Helsinki University of Technology Institute of Mathematics Research Reports
Espoo 2008 A552

A DISCRETE MAXIMUM PRINCIPLE IN HILBERT SPACE WITH APPLICATIONS TO NONLINEAR COOPERATIVE ELLIPTIC SYSTEMS

János Karátson Sergey Korotov

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology Institute of Mathematics Research Reports
Espoo 2008 A552

A DISCRETE MAXIMUM PRINCIPLE IN HILBERT SPACE WITH APPLICATIONS TO NONLINEAR COOPERATIVE ELLIPTIC SYSTEMS

János Karátson Sergey Korotov

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Mathematics and Systems Analysis

János Karátson, Sergey Korotov: *A discrete maximum principle in Hilbert space with applications to nonlinear cooperative elliptic systems*; Helsinki University of Technology Institute of Mathematics Research Reports A552 (2008).

Abstract: *Discrete maximum principles are derived for finite element discretizations of nonlinear elliptic systems with cooperative and weakly diagonally dominant coupling. The results are achieved via a Hilbert space background and, in the case of simplicial elements, are obtained under weakened acute type conditions for the FEM meshes.*

AMS subject classifications: 65N30, 65N50

Keywords: elliptic system, maximum principles, finite element method, discrete maximum principle

Correspondence

János Karátson
Department of Applied Analysis and Computational Mathematics, Eötvös Loránd University; H-1518, Budapest, Pf. 120, Hungary

Sergey Korotov
Institute of Mathematics, Helsinki University of Technology
P.O. Box 1100, FI-02015 TKK, Finland

karatson@cs.elte.hu, sergey.korotov@hut.fi

ISBN 978-951-22-9514-2 (print)
ISBN 978-951-22-9515-9 (PDF)
ISSN 0784-3143 (print)
ISSN 1797-5867 (PDF)

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Mathematics and Systems Analysis
P.O. Box 1100, FI-02015 TKK, Finland
email: math@tkk.fi http://math.tkk.fi/

1 Introduction

The maximum principle forms an important qualitative property of second order elliptic equations [24, 28], hence its discrete analogues (so-called discrete maximum principles, DMPs) have drawn much attention. Various DMPs, including geometric conditions on the computational meshes for FEM solutions, have been given e.g. in [6, 8, 9, 10, 17, 21, 25, 29, 31, 33] for linear and [18, 19, 22] for nonlinear equations. For elliptic operators with only principal part, if the discretized operator L_h and the FEM solution u_h satisfy $L_h u_h \leq 0$, then the DMP has the simple form $\max\limits_{\overline{\Omega}} u_h = \max\limits_{\partial\Omega} u_h$. On the other hand, for operators with lower order terms as well, one has the weaker statement

$$\max_{\overline{\Omega}} u_h = \max\{0, \max_{\partial\Omega} u_h\}. \tag{1}$$

There are also typical differences in the corresponding conditions on the meshes, as already pointed out in [10]. Namely, for operators with only principal part, the DMP holds for all meshes under sufficient conditions that express nonobtuseness when linear finite elements and simplicial meshes are considered; on the other hand, for operators with lower order terms as well, one can only provide the DMP for sufficiently fine mesh and needs stronger acuteness type conditions in the case of simplicial FEM meshes.

The extension of the (continuous) maximum principle, CMP, to elliptic systems has attracted much interest, but it has been proved to hold under strong restrictions only. The main class of problems where the CMP is generally valid is that of cooperative systems: roughly speaking, writing the system in the form

$$L\mathbf{u} = M\mathbf{u} + \mathbf{f}$$

(where L is an s-tuple of minus Laplacians or of more general elliptic operators, further, $\mathbf{u} = (u_1, \ldots, u_s)$ and M is an $s \times s$ matrix), this condition means $M_{ij} \geq 0$ for all $i \neq j$. In addition, one usually also assumes weak diagonal dominance of M. Important related results are found e.g. in [12, 26, 27], and some extensions to non-cooperative systems are also known, see [7] and references therein. These results are formulated either in terms of maximum principle or of nonnegativity preservation (i.e. $\mathbf{f} \geq 0$ implying $\mathbf{u} \geq 0$).

The goal of this paper is to provide DMPs for the FEM discretizations of some elliptic systems, which has not yet been done to our knowledge. Moreover, we consider nonlinear systems. We generalize our results in [18] to systems with cooperative and weakly diagonally dominant coupling, and under suitably weakened acuteness type conditions, we thus obtain an analogue of (1) for sufficiently fine meshes. We include the lower order terms in the elliptic operator, hence cooperativity will mean nonpositive cross-signs instead of nonnegative.

The main technical difficulties encountered are as follows. First, one has to get round the irreducibility criterion that is assumed in the basic algebraic background statement. Second, when lower order terms of polynomial growth are involved, one needs careful estimates using embedding results and quasi-regular meshes to ensure the required algebraic properties of the stiffness matrix. In order to provide a clean line of thoughts, we therefore first state a DMP in a Hilbert space setting, which helps to derive the corresponding results under the considered different conditions.

The paper is organized as follows. The required algebraic background is briefly summarized in section 2. A DMP for suitable operator equations in a Hilbert space setting is given in section 3. The main results are given in section 4, where DMPs are derived for three types of nonlinear elliptic systems with cooperative and weakly diagonally dominant coupling. The considered types of problems are systems with nonlinear coefficients, systems with lower order terms of sublinear and polynomial growth, respectively. An analogue of (1) is proved under suitably weakened acuteness type conditions for the FEM mesh, which in principle allow the angles to reach asymptotically 90°. The latter and other geometric issues are also discussed. Finally some applications are sketched in section 5.

2 Some algebraic background

First we recall a basic definition in the study of DMP (cf. [32, p. 23]):

Definition 2.1 A square $k \times k$ matrix $\mathbf{A} = (a_{ij})_{i,j=1}^k$ is called *irreducibly diagonally dominant* if it satisfies the following conditions:

(i) $\mathbf{A}$ is irreducible, i.e., for any $i \neq j$ there exists a sequence of nonzero entries $\{a_{i,i_1}, a_{i_1,i_2}, \dots, a_{i_s,j}\}$ of A, where $i, i_1, i_2, \dots, i_s, j$ are distinct indices,

(ii) $\mathbf{A}$ is diagonally dominant, i.e., $|a_{ii}| \geq \sum\limits_{\substack{j=1\\ j\neq i}}^{k} |a_{ij}|,\ i = 1, ..., k,$

(iii) for at least one index $i_0 \in \{1, ..., k\}$ the above inequality is strict, i.e.,

$$|a_{i_0,i_0}| > \sum_{\substack{j=1\\ j\neq i_0}}^{k} |a_{i_0,j}|.$$

Definition 2.2 Let $\mathbf{A}$ be an arbitrary $k \times k$ matrix. The *irreducible blocks* of $\mathbf{A}$ are the matrices $\mathbf{A}^{(l)}$ $\quad (l = 1, \dots, q)$ defined as follows.

Let us call the indices $i, j \in \{1, \dots, k\}$ *connectible* if there exists a sequence of nonzero entries $\{a_{i,i_1}, a_{i_1,i_2}, \dots, a_{i_s,j}\}$ of $\mathbf{A}$, where $i, i_1, i_2, \dots, i_s, j \in \{1, \dots, k\}$ are distinct indices. Further, let us call the indices i, j mutually connectible if both i, j and j, i are connectible in the above sense. (Clearly, mutual connectibility is an equivalence relation.) Let $N_1, \dots, N_q$ be the equivalence classes, i.e. the maximal sets of mutually connectible indices. (Clearly, $\mathbf{A}$ is irreducible iff $q = 1$.) Letting

$$N_l = \{s_1^{(l)}, \dots, s_{k_l}^{(l)}\}$$

for $l = 1, \ldots, q$, we have $k_1 + \ldots + k_q = k$. Then we define for all $l = 1, \ldots, q$ the $k_l \times k_l$ matrix $\mathbf{A}^{(l)}$ by

$$\mathbf{A}^{(l)}_{pq} := a_{s_p^{(l)}, s_q^{(l)}} \qquad (p, q = 1, \ldots, k_l).$$

Remark 2.1 One may prove (cf. [2, Th. 4.2]) that by a proper permutation of indices, $\mathbf{A}$ becomes a block lower triangular matrix with the irreducible diagonal blocks $\mathbf{A}^{(l)}$.

Let us now consider a system of equations of order $(k+m) \times (k+m)$:

$$\bar{\mathbf{A}}\bar{\mathbf{c}} = \bar{\mathbf{b}}, \tag{2}$$

where the matrix $\bar{\mathbf{A}}$ and the vectors $\bar{\mathbf{b}}$, $\bar{\mathbf{c}}$ have the following structure:

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \tilde{\mathbf{A}} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \qquad \bar{\mathbf{b}} = \begin{bmatrix} \mathbf{b} \\ \tilde{\mathbf{b}} \end{bmatrix}, \qquad \bar{\mathbf{c}} = \begin{bmatrix} \mathbf{c} \\ \tilde{\mathbf{c}} \end{bmatrix} \tag{3}$$

where $\mathbf{I}$ is the $m \times m$ identity matrix and $\mathbf{0}$ is the $m \times k$ zero matrix. Then (2) becomes

$$\begin{bmatrix} \mathbf{A} & \tilde{\mathbf{A}} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \tilde{\mathbf{c}} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \tilde{\mathbf{b}} \end{bmatrix}. \tag{4}$$

Following [9], we introduce

Definition 2.3 A $(k+m) \times (k+m)$ matrix $\bar{\mathbf{A}}$ with the structure (3) is said to be of *generalized nonnegative type* if the following properties hold:

(i) $a_{ii} > 0, \quad i = 1, ..., k,$

(ii) $a_{ij} \leq 0, \quad i = 1, ..., k, \ j = 1, ..., k+m \quad (i \neq j),$

(iii) $\sum_{j=1}^{k+m} a_{ij} \geq 0, \quad i = 1, ..., k,$

(iv) There exists an index $i_0 \in \{1, \ldots, k\}$ for which

$$\sum_{j=1}^{k} a_{i_0,j} > 0. \tag{5}$$

Remark 2.2 In the original definition in [9, p. 343], it is assumed instead of the above property (iv) that the principal block $\mathbf{A}$ is irreducibly diagonally dominant. However, if we assume that $\mathbf{A}$ is also irreducible, as will be done in Theorem 2.1, then its irreducibly diagonal dominance follows directly from Definition 2.3 under the given sign conditions on a_{ij}. We also note that a well-known theorem [32, p. 85] implies in this case that $\mathbf{A^{-1}} > 0$, i.e., the entries of the matrix $\mathbf{A^{-1}}$ are positive.

Many known results on various discrete maximum principles are based on the following theorem, considered as 'matrix maximum principle' (for a proof, see e.g. [9, Th. 3]).

Theorem 2.1 *Let $\bar{\mathbf{A}}$ be a $(k+m)\times(k+m)$ matrix with the structure (3), and assume that $\bar{\mathbf{A}}$ is of generalized nonnegative type in the sense of Definition 2.3, further, that $\mathbf{A}$ is irreducible.*

If the vector $\bar{\mathbf{c}} = (c_1, ..., c_{k+m})^T \in \mathbf{R}^{k+m}$ (where $(\,.\,)^T$ denotes the transposed) is such that $(\bar{\mathbf{A}}\bar{\mathbf{c}})_i \le 0,\ i = 1, ..., k,$ then

$$\max_{i=1,...,k+m} c_i \le \max\{0, \max_{i=k+1,...,k+m} c_i\}. \tag{6}$$

The irreducibility of $\mathbf{A}$ is a technical condition which is sometimes difficult to check in applications, see e.g. [11]. We now show that it can be omitted from the assumptions if (iv) is suitably strengthened. For convenient formulations, we will hence use the following

Definition 2.4 A $(k+m)\times(k+m)$ matrix $\bar{\mathbf{A}}$ with the structure (3) is said to be of *generalized nonnegative type with irreducible blocks* if properties (i)-(iii) of Definition 2.3 hold, further, property (iv) therein is replaced by the following stronger one:

(iv') For each irreducible component of $\mathbf{A}$ there exists an index $i_0 = i_0(l) \in N_l = \{s_1^{(l)}, \ldots, s_{k_l}^{(l)}\}$ for which $\sum\limits_{j=1}^{k} a_{i_0,j} > 0$.

Remark 2.3 Let assumptions (i)-(iii) hold in Definitions 2.3 or 2.4. Then for a given index $i_0 \in \{1, \ldots, k\}$, a sufficient condition for (5) to hold is that:

there exists an index $j_0 \in \{k+1, \ldots, k+m\}$ for which $a_{i_0,j_0} < 0$.

Namely, using also assumptions (ii) and (iii), respectively, we then have

$$\sum_{j=1}^{k} a_{i_0,j} > \sum_{j=1}^{k} a_{i_0,j} + a_{i_0,j_0} \ge \sum_{j=1}^{k} a_{i_0,j} + a_{i_0,j_0} + \sum_{\substack{j=k+1\\ j\neq j_0}}^{k+m} a_{i_0,j} = \sum_{j=1}^{k+m} a_{i_0,j} \ge 0.$$

Theorem 2.2 *Let $\bar{\mathbf{A}}$ be a $(k+m)\times(k+m)$ matrix with the structure (3), and assume that $\bar{\mathbf{A}}$ is of generalized nonnegative type with irreducible blocks in the sense of Definition 2.4.*

If the vector $\bar{\mathbf{c}} = (c_1, ..., c_{k+m})^T \in \mathbf{R}^{k+m}$ is such that $(\bar{\mathbf{A}}\bar{\mathbf{c}})_i \le 0,\ i = 1, ..., k,$ then (6) holds.

Proof. We may assume that $\mathbf{A}$ has the lower block triangular form mentioned in Remark 2.1. (Otherwise we can permute the indices to have this form, since the desired result is independent of the ordering of indices in the block $\mathbf{A}$.) That is, the block $\mathbf{A}$ in (3) has the irreducible diagonal blocks $\mathbf{A}^{(l)}$ (i.e. the irreducible components defined in Definition 2.2), and the corresponding blocks in $\mathbf{A}$ vanish in the upper block triangular part, further, we can use an analogous column decomposition of the block $\tilde{\mathbf{A}}$ in (3) to blocks $\tilde{\mathbf{A}}^{(l)}$ ($l = 1, \ldots, q$). Using an analogous decomposition of the vectors $\mathbf{c}$ and $\mathbf{b}$, system (4) can be written as

$$\begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{0} & \mathbf{0} & \ldots & \tilde{\mathbf{A}}^{(1)} \\ \mathbf{A}^{(21)} & \mathbf{A}^{(2)} & \mathbf{0} & \ldots & \tilde{\mathbf{A}}^{(2)} \\ \ldots & & & & \ldots \\ \mathbf{A}^{(q1)} & \mathbf{A}^{(q2)} & \ldots & \mathbf{A}^{(q)} & \tilde{\mathbf{A}}^{(q)} \\ \mathbf{0} & \ldots & \ldots & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{(1)} \\ \mathbf{c}^{(2)} \\ \ldots \\ \mathbf{c}^{(q)} \\ \tilde{\mathbf{c}} \end{bmatrix} = \begin{bmatrix} \mathbf{b}^{(1)} \\ \mathbf{b}^{(2)} \\ \ldots \\ \mathbf{b}^{(q)} \\ \tilde{\mathbf{b}} \end{bmatrix} \tag{7}$$

We must prove that if $\mathbf{b}^{(1)}, ..., \mathbf{b}^{(q)} \le 0$, then (6) holds, i.e. $\bar{\mathbf{c}} \le \max \tilde{\mathbf{c}}$.

