

# On modifications of continuous and discrete maximum principles for reaction-diffusion problems

István Faragó\*, Sergey Korotov†, Tamás Szabó‡

October 1, 2010

Department of Applied Analysis and Computational Mathematics  
Eötvös Loránd University  
H-1117, Budapest, Pázmány P. s. 1/c., Hungary  
e-mail: faragois@cs.elte.hu, smkorotov@gmail.com, szabot@cs.elte.hu

BCAM - Basque Center for Applied Mathematics  
Bizkaia Technology Park, Building 500, E – 48160 Derio  
Basque Country, Spain  
e-mail: korotov@bcamath.org

**Abstract:** In this work, we present and discuss some modifications, in the form of two-sided estimation (and also for arbitrary source functions instead of usual sign-conditions), of continuous and discrete maximum principles for the reaction-diffusion problems solved by the finite element and finite difference methods.

**Keywords:** reaction-diffusion problem, maximum principle, discrete maximum principle, monotone matrix, two-sided a priori estimation, finite element method, finite difference method, code validation

**Mathematical Subject Classification:** 35B50, 65N06, 65N30, 65N50

---

\*The first author was supported by Hungarian National Research Fund OTKA No. K67819.

†The second author was partially supported by Hungarian National Research Fund OTKA No. K67819.

‡The first and the third authors were supported by Jedlik project "ReCoMend" 2008–2011.

# 1 Motivation

Consider the following boundary-value problem of elliptic type: Find a function  $u \in C^2(\overline{\Omega})$  such that

$$-\Delta u + cu = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \quad (1)$$

where  $\Omega \subset \mathbf{R}^d$  is a bounded domain with Lipschitz continuous boundary  $\partial\Omega$  and the reactive coefficient  $c(x) \geq 0$  for all  $x \in \overline{\Omega}$ . We also assume that  $c$  and the right-hand side (be shortly called RHS, or source, in what follows) function  $f$  are both from  $C(\overline{\Omega})$ .

The classical solution of problem (1) is known to satisfy the so-called *maximum principle* (MP), which is often written as follows:

$$f(x) \geq 0 \quad \text{in } \overline{\Omega} \quad \implies \quad \max_{x \in \overline{\Omega}} u(x) \geq 0. \quad (2)$$

For more general forms of MPs (for example in case of nonhomogeneous Dirichlet or mixed boundary conditions) see e.g. [18, 13], however it is important to emphasize that MPs often considered for problems of elliptic type in the numerical community have a form of implication involving certain sign-type conditions only (like in (2)). The problem of construction (and proofs of validity) of suitable discrete analogues (the discrete maximum principles, or DMPs in short) of MPs for various types of numerical discretizations has also attracted a lot of attention by numerical scientists during last decades, see [4, 6, 7, 13, 16, 21, 24, 25, 26] and references therein.

However, from (2) we can only get an information on the sign of the unknown solution  $u$ , which can be often important to know (and provide on the discrete level in practical calculations especially if  $u$  models some physical quantity which is nonnegative by definition - absolute temperature, density, concentration, etc. [17]). At the same time there exist various a priori (upper and lower) estimates (see e.g. [18] and references therein) on the magnitude of solutions of certain elliptic problems in the PDE community, which can also be of interest both theoretically and practically. Also, many real-life problems have function  $f$  which may easily change its sign in the solution domain, and, in addition, most of MPs used (and imitated in numerics) e.g. for parabolic equations are formulated independently of signs of the source functions, see [8] and references therein. The goal of this work is to combine several available theoretical estimates in order to get a priori two-sided bounds for the (classical) solutions of elliptic problems (1) (with positive reactive terms) for arbitrary source functions and show how to provide the validity of their natural discrete analogues if some popular numerical technique (e.g. the finite element method (FEM) or the finite difference method (FDM)) is used for the discretization.

First, we shall present the following key result on continuous level.

