elements with RadauIIA with one stage variable in a concrete case with a smooth solution. We do not repeat such experiments here. More numerical examples involving the efficiency of the AMG preconditioner for higher–order elements, the increased efficient of various block Gauss–Seidel preconditioners can also be found here. Finally, in [19] we discussed techniques to lessen the s–dependency considerably for parabolic PDEs. In [18], we did not manage the same performance boost for the bidomain equations, probably because the conductivity tensors and AMG makes the problem more complicated. We do however believe that it is possible to improve the dependency of s considerably by taking advantage of the structure of the Runge–Kutta coefficient matrix.

The numerical experiments in this paper concern the RadauIIA method. Similar results have been obtained with the LobattoIIIC method, but these are not presented



(a) s=1

(b) s=2

(c) s=3

(d) s=4

FIGURE 6.1. *The number of iterations required for BiCGStab to reach a relative tolerance of* 10^{-5} *for the 2D problem* (1.6)–(1.7) *using linear finite elements in space and the RadauIIA scheme with one to four nodes in time.*

in order to limit the number of figures.

7. Final remarks. In this paper we have presented a block diagonal preconditioner for the fully implicit Runge–Kutta discretization of the bidomain equations. We have also shown that if the Runge–Kutta scheme is irreducible, A–stable and has an invertible A–matrix, then an order–optimal block diagonal preconditioner can be constructed based on an order–optimal preconditioner for the implicit Euler discretization. Such order–optimal preconditioner was presented in [13]. The order–optimality is confirmed by numerical experiments using algebraic multigrid.

REFERENCES

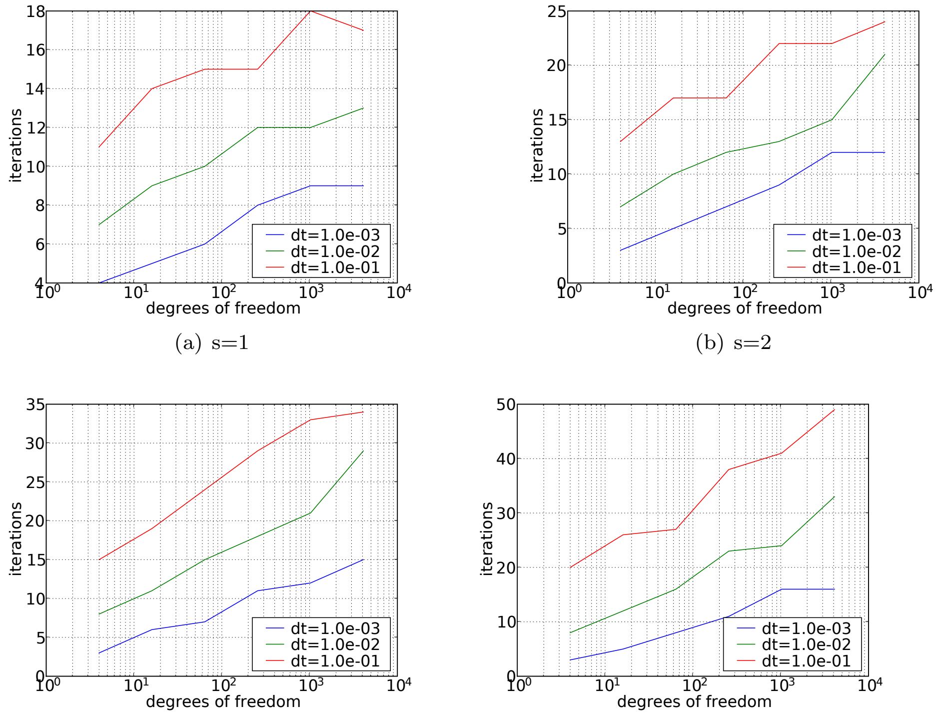
[1] ML – Multilevel Preconditioning Package. http://software.sandia.gov/trilinos/packages/ml/index.html.

[2] PyCC – Python Cardiac Computation. http://heim.ifi.uio.no/∼skavhaug/heart_simulations.html.

[3] SyFi – Symbolic Finite Elements. http://www.fenics.org/syfi.

[4] D.N. Arnold, R.S. Falk, and R. Winther. Preconditioning discrete approximations of the Reissner–Mindlin plate model. *Math. Modelling Numer. Anal.*, 31:517–557, 1997.

[5] I. Babuska and A. K. Aziz. Survey lectures on the mathematical foundations on the finite element method. In A. K. Aziz, editor, *The Mathematical Foundation of the Finite Element*



(a) s=1 (b) s=2

(c) s=3 (d) s=4

FIGURE 6.2. *The number of iterations required for BiCGStab to reach a relative tolerance of* 10^{-5} *for the 2D problem* (1.6)–(1.7) *using third order finite elements in space and the RadauIIA scheme with one to four nodes in time.*

Method with Applications to Partial Differential Equations, pages 3–363. Academic Press, New York - London, 1972.

[6] J. Bergh and J. Löfström. *Interpolation spaces.* Springer-Verlag, 1976.

[7] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods.* Number 15 in Texts in Applied Mathematics. Springer, 2nd edition, 2002.

[8] A. Graham. *Kronecker Products and Matrix Calculus With Applications.* Halsted Press,John Wiley and Sons, 1981.

[9] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I - Nonstiff Problems.* Springer Verlag, 2nd edition, 1992.

[10] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II - Stiff and Differential-Algebraic Problems.* Springer Verlag, 2nd edition, 1996.

[11] E. Haug and R. Winther. A domain embedding preconditioner for the Lagrange multiplier system. *Math. Comp*, 69:65–82, 1999.