Step 1. First we consider the special case when $\tilde{\mathbf{b}} \le 0$. Then $\tilde{\mathbf{c}} = \tilde{\mathbf{b}} \le 0$, hence the statement (6) becomes $\bar{\mathbf{c}} \le 0$. Since $\bar{\mathbf{c}} = [\mathbf{c}, \tilde{\mathbf{c}}]^T$, we in fact need to prove $\mathbf{c} \le 0$. We prove by induction that $\mathbf{c}^{(1)}, ..., \mathbf{c}^{(q)} \le 0$.

Note first that $\mathbf{A}^{(l)}$ ($l = 1, \ldots, q$) are of generalized nonnegative type, since they inherit Assumptions (i)-(iv') in Definition 2.3 from $\mathbf{A}$. Namely, this is obvious for Assumptions (i)-(ii). The nonnegativity in Assumption (iii) holds for $\mathbf{A}^{(l)}$ since we drop nonpositive elements in the row sum for $\mathbf{A}^{(l)}$ compared to the row sum for $\mathbf{A}$. Finally, Assumption (iv') for $\mathbf{A}$ just means that the original Assumption (iv) holds for each $\mathbf{A}^{(l)}$. Also, $\mathbf{A}^{(l)}$ are irreducible by definition, hence Theorem 2.1 can be applied to systems of the form (3) with left upper block $\mathbf{A}^{(l)}$. We will do this repeatedly for the case $\tilde{\mathbf{c}} \le 0$ to obtain nonpositive solution vectors.

The first and last rows of (7) yield the system

$$\begin{bmatrix} \mathbf{A}^{(1)} & \tilde{\mathbf{A}}^{(1)} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{(1)} \\ \tilde{\mathbf{c}} \end{bmatrix} = \begin{bmatrix} \mathbf{b}^{(1)} \\ \tilde{\mathbf{b}} \end{bmatrix}. \tag{8}$$

Here $\mathbf{b}^{(\mathbf{1})} \le 0$ and $\tilde{\mathbf{c}} = \tilde{\mathbf{b}} \le 0$, hence Theorem 2.1 yields $\mathbf{c}^{(1)} \le 0$.

Now let $l \in \{2, \ldots, q\}$ and assume that $\mathbf{c}^{(1)}, ..., \mathbf{c}^{(l-1)} \le 0$. The lth and last rows of (7) yield the system

$$\begin{bmatrix} \mathbf{A}^{(l)} & \tilde{\mathbf{A}}^{(l)} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c}^{(l)} \\ \tilde{\mathbf{c}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{b}}^{(l)} \\ \tilde{\mathbf{b}} \end{bmatrix}, \tag{9}$$

where $\hat{\mathbf{b}}^{(l)} := \mathbf{b}^{(l)} - \sum_{s=1}^{l-1} \mathbf{A}^{(ls)} \mathbf{c}^{(s)}$. Here $\mathbf{b}^{(l)} \le 0$ by assumption, $\mathbf{c}^{(s)} \le 0$ ($s = 1, \ldots, l-1$) from the inductional assumption and $\mathbf{A}^{(ls)} \le 0$ elementwise from property (ii) of Definition 2.3, therefore $\hat{\mathbf{b}}^{(l)} \le 0$. Using $\tilde{\mathbf{c}} = \tilde{\mathbf{b}} \le 0$ and applying Theorem 2.1 again, we obtain $\mathbf{c}^{(l)} \le 0$.

Step 2. Let us consider the case when $\max \tilde{\mathbf{b}} = \max \tilde{\mathbf{c}} > 0$. We must prove that if $\mathbf{b}^{(1)}, ..., \mathbf{b}^{(q)} \le 0$ (i.e. $\mathbf{b} \le 0$) then (6) holds, i.e. that $\bar{\mathbf{c}} \le \max \tilde{\mathbf{c}}$.

Let $\mathbf{c}^* := \bar{\mathbf{c}} - (\max \tilde{\mathbf{c}}) \cdot \mathbf{1}_{k+m}$, where $\mathbf{1}_{k+m}$ is the constant 1 vector of length $k+m$. Since $\bar{\mathbf{A}}\bar{\mathbf{c}} = \bar{\mathbf{b}}$, therefore $\mathbf{c}^*$ is the solution of the linear system $\bar{\mathbf{A}}\mathbf{c}^* = \mathbf{b}^*$, where

$$\mathbf{b}^* := \bar{\mathbf{b}} - (\max \tilde{\mathbf{c}}) \cdot \bar{\mathbf{A}} \mathbf{1}_{k+m} = \begin{bmatrix} \mathbf{b} \\ \tilde{\mathbf{b}} \end{bmatrix} - (\max \tilde{\mathbf{c}}) \cdot \begin{bmatrix} \mathbf{A} & \tilde{\mathbf{A}} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{1}_k \\ \mathbf{1}_m \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{b} - (\max \tilde{\mathbf{c}}) \cdot [\, \mathbf{A} \quad \tilde{\mathbf{A}} \,]\, \mathbf{1}_{k+m} \\ \tilde{\mathbf{b}} - (\max \tilde{\mathbf{c}}) \cdot \mathbf{1}_m \end{bmatrix} . \tag{10}$$

Here the first component in (10) is nonpositive, since $\mathbf{b} \le 0$ and $\max \tilde{\mathbf{c}} > 0$ by assumption, further, $[\, \mathbf{A} \quad \tilde{\mathbf{A}} \,]\, \mathbf{1}_{k+m} \ge 0$ by item (iii) of Definition 2.3. The second component in (10) is also nonpositive, since obviously $\tilde{\mathbf{b}} = \tilde{\mathbf{c}} \le \max \tilde{\mathbf{c}}$. Therefore $\mathbf{b}^* \le 0$. Thus, applying step 1 to system $\bar{\mathbf{A}}\mathbf{c}^* = \mathbf{b}^*$, we obtain $\mathbf{c}^* \le 0$, i.e. $\bar{\mathbf{c}} - (\max \tilde{\mathbf{c}}) \cdot \mathbf{1}_{k+m} \le 0$, which was to be proved. ■

Consequently, in what follows, our main goal is to show that the stiffness matrix of the problems considered is of generalized nonnegative type with irreducible blocks in the sense of Definition 2.4..

3 A discrete maximum principle in Hilbert space

3.1 Formulation of the problem

Let H be a real Hilbert space and $H_0 \subset H$ a given subspace. We consider the following operator equation: for given vectors $\psi, g^* \in H$, find $u \in H$ such that

$$\langle A(u), v \rangle = \langle \psi, v \rangle \qquad (v \in H_0) \tag{11}$$

$$\text{and} \quad u - g^* \in H_0 \tag{12}$$

with an operator $A : H \to H$ satisfying the following conditions:

Assumptions 3.1.

(i) The operator $A : H \to H$ has the form

$$A(u) = B(u)u + R(u)u \tag{13}$$

where B and R are given operators mapping from H to $B(H)$. (Here $B(H)$ denotes the set of bounded linear operators in H.)

(ii) There exists a constant $m > 0$ such that

$$\langle B(u)v, v \rangle \ge \; m \, \|v\|^2 \qquad (u \in H, \; v \in H_0). \tag{14}$$

(iii) There exist subsets of 'positive vectors' $D, P \subset H$ such that

$$\langle R(u)w, v \rangle \ge 0 \qquad (u \in H, \; v \in D, \; w \in P \text{ or } w = v). \tag{15}$$

(iv) There exists a continuous function $M_R : \mathbf{R}^+ \to \mathbf{R}^+$ and another norm $|||.|||$ on H such that

$$\langle R(u)w, v \rangle \le M_R(\|u\|) \, |||w||| \, |||v||| \qquad (u, w, v \in H). \tag{16}$$

In practice for PDE problems (considered in section 4), g^* plays the role of boundary condition and H_0 will be the subspace corresponding to homogeneous boundary conditions, further, $B(u)$ is the principal part of A.

Assumptions 3.1 are not in general known to imply existence and uniqueness for (11)-(12). The following extra conditions already ensure well-posedness:

Assumptions 3.2.

(i) Let

$$F(u) := B(u)u, \qquad G(u) := R(u)u \qquad (u \in H). \tag{17}$$

The operators $F, G : H \to H$ are Gateaux differentiable, further, F' and G' are bihemicontinuous (i.e. mappings $(s,t) \mapsto F'(u+sk+tw)h$ are continuous from $\mathbf{R}^2$ to H, and similarly for G').

(ii) There exists a continuous function $M_A : \mathbf{R}^+ \to \mathbf{R}^+$ such that

$$\langle A'(u)w, v\rangle \le M_A(\|u\|)\,\|w\|\,\|v\| \qquad (u \in H,\ w, v \in H_0). \tag{18}$$

(iii) There exists a constant $m > 0$ such that

$$\langle F'(u)v, v\rangle \ge\ m\,\|v\|^2 \qquad (u \in H,\ v \in H_0). \tag{19}$$

(iv) We have

$$\langle G'(u)v, v\rangle \ge\ 0 \qquad (u \in H,\ v \in H_0). \tag{20}$$

Proposition 3.1 *If Assumptions 3.1–3.2 hold, then problem (11)-(12) is well-posed.*

Proof. Problem (11)-(12) can be rewritten as follows:

$$\text{find}\ \ u_0 \in H :\ \langle \tilde{A}(u_0), v\rangle \equiv \langle A(u_0+g^*), v\rangle = \langle \psi, v\rangle \qquad (v \in H_0), \tag{21}$$

$$\text{and}\ \ \text{let}\ \ u := u_0 + g^*. \tag{22}$$

From (19) and (20) we have

$$\langle A'(u)v, v\rangle \ge\ m\,\|v\|^2 \qquad (u \in H,\ v \in H_0) \tag{23}$$

whence A is uniformly monotone on H_0, further, from (18), A is locally Lipschitz continuous on H_0. These properties of A are inherited by $\tilde{A}$ by the definition of the latter: that is, for all $u, v \in H_0$, we obtain

$$m\,\|u-v\|^2 \le \langle \tilde{A}(u)-\tilde{A}(v), u-v\rangle, \qquad \|\tilde{A}(u)-\tilde{A}(v)\| \le M_A(\max\{\|u\|, \|v\|\})\,\|u-v\|\,. \tag{24}$$

These imply well-posedness for (21), see, e.g., [13, 23]. ■

3.2 Galerkin type discretization

Let $n_0 \le n$ be positive integers and $\phi_1, ..., \phi_n \in H$ be given linearly independent vectors such that $\phi_1, ..., \phi_{n_0} \in H_0$. We consider the finite dimensional subspaces

$$V_h = \text{span}\{\phi_1, ..., \phi_n\} \subset H, \qquad V_h^0 = \text{span}\{\phi_1, ..., \phi_{n_0}\} \subset H_0 \tag{25}$$

with a real positive parameter $h > 0$. In practice, as is usual for FEM, h is inversely proportional to n, and one will consider a family of such subspaces, see Definition 3.2 later.

Let $g^h = \sum\limits_{j=n_0+1}^{n} g_j\phi_j \in V_h$ be a given approximation of the component of g^* in $H \setminus H_0$. To find the Galerkin solution of (11)-(12) in V_h, we solve the following problem: find $u^h \in V_h$ such that

$$\langle A(u^h), v\rangle = \langle \psi, v\rangle \qquad (v \in V_h^0) \tag{26}$$

$$\text{and} \quad u^h - g^h \in V_h^0. \tag{27}$$

Using (13), we can rewrite (26) as

$$\langle B(u^h)u^h, v\rangle + \langle R(u^h)u^h, v\rangle = \langle \psi, v\rangle \qquad (v \in V_h^0). \tag{28}$$

Let us now formulate the nonlinear algebraic system corresponding to (28). We set

$$u^h = \sum_{j=1}^{n} c_j\phi_j, \tag{29}$$

and look for the coefficients $c_1, \ldots, c_n$. For any $\bar{\mathbf{c}} = (c_1, ..., c_n)^T \in \mathbf{R}^n$, $i = 1, ..., n_0$ and $j = 1, ..., n$, we set

$$b_{ij}(\bar{\mathbf{c}}) := \langle B(u^h)\phi_j, \phi_i\rangle \qquad r_{ij}(\bar{\mathbf{c}}) := \langle R(u^h)\phi_j, \phi_i\rangle, \qquad d_i := \langle \psi, \phi_i\rangle,$$

$$a_{ij}(\bar{\mathbf{c}}) := b_{ij}(\bar{\mathbf{c}}) + r_{ij}(\bar{\mathbf{c}}). \tag{30}$$

Putting (29) and $v = \phi_i$ into (28), we obtain the $n_0 \times n$ system of algebraic equations

$$\sum_{j=1}^{n} a_{ij}(\bar{\mathbf{c}})\, c_j = d_i \qquad (i = 1, ..., n_0). \tag{31}$$

Using the notations

$$\mathbf{A}(\bar{\mathbf{c}}) := \{a_{ij}(\bar{\mathbf{c}})\},\ i, j = 1, ..., n_0, \qquad \tilde{\mathbf{A}}(\bar{\mathbf{c}}) := \{a_{ij}(\mathbf{c})\},\ i = 1, ..., n_0;\ j = n_0+1, ..., n,$$

$$\mathbf{d} := \{d_j\},\ \mathbf{c} := \{c_j\}, \quad j = 1, ..., n_0, \quad \text{and} \quad \tilde{\mathbf{c}} := \{c_j\}, \quad j = n_0 + 1, ..., n, \tag{32}$$

system (31) turns into

$$\mathbf{A}(\bar{\mathbf{c}})\mathbf{c} + \tilde{\mathbf{A}}(\bar{\mathbf{c}})\tilde{\mathbf{c}} = \mathbf{d}. \tag{33}$$

In order to obtain a system with a square matrix, we enlarge our system to an $n \times n$ one. Since $u^h - g^h \in V_h^0$, the coordinates c_i with $n_0 + 1 \le i \le n$ satisfy automatically $c_i = g_i$, i.e.,

$$\tilde{\mathbf{c}} = \tilde{\mathbf{g}} := \{g_j\}, \quad j = n_0 + 1, ..., n,$$

hence we can replace (33) by the equivalent system

$$\begin{bmatrix} \mathbf{A}(\bar{\mathbf{c}}) & \tilde{\mathbf{A}}(\bar{\mathbf{c}}) \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \tilde{\mathbf{c}} \end{bmatrix} = \begin{bmatrix} \mathbf{d} \\ \tilde{\mathbf{g}} \end{bmatrix} . \tag{34}$$

Defining further

$$\bar{\mathbf{A}}(\bar{\mathbf{c}}) := \begin{bmatrix} \mathbf{A}(\bar{\mathbf{c}}) & \tilde{\mathbf{A}}(\bar{\mathbf{c}}) \\ 0 & \mathbf{I} \end{bmatrix}, \qquad \bar{\mathbf{c}} := \begin{bmatrix} \mathbf{c} \\ \tilde{\mathbf{c}} \end{bmatrix}, \tag{35}$$

we rewrite (33) as follows:

$$\bar{\mathbf{A}}(\bar{\mathbf{c}})\bar{\mathbf{c}} = \mathbf{d}. \tag{36}$$

3.3 Maximum principle for the abstract discretized problem

When formulating a discrete maximum principle for system (36), we will use the following notion:

Definition 3.1 Certain pairs $\{\phi_i, \phi_j\} \in V_h \times V_h$ are called '*neighbouring basis vectors*', and then i, j are called '*neighbouring indices*', such that the following property holds: the set $\{1, \ldots, n\}$ can be partitioned into disjoint sets $S_1, \ldots, S_r$ such that for each $k = 1, \ldots, r$,

(i) both $S_k^0 := S_k \cap \{1, \ldots, n_0\}$ and $\tilde{S}_k := S_k \cap \{n_0 + 1, \ldots, n\}$ are nonempty;

(ii) the graph of all neighbouring indices in S_k^0 is connected;

(iii) the graph of all neighbouring indices in S_k is connected.