**Theorem 1.** *Let functions  $c$  and  $f$  in (1) be from  $C(\overline{\Omega})$ , and let, additionally,*

$$c(x) \geq c_0 > 0 \quad \text{for all } x \in \overline{\Omega}. \quad (3)$$

*Then the following (a priori) two-sided estimates for the classical solution of problem (1) are valid:*

$$\min \left\{ 0, \min_{s \in \overline{\Omega}} \frac{f(s)}{c(s)} \right\} \leq u(x) \leq \max \left\{ 0, \max_{s \in \overline{\Omega}} \frac{f(s)}{c(s)} \right\} \quad \text{for any } x \in \overline{\Omega}. \quad (4)$$

*Proof.* To prove the upper estimate for  $u$  in above, one notices first that it is clearly valid if  $u \leq 0$  everywhere in  $\overline{\Omega}$ , i.e. when  $u$  attains its maximum on the boundary  $\partial\Omega$ . Further, if  $u$  attains its positive maximum at some interior point  $x_0 \in \Omega$ , then all the first order partial derivatives  $u'_{x_i}(x_0) = 0$ , and all the second order partial derivatives  $u''_{x_i x_i}(x_0) \leq 0$ , therefore from (1) and (3) we observe that  $u(x_0) \leq \frac{f(x_0)}{c(x_0)}$ , from which (4) follows immediately. The lower estimate in (4) can be proved similarly.  $\square$

*Remark 1.* The cases of nonhomogeneous boundary conditions (which are dropped in what follows for simplicity of presentation only) can be treated in the same manner (cf. [18, Chapt. 3]). E.g. in the case of nonvanishing Dirichlet boundary condition, the following estimates similar to (4) hold

$$\min \left\{ 0, \min_{s \in \partial\Omega} u(s), \min_{s \in \overline{\Omega}} \frac{f(s)}{c(s)} \right\} \leq u(x) \leq \max \left\{ 0, \max_{s \in \partial\Omega} u(s), \max_{s \in \overline{\Omega}} \frac{f(s)}{c(s)} \right\}, \quad x \in \overline{\Omega}.$$

It is also interesting to notice in this respect that the very first published paper “purely devoted” to DMPs, by R. Varga [24], is considering the case of arbitrary Dirichlet boundary conditions, but it does not analyse another important case of nonzero source functions.

*Remark 2.* From (4) one immediately derives the following implication:

$$f(x) \geq 0 \quad \text{in } \overline{\Omega} \quad \implies \quad 0 \leq u(x) \leq \max_{s \in \overline{\Omega}} \frac{f(s)}{c(s)}, \quad (5)$$

which forms a sharper (two-sided) estimation of the behaviour of  $u$  (provided (3) is valid) than the standard (of one-sided nature) MP (2) guaranteeing only the sign of  $u$ . Therefore, we shall call the estimates in (4) the *modified maximum principle* (or MMP in short) in what follows as (4) makes both *sharpening* and also *generalizing* of the standard maximum principle (2).

*Remark 3.* We mention that DMPs have been widely used for proving stability and finding the rate of convergence of FD approximations, see e.g. [2, 3, 6], and for proving the convergence of FE approximations in the maximum norm, see e.g. [2, 7].

## 2 On algebraic analogues of estimates (2) and (4)

After discretization of (1) by most of popular numerical techniques (e.g. by FEM or FDM) we arrive at the problem of solving  $n \times n$  system of linear algebraic equations

$$\mathbf{A}\mathbf{u} = \mathbf{F}, \quad (6)$$

where the vector of unknowns  $\mathbf{u} = [u_1, \dots, u_n]^T$  approximates the unknown solution  $u$  at certain selected points  $B_1, \dots, B_n$  of the solution domain  $\Omega$ , and the vector  $\mathbf{F} = [F_1, \dots, F_n]^T$  approximates (in the sense related to the nature of a concrete numerical method used, see Section 3 for more details on this) the values  $f(B_i)$ ,  $i = 1, \dots, n$ .

In what follows, the entries of matrix  $\mathbf{A}$  will be denoted by  $a_{ij}$ , and all matrix and vector inequalities appearing in the text are always understood component-wise.

**Definition 1.** The square  $n \times n$  matrix  $\mathbf{M}$  is called *monotone* if

$$\mathbf{M}\mathbf{z} \geq 0 \implies \mathbf{z} \geq 0. \quad (7)$$

The following theorem is well-known, see e.g. [3, p. 119] for its proof.