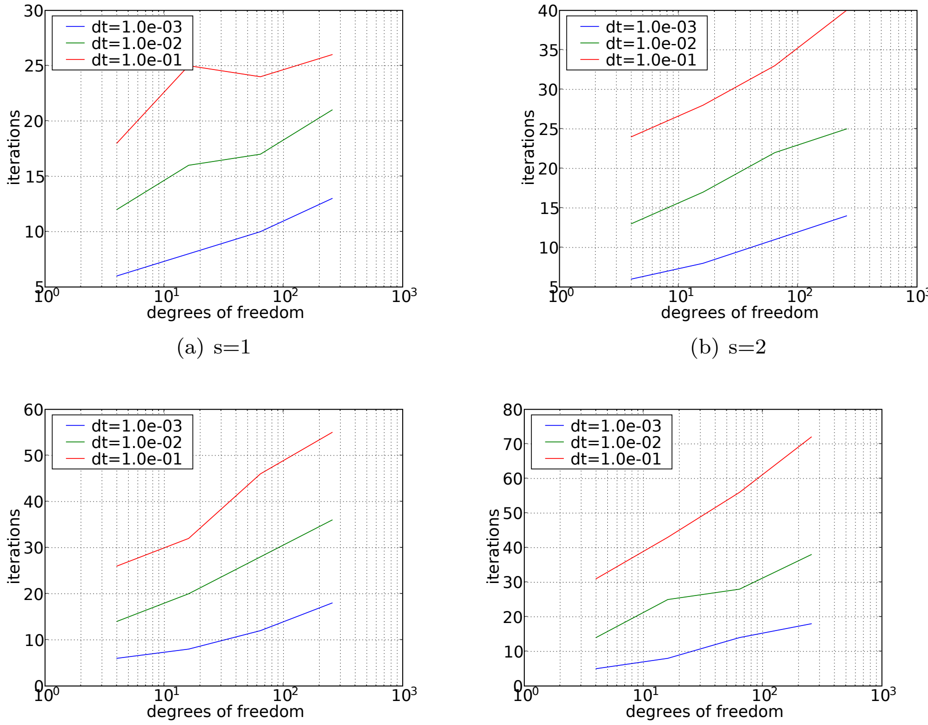
[12] H. Hu. A projective method for rescaling a diagonally stable matrix to be positive definite. *SIAM J. Matrix Anal. Appl.*, 13(4):1255–1263, 1992.

[13] K.-A. Mardal, B. F. Nielsen, X. Cai, and A. Tveito. An order optimal solver for the discretized Bidomain equations. *Numerical Linear Algebra with Applications*, 14(2):83–98, 2007.

[14] K.-A. Mardal, T. K. Nilssen, and G. A. Staff. Order optimal preconditioners for implicit Runge-Kutta schemes applied to parabolic PDE's. *SIAM Journal of Scientific Computing*, 29(1):361–375, 2007.

[15] K.-A. Mardal and R. Winther. Uniform preconditioners for the time dependent Stokes problem. *Numer. Math.*, 98(2):305–327, 2004.

[16] T. K. Nilssen. Weakly positive definite matrices. Simula Research Laboratory, Research Report 07–2005.
URL: http://www.simula.no/departments/scientific/publications/Nilssen.2005.1.



(a) s=1 (b) s=2

(c) s=3 (d) s=4

FIGURE 6.3. *The number of iterations required for BiCGStab to reach a relative tolerance of* 10^{-5} *for the 2D problem* (1.6)–(1.7) *using fifth order finite elements in space and the RadauIIA scheme with one to four nodes in time.*

[17] G. Plank, R. Weber dos Santos M. Liebmann, E. J. Vigmond, and G. Haase. Algebraic multigrid preconditioner for the cardiac bidomain model. *IEEE Trans Biomed Eng.*, 54(4):585–596, 2007.

[18] G. A. Staff, K.-A. Mardal, and T. K. Nilssen. Block preconditioners for fully implicit Runge-Kutta schemes applied to the Bidomain equations. In J. Priaux P. Wesseling, E. Oate, editor, *ECCOMAS CFD 06*, pages 1–15. European Community on Computational Methods in Applied Sciences, 2006. The proceeding is published electronically.

[19] G. A. Staff, K.-A. Mardal, and T. K. Nilssen. Preconditioning of fully implicit Runge-Kutta schemes for parabolic PDEs. *Modeling, Identification and Control*, 27(2):109–123, 2006.

[20] J. Sundnes, G. T. Lines, X. Cai, B. F. Nielsen, K.-A. Mardal, and A. Tveito. *Computing the electrical activity in the heart*. Springer-Verlag, 2006.

[21] J. Sundnes, G. T. Lines, and A. Tveito. An operator splitting method for solving the Bidomain equations coupled to a volume conductor model for the torso. *Mathematical biosciences*, 194(2):233–248, 2005.

[22] J. Sundnes, B. F. Nielsen, K.-A. Mardal, X. Cai, G. T. Lines, and A. Tveito. On the computational complexity of the bidomain and the monodomain models of electrophysiology. *Annals of biomedical engineering*, 34(7):1088–1097, 2006.

[23] L. N. Trefethen and D. Bau III. *Numerical Linear Algebra*. SIAM, 1997.

[24] L. Tung. *A Bi-domain model for describing ischemic myocardial D-C potentials*. PhD thesis, MIT, Cambridge, 1978.

[25] R. L. Winslow, J. Rice, S. Jafri, E Marban, and B O'Rourke. Mechanisms of altered excitation-contraction coupling in canine tachycardia-induced heart failure, II: Model studies. *Circ Res*, 84:571–586, 1999.