In later PDE applications, this notion is meant to express that the supports of basis functions cover the domain, both its interior and the boundary.

The following notion will be crucial for our study:

Definition 3.2 A set of subspaces $\mathcal{V} = \{V_h\}_{h \to 0}$ in H is said to be a *family of subspaces* if for any $\varepsilon > 0$ there exists $V_h \in \mathcal{V}$ with $h < \varepsilon$.

First we give sufficient conditions for the generalized nonnegativity of the matrix $\bar{\mathbf{A}}(\bar{\mathbf{c}})$.

Theorem 3.1 *Let Assumptions 3.1 hold. Let us consider the discretization of operator equation (11)-(12) in a family of subspaces $\mathcal{V} = \{V_h\}_{h \to 0}$ with bases as in (25). Let $u^h \in V_h$ be the solution of (28) and let the following properties hold:*

(a) *For all $\phi_i \in V_h^0$ and $\phi_j \in V_h$, one of the following holds: either*

$$\langle B(u^h)\phi_j, \phi_i\rangle = 0 \quad \text{and} \quad \langle R(u^h)\phi_j, \phi_i\rangle \leq 0, \tag{37}$$

or

$$\langle B(u^h)\phi_j, \phi_i\rangle \leq -M_B(h) \tag{38}$$

with a proper function $M_B : \mathbf{R}^+ \to \mathbf{R}^+$ (independent of h, ϕ_i, ϕ_j) such that, defining

$$T(h) := \sup\{|||\phi_i||| : \ \phi_i \in V_h)\}\,, \tag{39}$$

we have

$$\lim_{h\to 0} \frac{M_B(h)}{T(h)^2} = +\infty. \tag{40}$$

(b) *If, in particular, $\phi_i \in V_h^0$ and $\phi_j \in V_h$ are neighbouring basis vectors (as defined in Definition 3.1), then (38)-(40) hold.*

(c) *$M_R(\|u^h\|)$ is bounded as $h \to 0$, where M_R is the function in Assumption 3.1 (iv).*

(d) *For all $u \in H$ and $h > 0$, $\sum\limits_{j=1}^{n} \phi_j \in \ker B(u)$.*

(e) *For all $h > 0$, $i = 1, ..., n$, we have $\phi_i \in D$ and $\sum\limits_{j=1}^{n} \phi_j \in P$ for the sets D, P introduced in Assumption 3.1 (iii).*

Then for sufficiently small h, the matrix $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ defined in (35) is of generalized nonnegative type with irreducible blocks in the sense of Definition 2.4.

Proof. Our task is to check properties (i)-(iv') of Definition 2.4 for

$$a_{ij}(\bar{\mathbf{c}}) = \langle B(u^h)\phi_j, \phi_i\rangle + \langle R(u^h)\phi_j, \phi_i\rangle \qquad (i, j = 1, ..., n). \tag{41}$$

(i) For any $i = 1, ..., n_0$, we have $\phi_i \in V_h^0 \subset H_0$ from (25), hence we can set $v = \phi_i$ in (14). Further, by assumption (e), we have $\phi_i \in D$, hence we can set $v = w = \phi_i$ in (15). These imply

$$a_{ii}(\bar{\mathbf{c}}) = \langle B(u^h)\phi_i, \phi_i\rangle + \langle R(u^h)\phi_i, \phi_i\rangle \geq m\,\|\phi_i\|^2 > 0.$$

(ii) Let $i = 1, ..., n_0,\ j = 1, ..., n$ with $i \neq j$. If (37) holds then $a_{ij}(\bar{\mathbf{c}}) \leq 0$ by (41). If (38) holds then, using also (41), (16), respectively, and letting $\tilde{M} := \sup M_R(\|u^h\|)$, we obtain

$$a_{ij}(\bar{\mathbf{c}}) \leq -M_B(h) + M_R(\|u^h\|)\,|||\phi_i|||\,|||\phi_j||| \leq -M_B(h) + M_R(\|u^h\|)\,T(h)^2$$

$$\leq T(h)^2 \Big(-\frac{M_B(h)}{T(h)^2} + \tilde{M}\Big) < 0 \tag{42}$$

for sufficiently small h, since by (40) the expression in brackets tends to $-\infty$ as $h \to 0$.

(iii) For any $i = 1, ..., n_0$,

$$\sum_{j=1}^{n} a_{ij}(\bar{\mathbf{c}}) = \Big\langle B(u^h)\Big(\sum_{j=1}^{n}\phi_j\Big),\, \phi_i\Big\rangle + \Big\langle R(u^h)\Big(\sum_{j=1}^{n}\phi_j\Big),\, \phi_i\Big\rangle \geq 0,$$

since the first term equals zero by assumption (d), further, by assumption (e) we can set $w = \sum\limits_{j=1}^{n}\phi_j$ and $v = \phi_i$ in (15), hence the second term is nonnegative.

(iv') We must prove that for each irreducible component of $\mathbf{A}(\bar{\mathbf{c}})$ there exists an index $i_0 \in N_l = \{s_1^{(l)}, \ldots, s_{k_l}^{(l)}\}$ for which $\sum\limits_{j=1}^{n_0} a(\bar{\mathbf{c}})_{i_0,j} > 0$. Here, with the notations of Definition 2.2, the matrix $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ has q irreducible blocks $\mathbf{A}^{(l)}(\bar{\mathbf{c}})$ $(l = 1, \ldots, q)$, and N_l denotes the indices arising in $\mathbf{A}^{(l)}(\bar{\mathbf{c}})$. Then $k_1 + \ldots + k_q = n_0$. Using Remark 2.3, we must prove that for all $l = 1, \ldots, q$ there exist indices $i_0 \in N_l$ and $j_0 \in \{n_0+1, \ldots, n\}$ such that $a(\bar{\mathbf{c}})_{i_0,j_0} < 0$.

From now, let $N_0 := \{1, \ldots, n_0\}$, $\tilde{N} := \{n_0+1, \ldots, n\}$ and $N := \{1, \ldots, n\} = N_0 \cup \tilde{N}$.

First note that if $i \in N_0, j \in N$ are neighbouring indices then $a_{ij}(\bar{\mathbf{c}}) < 0$. Namely, (38) holds by assumption (b), whence (42) yields $a_{ij}(\bar{\mathbf{c}}) < 0$. Hence, it suffices to find $i_0 \in N_l$ and $j_0 \in \tilde{N}$ such that i_0, j_0 are neighbouring indices.

Now we observe that each N_l contains entire sets S_k^0, introduced in Definition 3.1. Namely, by item (ii) of Definition 3.1, the graph of all neighbouring indices in S_k^0 is connected, i.e. for all $i, j \in S_k^0$ there exists a chain (i, i_1), $(i_1, i_2), \ldots, (i_r, j)$ of neighbouring indices (with all $i_m \in S_k^0$), whence by the above $a_{i,i_1}(\bar{\mathbf{c}}) < 0$, $a_{i_1,i_2}(\bar{\mathbf{c}}) < 0, \ldots, a_{i_r,j}(\bar{\mathbf{c}}) < 0$. Therefore the entries of $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ with indices in S_k^0 belong to the same irreducible component, i.e. S_k^0 lies entirely in one of the sets N_l.

Consequently, it suffices to prove that for all $k = 1, \ldots, r$ there exist indices $i_0 \in S_k^0$ and $j_0 \in \tilde{N}$ such that i_0, j_0 are neighbouring indices. By item (i) of Definition 3.1, there exists $i \in S_k^0$ and $j \in \tilde{S}_k$. Using that $i, j \in S_k$, by item (iii) of Definition 3.1, there exists a chain (i, i_1), $(i_1, i_2), \ldots, (i_r, j)$ of neighbouring indices with all $i_m \in S_k$. If $i_1 \in \tilde{S}_k$ then we let $i_0 := i (\in S_k^0)$ and $j_0 := i_1 (\in \tilde{N})$. Otherwise, since $j \in \tilde{S}_k$, there exists a first index k in the chain such that $i_k \in S_k^0$ and $i_{k+1} \in \tilde{S}_k$, and then we let $i_0 := i_k (\in S_k^0)$ and $j_0 := i_{k+1} (\in \tilde{N})$. ■

By Theorem 2.2, we immediately obtain the corresponding algebraic *discrete maximum principle:*

Corollary 3.1 *Let the assumptions of Theorem 3.1 hold. For sufficiently small h, if $d_i \leq 0$ $(i = 1, ..., n_0)$ in (32) and $\bar{\mathbf{c}} = (c_1, ..., c_n)^T \in \mathbf{R}^n$ is the solution of (36), then*

$$\max_{i=1,\ldots,n} c_i \leq \max\{0, \max_{i=n_0+1,\ldots,n} c_i\}. \tag{43}$$

Remark 3.1 Assumption (c) of Theorem 3.1 follows in particular if Assumptions 3.2 are added to Assumptions 3.1 as done in Proposition 3.1, provided that the functions $g^h \in V_h$ in (27) are bounded in H-norm as $h \to 0$. (In practice, the usual choices for g^h even produce $g^h \to g^*$ in H-norm.) In fact, in this case $\|u^h\|$ is bounded as $h \to 0$; then the continuity of M_R yields that $M_R(\|u^h\|)$ is bounded too.

Namely, using (23),

$$\langle A(u^h)-A(g^h), u^h-g^h\rangle = \langle A'(\theta u^h+(1-\theta)g^h)(u^h-g^h), u^h-g^h\rangle \geq m\,\|u^h-g^h\|^2$$

(where $\theta \in [0,1]$). From (26)

$$\langle A(u^h) - A(g^h), u^h - g^h\rangle = \langle f - A(g^h), u^h - g^h\rangle \tag{44}$$

and from (18)

$$\begin{aligned}\langle A(g^*) - A(g^h), u^h - g^h\rangle &= \langle A'(\theta g^* + (1-\theta)g^h)(g^* - g^h), u^h - g^h\rangle \\ &\leq M_A(\max\{\|g^*\|, \|g^h\|\})\,\|g^* - g^h\|\,\|u^h - g^h\|\end{aligned} \tag{45}$$

(where $\theta \in [0,1]$). From the above,

$$\begin{aligned}m\,\|u^h-g^h\|^2 &\leq \langle f-A(g^*), u^h-g^h\rangle + M_A(\max\{\|g^*\|, \|g^h\|\})\,\|g^*-g^h\|\,\|u^h-g^h\| \\ &\leq (\|f - A(g^*)\| + M_A(\max\{\|g^*\|, \|g^h\|\})\,\|g^* - g^h\|)\,\|u^h - g^h\|\,.\end{aligned}$$

Using the notation $\gamma := \sup_{h>0} \|g^* - g^h\|$, we obtain

$$\|u^h\| \leq \|g^h\| + \|u^h - g^h\| \leq \|g^*\| + \gamma + \tfrac{1}{m}\Big(\|f - A(g^*)\| + M_A(\|g^*\| + \gamma)\,\gamma\Big),$$

i.e. $\|u^h\|$ is bounded as $h \to 0$.

4 Discrete maximum principles for cooperative and weakly diagonally dominant elliptic systems

4.1 Systems with nonlinear coefficients

4.1.1 Formulation of the problem

We consider nonlinear elliptic systems of the form

$$\left.\begin{aligned} -\mathrm{div}\,\Big(b_k(x,u,\nabla u)\,\nabla u_k\Big) + \sum_{l=1}^{s} V_{kl}(x,u,\nabla u)\,u_l \;&=\; f_k(x) \quad \text{in } \Omega, \\ b_k(x,u,\nabla u)\tfrac{\partial u_k}{\partial \nu} \;&=\; \gamma_k(x) \quad \text{on } \Gamma_N, \\ u_k \;&=\; g_k(x) \quad \text{on } \Gamma_D \end{aligned}\right\} \qquad (k = 1,\ldots,s) \tag{46}$$

with unknown function $u = (u_1,\ldots,u_s)$, under the following assumptions, where inequalities for functions are understood pointwise for all possible arguments:

Assumptions 4.1.

(i) $\Omega \subset \mathbf{R}^d$ is a bounded piecewise C^1 domain; Γ_D, Γ_N are disjoint open measurable subsets of $\partial\Omega$ such that $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$ and $\Gamma_D \neq \emptyset$.

(ii) (Smoothness and boundedness.) For all $k, l = 1, \ldots, s$ we have $b_k \in (C^1 \cap L^\infty)(\Omega \times \mathbf{R}^s \times \mathbf{R}^{s\times d})$ and $V_{kl} \in L^\infty(\Omega \times \mathbf{R}^s \times \mathbf{R}^{s\times d})$.

(iii) (Ellipticity.) There exists $m > 0$ such that $b_k \geq m$ holds for all $k = 1, \ldots, s$.

(iv) (Cooperativity.) We have

$$V_{kl} \leq 0 \qquad (k, l = 1, \ldots, s,\ k \neq l). \tag{47}$$

(v) (Weak diagonal dominance.) We have

$$\sum_{l=1}^{s} V_{kl} \geq 0 \qquad (k = 1, \ldots, s). \tag{48}$$

(vi) For all $k = 1, \ldots, s$ we have $f_k \in L^2(\Omega)$, $\gamma_k \in L^2(\Gamma_N)$, $g_k = g^*_{k|\Gamma_D}$ with $g^*_k \in H^1(\Omega)$.

Remark 4.1 (i) Assumptions (47)-(48) imply

$$V_{kk} \geq 0 \qquad (k = 1, \ldots, s). \tag{49}$$

(ii) One may consider additional terms on the Neumann boundary, see Remark 4.4 later.

For the weak formulation of such problems, we define the Sobolev space

$$H^1_D(\Omega) := \{z \in H^1(\Omega) :\ z_{|\Gamma_D} = 0\}. \tag{50}$$

The weak formulation of problem (46) then reads as follows: find $u \in H^1(\Omega)^s$ such that

$$\langle A(u), v\rangle = \langle \psi, v\rangle \qquad (\forall v \in H^1_D(\Omega)^s) \tag{51}$$

$$\text{and} \quad u - g^* \in H^1_D(\Omega)^s, \tag{52}$$

where

$$\langle A(u), v\rangle = \int_\Omega \Big(\sum_{k=1}^{s} b_k(x, u, \nabla u)\, \nabla u_k \cdot \nabla v_k + \sum_{k,l=1}^{s} V_{kl}(x, u, \nabla u)\, u_l\, v_k\Big) \tag{53}$$

for given $u = (u_1, \ldots, u_s) \in H^1(\Omega)^s$ and $v = (v_1, \ldots, v_s) \in H^1_D(\Omega)^s$, further,

$$\langle \psi, v\rangle = \int_\Omega \sum_{k=1}^{s} f_k v_k + \int_{\Gamma_N} \sum_{k=1}^{s} \gamma_k v_k \tag{54}$$

for given $v = (v_1, \ldots, v_s) \in H^1_D(\Omega)^s$, and $g^* := (g^*_1, \ldots, g^*_s)$.