**Theorem 2.** *The square  $n \times n$  matrix  $\mathbf{M}$  is monotone if and only if  $\mathbf{M}$  is nonsingular and  $\mathbf{M}^{-1} \geq 0$ .*

Further, if one provides  $\mathbf{A}$  in (6) be monotone then  $\mathbf{A}^{-1} \geq 0$  and using assumption that  $\mathbf{F} \geq 0$  (usually trivially guaranteed by  $f \geq 0$  from MP (2), e.g. for linear FEM and FDM) we immediately get that  $\mathbf{u} = \mathbf{A}^{-1}\mathbf{F} \geq 0$ . These arguments describe in short a standard scheme for proving the following DMP

$$\mathbf{F} \geq 0 \implies \mathbf{u} \geq 0, \quad (8)$$

which naturally imitates the MP (2) (cf. [6, 7, 16, 13, 4]).

*Remark 4.* If we provide with more information on entries of  $\mathbf{A}^{-1}$  (besides  $\mathbf{A}^{-1} \geq 0$ ), we can estimate the vector  $\mathbf{u}$  (e.g. not only signs of its entries, but also their magnitudes, etc) more precisely. For example, in [23, p. 85] it is shown that when  $\mathbf{A}$  is irreducibly diagonally dominant with positive diagonal entries and nonpositive off-diagonal entries, or irreducible Stieltjes matrix, (which often happens after discretizing problem (1) by various numerical techniques), then we have even a stronger result  $\mathbf{A}^{-1} > 0$ , which can be useful to get a better estimation of behaviour of numerical approximations.

However, the property of irreducibility is not so easy to guarantee e.g. for FE approximations, see [10] for several examples on that. Even more sharp estimates for the entries of  $\mathbf{u}$  can be sometimes derived, if we use formulae for exact computing  $\mathbf{A}^{-1}$ , which are available e.g. for tridiagonal matrices appearing in numerical solution of certain one-dimensional problems, see [20].

**Definition 2.** The *infinity norm*  $\|\cdot\|_\infty$  of the square  $n \times n$  matrix  $\mathbf{M}$  (with entries  $m_{ij}$ ) is defined as

$$\|\mathbf{M}\|_\infty := \max_{i=1,\dots,n} \sum_{j=1}^n |m_{ij}|. \quad (9)$$

**Definition 3.** The square  $n \times n$  matrix  $\mathbf{M}$  (with entries  $m_{ij}$ ) is called *strictly row diagonally dominant* (or SDD in short) if the values

$$\alpha_i(\mathbf{M}) := |m_{ii}| - r_i > 0 \quad \text{for all } i = 1, \dots, n, \quad (10)$$

where  $r_i$  is the sum of absolute values of all off-diagonal entries in the  $i$ -th row of  $\mathbf{M}$ , i.e.  $r_i := \sum_{j=1, j \neq i}^n |m_{ij}|$ .

The following result, see [1, 22] for proofs, is useful for our purposes.

**Theorem 3.** *Let the square  $n \times n$  matrix  $\mathbf{M}$  be SDD. Then*

$$\|\mathbf{M}^{-1}\|_\infty \leq \frac{1}{\min_{i=1,\dots,n} \alpha_i(\mathbf{M})}. \quad (11)$$

*Remark 5.* It is worth to mention that diffusion-reaction problems with nonzero reaction terms often lead to SDD matrices in system (6) (see again Section 3), which gives a hope that we can often prove suitable discrete analogues of the estimates in (4) for such a type of problems.

**Theorem 4.** *Let matrix  $\mathbf{A}$  in system (6) be SDD and monotone. Then the following two-sided estimates for the entries of the solution  $\mathbf{u}$  are valid*

$$\min\left\{0, \min_{j=1,\dots,n} \frac{F_j}{\alpha_j(\mathbf{A})}\right\} \leq u_i \leq \max\left\{0, \max_{j=1,\dots,n} \frac{F_j}{\alpha_j(\mathbf{A})}\right\}, \quad i = 1, \dots, n. \quad (12)$$