4.1.2 Finite element discretization

We define the finite element discretization of problem (46) in the following way. First, let $\bar{n}_0 \le \bar{n}$ be positive integers and let us choose basis functions

$$\varphi_1, \dots, \varphi_{\bar{n}_0} \in H^1_D(\Omega), \qquad \varphi_{\bar{n}_0+1}, \dots, \varphi_{\bar{n}} \in H^1(\Omega) \setminus H^1_D(\Omega), \tag{55}$$

which correspond to homogeneous and inhomogeneous boundary conditions on Γ_D, respectively. (For simplicity, we will refer to them as 'interior basis functions' and 'boundary basis functions', respectively, thus adopting the terminology of Dirichlet problems even in the general case.) These basis functions are assumed to be continuous and to satisfy

$$\varphi_p \ge 0 \quad (p = 1, \dots, \bar{n}), \qquad \sum_{p=1}^{\bar{n}} \varphi_p \equiv 1, \tag{56}$$

further, that there exist node points $B_p \in \Omega$ ($p = 1, \dots, \bar{n}_0$) and $B_p \in \Gamma_D$ ($p = \bar{n}_0 + 1, \dots, \bar{n}$) such that

$$\varphi_p(B_q) = \delta_{pq} \tag{57}$$

where δ_{pq} is the Kronecker symbol. (These conditions hold e.g. for standard linear, bilinear or prismatic finite elements.) Finally, we assume that any two interior basis functions can be connected with a chain of interior basis functions with overlapping support. By its geometric meaning, this assumption obviously holds for any reasonable FE mesh.

We in fact need a basis in the corresponding product spaces, which we define by repeating the above functions in each of the s coordinates and setting zero in the other coordinates. That is, let $n_0 := s\bar{n}_0$ and $n := s\bar{n}$. First, for any $1 \le i \le n_0$,

$$\text{if } \ i = (k-1)\bar{n}_0 + p \quad \text{for some } 1 \le k \le s \text{ and } 1 \le p \le \bar{n}_0, \ \text{ then}$$

$$\phi_i := (0, \dots, 0, \varphi_p, 0, \dots, 0) \qquad \text{where } \ \varphi_p \ \text{ stands at the } k\text{-th entry}, \tag{58}$$

that is, $(\phi_i)_m = \varphi_p$ if $m = k$ and $(\phi_i)_m = 0$ if $m \ne k$. From these, we let

$$V_h^0 := \text{span}\{\phi_1, ..., \phi_{n_0}\} \ \subset H^1_D(\Omega)^s. \tag{59}$$

Similarly, for any $n_0 + 1 \le i \le n$,

$$\text{if } \ i = n_0 + (k-1)(\bar{n} - \bar{n}_0) + p - \bar{n}_0 \quad \text{for some } 1 \le k \le s \text{ and } \bar{n}_0 + 1 \le p \le \bar{n}, \ \text{ then}$$

$$\phi_i := (0, \dots, 0, \varphi_p, 0, \dots, 0) \qquad \text{where } \ \varphi_p \ \text{ stands at the } k\text{-th entry}, \tag{60}$$

that is, $(\phi_i)_m = \varphi_p$ if $m = k$ and $(\phi_i)_m = 0$ if $m \ne k$. From (59) and these, we let

$$V_h := \text{span}\{\phi_1, ..., \phi_n\} \ \subset H^1(\Omega)^s. \tag{61}$$

Using the above FEM subspaces, the finite element discretization of problem (46) leads to the task of finding $u^h \in V_h$ such that

$$\langle A(u^h), v\rangle = \langle \psi, v\rangle \qquad (\forall v \in V_h^0) \tag{62}$$

$$\text{and} \quad u^h - g^h \in V_h^0, \quad \text{i.e.,} \;\; u^h = g^h \text{ on } \Gamma_D \tag{63}$$

(where $g^h = \sum\limits_{j=n_0+1}^{n} g_j\phi_j \in V_h$ is the approximation of g^* on Γ_D). Then, setting $u^h = \sum\limits_{j=1}^{n} c_j\phi_j$ and $v = \phi_i$ $(i = 1,\ldots,n_0)$ in (51) (just as in (29)-(31)), we obtain the $n_0 \times n$ system of algebraic equations

$$\sum_{j=1}^{n} a_{ij}(\bar{\mathbf{c}})\, c_j = d_i \qquad (i = 1, ..., n_0), \tag{64}$$

where for any $\bar{\mathbf{c}} = (c_1, ..., c_n)^T \in \mathbf{R}^n$ $\;(i = 1, ..., n_0, \;\; j = 1, ..., n)$,

$$a_{ij}(\bar{\mathbf{c}}) := \int_\Omega \Big(\sum_{k=1}^{s} b_k(x, u^h, \nabla u^h)\, (\nabla\phi_j)_k \cdot (\nabla\phi_i)_k + \sum_{k,l=1}^{s} V_{kl}(x, u^h, \nabla u^h)\, (\phi_j)_l\, (\phi_i)_k \Big) \tag{65}$$

$$\text{and} \qquad d_i := \int_\Omega \sum_{k=1}^{s} f_k(\phi_i)_k + \int_{\Gamma_N} \sum_{k=1}^{s} \gamma_k(\phi_i)_k \,. \tag{66}$$

In the same way as for (36), we enlarge system (64) to a square one by adding an identity block, and write it briefly as

$$\bar{\mathbf{A}}(\bar{\mathbf{c}})\bar{\mathbf{c}} = \mathbf{d}\,. \tag{67}$$

That is, for $i = 1, ..., n_0$ and $j = 1, ..., n$, the matrix $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ has the entry $a_{ij}(\bar{\mathbf{c}})$ from (65).

In what follows, we will need notions of (patch-)regularity of the considered meshes, cf. [4].

Definition 4.1 Let $\Omega \subset \mathbf{R}^d$ and let us consider a family of FEM subspaces $\mathcal{V} = \{V_h\}_{h\to 0}$ constructed as above. The corresponding family of meshes will be called

(a) *regular from above* if there exists a constant $c_0 > 0$ such that for any $V_h \in \mathcal{V}$ and basis function $\varphi_p \in V_h$,

$$meas(\operatorname{supp}\varphi_p) \le c_0 h^d \tag{68}$$

(where *meas* denotes d-dimensional measure and supp denotes the support, i.e. the closure of the set where the function does not vanish);

(b) *regular* if there exist constants $c_1, c_2 > 0$ such that for any $V_h \in \mathcal{V}$ and basis function $\varphi_p \in V_h$,

$$c_1 h^d \le meas(\operatorname{supp}\varphi_p) \le c_2 h^d\,; \tag{69}$$

(c) *quasi-regular* w.r.t. problem (46) if (69) is replaced by

$$c_1 h^\gamma \le meas(\operatorname{supp}\varphi_p) \le c_2 h^d \tag{70}$$

for some fixed constant

$$d \le \gamma < d + 2. \tag{71}$$

4.1.3 Discrete maximum principle for systems with nonlinear coefficients

Our goal is to apply Theorem 3.1 to derive a DMP for problem (46). For this, we first define the underlying operators as in subsection 3.1, and check Assumptions 3.1.

Lemma 4.1 *For any $u \in H^1(\Omega)^s$, let us define the operators $B(u)$ and $R(u)$ via*

$$\langle B(u)w, v\rangle = \int_\Omega \sum_{k=1}^{s} b_k(x,u,\nabla u)\, \nabla w_k \cdot \nabla v_k, \quad \langle R(u)w, v\rangle = \int_\Omega \sum_{k,l=1}^{s} V_{kl}(x,u,\nabla u)\, w_l\, v_k \tag{72}$$

($w \in H^1(\Omega)^s$, $v \in H^1_D(\Omega)^s$). Together with the operator A, defined in (53), the operators $B(u)$ and $R(u)$ satisfy Assumptions 3.1 in the spaces $H = H^1(\Omega)^s$ and $H_0 = H^1_D(\Omega)^s$.

Proof. We define

$$\|v\|^2 := \sum_{k=1}^{s} \Big(\int_\Omega |\nabla v_k|^2 + \int_{\Gamma_D} |v_k|^2\Big) \tag{73}$$

on $H^1(\Omega)^s$, which is a norm since $\Gamma_D \neq \emptyset$. Then for $v \in H^1_D(\Omega)^s$ we have $\|v\|^2 = \sum_{k=1}^{s} \int_\Omega |\nabla v_k|^2$.

(i) It is obvious from (53) and (72) that $A(u) = B(u)u + R(u)u$.

(ii) Assumption 4.1 (iii) implies for all $u \in H^1(\Omega)^s$, $v \in H^1_D(\Omega)^s$ that

$$\langle B(u)v, v\rangle = \int_\Omega \sum_{k=1}^{s} b_k(x,u,\nabla u)\, |\nabla v_k|^2 \geq m \int_\Omega \sum_{k=1}^{s} |\nabla v_k|^2 = m\, \|v\|^2 . \tag{74}$$

(iii) Let $D \subset H^1(\Omega)^s$ consist of the functions that have only one nonzero coordinate that is nonnegative, i.e. $v \in D$ iff $v = (0, \dots, 0, z, 0, \dots, 0)$ with z at the k-th entry for some $1 \leq k \leq s$ and $z \in H^1(\Omega)$, $z \geq 0$. Further, let $P \subset H^1(\Omega)^s$ consist of the functions that have identical nonnegative coordinates, i.e. $v \in P$ iff $v = (y, \dots, y)$ for some $y \in H^1(\Omega)$, $y \geq 0$. Now let $u \in H^1(\Omega)^s$ and $v \in D$. If $w \in P$, then

$$\langle R(u)w, v\rangle = \int_\Omega \Big(\sum_{l=1}^{s} V_{kl}(x,u,\nabla u)\Big)\, yz \geq 0$$

by (48) and that $y, z \geq 0$. If $w = v$, then

$$\langle R(u)w, v\rangle = \int_\Omega V_{kk}(x,u,\nabla u)\, z^2 \geq 0$$

by (49).

(iv) Let $\tilde{V} := \max_{k,l} \|V_{kl}\|_{L^\infty}$, which is finite by Assumption 4.1 (ii), and let us define the new norm

$$|||v|||^2 := \|v\|^2_{L^2(\Omega)^s} = \int_\Omega \sum_{k=1}^s v_k^2 \tag{75}$$

on $H^1(\Omega)^s$. Then we have for all $u, w, v \in H^1(\Omega)^s$

$$\langle R(u)w, v\rangle \le \tilde{V} \int_\Omega \sum_{k,l=1}^s |w_l|\,|v_k| \le s\tilde{V} \int_\Omega (\sum_{k=1}^s |v_k|^2)^{1/2} (\sum_{l=1}^s |w_l|^2)^{1/2} \le s\tilde{V}\, |||w|||\, |||v|||,$$

i.e. (16) holds with the constant function

$$M_R(r) \equiv s\tilde{V} \qquad (r \ge 0). \qquad \blacksquare \tag{76}$$

Now we consider a finite element discretization for problem (46), developed as in subsection 4.1.2. We can then prove the following nonnegativity result for the stiffness matrix:

Theorem 4.1 *Let problem (46) satisfy Assumptions 4.1. Let us consider a family of finite element subspaces $\mathcal{V} = \{V_h\}_{h\to 0}$ satisfying the following property: there exists a real number γ satisfying*

$$d \le \gamma < d + 2$$

(where d is the space dimension) such that for any $p = 1, ..., \bar{n}_0,\ t = 1, ..., \bar{n}$ ($p \ne t$), if $meas(\operatorname{supp}\varphi_p \cap \operatorname{supp}\varphi_t) > 0$ then

$$\nabla\varphi_t \cdot \nabla\varphi_p \le 0 \ \text{ on } \Omega \quad \text{and} \quad \int_\Omega \nabla\varphi_t \cdot \nabla\varphi_p \le -K_0\, h^{\gamma-2} \tag{77}$$

with some constant $K_0 > 0$ independent of p, t and h. Further, let the family of associated meshes be regular from above, according to Definition 4.1.

Then for sufficiently small h, the matrix $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ defined in (65) is of generalized nonnegative type with irreducible blocks in the sense of Definition 2.4.

Proof. We wish to apply Theorem 3.1. With the operator A defined in (53), our problem (51)-(52) coincides with (11)-(12). The FEM subspaces (59) and (61) fall into the class (25). Using the operators $B(u)$ and $R(u)$ in (72), the discrete problem (62)-(63) turns into the form (28) such that by Lemma 4.1, $B(u)$ and $R(u)$ satisfy Assumptions 3.1 in the spaces $H = H^1(\Omega)^s$ and $H_0 = H^1_D(\Omega)^s$. Our remaining task is therefore to check items (a)-(e) of Theorem 3.1.

For this, we first need to define neighbouring basis functions as required in Definition 3.1. Let $\phi_i, \phi_j \in V_h$. Using definitions (58) and (60), assume that ϕ_i has φ_p at its k-th entry and ϕ_j has φ_t at its l-th entry. Then we call ϕ_i and

ϕ_j neighbouring basis functions if $k = l$ and $meas(\operatorname{supp}\varphi_p \cap \operatorname{supp}\varphi_t) > 0$. Let $N := \{1, \dots, n\}$ as before. For any $k = 1, \dots, s$ let

$$S_k^0 := \{i \in N : \ i = (k-1)\bar{n}_0 + p \text{ for some } 1 \leq p \leq \bar{n}_0\},$$

$$\tilde{S}_k := \{i \in N : \ i = n_0 + (k-1)(\bar{n} - \bar{n}_0) + p - \bar{n}_0 \text{ for some } \bar{n}_0 + 1 \leq p \leq \bar{n}\},$$

$$S_k := S_k^0 \cup \tilde{S}_k \,,$$

i.e. by (58) and (60), the basis functions ϕ_i with index $i \in S_k$ have a nonzero coordinate φ_p for some p at the k-th entry, and in particular, $i \in S_k^0$ if this φ_p is an 'interior' basis function (i.e. $1 \leq p \leq \bar{n}_0$) and $i \in \tilde{S}_k$ if this φ_p is a 'boundary' basis function (i.e. $\bar{n}_0 + 1 \leq p \leq \bar{n}$). Clearly, the set $N = \{1, \dots, n\}$ can be partitioned into the disjoint sets $S_1, \dots, S_s$, and we have to check items (i)-(iii) of Definition 3.1. Let $k \in \{1, \dots, s\}$. By definition $S_k^0 = S_k \cap \{1, \dots, n_0\}$ and $\tilde{S}_k = S_k \cap \{n_0 + 1, \dots, n\}$, and both S_k^0 and $\tilde{S}_k$ are nonempty, hence item (i) holds. We have assumed in subsection 4.1.2 that any two 'interior' basis functions φ_p, φ_t can be connected with a chain of interior basis functions with overlapping support. Writing the terms of this chain at the k-th entry of the vector basis function, this just means that the graph of all neighbouring indices in S_k^0 is connected, i.e. item (ii) holds. Finally, it follows from (56) that arbitrary two basis functions φ_p, φ_t can be connected with a chain of basis functions with overlapping support. (Namely, take the union of the supports of the basis functions in all possible chains with overlapping supports from φ_p. If the obtained set Ω_p were not the entire Ω, then we would have $\sum_{p=1}^{\bar{n}} \varphi_p(x) = 0$ for $x \in \partial\Omega$ in contrast to (56). Therefore $\Omega_p = \Omega$, hence one of the chains reaches φ_t as well.) Writing the terms of this chain at the k-th entry of the vector basis function, this just means as above that the graph of all neighbouring indices in S_k is connected, i.e. item (iii) holds.

Now we are in the position to check assumptions (a)-(e) of Theorem 3.1.