*Proof.* First of all we notice that  $\alpha_i(\mathbf{A}) > 0$  for all  $i = 1, \dots, n$  as  $\mathbf{A}$  is SDD. Further, it is clear that the solution  $\mathbf{u}$  of system (6) is equivalently a solution of the following system:

$$\bar{\mathbf{A}}\mathbf{u} = \bar{\mathbf{F}}, \quad (13)$$

where  $\bar{\mathbf{A}} = \mathbf{D}\mathbf{A}$  and  $\bar{\mathbf{F}} = \mathbf{D}\mathbf{F}$ . Here  $\mathbf{D}$  is a diagonal matrix with strictly positive numbers  $1/\alpha_i(\mathbf{A})$ ,  $i = 1, \dots, n$  on its diagonal. Obviously,  $\bar{\mathbf{A}}$  is also SDD with  $\alpha_i(\bar{\mathbf{A}}) = 1$  for all  $i = 1, \dots, n$ . In addition,  $\bar{\mathbf{A}}$  is monotone as  $\mathbf{A}$  is monotone.

Let  $\mathbf{G} := \bar{\mathbf{A}}^{-1}$  have nonnegative entries denoted by  $g_{ij}$ . As  $\mathbf{u} = \bar{\mathbf{A}}^{-1}\bar{\mathbf{F}}$ , we observe that

$$\|\mathbf{G}\|_\infty \min\left\{0, \min_{j=1, \dots, n} \frac{F_j}{\alpha_j(\mathbf{A})}\right\} \leq u_i = \sum_{j=1}^n g_{ij} \frac{F_j}{\alpha_j(\mathbf{A})} \leq \|\mathbf{G}\|_\infty \max\left\{0, \max_{j=1, \dots, n} \frac{F_j}{\alpha_j(\mathbf{A})}\right\}.$$

Now, applying Theorem 3 to  $\bar{\mathbf{A}}$ , for which  $\alpha_i(\bar{\mathbf{A}}) = 1$ , we see that  $\|\mathbf{G}\|_\infty = \|\bar{\mathbf{A}}^{-1}\|_\infty \leq 1$ . From this and the above inequalities, we finally get the required estimates (12).  $\square$

*Remark 6.* It is clear that estimates (12) immediately imply DMP (8) provided  $\mathbf{F} \geq 0$ . Also, estimation (12) is considerably sharper than the obvious, but very rough, bounds  $|u_i| \leq \|\mathbf{A}^{-1}\|_\infty \|\mathbf{F}\|_\infty$ .

*Remark 7.* We notice that estimates close to (12) were obtained earlier by Windisch in [26] (however, in a more complicated way), but only for a more restrictive case of strictly row diagonally dominant  $M$ -matrices (cf. Remark 12). We also notice that we could easily get in the proof of the above theorem even sharper estimation, dropping zeros in (12) (cf. [26, Th. 1]). However, as we link our results to the continuous case, i.e. to (4), containing zeros, it is not actually necessary to do so in what follows.

*Remark 8.* In the recent work by Smelov [21], a very general case of DMP with an arbitrary SDD matrix  $\mathbf{A}$  has been considered and two-sided estimation similar to (12) has also been presented. However, adding a quite natural (and rather standard nowadays) requirement of monotonicity for matrices appearing in (6) (as we do in this work) leads to a more sharp estimation (12), which moreover imitates its continuous counterpart (4).

As (12) actually resembles the estimates (4), it is natural to give the following definition.

**Definition 4.** We say that the solution  $\mathbf{u}$  of system (6) with SDD matrix  $\mathbf{A}$  satisfies the *modified discrete maximum principle* (or MDMP, in short),

corresponding to MMP (4), if estimates (12) are valid and if, in addition,

$$\max_{j=1,\dots,n} \frac{F_j}{\alpha_j(\mathbf{A})} \leq \max\left\{0, \max_{s \in \bar{\Omega}} \frac{f(s)}{c(s)}\right\}, \quad \min_{j=1,\dots,n} \frac{F_j}{\alpha_j(\mathbf{A})} \geq \min\left\{0, \min_{s \in \bar{\Omega}} \frac{f(s)}{c(s)}\right\}. \quad (14)$$

*Remark 9.* The conditions (14) are really important in order to produce reliable (i.e. controllable) numerical approximations as, for example, linear FE and FD approximations do stay within the same (a priori known) limits then as the exact solutions they do approximate.

*Remark 10.* While the SDD-property of  $\mathbf{A}$  is almost automatically guaranteed after discretization by the nature of the reaction-diffusion problems (namely, due to the presence of nonvanishing reactive terms), its monotonicity, required in Theorem 4, should be provided a priori (or proved separately in each concrete case). One common approach for this in FEM is to impose certain a priori geometric requirements on the FE meshes employed so that all the off-diagonal entries  $a_{ij} \leq 0$  (see e.g. [4, 5, 7, 11, 13, 15, 16, 25] for more details on this subject). As far it concerns FDM, this property for the off-diagonal entries of  $\mathbf{A}$  is often guaranteed a priori by many standard FD schemes producing the so-called  $M$ -matrices [9].

*Remark 11.* One of advantages for dealing namely with the property  $a_{ij} \leq 0$  ( $i \neq j$ ) is an easy calculation (or estimation) of values  $\alpha_i(\mathbf{A})$  and their relation to  $c$  in general (see the next section for several examples in this respect).

### 3 Applications

In this section we demonstrate how theoretical results of the previous sections can be used for proving MDMPs for several popular numerical schemes (of FEM and FDM types), thus increasing the level of reliability of practical calculations by these techniques. For definitions of functional spaces employed in below see e.g. [18].

**FE approximations:** The standard FE scheme is based on the so-called variational formulation of (1), which reads: Find  $u \in H_0^1(\Omega)$  such that

$$a(u, v) = \mathcal{F}(v) \quad \forall v \in H_0^1(\Omega), \quad (15)$$

where

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx + \int_{\Omega} cuv dx \quad \text{and} \quad \mathcal{F}(v) = \int_{\Omega} f v dx. \quad (16)$$

The existence and uniqueness of the (weak) solution  $u$  is provided by the standard Lax-Milgram lemma. (Actually, for the well-posedness in above, one can only require that  $c \in L^\infty(\Omega)$  and  $f \in L^2(\Omega)$ , but we shall need more smoothness from these functions in what follows.)

Let  $\mathcal{T}_h$  be a FE mesh of  $\bar{\Omega}$  with interior nodes  $B_1, \dots, B_n$  lying in  $\Omega$  and boundary nodes  $B_{n+1}, \dots, B_{n+n^\partial}$  lying on  $\partial\Omega$ . Further, let the basis functions  $\phi_1, \phi_2, \dots, \phi_{n+n^\partial}$ , associated with these nodes, have the following properties (easily met if e.g. simplicial, block or prismatic FE meshes are used):

$$\phi_i(B_j) = \delta_{ij}, \quad i, j = 1, \dots, n+n^\partial, \quad \phi_i \geq 0 \text{ in } \bar{\Omega}, \quad i = 1, \dots, n+n^\partial, \quad \sum_{i=1}^{n+n^\partial} \phi_i \equiv 1 \text{ in } \bar{\Omega}, \quad (17)$$

where  $\delta_{ij}$  is the Kronecker delta. We also assume that the basis functions  $\phi_1, \phi_2, \dots, \phi_n$  vanish on the boundary  $\partial\Omega$ , thus spanning a finite-dimensional subspace  $V_h^0$  of  $H_0^1(\Omega)$ .

The FE approximation of  $u$  is defined as a function  $u_h \in V_h^0$  such that

$$a(u_h, v_h) = \mathcal{F}(v_h) \quad \forall v_h \in V_h^0, \quad (18)$$

whose existence and uniqueness are also provided by the Lax-Milgram lemma.

Algorithmically,  $u_h = \sum_{i=1}^n u_i \phi_i$ , where the coefficients  $u_i$  are the entries of the solution  $\mathbf{u}$  of system (6) with  $a_{ij} = a(\phi_i, \phi_j)$  and  $F_i = \mathcal{F}(\phi_i)$ . It is clear that, if (17) hold, the FE approximation  $u_h$  satisfies then the bounds from (12) at each point of  $\bar{\Omega}$  if all values  $u_i$  do satisfy them.