(a) Let $\phi_i \in V_h^0$, $\phi_j \in V_h$, and let ϕ_i have φ_p at its k-th entry and ϕ_j have φ_t at its l-th entry. We must prove that either (37) or (38)-(40) holds. If $k \neq l$ then ϕ_i and ϕ_j have no common nonzero coordinates, hence $\langle B(u^h)\phi_j, \phi_i \rangle = 0$; further, by (47) and (56),

$$\langle R(u^h)\phi_j, \phi_i \rangle = \int_\Omega V_{kl}(x, u^h, \nabla u^h)\, \varphi_t\, \varphi_p \leq 0 \tag{78}$$

i.e. (37) holds. If $k = l$, then Assumption 4.1 (iii) and (77) yield

$$\langle B(u^h)\phi_j, \phi_i \rangle = \int_\Omega b_k(x, u^h, \nabla u^h)\, \nabla\varphi_t \cdot \nabla\varphi_p \leq m \int_{\Omega_{pt}} \nabla\varphi_t \cdot \nabla\varphi_p \tag{79}$$

where $\Omega_{pt} := \operatorname{supp}\varphi_p \cap \operatorname{supp}\varphi_t$. If $meas(\Omega_{pt}) = 0$ then $\langle B(u^h)\phi_j, \phi_i \rangle = 0$ and we have (78) similarly as before, hence (37) holds again. If $meas(\Omega_{pt}) > 0$ then (77) implies

$$\langle B(u^h)\phi_j, \phi_i \rangle \leq -mK_0 h^{\gamma-2} \equiv -\hat{c}_1\, h^{\gamma-2} =: -M_B(h) \tag{80}$$

and we must check (40). Here the norm (75) of the basis functions satisfies the following estimate, where ϕ_j has φ_t at its l-th entry as before, and we use (68) and that (56) implies $\varphi_t \le 1$:

$$|||\phi_j|||^2 = \|\phi_j\|^2_{L^2(\Omega)^s} = \|\varphi_t\|^2_{L^2(\Omega)} \le \int_{\operatorname{supp}\varphi_t} 1 \; = meas(\operatorname{supp}\varphi_t) \le c_2 h^d \,, \tag{81}$$

hence (39) gives $T(h)^2 \le h^d$. From this, using (80) and that $\gamma < d+2$ (as defined for (70)), we obtain

$$\lim_{h\to 0} \frac{M_B(h)}{T(h)^2} \ge \frac{\hat{c}_1}{c_2} \lim_{h\to 0} \; h^{\gamma-2-d} = +\infty. \tag{82}$$

(b) Let $\phi_i \in V_h^0$ and $\phi_j \in V_h$ be neighbouring basis vectors, i.e, as defined before in the proof, $k = l$ and $meas(\operatorname{supp}\varphi_p \cap \operatorname{supp}\varphi_t) > 0$. Then, as seen just above, we obtain (80) and (82), which coincide with (38)-(40).

(c) We have obtained the constant bound $M_R(r) \equiv s\tilde{V}$ in (76) for Assumption 3.1 (iii), hence $M_R(\|u^h\|) \equiv s\tilde{V}$ is trivially bounded as $h \to 0$.

(d) For all $u \in H^1(\Omega)^s$ and $h > 0$, the definition of the functions ϕ_j and assumption (56) imply

$$\sum_{j=1}^{n} \phi_j = \begin{pmatrix} \sum\limits_{p=1}^{\bar{n}} \varphi_p \\ \sum\limits_{p=1}^{\bar{n}} \varphi_p \\ \dots \\ \sum\limits_{p=1}^{\bar{n}} \varphi_p \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} =: \mathbf{1}\,. \tag{83}$$

Then by (72)

$$\langle B(u)(\sum_{j=1}^{n} \phi_j), v\rangle = \langle B(u)\mathbf{1}, v\rangle = \int_\Omega \sum_{k=1}^{s} b_k(x,u,\nabla u)\, \nabla 1 \cdot \nabla v_k = 0$$

for all $v \in H^1_D(\Omega)^s$, i.e. $\sum\limits_{j=1}^{n} \phi_j \in \ker B(u)$.

(e) Let $h > 0$ and $i = 1, ..., n$ be arbitrary. We must prove that $\phi_i \in D$ and $\sum\limits_{j=1}^{n} \phi_j \in P$ for the sets D, P defined in the proof of Lemma 4.1, paragraph (iii). First, by definition, ϕ_i has only one nonzero coordinate function φ_p that is nonnegative by (56), i.e. $\phi_i \in D$. Second, as seen in (83), we have $\sum\limits_{j=1}^{n} \phi_j = \mathbf{1}$ which belongs to P. ■

We immediately obtain

Corollary 4.1 *Let the assumptions of Theorem 4.1 hold and let $f_k \le 0$, $\gamma_k \le 0$ ($k = 1, \dots, s$). For sufficiently small h, if $\bar{\mathbf{c}} = (c_1, ..., c_n)^T \in \mathbf{R}^n$ is the solution of (64) with matrix $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ defined in (65), then*

$$\max_{i=1,...,n} c_i \le \max\{0, \max_{i=n_0+1,...,n} c_i\}. \tag{84}$$

Proof. By (66) we have $d_i \le 0$ ($i = 1, ..., n_0$), hence Corollary 3.1 can be used. ■

The meaning of (84) is as follows. Let us split the vector $\bar{\mathbf{c}} = (c_1, ..., c_n)^T \in \mathbf{R}^n$ as in (35), i.e. $\bar{\mathbf{c}} = [\mathbf{c}, \tilde{\mathbf{c}}]^T$, where $\mathbf{c} = (c_1, ..., c_{n_0})^T$ and $\tilde{\mathbf{c}} = (c_{n_0+1}, ..., c_n)^T$. Following the notions introduced after (55), the vectors $\mathbf{c}$ and $\tilde{\mathbf{c}}$ contain the coefficients of the 'interior basis functions' and 'boundary basis functions', respectively. Then (84) states that the maximal coordinate is nonpositive or arises for a boundary basis function.

Our main interest is the meaning of Corollary 4.1 for the FEM solution $u^h = (u_1^h, \dots, u_s^h)$ itself.

Theorem 4.2 *Let the basis functions satisfy (56)-(57). If (84) holds for the FEM solution $u^h = (u_1^h, \dots, u_s^h)$, then u^h satisfies*

$$\max_{k=1,...,s} \max_{\overline{\Omega}} u_k^h \le \max_{k=1,...,s} \max\{0, \max_{\Gamma_D} g_k^h\}. \tag{85}$$

Proof. Let us refine the above splitting $\bar{\mathbf{c}} = [\mathbf{c}, \tilde{\mathbf{c}}]^T$ of the vector $\bar{\mathbf{c}} = (c_1, ..., c_n) \in \mathbf{R}^n$ as

$$\bar{\mathbf{c}} = (\, c_1^{(1)}, \dots, c_{\bar{n}_0}^{(1)};\, c_1^{(2)}, \dots, c_{\bar{n}_0}^{(2)}; \dots;\, c_1^{(s)}, \dots, c_{\bar{n}_0}^{(s)}; \\ c_{\bar{n}_0+1}^{(1)}, \dots, c_{\bar{n}}^{(1)};\, c_{\bar{n}_0+1}^{(2)}, \dots, c_{\bar{n}}^{(2)}; \dots;\, c_{\bar{n}_0+1}^{(s)}, \dots, c_{\bar{n}}^{(s)}), \tag{86}$$

that is, $\mathbf{c}$ has the $n_0 = s\bar{n}_0$ entries from $c_1^{(1)}$ to $c_{\bar{n}_0}^{(s)}$ belonging to the interior points, and $\tilde{\mathbf{c}}$ has the $n - n_0 = s(\bar{n} - \bar{n}_0)$ entries from $c_{\bar{n}_0+1}^{(1)}$ to $c_{\bar{n}}^{(s)}$ belonging to the boundary points, such that the upper index from 1 to s gives the number of coordinate in the elliptic system. Here for all $k = 1, \dots, s$ we have

$$u_k^h = \sum_{p=1}^{\bar{n}} c_p^{(k)} \varphi_p \,.$$

Now let $k^* \in \{1, \dots, s\}$ and $p^* \in \{1, \dots, \bar{n}\}$ be indices such that

$$c_{p^*}^{(k^*)} = \max_{i=1,...,n} c_i \,.$$

For all $k = 1, \dots, s$, using (56),

$$\max_{\overline{\Omega}} u_k^h = \max_{\overline{\Omega}} \sum_{p=1}^{\bar{n}} c_p^{(k)} \varphi_p \le c_{p^*}^{(k^*)} \sum_{p=1}^{\bar{n}} \varphi_p = c_{p^*}^{(k^*)} \,,$$

further, using (57),

$$u^h_{(k^*)}(B_{p^*}) = \sum_{p=1}^{\bar{n}} c_p^{(k^*)} \varphi_p(B_{p^*}) = \sum_{p=1}^{\bar{n}} c_p^{(k^*)} \delta_{p,p^*} = c_{p^*}^{(k^*)} \ .$$

These together mean that

$$\max_{k=1,\dots,s} \max_{\overline{\Omega}} u^h_k = u^h_{(k^*)}(B_{p^*}).$$

By (84), either $c_{p^*}^{(k^*)} \le 0$ or $p^* \in \{n_0+1, \dots, \bar{n}\}$ (i.e. p^* is a 'boundary index', for which $B_{p^*} \in \Gamma_D$). In the first case

$$\max_{k=1,\dots,s} \max_{\overline{\Omega}} u^h_k = u^h_{(k^*)}(B_{p^*}) = c_{p^*}^{(k^*)} \le 0,$$

and in the second case

$$\max_{k=1,\dots,s} \max_{\overline{\Omega}} u^h_k = u^h_{(k^*)}(B_{p^*}) \le \max_{\Gamma_D} u^h_{(k^*)} \le \max_{k=1,\dots,s} \max_{\Gamma_D} u^h_k = \max_{k=1,\dots,s} \max_{\Gamma_D} g^h_k \ .$$

(In fact, there is of course equality in the above estimate.) These two relations just mean that (85) holds. ■

Thus we obtain the *discrete maximum principle* for system (46):

Corollary 4.2 *Let the assumptions of Theorem 4.1 hold and let*

$$f_k \le 0, \qquad \gamma_k \le 0 \qquad (k = 1, \dots, s).$$

Let the basis functions satisfy (56)-(57). Then for sufficiently small h, if $u^h = (u^h_1, \dots, u^h_s)$ *is the FEM solution of system (46), then*

$$\max_{k=1,\dots,s} \max_{\overline{\Omega}} u^h_k \le \max_{k=1,\dots,s} \max\{0, \max_{\Gamma_D} g^h_k\}. \tag{87}$$

Remark 4.2 (i) Let $f_k \le 0$, $\gamma_k \le 0$ for all k. The result (87) can be divided in two cases, both of which are remarkable: if at least one of the functions g^h_k has positive values on Γ_D then

$$\max_{k=1,\dots,s} \max_{\overline{\Omega}} u^h_k = \max_{k=1,\dots,s} \max_{\Gamma_D} g^h_k \tag{88}$$

(which can be called more directly a discrete maximum principle than (87)), and if $g_k \le 0$ on Γ_D for all k, then we obtain the nonpositivity property

$$u^h_k \le 0 \quad \text{on } \Omega \text{ for all } k\,. \tag{89}$$

(ii) Analogously, if $f_k \ge 0$, $\gamma_k \ge 0$ for all k, then (by reversing signs) we can derive the corresponding discrete minimum principles instead of (87) and (88), or the corresponding nonnegativity property instead of (89).

Remark 4.3 The key assumption for the meshes in the above results is property (77). A simple but stronger sufficient condition to satisfy (77) is the following: if for any $p = 1, ..., \bar{n}_0,\ t = 1, ..., \bar{n}\ (p \neq t)$

$$\nabla\varphi_t \cdot \nabla\varphi_p \leq -\frac{\sigma}{h^2} < 0 \tag{90}$$

on $supp\ \varphi_p \cap supp\ \varphi_t$ with some constant $\sigma > 0$ independent of p, t and V_h, and in addition, if the family of meshes is quasi-regular according to Definition 4.1, then (77) is satisfied. For simplicial FEM, assumption (90) corresponds to acute triangulations. These properties and less strong assumptions to satisfy (77) will be discussed in subsection 4.4.

Remark 4.4 The results of this section may hold as well if there are additional terms $\sum_{l=1}^{s} \omega_{kl}(x, u, \nabla u)\, u_l$ on the Neumann boundary Γ_N, which we did not include for technical simplicity. Then ω_{kl} must satisfy similar properties as assumed for V_{kl} in (47)-(48).

4.2 Systems with general reaction terms of sublinear growth

It seems somewhat restrictive in (46) that both the principal and lower-order parts of the equations are given as containing products of coefficients with ∇u_k and u_l, respectively. Whereas this is widespread in real models for the principal part (and often the coefficient of ∇u_k depends only on x, or x and $|\nabla u|$), on the contrary, the lower order terms are usually not given in such a coefficient form. Now we consider problems where the dependence on the lower order terms is given as general functions of x and u. In this section these functions are allowed to grow at most linearly, in which case one can reduce the problem to the previous one (46) directly. (Superlinear growth of q_k will be dealt with in the next section.)

Accordingly, let us now consider the system

$$\left.\begin{aligned} -\mathrm{div}\left(b_k(x, u, \nabla u)\, \nabla u_k\right) + \ q_k(x, u_1, \ldots, u_s) \ &= \ f_k(x) \quad \text{in } \Omega, \\ b_k(x, u, \nabla u)\tfrac{\partial u_k}{\partial \nu} \ &= \gamma_k(x) \quad \text{on } \Gamma_N, \\ u_k \ &= \ g_k(x) \quad \text{on } \Gamma_D \end{aligned}\right\} \quad (k = 1, \ldots, s) \tag{91}$$

under the following assumptions:

Assumptions 4.2.

(i) $\Omega \subset \mathbf{R}^d$ is a bounded piecewise C^1 domain; Γ_D, Γ_N are disjoint open measurable subsets of $\partial\Omega$ such that $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$.

(ii) (Smoothness and boundedness.) For all $k, l = 1, \ldots, s$ we have $b_k \in (C^1 \cap L^\infty)(\Omega \times \mathbf{R}^s \times \mathbf{R}^{s\times d})$ and $q_k \in W^{1,\infty}(\Omega \times \mathbf{R}^s)$.

(iii) (Ellipticity.) There exists $m > 0$ such that $b_k \geq m$ holds for all $k = 1, \dots, s$.

(iv) (Cooperativity.) We have

$$\frac{\partial q_k}{\partial \xi_l}(x, \xi) \leq 0 \qquad (k, l = 1, \dots, s,\ k \neq l;\ x \in \Omega,\ \xi \in \mathbf{R}^s). \tag{92}$$

(v) (Weak diagonal dominance for the Jacobians.) We have

$$\sum_{l=1}^{s} \frac{\partial q_k}{\partial \xi_l}(x, \xi) \geq 0 \qquad (k = 1, \dots, s;\ x \in \Omega,\ \xi \in \mathbf{R}^s). \tag{93}$$

(vi) For all $k = 1, \dots, s$ we have $f_k \in L^2(\Omega)$, $\gamma_k \in L^2(\Gamma_N)$, $g_k = g^*_{k|\Gamma_D}$ with $g^* \in H^1(\Omega)$.