Further, all the diagonal entries  $a_{ii} = a(\phi_i, \phi_i) > 0$ ,  $i = 1, \dots, n$ . Assume that all the FE meshes used are such that  $a_{ij} \leq 0$  ( $i \neq j$ ), therefore  $\alpha_i(\mathbf{A}) = \sum_{j=1}^n a_{ij}$ . Moreover, for any  $i = 1, \dots, n$ , we calculate that

$$\begin{aligned} \alpha_i(\mathbf{A}) &= \sum_{j=1}^n a_{ij} = a(\phi_i, \sum_{j=1}^n \phi_j) = a(\phi_i, \sum_{j=1}^{n+n^\partial} \phi_j) - a(\phi_i, \sum_{j=n+1}^{n+n^\partial} \phi_j) = \\ &= a(\phi_i, 1) - a(\phi_i, \sum_{j=n+1}^{n+n^\partial} \phi_j) = \int_{\Omega} c\phi_i dx - \sum_{j=n+1}^{n+n^\partial} a_{ij} \geq \int_{\Omega} c\phi_i dx > 0, \end{aligned}$$

where the very last strict inequality holds due to condition (3). Therefore, the matrix  $\mathbf{A}$  is always SDD for our type of problems. Moreover  $\mathbf{A}$  is the Minkowski matrix, and therefore it is monotone (cf. [3, p. 119]). Hence, estimate (12) is valid.



The proofs of estimates (14) strongly depend much on how we compute  $a_{ij}$  and  $F_j$  in real FEM calculations. First, we consider the simplest case when  $c$  is constant and  $f$  is e.g. piecewise polynomial so that all  $a_{ij}$  and  $F_j$  are computed exactly while implemented. We see immediately that, if all  $F_i \leq 0$  then the first inequality in (14) holds. Let now some  $F_i > 0$  (i.e.  $\int_{\Omega} f \phi_i dx > 0$ ). Then

$$\frac{F_i}{\alpha_i(\mathbf{A})} = \frac{\int_{\Omega} f \phi_i dx}{\alpha_i(\mathbf{A})} \leq \frac{\int_{\Omega} f \phi_i dx}{\int_{\Omega} c \phi_i dx} \leq \frac{\int_{\Omega} \max\{0, \max_{s \in \Omega} f(s)\} \phi_i dx}{c \int_{\Omega} \phi_i dx} = \max\left\{0, \max_{s \in \Omega} \frac{f(s)}{c}\right\},$$

and, similarly, if all  $F_i \geq 0$  then the second inequality in (14) holds. Let now some  $F_i < 0$  (i.e.  $\int_{\Omega} f \phi_i dx < 0$ ). Then

$$\frac{F_i}{\alpha_i(\mathbf{A})} = \frac{\int_{\Omega} f \phi_i dx}{\alpha_i(\mathbf{A})} \geq \frac{\int_{\Omega} f \phi_i dx}{\int_{\Omega} c \phi_i dx} \geq \frac{\int_{\Omega} \min\{0, \min_{s \in \Omega} f(s)\} \phi_i dx}{c \int_{\Omega} \phi_i dx} = \min\left\{0, \min_{s \in \Omega} \frac{f(s)}{c}\right\},$$

i.e. the estimates from (14) hold true in this case.

If  $c$  is not constant and  $f$  is not necessarily piecewise polynomial, then for computations of integrals  $\int_{\Omega} c \phi_i \phi_j dx$  and  $\int_{\Omega} f \phi_j dx$  in practice, we should use certain quadrature rules, and, thus, each such case requires a separate analysis. Here, we only demonstrate how to prove the required estimates if the simplest quadrature rule

$$\int_S g(x) dx \approx \frac{\text{meas}_d S}{N_S} \sum_{i=1}^{N_S} g(\xi_i) \quad (19)$$

is used, where  $S$  is a finite element (e.g. simplex, block or prisms) from the mesh  $\mathcal{T}_h$  and  $\xi_1, \dots, \xi_{N_S}$  are its  $N_S$  vertices. Then we observe that

$$a_{ij} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j dx + \int_{\Omega} c \phi_i \phi_j dx \approx \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j dx$$

for  $i \neq j$ , and

$$a_{ii} \approx \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_i dx + \frac{c(B_i) \text{meas}_d(\text{supp } \phi_i)}{N_S} > 0,$$

where  $\text{supp } \phi_i$  denotes the support of the function  $\phi_i$ . The sign  $\approx$  means that the value on the left-hand side of it is replaced in actual calculations by the

value on the right-hand side of it, for computing of which we use quadrature (19) and (17).