Remark 4.5 Similarly to (49), assumptions (92)-(93) now imply

$$\frac{\partial q_k}{\partial \xi_k}(x, \xi) \geq 0 \qquad (k = 1, \dots, s;\ x \in \Omega,\ \xi \in \mathbf{R}^s). \tag{94}$$

The basic idea to deal with problem (91) is to reduce it to (46) via suitably defined functions $V_{kl} : \Omega \times \mathbf{R}^s \to \mathbf{R}$. Namely, let

$$V_{kl}(x, \xi) := \int_0^1 \frac{\partial q_k}{\partial \xi_l}(x, t\xi)\, dt \qquad (k, l = 1, \dots, s;\ x \in \Omega,\ \xi \in \mathbf{R}^s). \tag{95}$$

Then the Newton-Leibniz formula yields

$$q_k(x, \xi) = q_k(x, 0) + \sum_{l=1}^{s} V_{kl}(x, \xi)\, \xi_l \qquad (k = 1, \dots, s;\ x \in \Omega,\ \xi \in \mathbf{R}^s). \tag{96}$$

Defining

$$\hat{f}_k(x) := f_k(x) - q_k(x, 0) \qquad (k = 1, \dots, s), \tag{97}$$

problem (91) then becomes

$$\left.\begin{aligned} -\mathrm{div}\left(b_k(x, u, \nabla u)\, \nabla u_k\right) + \sum_{l=1}^{s} V_{kl}(x, u)\, u_l &= \hat{f}_k(x) \quad \text{in } \Omega, \\ b_k(x, u, \nabla u)\tfrac{\partial u_k}{\partial \nu} &= \gamma_k(x) \quad \text{on } \Gamma_N, \\ u_k &= g_k(x) \quad \text{on } \Gamma_D \end{aligned}\right\} \quad (k = 1, \dots, s), \tag{98}$$

which is a special case of (46). Here the assumption $q_k \in W^{1,\infty}(\Omega \times \mathbf{R}^s)$ yields that $V_{kl} \in L^\infty(\Omega \times \mathbf{R}^s)$ $(k, l = 1, \dots, s)$. Clearly, assumptions (92) and (93) imply that the functions V_{kl} defined in (95) satisfy (47) and (48), respectively. The remaining items of Assumptions 4.1 and 4.2 coincide, therefore system (98) satisfies Assumptions 4.2.

Consequently, for a finite element discretization developed as in subsection 4.1.2, Theorem 4.2 yields the discrete maximum principle (85) for suitable discretizations of (98), provided $\hat{f}_k \leq 0$ and $\gamma_k \leq 0$ $(k = 1, \dots, s)$. For the original system (91), we thus obtain

Corollary 4.3 *Let problem (91) satisfy Assumptions 4.2, and let its FEM discretization satisfy the corresponding conditions of Theorem 4.1. If*

$$f_k \le q_k(x,0), \qquad \gamma_k \le 0 \qquad (k=1,\dots,s)$$

and $u^h = (u_1^h,\dots,u_s^h)$ is the FEM solution of system (91), then for sufficiently small h,

$$\max_{k=1,\dots,s} \max_{\overline{\Omega}} u_k^h \le \max_{k=1,\dots,s} \max\{0, \max_{\Gamma_D} g_k^h\}. \tag{99}$$

4.3 Systems with general reaction terms of superlinear growth

In the previous section we have required the functions q_k to grow at most linearly via the condition $q_k \in W^{1,\infty}(\Omega \times \mathbf{R}^s)$. However, this is a strong restriction and is not satisfied even by (nonlinear) polynomials of u_k that often arise in reaction-diffusion problems. In this section we extend the previous results to problems where the functions q_k may grow polynomially. This generalization, however, needs stronger assumptions in other parts of the problem, because we now need the monotonicity of the corresponding operator in the proof of the DMP. For this to hold, the row-diagonal dominance for the Jacobians.in assumption 4.2 (v) must be strengthened to diagonal dominance w.r.t. both rows and columns. (In addition, the principal part must be more specific too, but this is not so much restrictive since in practice it is even linear.)

Accordingly, let us now consider the system

$$\left.\begin{aligned} -\mathrm{div}\left(b_k(x,\nabla u_k)\,\nabla u_k\right) + \; q_k(x,u_1,\dots,u_s) \; &= \; f_k(x) \quad \text{in } \Omega, \\ b_k(x,\nabla u_k)\tfrac{\partial u_k}{\partial \nu} &= \gamma_k(x) \quad \text{on } \Gamma_N, \\ u_k \; &= \; g_k(x) \quad \text{on } \Gamma_D \end{aligned}\right\} \qquad (k=1,\dots,s) \tag{100}$$

under the following assumptions:

Assumptions 4.3.

(i) $\Omega \subset \mathbf{R}^d$ is a bounded piecewise C^1 domain; Γ_D, Γ_N are disjoint open measurable subsets of $\partial\Omega$ such that $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$.

(ii) (Smoothness and growth.) For all $k,l = 1,\dots,s$ we have $b_k \in (C^1 \cap L^\infty)(\Omega \times \mathbf{R}^d)$ and $q_k \in C^1(\Omega \times \mathbf{R}^s)$. Further, let

$$2 \le p < p^*, \quad \text{where } p^* := \tfrac{2d}{d-2} \;\text{ if } d \ge 3 \;\text{ and }\; p^* := +\infty \text{ if } d=2; \tag{101}$$

then there exist constants $\beta_1, \beta_2 \ge 0$ such that

$$\left|\frac{\partial q_k}{\partial \xi_l}(x,\xi)\right| \le \beta_1 + \beta_2 |\xi|^{p-2} \qquad (k,l=1,\dots,s;\; x \in \Omega,\; \xi \in \mathbf{R}^s). \tag{102}$$

(iii) (Ellipticity.) There exists $m > 0$ such that $b_k \geq m$ holds for all $k = 1, \ldots, s$. Further, defining $a_k(x, \eta) := b_k(x, \eta)\eta$ for all k, the Jacobian matrices $\frac{\partial}{\partial \eta} a_k(x, \eta)$ are uniformly spectrally bounded from both below and above.

(iv) (Cooperativity.) We have

$$\frac{\partial q_k}{\partial \xi_l}(x, \xi) \leq 0 \qquad (k, l = 1, \ldots, s, \ k \neq l; \ x \in \Omega, \ \xi \in \mathbf{R}^s). \tag{103}$$

(v) (Weak diagonal dominance for the Jacobians w.r.t. rows and columns.) We have

$$\sum_{l=1}^{s} \frac{\partial q_k}{\partial \xi_l}(x, \xi) \geq 0, \qquad \sum_{l=1}^{s} \frac{\partial q_l}{\partial \xi_k}(x, \xi) \geq 0 \qquad (k = 1, \ldots, s; \ x \in \Omega, \ \xi \in \mathbf{R}^s). \tag{104}$$

(vi) For all $k = 1, \ldots, s$ we have $f_k \in L^2(\Omega)$, $\gamma_k \in L^2(\Gamma_N)$, $g_k = g^*_{k|\Gamma_D}$ with $g^* \in H^1(\Omega)$.

Remark 4.6 (i) Similarly to (49), assumptions (103)-(104) now imply

$$\frac{\partial q_k}{\partial \xi_k}(x, \xi) \geq 0 \qquad (k = 1, \ldots, s; \ x \in \Omega, \ \xi \in \mathbf{R}^s). \tag{105}$$

(ii) Similarly to Remark 4.4, one may include additional terms $s_k(x, u_1, \ldots, u_s)$ on the Neumann boundary Γ_N, which we omit here for technical simplicity; then s_k must satisfy similar properties as assumed for q_k.

To handle system (100), we start as in the previous subsection by reducing it to a system with nonlinear coefficients: if the functions V_{kl} and $\hat{f}_k$ $(k, l = 1, \ldots, s)$ are defined as in (95) and (97), respectively, then (100) takes a form similar to (98):

$$\left.\begin{aligned} -\mathrm{div}\left(b_k(x, \nabla u)\, \nabla u_k\right) + \sum_{l=1}^{s} V_{kl}(x, u)\, u_l &= \hat{f}_k(x) \quad \text{in } \Omega, \\ b_k(x, u, \nabla u) \tfrac{\partial u_k}{\partial \nu} &= \gamma_k(x) \quad \text{on } \Gamma_N, \\ u_k &= g_k(x) \quad \text{on } \Gamma_D \end{aligned}\right\} \qquad (k = 1, \ldots, s). \tag{106}$$

The difference compared to the previous subsection is the superlinear growth allowed in (102), which does not let us apply Theorem 4.2 directly as we did for system (91). Instead, we must reprove Theorem 4.1 under Assumptions 4.3.

First, when considering a finite element discretization developed as in subsection 4.1.2, we need a strengthened assumption for the quasi-regularity of the mesh.

Definition 4.2 Let $\Omega \subset \mathbf{R}^d$ and let us consider a family of FEM subspaces $\mathcal{V} = \{V_h\}_{h\to 0}$ constructed as in subsection 4.1.2. The corresponding mesh will be called *quasi-regular* w.r.t. problem (100) if

$$c_1 h^\gamma \le meas(\operatorname{supp} \varphi_p) \le c_2 h^d\,, \tag{107}$$

where the positive real number γ satisfies

$$d \le \gamma < \ \gamma_d^*(p) :=\ 2d - \frac{(d-2)p}{2} \tag{108}$$

with p from Assumption 4.3 (ii).

Remark 4.7 Assumption (108) makes sense for γ since by (101),

$$d < d + d(1 - \tfrac{p}{p^*}) = \gamma_d^*(p)\,. \tag{109}$$

Note on the other hand that $\gamma_d^*(p) \le \gamma_d^*(2) = d+2$, which is in accordance with (71). Further, we have, in particular, in 2D: $\gamma_2^*(p) \equiv 4$ for all $2 \le p < \infty$, and in 3D: $\gamma_3^*(p) = 6 - (p/2)$ (where $2 \le p \le 6$, and accordingly $3 \le \gamma_3^*(p) \le 5$).

Next, as an analogue of Lemma 4.1, we need a technical result for problem (100):

Lemma 4.2 *Let Assumptions 4.3 hold. Analogously to (72), for any $u \in H^1(\Omega)^s$ let us define the operators $B(u)$ and $R(u)$ via*

$$\langle B(u)w, v\rangle = \int_\Omega \sum_{k=1}^s b_k(x, \nabla u)\, \nabla w_k \cdot \nabla v_k, \quad \langle R(u)w, v\rangle = \int_\Omega \sum_{k,l=1}^s V_{kl}(x,u)\, w_l\, v_k$$

($w \in H^1(\Omega)^s$, $v \in H_D^1(\Omega)^s$). Together with $A(u) := B(u)u + R(u)u$, the operators $B(u)$ and $R(u)$ satisfy Assumptions 3.1-Assumptions 3.2.

Proof. First, we must verify Assumptions 3.1. The stronger growth (102) causes a difference only in proving Assumption 3.1 (iv), i.e. to fulfil (16). Hence we only verify this property, the proof of the other items of Assumption 3.1 is the same as in Lemma 4.1.

Consider p^* as defined in (101). Then by [1]) we have the Sobolev embedding estimate

$$\|h\|_{L^{p^*}(\Omega)} \le k_1 \|h\|_{H^1} \quad (h \in H^1(\Omega)) \tag{110}$$

with a constant $k_1 > 0$, where $\|h\|_{H^1}^2 := \int_\Omega |\nabla h|^2 + \int_{\Gamma_D} |h|^2$. This is inherited for $v \in H^1(\Omega)^s$ too under the product norm $\|.\|$ on $H^1(\Omega)^s$ defined in (73). Here, by (95) and (102),

$$|\langle R(u)w, v\rangle| = |\int_\Omega \sum_{k,l=1}^s V_{kl}(x,u)\, w_l\, v_k| \le \int_\Omega \sum_{k,l=1}^s (\beta_1 + \beta_2 |u|^{p-2})\, |w_l|\, |v_k|\,. \tag{111}$$

Letting $|v|^2 := \sum_{k=1}^{s} v_k^2 \quad (v \in H^1(\Omega)^s)$, we have $\sum_{k,l=1}^{s} |w_l|\,|v_k| \le s|w|\,|v|$, hence

$$|\langle R(u)w, v\rangle| \le s \int_\Omega (\beta_1 + \beta_2 |u|^{p-2})\,|w|\,|v|\,. \tag{112}$$

For vector functions $v \in L^p(\Omega)^s$, we define

$$\|v\|_{L^p} := \big\||v|\big\|_{L^p(\Omega)} \tag{113}$$

with $|v|$ defined as above. Let us now fix a real number r satisfying

$$1 < r \le \frac{p^*}{p-2}\,. \tag{114}$$

If $q > 1$ is chosen to have

$$\frac{1}{r} + \frac{1}{q} = 1, \tag{115}$$

then Hölder's inequality implies

$$\int_\Omega |u|^{p-2}\,|w|\,|v| \;\le \big\||u|^{p-2}\big\|_{L^r(\Omega)} \|w\|_{L^{2q}} \|v\|_{L^{2q}} = \|u\|_{L^{(p-2)r}}^{p-2} \|w\|_{L^{2q}} \|v\|_{L^{2q}}. \tag{116}$$

Here $(p-2)r \le p^*$ and (110) imply

$$\|u\|_{L^{(p-2)r}(\Omega)}^{p-2} \le k_2 \|u\|_{L^{p^*}(\Omega)}^{p-2} \le k_3 \|u\|^{p-2} \tag{117}$$

with some constants $k_2, k_3 > 0$. Setting $u \equiv 1$ in (116) and using (117), we obtain

$$\int_\Omega |w|\,|v| \;\le k_4\,\|w\|_{L^{2q}}\,\|v\|_{L^{2q}} \tag{118}$$

with some constant $k_4 > 0$. Then (112), (118) and (116) imply

$$|\langle R(u)w, v\rangle| \le s\,(\beta_1\, k_4 + \beta_2\, k_3\, \|u\|^{p-2})\,\|w\|_{L^{2q}}\,\|v\|_{L^{2q}}. \tag{119}$$

That is, if we define the new norm $|||\,.\,|||$ as

$$|||v||| := \|v\|_{L^{2q}} \qquad (v \in H^1(\Omega)^s), \tag{120}$$

then (16) holds with

$$M_R(t) := s(\beta_1\, k_4 + \beta_2\, k_3\, t^{p-2}) \qquad (t \ge 0). \tag{121}$$

Now we have to verify Assumptions 3.2. Note first that we have

$$\langle A(u), v\rangle = \int_\Omega \Big(\sum_{k=1}^{s} b_k(x, \nabla u)\, \nabla u_k \cdot \nabla v_k + \sum_{k,l=1}^{s} V_{kl}(x,u)\, u_l\, v_k \Big) \tag{122}$$

($u \in H^1(\Omega)^s$, $v \in H_D^1(\Omega)^s$). Using the notation a_k in Assumption 4.3 (iii) and from (96), we obtain

$$\langle A(u), v\rangle = \langle F(u), v\rangle + \langle G(u), v\rangle \tag{123}$$

where

$$\langle F(u), v\rangle = \int_\Omega \sum_{k=1}^s a_k(x, \nabla u)\cdot\nabla v_k, \qquad \langle G(u), v\rangle = \int_\Omega \Big(\sum_{k=1}^s q_k(x,u)\, v_k - \sum_{k=1}^s q_k(x,0)\, v_k\Big) \tag{124}$$

($u \in H^1(\Omega)^s$, $v \in H^1_D(\Omega)^s$). Here, by Assumption 4.3 (iii), there exist constants $M \geq m > 0$ such that

$$\frac{\partial a_k}{\partial \eta}(x,\eta)\ \xi\cdot\xi \geq\ m|\xi|^2, \qquad \frac{\partial a_k}{\partial \eta}(x,\eta)\ \xi\cdot\zeta \leq\ M|\xi|\,|\zeta| \tag{125}$$

($x \in \Omega,\ \ \eta, \xi, \zeta \in \mathbf{R}^d$). We can now check properties (i)-(iv) of Assumptions 3.2.

(i) Under Assumptions 4.3, it follows e.g. from [13, Theorem 6.2] that the operators F, G in (124) are Gateaux differentiable, further, that F' and G' are bihemicontinuous. In fact, the latter have the form

$$\langle F'(u)w, v\rangle = \int_\Omega \sum_{k=1}^s \frac{\partial a_k}{\partial \eta}(x, \nabla u)\, \nabla w_k\cdot\nabla v_k, \quad \langle G'(u)w, v\rangle = \int_\Omega \sum_{k,l=1}^s \frac{\partial q_k}{\partial \xi_l}(x,u)\, w_l\, v_k\,. \tag{126}$$

(ii) Let $u \in H^1(\Omega)^s$, $w, v \in H^1_D(\Omega)^s$. We obtain from (125) and (126) that

$$\langle F'(u)w, v\rangle \leq M\|w\|\,\|v\| \tag{127}$$

where $\|h\|^2 := \sum\limits_{k=1}^s \int_\Omega |\nabla h_k|^2$ is the product norm $\|.\|$ on $H^1_D(\Omega)^s$. Further, by (126) and (102),

$$|\langle G'(u)w, v\rangle| \leq \int_\Omega \sum_{k,l=1}^s (\beta_1 + \beta_2|u|^{p-2})\,|w_l|\,|v_k|\,. \tag{128}$$

This means that $G'(u)$ has the same bound as $R(u)$ in (111), but the latter has been estimated above by (119), hence $G'(u)$ also has the bound (119). If we now choose $r = \frac{p}{p-2}$ in (114), then (115) yields $q = \frac{p}{2}$, and setting the latter in the bound in (119) thus gives

$$|\langle G'(u)w, v\rangle| \leq (\beta_1\, k_4 + \beta_2\, k_3\, \|u\|^{p-2})\, \|w\|_{L^p}\, \|v\|_{L^p}\,. \tag{129}$$

Using (110) and that $p < p^*$, we obtain $\|w\|_{L^p} \leq k_5\|w\|_{L^{p^*}} \leq k_6\|w\|$ on $H^1_D(\Omega)^s$, hence

$$|\langle G'(u)w, v\rangle| \leq k_6(\beta_1\, k_4 + \beta_2\, k_3\, \|u\|^{p-2})\, \|w\|\, \|v\|\,. \tag{130}$$

Finally, from $A'(u) = F'(u) + G'(u)$, using (127) and (130), we obtain

$$|\langle A'(u)w, v\rangle| \leq \Big(M + k_6(\beta_1\, k_4 + \beta_2\, k_3\, \|u\|^{p-2})\Big)\, \|w\|\, \|v\|\,,$$

i.e. the required estimate (18) with $M_A(t) := M + k_6(\beta_1\, k_4 + \beta_2\, k_3\, t^{p-2})$ ($t \geq 0$).

(iii) We obtain immediately from (125) and (126) that

$$\langle F'(u)v,v\rangle \geq m\|v\|^2 \qquad (u\in H^1(\Omega)^s,\ v\in H^1_D(\Omega)^s). \tag{131}$$

(iv) By Assumptions 4.3 (iv)-(v), for all $x\in\Omega$ and $\xi\in\mathbf{R}^s$ the Jacobians $\frac{\partial q_k}{\partial \xi_l}(x,\xi)$ are M-matrices and weakly diagonally dominant w.r.t. both rows and columns. It is well-known that such matrices are positive semidefinite. Therefore

$$\langle G'(u)v,v\rangle = \int_\Omega \sum_{k,l=1}^{s} \frac{\partial q_k}{\partial \xi_l}(x,u)\, v_l\, v_k \geq\ 0 \qquad (u\in H,\ v\in H_0). \quad \blacksquare \tag{132}$$

Now we can prove the desired nonnegativity result for the stiffness matrix, i.e. the analogue of Theorem 4.1 for system (100). Here the entries of $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ are

$$a_{ij}(\bar{\mathbf{c}}) = \int_\Omega \Big(\sum_{k=1}^{s} b_k(x,\nabla u^h)\,(\nabla\phi_j)_k\cdot(\nabla\phi_i)_k + \sum_{k,l=1}^{s} V_{kl}(x,u^h)\,(\phi_j)_l\,(\phi_i)_k\Big), \tag{133}$$

where by (95),

$$V_{kl}(x,u^h(x)) = \int_0^1 \frac{\partial q_k}{\partial \xi_l}(x,tu^h(x))\,dt \qquad (k,l=1,\dots,s;\ x\in\Omega). \tag{134}$$

Theorem 4.3 *Let problem (100) satisfy Assumptions 4.3. Let us consider a family of finite element subspaces V_h ($h\to 0$) satisfying the following property: there exists a real number γ satisfying (108) such that for any indices $p=1,...,\bar{n}_0,\ t=1,...,\bar{n}$ ($p\neq t$), if meas$(\operatorname{supp}\varphi_p\cap\operatorname{supp}\varphi_t)>0$ then*

$$\nabla\varphi_t\cdot\nabla\varphi_p \leq 0 \ \text{ on } \Omega \quad \text{and} \quad \int_\Omega \nabla\varphi_t\cdot\nabla\varphi_p \leq -K_0\,h^{\gamma-2} \tag{135}$$

with some constant $K_0>0$ independent of p,t and h. Further, let the family of meshes be regular from above, according to Definition 4.1.

Then for sufficiently small h, the matrix $\bar{\mathbf{A}}(\bar{\mathbf{c}})$ defined in (133) is of generalized nonnegative type with irreducible blocks in the sense of Definition 2.4.

Proof. We follow the proof of Theorem 4.1 and wish to apply Theorem 3.1. Most of the arguments are identical, corresponding to the conditions that coincide in Assumptions 4.1 and 4.3. We will concentrate on the different parts. Since Assumptions 3.1 hold by Lemma 4.1, we are left to check assumptions (a)-(e) of Theorem 3.1.

(a) Let $\phi_i\in V_h^0,\ \ \phi_j\in V_h$, and let ϕ_i have φ_p at its k-th entry and ϕ_j have φ_t at its l-th entry. We obtain similarly as in the proof of Theorem 4.1 that (37) holds if either $k\neq l$, or $k=l$ and $meas(\Omega_{pt})=0$, where

$\Omega_{pt} := \mathrm{supp}\,\varphi_p \cap \mathrm{supp}\,\varphi_t$. The stronger growth (102) causes a difference only in verifying (38)-(40) in the case $k = l$ and $meas(\Omega_{pt}) > 0$. Here, in the same way as in (80), we obtain

$$\langle B(u^h)\phi_j, \phi_i\rangle \leq -\hat{c}_1\, h^{\gamma-2} =: -M_B(h) \tag{136}$$

and we must check (40). Let us now choose a real number r satisfying

$$\frac{d}{2+d-\gamma} < r \leq \frac{p^*}{p-2}\,. \tag{137}$$

Here $\gamma \geq 2$ implies $d/(2+d-\gamma) \geq 1$, hence (137) is a special case of (114). Such an r exists since

$$\frac{d}{2+d-\gamma} < \frac{p^*}{p-2}\,, \tag{138}$$

which holds for the following reason. If $d = 2$ then $p^* = +\infty$, hence there is nothing to prove. If $d \geq 3$ then we first observe that the fact $p \geq 2$ and (108) imply

$$\gamma < 2d - \tfrac{(d-2)p}{2} = d + 2 - \tfrac{(d-2)(p-2)}{2} \leq d + 2\,, \tag{139}$$

hence the denominator of $d/(2+d-\gamma)$ is positive. Hence we can take the reciprocal of (138) and use the definition $p^* := \frac{2d}{d-2}$ to obtain

$$2(2+d-\gamma) > (d-2)(p-2)$$

to be proved, but this just follows from the first inequality of (139). Now let $q > 1$ be chosen to satisfy (115), and by (120), let us define the corresponding norm

$$|||v||| := \|v\|_{L^{2q}} \qquad (v \in H^1(\Omega)^s), \tag{140}$$

for which, as seen in Lemma 4.2, estimate (16) holds with (121). Here (113) yields

$$|||v|||^2 = \Big\|\sum_{k=1}^{s} v_k^2\Big\|_{L^q(\Omega)} \qquad (v \in H^1(\Omega)^s). \tag{141}$$

Hence we obtain the following estimate, where ϕ_j has φ_t at its l-th entry as before, and we use (69) and that (56) implies $\varphi_t \leq 1$:

$$|||\phi_j|||^2 = \Big\|\,|\varphi_t|^2\Big\|_{L^q(\Omega)} = \|\,\varphi_t\|^2_{L^q(\Omega)} \leq \Big(\int_{\mathrm{supp}\,\varphi_t} 1\Big)^{1/q} = meas(\mathrm{supp}\,\varphi_t)^{1/q} \leq c_2 h^{d/q}\,, \tag{142}$$

hence (39) gives $T(h)^2 \leq h^{d/q}$. Here (115) and (137) imply $\gamma - 2 - (d/q) = \gamma - 2 - d + (d/r) < 0$. From this, using (136) we obtain

$$\lim_{h\to 0} \frac{M_B(h)}{T(h)^2} \geq \frac{\hat{c}_1}{c_2} \lim_{h\to 0}\, h^{\gamma-2-(d/q)} = +\infty. \tag{143}$$

(b) This assumption is proved identically to that in Theorem 4.1, using the same definition of neighbouring basis vectors.

(c) By (121), we must verify that $M_R(\|u^h\|) = s(\beta_1 k_4 + \beta_2 k_3 \|u^h\|^{p-2})$ is bounded as $h \to 0$. Note that Assumptions 3.2 hold by Lemma 4.1, and the functions $g^h \in V_h$ in (63) (that are the V_h-interpolants of g on Γ_D) are bounded in $H^1(\Omega)^s$-norm as $h \to 0$. From these two properties, as pointed out in Remark 3.1, it follows that $\|u^h\|$ is bounded as $h \to 0$, and then obviously $M_R(\|u^h\|)$ is bounded too.

(d)-(e) These assumptions are independent of the growth conditions on q_k, and are proved identically to those in Theorem 4.1. ∎

Similarly as in Corollary 4.3, using Theorem 4.3, Corollary 3.1 and Theorem 4.2, respectively, we obtain the *discrete maximum principle* for system (100):

Corollary 4.4 *Let problem (100) satisfy Assumptions 4.3, and let its FEM discretization satisfy the conditions of Theorem 4.3. If*

$$f_k \le q_k(x,0), \qquad \gamma_k \le 0 \qquad (k = 1,\ldots,s)$$

then for sufficiently small h, the FEM solution $u^h = (u_1^h, \ldots, u_s^h)$ of system (100) satisfies

$$\max_{k=1,\ldots,s} \max_{\overline{\Omega}} u_k^h \le \max_{k=1,\ldots,s} \max\{0, \max_{\Gamma_D} g_k^h\}. \tag{144}$$

Remark 4.8 As pointed out in Remark 4.2, the result (144) can be divided in two cases: a 'more direct' DMP (88) or the nonpositivity property (89). Further, if $f_k \ge q_k(x,0)$, $\gamma_k \ge 0$ for all k, then (by reversing signs) one can derive the corresponding discrete minimum principle or nonnegativity property. We formulate the latter below for its practical importance.

Corollary 4.5 *Let problem (100) satisfy Assumptions 4.3, and let its FEM discretization satisfy the conditions of Theorem 4.3. If*

$$f_k \ge q_k(x,0), \quad \gamma_k \ge 0, \quad g_k \ge 0 \qquad (k = 1,\ldots,s)$$

then for sufficiently small h, the FEM solution $u^h = (u_1^h, \ldots, u_s^h)$ of system (100) satisfies

$$u_k^h \ge 0 \quad \text{on } \Omega \qquad (k = 1,\ldots,s). \tag{145}$$

4.4 Sufficient conditions and their geometric meaning

The key assumption for the FEM subspaces V_h and the associated meshes in the above results has been the following property, see (77) in Theorem 4.1 and (135) in Theorem 4.3. There exists a real number γ satisfying (71) or

(108), respectively, such that for any indices $p = 1, ..., \bar{n}_0,\ t = 1, ..., \bar{n}\ (p \neq t)$, if $meas(\mathrm{supp}\,\varphi_p \cap \mathrm{supp}\,\varphi_t) > 0$ then

$$\nabla\varphi_t \cdot \nabla\varphi_p \leq 0 \quad \text{on } \Omega \quad \text{and} \tag{146}$$

$$\int_\Omega \nabla\varphi_t \cdot \nabla\varphi_p \leq -K_0\, h^{\gamma-2} \tag{147}$$

with some constant $K_0 > 0$ independent of p, t and h. (The family of meshes must also be regular from above as in (68), but that requirement obviously holds for the usual definition of the mesh parameter h as the maximal diameter of elements.)

A classical way to satisfy such conditions is a pointwise inequality like (90) together with suitable mesh regularity, see Remark 4.3. However, one can ensure (146)-(147) with less strong conditions as well. We summarize some possibilities below.

Proposition 4.1 *Let the family of FEM discretizations $\mathcal{V} = \{V_h\}_{h\to 0}$ satisfy either of the following conditions, where φ_t, φ_p are arbitrary basis functions such that $p = 1, ..., \bar{n}_0,\ t = 1, ..., \bar{n},\ p \neq t$, we let*

$$\Omega_{pt} := \mathrm{supp}\,\varphi_p \cap \mathrm{supp}\,\varphi_t\,,$$

further, let

$$\sigma > 0 \quad \textit{and} \quad c_1, c_2, c_3 > 0$$

denote constants independent of the indices p, t and the mesh parameter h, and finally, d is the space dimension and γ satisfies (108).

(i) Let the basis functions satisfy

$$\nabla\varphi_t \cdot \nabla\varphi_p \leq -\frac{\sigma}{h^2} < 0 \quad \textit{on } \Omega_{pt}, \tag{148}$$

and the family of meshes be quasi-regular as in (107):

$$c_1 h^\gamma \leq meas(\mathrm{supp}\,\varphi_p) \leq c_2 h^d\,. \tag{149}$$

(ii) Let there exist $0 < \varepsilon \leq \gamma - d$ such that the basis functions satisfy

$$\nabla\varphi_t \cdot \nabla\varphi_p \leq -\frac{\sigma}{h^{2-\varepsilon}} < 0 \quad \textit{on } \Omega_{pt}, \tag{150}$$

but let the quasi-regularity (107) of the family of meshes be now strengthened to

$$c_1 h^{\gamma-\varepsilon} \leq meas(\mathrm{supp}\,\varphi_p) \leq c_2 h^d\,. \tag{151}$$

(iii) Let there exist subsets $\Omega_{pt}^+ \subset \Omega_{pt}$ for all p, t such that the basis functions satisfy

$$\nabla\varphi_t \cdot \nabla\varphi_p \leq -\frac{\sigma}{h^2} < 0 \quad \textit{on } \Omega_{pt}^+ \quad \textit{and} \quad \nabla\varphi_t \cdot \nabla\varphi_p \leq 0 \quad \textit{on } \Omega_{pt} \setminus \Omega_{pt}^+ \tag{152}$$

and we have

$$\frac{meas(\Omega_{pt}^{+})}{meas(\Omega_{pt})} \geq c_3 > 0\,, \tag{153}$$

further, let the family of meshes be quasi-regular as in (107):

$$c_1 h^{\gamma} \leq meas(\operatorname{supp} \varphi_p) \leq c_2 h^{d}\,. \tag{154}$$

Then (146)-(147) holds.