Therefore, if FE meshes and basis functions are such that  $\nabla\phi_i \cdot \nabla\phi_j \leq 0$  (which is less stringent than the condition  $\nabla\phi_i \cdot \nabla\phi_j \leq -\varepsilon < 0$  actually required in the previous case (cf. [4, 7, 13])), we calculate that, if quadrature (19) is used for practical computing of entries of  $\mathbf{A}$ , then the actual values of numbers  $\alpha_i(\mathbf{A})$  are such that

$$\alpha_i(\mathbf{A}) \approx \frac{c(B_i) \text{meas}_d(\text{supp } \phi_i)}{N_S}.$$

Further, using (19) again, now for actual computing of the RHS of system (6), we see that

$$F_i \approx \frac{f(B_i) \text{meas}_d(\text{supp } \phi_i)}{N_S},$$

and it is now easy to show that the estimates in (14) do hold in this case, too.

**FD approximations:** If, for example, the standard  $(2^d + 1)$ -nodes stencils [9] with the same mesh size  $h$  (or different ones) in the available space directions are used, we get system (6) with the matrix which has strictly positive diagonal and nonpositive off-diagonal entries and which is always SDD matrix, i.e. monotone [3, p. 119]. As an illustration, such type of schemes for the one-dimensional case and uniform mesh leads to the tridiagonal matrix of the following form

$$\mathbf{A} = \text{tridiag} \left[ -\frac{1}{h^2}, \frac{2}{h^2} + c(B_i), -\frac{1}{h^2} \right]. \quad (20)$$

For such FD schemes (also in any dimensions), we always have  $\alpha_i(\mathbf{A}) \geq c(B_i)$  and  $F_i = f(B_i)$ , therefore, estimates (14) can be proved very easily again.

*Remark 12.* More complicated FEM and FDM schemes, e.g. those leading to some positive off-diagonal entries but still to monotone matrices (see [2, 3, 6, 14, 19] for some examples) can be analysed in the above manner. It is worth to mention here one interesting case, not covered by Windisch's results, but provable due to Theorem 4 and analysis of MDMP-validity as in the previous examples of this section. Imagine that in some part of the solution domain  $\Omega_0 \subset \Omega$  the RHS function  $f$  is zero. Then it is natural that all entries  $F_j$  associated with nodes  $B_j$  lying in  $\Omega_0$  are zeros in any meaningful (FEM or FDM) schemes. Therefore, in estimates (12) the corresponding fractions  $\frac{F_j}{\alpha_j(\mathbf{A})}$  are zeros a priori, independently of values  $\alpha_j(\mathbf{A})$ . It means

that appearance of certain positive off-diagonal entries (with indices associate to  $\Omega_0$ ) can be easily allowed without any effect on the desired two-sided estimations provided the resulting matrix  $\mathbf{A}$  remains SDD and monotone.

## 4 Conclusions

In this paper, we have analysed some modifications, in the form of two-sided estimation and for arbitrary source functions, of continuous and discrete maximum principles for the reaction-diffusion problems (with non-vanishing reactive terms) solved by the finite element and finite difference methods.

The main results and subjects for the future research work can be generally summarized as follows:

- From the standard monotonicity requirements on the parameters (mesh shape, mesh size, type of finite difference stencil used) of the FEM and FDM schemes we can get more precise information on the behaviour of numerical approximations than just usual DMP-property of the sign-type (8), and also for the case of an arbitrary source function.
- A priori guaranteeing that the approximations fit some easily computable two-sided bounds can be used, for example, for checking computer software for possible bugs.
- It would be interesting to see what is going around the case  $c \equiv 0$ , or for the case when  $c = 0$  in some parts of  $\Omega$ . However, there is no much sense to analyse the case  $c = 0$  for one-dimensional problems and FEM/FDM, as it is well-known (see e.g. [12]) that in this case most of popular finite element and finite difference schemes produce solutions which coincide with the exact solution at all the nodal points, thus fitting a priori any meaningful (upper and lower) estimates in our context.
- The approach proposed seems to be easy to apply for close elliptic problems with convective terms. Also the case of mixed boundary conditions can be analysed similarly.

## References

- [1] Ahlberg, J. H., Nilson, E. N., Convergence properties of the spline fit, *J. SIAM* 11 (1963), 95–104.
- [2] Axelsson, O., Kolotilina, L., Monotonicity and discretization error estimates, *SIAM J. Numer. Anal.* 27 (1990), 1591–1611.