Proof is obvious.

In view of well-known results (see e.g. [3, 10, 22, 33]), the conditions (146) and (148) have a nice geometric interpretations. Namely, in order to satisfy condition (146) in the case of a simplicial mesh, it is sufficient if the employed mesh is nonobtuse, further, condition (148) is satisfied if the employed family of simplicial meshes is uniformly acute [22, 5]. We note that these conditions are sufficient but not necessary: as shown by [18, Remark 6] , the DMP may still hold if some obtuse interior angles occur in the simplices of the meshes. This is analogous to the case of linear problems [21, 29]. In the case of bilinear elements, condition (146) is equivalent to the so-called condition of non-narrow mesh, see [8]. The case of prismatic finite elements is treated in [14].

The weaker conditions (150) and (152) allow in theory easier refinement procedures. First, (150) may allow the acute angles to deteriorate (i.e. tend to 90°) as $h \to 0$. Namely, if a family of simplicial meshes is regular then $|\nabla\varphi_t| = O(h^{-1})$ for all linear basis functions: hence, considering two basis functions φ_p, φ_t and letting α denote the angle of their gradients on a given simplex, the sufficient condition

$$\cos\alpha \leq -\sigma h^{\varepsilon} \tag{155}$$

(with some constant $\sigma > 0$ independent of h) implies

$$\nabla\varphi_t \cdot \nabla\varphi_p = |\nabla\varphi_t|\ |\nabla\varphi_p| \cos\alpha \leq -\frac{\sigma\, h^{\varepsilon}}{h^2}\,,$$

i.e. (150) holds. Clearly, if $h \to 0$ then (155) allows $\cos\alpha \to 0$, i.e. $\alpha \to 90°$. (In particular, for problem (46), when (108) coincides with $d \leq \gamma < d+2$ as in (71), then $\gamma - d$ can be chosen arbitrarily close to 2. Hence the exponent $2-\varepsilon$ in (150) can be arbitrarily close to 0, i.e. the decay of angles to 90° may be fast as $h \to 0$.)

Second, (152) means that one can allow some right angles, but each Ω_{pt}, which consists of a finite number of elements, must contain some elements with acute angles and the measure of these must not asymptotically vanish.

5 Some applications

5.1 Reaction-diffusion systems in chemistry

The steady states of certain reaction-diffusion processes in chemistry are described by systems of the following form:

$$\left.\begin{aligned} -b_k\Delta u_k + \ P_k(x,u_1,\dots,u_s) \ &= \ f_k(x) \quad \text{in } \Omega, \\ b_k \tfrac{\partial u_k}{\partial \nu} \ &= \gamma_k(x) \quad \text{on } \Gamma_N, \\ u_k \ &= \ g_k(x) \quad \text{on } \Gamma_D \end{aligned}\right\} \qquad (k=1,\dots,s). \tag{156}$$

Here, for all k, the quantity u_k describes the concentration of the kth species, and P_k is a polynomial which characterizes the rate of the reactions involving the k-th species. A common way to describe such reactions is the so-called mass action type kinetics [15, 16], which implies that P_k has no constant term for any k, in other words, $P_k(x,0) \equiv 0$ on Ω for all k. Further, the reaction between different species is often proportional to the product of their concentration, in which case

$$P_k(x,u_1,\dots,u_s) = a_{kk}(x)u_k^\alpha + \sum_{k\neq l} a_{kl}(x)u_k u_l\,.$$

The function $f_k \geq 0$ describes a source independent of concentrations.

We consider system (156) under the following conditions, such that it becomes a special case of system (100). As pointed out later, such chemical models describe processes with cross-catalysis and strong autoinhibiton.

Assumptions 5.1.

(i) $\Omega \subset \mathbf{R}^d$ is a bounded piecewise C^1 domain, where $d=2$ or 3, and Γ_D, Γ_N are disjoint open measurable subsets of $\partial\Omega$ such that $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$.

(ii) (Smoothness and growth.) For all $k,l=1,\dots,s$, the functions P_k are polynomials of arbitrary degree if $d=2$ and of degree at most 4 if $d=3$, further, $P_k(x,0)\equiv 0$ on Ω.

(iii) (Ellipticity.) $b_k > 0$ $(k=1,\dots,s)$ are given numbers.

(iv) (Cooperativity.) We have

$$\frac{\partial P_k}{\partial \xi_l}(x,\xi) \leq 0 \qquad (k,l=1,\dots,s,\ k\neq l;\ x\in\Omega,\ \xi\in\mathbf{R}^s). \tag{157}$$

(v) (Weak diagonal dominance for the Jacobians w.r.t. rows and columns.) We have

$$\sum_{l=1}^s \frac{\partial P_k}{\partial \xi_l}(x,\xi) \geq 0, \qquad \sum_{l=1}^s \frac{\partial P_l}{\partial \xi_k}(x,\xi) \geq 0 \qquad (k=1,\dots,s;\ x\in\Omega,\ \xi\in\mathbf{R}^s). \tag{158}$$

(vi) For all $k = 1, \ldots, s$ we have $f_k \in L^2(\Omega)$, $\gamma_k \in L^2(\Gamma_N)$, $g_k = g^*_{k|\Gamma_D}$ with $g^* \in H^1(\Omega)$.

Similarly to (94), assumptions (157)-(158) now imply

$$\frac{\partial P_k}{\partial \xi_k}(x, \xi) \geq 0 \qquad (k = 1, \ldots, s;\ x \in \Omega,\ \xi \in \mathbf{R}^s). \tag{159}$$

Returning to the model described by system (156), the chemical meaning of the cooperativity (157) is cross-catalysis, whereas (159) means autoinhibiton. Cross-catalysis arises e.g. in gradient systems [30]. Condition (158) means that autoinhibition is strong enough to ensure both weak diagonal dominances.

By definition, the concentrations u_k are nonnegative, therefore a proper numerical model must produce such numerical solutions. We can use Corollary 4.5 to obtain the required property:

Corollary 5.1 *Let problem (156) satisfy Assumptions 5.1, and let its FEM discretization satisfy the conditions of Theorem 4.3. If*

$$f_k \geq 0, \quad \gamma_k \geq 0, \quad g_k \geq 0 \qquad (k = 1, \ldots, s)$$

then for sufficiently small h, the FEM solution u^h of system (156) satisfies

$$u_k^h \geq 0 \quad \text{on } \Omega \qquad (k = 1, \ldots, s). \tag{160}$$

5.2 Linear elliptic systems

Maximum principles or nonnegativity preservation for linear elliptic systems have attracted great interest, as mentioned in the introduction. Hence it is worthwile to derive the corresponding DMPs from the previous results. Let us therefore consider linear elliptic systems of the form

$$\left.\begin{aligned} -\mathrm{div}\,(b_k(x)\,\nabla u_k) + \sum_{l=1}^{s} V_{kl}(x)\,u_l &= f_k(x) \quad \text{in } \Omega, \\ b_k(x)\tfrac{\partial u_k}{\partial \nu} &= \gamma_k(x) \quad \text{on } \Gamma_N, \\ u_k &= g_k(x) \quad \text{on } \Gamma_D \end{aligned}\right\} \qquad (k = 1, \ldots, s) \tag{161}$$

where for all $k, l = 1, \ldots, s$ we have $b_k \in W^{1,\infty}(\Omega)$ and $V_{kl} \in L^\infty(\Omega)$.

Let Assumptions 4.1 hold (where in fact we do not need assumption (ii)). Then (161) is a special case of (46), hence Corollary 4.2 holds, as well as the analogous results mentioned in Remark 4.2. Here we formulate two of these that follow the most studied CMP results:

Corollary 5.2 *Let problem (161) satisfy Assumptions 4.1, let its FEM discretization satisfy the conditions of Theorem 4.1 and let h be sufficiently* *small. If $u^h = (u_1^h, \ldots, u_s^h)$ is the FEM solution of system (161), then the following properties hold.*

(1) If $f_k \leq 0$, $\gamma_k \leq 0$ $(k = 1, \ldots, s)$ and $\max\limits_{k=1,\ldots,s} \max\limits_{\Gamma_D} g_k^h > 0$, then

$$\max_{k=1,\ldots,s} \max_{\overline{\Omega}} u_k^h = \max_{k=1,\ldots,s} \max_{\Gamma_D} g_k^h. \tag{162}$$

(2) If $f_k \geq 0$, $\gamma_k \geq 0$ and $g_k \geq 0$ $(k = 1, \ldots, s)$, then

$$u_k^h \geq 0 \quad \text{on } \Omega \qquad (k = 1, \ldots, s). \tag{163}$$

Acknowledgement. We hereby thank János Tóth for his valuable help to provide the chemical applications to reaction-diffusion models.

References

[1] Adams, R.A., *Sobolev spaces*, Academic Press, New York-London, 1975.

[2] Axelsson, O., *Iterative Solution Methods,* Cambridge University Press, 1994.

[3] Brandts, J., Korotov, S., Křížek, M., Dissection of the path-simplex in $\mathbf{R^n}$ into n path-subsimplices, *Linear Algebra Appl.* 421 (2007), no. 2-3, 382–393.

[4] Brandts, J., Korotov, S., Křížek, M., On the equivalence of regularity criteria for triangular and tetrahedral finite element partitions, *Comput. Math. Appl.* 55 (2008), 2227–2233.

[5] Brandts, J., Korotov, S., Křížek, M., Šolc, J., On nonobtuse simplicial partitions, *SIAM Rev.*, to appear.

[6] Burman, E., Ern, A., Nonlinear diffusion and the discrete maximum principle for stabilized Galerkin approximations of the convection-diffusion-reaction equation, *Comput. Methods Appl. Mech. Engrg.* 191 (2002), 3833–3855.

[7] Caristi, G., Mitidieri, E., Further results on maximum principles for noncooperative elliptic systems, *Nonlinear Anal.* 17 (1991), no. 6, 547–558.

[8] Christie, I., Hall, C., The maximum principle for bilinear elements, *Internat. J. Numer. Methods Engrg.* 20 (1984), 549–553.

[9] Ciarlet, P. G., Discrete maximum principle for finite-difference operators, *Aequationes Math.* 4 (1970), 338–352.

[10] Ciarlet, P. G., Raviart, P.-A., Maximum principle and uniform convergence for the finite element method, *Comput. Methods Appl. Mech. Engrg.* 2 (1973), 17–31.

[11] Draganescu, A., Dupont, T. F., Scott, L. R., Failure of the discrete maximum principle for an elliptic finite element problem, *Math. Comp.* 74 (2005), no. 249, 1–23.

[12] de Figueiredo, D. G., Mitidieri, E., Maximum principles for cooperative elliptic systems, *C. R. Acad. Sci. Paris Sér. I Math.* 310 (1990), no. 2, 49–52.

[13] Faragó, I., Karátson, J., *Numerical solution of nonlinear elliptic problems via preconditioning operators. Theory and applications.* Advances in Computation, Volume 11, NOVA Science Publishers, New York, 2002.

[14] Hannukainen, A., Korotov, S., Vejchodský, T., Discrete maximum principles for FE solutions of the diffusion-reaction problem on prismatic meshes, *J. Comput. Appl. Math.* (in press), 1–16. Preprint 105(2008), Institute of Mathematics of the Academy of Sciences of the Czech Republic.

[15] Hárs, V., Tóth, J., On the inverse problem of reaction kinetics, In: *Qualitative Theory of Differential Equations* (Szeged, Hungary, 1979), Coll. Math. Soc. János Bolyai 30, ed. M. Farkas, North-Holland - János Bolyai Mathematical Society, Budapest, 1981, pp. 363-379.

[16] Horn, F., Jackson, R., General mass action kinetics, *Arch. Rat. Mech. Anal.* 47 (1972), 81–116.

[17] Ishihara, K., Strong and weak discrete maximum principles for matrices associated with elliptic problems, *Linear Algebra Appl.* 88/89 (1987), 431–448.

[18] Karátson, J., Korotov, S., Discrete maximum principles for finite element solutions of nonlinear elliptic problems with mixed boundary conditions, *Numer. Math.* 99 (2005), 669–698.

[19] Karátson J., Korotov, S., Křížek, M., On discrete maximum principles for nonlinear elliptic problems, *Math. Comput. Simul.* 76 (2007) pp. 99–108.

[20] Korotov, S., Křížek, M., Acute type refinements of tetrahedral partitions of polyhedral domains, *SIAM J. Numer. Anal.* 39 (2001), 724–733.

[21] Korotov, S., Křížek, M., Neittaanmäki, P., Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle, *Math. Comp.* 70 (2001), 107–119.

[22] Křížek, M., Lin Qun, On diagonal dominance of stiffness matrices in 3D, *East-West J. Numer. Math.* 3 (1995), 59–69.

[23] Křížek, M., Neittaanmäki, P., *Mathematical and numerical modelling in electrical engineering: theory and applications*, Kluwer Academic Publishers, 1996.

[24] Ladyzhenskaya, O. A., Ural'tseva, N. N., *Linear and quasilinear elliptic equations*, Leon Ehrenpreis Academic Press, New York-London, 1968.

[25] Liska, R., Shashkov, M., Enforcing the discrete maximum principle for linear finite element solutions of second order elliptic problems, *Commun. Comput. Phys.* 3 (2008), no. 4, 852–877.

[26] López-Gómez, J., Molina-Meyer, M., The maximum principle for cooperative weakly coupled elliptic systems and some applications, *Diff. Int. Equations* 7 (1994), no. 2, 383–398.

[27] Mitidieri, E., Sweers, G., Weakly coupled elliptic systems and positivity, *Math. Nachr.* 173 (1995), 259–286.

[28] Protter, M. H., Weinberger, H. F., *Maximum principles in differential equations*, Springer-Verlag, New York, 1984.

[29] Ruas Santos, V., On the strong maximum principle for some piecewise linear finite element approximate problems of non-positive type, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* 29 (1982), 473–491.

[30] Tóth, J., Gradient systems are cross-catalytic, *Reaction Kinetics and Catalysis Letters* 12 (3) (1979), 253–257.

[31] Vejchodský, T., Solin, P., Discrete maximum principle for higher-order finite elements in 1D, *Math. Comp.* 76 (2007), no. 260, 1833–1846.

[32] Varga, R., *Matrix iterative analysis*, Prentice Hall, New Jersey, 1962.

[33] Xu, J., Zikatanov, L., A monotone finite element scheme for convection-diffusion equations, *Math. Comp.* 68 (1999), 1429–1446.

(continued from the back cover)

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The reports are available at *http://math.tkk.fi/reports/* .
The list of reports is continued inside the back cover.

A551 István Faragó, Janos Karatson, Sergey Korotov
Discrete maximum principles for the FEM solution of some nonlinear parabolic problems
August 2008

A550 István Faragó, Róbert Horváth, Sergey Korotov
Discrete maximum principles for FE solutions of nonstationary diffusion-reaction problems with mixed boundary conditions
August 2008

A549 Antti Hannukainen, Sergey Korotov, Tomás Vejchodský
On weakening conditions for discrete maximum principles for linear finite element schemes
August 2008

A548 Kalle Mikkola
Weakly coprime factorization, continuous-time systems, and strong-H^p and Nevanlinna fractions
August 2008

A547 Wolfgang Desch, Stig-Olof Londen
A generalization of an inequality by N. V. Krylov
June 2008

ISBN 978-951-22-9514-2 (print)
ISBN 978-951-22-9515-9 (PDF)
ISSN 0784-3143 (print)
ISSN 1797-5867 (PDF)