- [3] Bramble, J. H., Hubbard, B. E., On a finite difference analogue of an elliptic boundary problem which is neither diagonally dominant nor of non-negative type, *J. Math. and Phys.* 43 (1964), 117–132.
- [4] Brandts, J., Korotov, S., Křížek, M., The discrete maximum principle for linear simplicial finite element approximations of a reaction-diffusion problem, *Linear Algebra Appl.* 429 (2008), 2344–2357.
- [5] Brandts, J., Korotov, S., Křížek, M., Šolc, J., On nonobtuse simplicial partitions, *SIAM Rev.* 51 (2009), 317–335.
- [6] Ciarlet, P.G., Discrete maximum principle for finite-difference operators, *Aequationes Math.* 4 (1970), 338–352.
- [7] 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.
- [8] Faragó, I., Horváth, R., Continuous and discrete parabolic operators and their qualitative properties, *IMA Numer. Anal.*, 29 (2009), 606–631.
- [9] Forsythe, G. E., Wasow, W. R., *Finite-Difference Methods for Partial Differential Equations*. Reprint of the 1960 original. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2004.
- [10] Hannukainen, A., Korotov, S., Vejchodský, T., *On weakening conditions for discrete maximum principles for linear finite element schemes*. In “Proc. of the Fourth Internat. Conf. on Numerical Analysis and Applications (NAA–2008), Rousse, Bulgaria” (ed. by S. Margenov et al.), LNCS 5434, Springer-Verlag, 2009, 297–304.
- [11] Hannukainen, A., Korotov, S., Vejchodský, T., Discrete maximum principle for FE solutions of the diffusion-reaction problem on prismatic meshes, *J. Comput. Appl. Math.* 226 (2009), 275–287.
- [12] Hlaváček, I., Křížek, M., On exact results in the finite element method, *Appl. Math.* 46 (2001), 467–478.
- [13] 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.
- [14] Korotov, S., Křížek, M., Neittaanmäki, P., Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle, *Math. Comp.* 70 (2001), 107–119.

- [15] Korotov, S., Vejchodský, T., *A comparison of simplicial and block finite elements*. In “Proc. of The Eighth Europ. Conf. on Numerical Mathematics and Advanced Applications (ENUMATH–2009), Uppsala, Sweden”, (ed. by G. Kreiss et al.), Springer-Verlag, 2010, 1–8 (in press).
- [16] Křížek, M., Qun Lin, On diagonal dominance of stiffness matrices in 3D, *East-West J. Numer. Math.*, 3 (1995), 59–69.
- [17] Křížek, M., Neittaanmäki, P., *Mathematical and Numerical Modelling in Electrical Engineering. Theory and Applications*, Kluwer Academic Publishers, Dordrecht, 1996.
- [18] Ladyzhenskaya, O.A., Ural'tseva, N.N., *Linear and Quasilinear Elliptic Equations*, Leon Ehrenpreis Academic Press, New York-London, 1968.
- [19] Lorenz, J., Zur Inversmonotonie diskreter Probleme, *Numer. Math.*, 27 (1976/77), 227–238.
- [20] Rózsa, P., *Linear Algebra and its Applications*, Muszaki Konyvkiado, 1976 (in Hungarian).
- [21] Smelov, V. V., Extension of the algebraic aspect of the discrete maximum principle, *Russian J. Numer. Anal. Math. Modelling* 22 (2007), 601–614.
- [22] Varah, J. M., A lower bound for the smallest singular value of a matrix, *Linear Algebra Appl.* 11 (1975), 3–5.
- [23] Varga, R., *Matrix Iterative Analysis*, Prentice Hall, 1962.
- [24] Varga, R., On discrete maximum principle, *J. SIAM Numer. Anal.* 3 (1966), 355–359.
- [25] Vejchodský, T., Šolín, P., Discrete maximum principle for higher-order finite elements in 1D, *Math. Comp.* 76 (2007), 1833–1846.
- [26] Windisch, G., A maximum principle for systems with diagonally dominant  $M$ -matrices. In: *Discretization in Differential Equations and Enclosures* (ed. E. Adams et al.), *Math. Res.*, 36, Akademie-Verlag, Berlin, 1987, 243–250.