An Overlapping Additive Schwarz-Richardson Method for Monotone Nonlinear Parabolic Problems

Marilena Munteanu and Luca F. Pavarino *

Department of Mathematics, University of Milan, Via C. Saldini 50, 20133 Milano, Italy. {munteanu,Luca.Pavarino}@mat.unimi.it

Summary. We construct a scalable overlapping Additive Schwarz-Richardson (ASR) algorithm for monotone nonlinear parabolic problems and we prove that the rate of convergence depends on the stable decomposition constant. Numerical experiments in the plane confirm the theoretical results.

1 Introduction

In the past decade, domain decomposition techniques have been increasingly employed to solve nonlinear problems. As a first approach, domain decomposition methods provide preconditioners for the Jacobian system in a Newton iteration. In this context, Schwarz-type preconditioners have been successfully used to solve problems from various applied fields, e.g. computational fluid dynamics [4, 7], full potential problems [3], cardiac electrical activity [9], unsteady nonlinear radiation diffusion [11]. Additive Schwarz-type methods have been used not only as inner iteration in a Newton-Krylov-Schwarz scheme, but also as outer iteration in nested solvers as ASPIN [5, 1] or nonlinear additive Schwarz [6]. We propose an iterative process based on the additive Schwarz algorithm applied to the nonlinear problem. The main idea of this paper can be traced back to [2], where a linear preconditioner for a nonlinear system arising from the discretization of a monotone elliptic problem is studied. Using the classical assumptions of the abstract theory of additive Schwarz methods (stable decomposition, strengthened Cauchy-Schwarz inequality and local stability, see [12]), we prove that the rate of convergence of the proposed algorithm depends on the stable decomposition constant C_0 and we construct a scalable Additive Schwarz-Richardson (ASR) method.

* This work was supported by Istituto Nazionale di Alta Matematica Francesco Severi, Roma and by grants of M.I.U.R (PRIN 2004014411).

2 Nonlinear Parabolic Problems

Let V be a Banach space and H a Hilbert space with a scalar product $(\cdot,\cdot)$ satisfying $V \subset H$ and V dense in H. Let V^* denote the dual space of V and $< \cdot,\cdot >$ the duality between V^* and V. The Riesz representation theorem and the density of V in H implies that for every $u \in H$ there exists an unique element in the dual space V^*, by convention still denoted by u, such that $(u,v) =< u,v > \ \forall\, v \in V$. Let Ω be a bounded domain of $\mathbb{R}^d, d = 2,3$ with the boundary $\partial\Omega$ polyhedral and Lipschitz continuous and $\mathcal{T}_h$ a triangulation of the domain Ω. In the following, we restrict to the case $V = \{v \in H^1(\Omega) : \gamma v = 0 \text{ on } \Gamma_1 \subset \partial\Omega,\ \mu(\Gamma_1) > 0\}$ and $H = L^2(\Omega)$. We consider the nonlinear form $b : H^1(\Omega) \times H^1(\Omega) \longrightarrow \mathbb{R}$ satisfying the following properties:

1. b is *Lipschitz continuous*:
   $\exists L > 0\ \forall v,\ w,\ z \in H^1(\Omega)\ |b(v,z) - b(w,z)| \leq L||v-w|| \cdot ||z||$
2. b is *bounded*:
   $\exists C > 0$ such that $|b(v,w)| \leq C(1 + ||v||)||w||,\ \forall v,\ w \in H^1(\Omega)$
3. b is *hemicontinuous*:
   $\forall u,\ v,\ w \in H^1(\Omega),\ f : [0,1] \longrightarrow \mathbb{R},\ f(\alpha) = b(u + \alpha v, w)$ is continuous
4. b is *strictly monotone*: $b(v, v-w) - b(w, v-w) \geq 0,\ \forall v, w \in H^1(\Omega)$
   and the equality holds only for $v = w$
5. $b\left(v, \sum\limits_{i=1}^{n} \alpha_i w_i\right) = \sum\limits_{i=1}^{n} \alpha_i b(v, w_i)\ \forall v, w_i \in H^1(\Omega),\ \forall \alpha_i \in \mathbb{R},\ i = 1,\dots,n$
6. $b(v,v) \geq c||v||^2_{H^1(\Omega)} - c_0||v||_1 - c_1||v||^2_{L^2(\Omega)} - c_2,\ \forall v \in H^1(\Omega)$,
   where $c > 0, c_0 > 0, c_1 \geq 0, c_2 \geq 0$ are constants.

We consider the following nonlinear parabolic problem: given $u_0 \in L^2(\Omega)$ and $f \in L^2((0,T);V^*)$ find $u \in W \equiv \{u \in L^2((0,T);V),\ u' \in L^2((0,T);V^*)\}$ such that

$$\begin{cases} < u'(t), w > + b(u(t), w) =< f(t), v >,\ \forall t \in (0,T) \setminus E_w,\ \forall w \in V \\ u(0) = u_0 \end{cases} \tag{1}$$

where $E_w \subset (0,T)$ is a set of measure zero that depends on the function w.
The continuous problem (1) is discretized in time by the backward Euler method and in space by the finite element method. Consequently, we obtain the fully discrete problem: given an arbitrary sequence $\{u_h^0\} \subset L^2(\Omega)$ of approximations of u^0 such that $\lim\limits_{h\to 0} ||u_h^0 - u^0|| = 0$, find $u_h^m \in V_h$ such that

$$\left(\frac{u_h^m - u_h^{m-1}}{\tau}, v\right) + b(u_h^m, v) =< f^m, v >,\ \forall v \in V_h \tag{2}$$

where $V_h = \{v|\ v = 0 \text{ on } \overline{\Gamma}_1, v \text{ is continuous on } \overline{\Omega},\ v|_T \text{ is linear } \forall T \in \mathcal{T}_h\}$ is the standard piecewise linear finite element space, $\tau = T/M$ and u_h^m is the value of the discrete function u_h at time $t^m = m\tau$.
Results on the existence and uniqueness of the solution of the discrete and continuous parabolic problems can be found e.g. in [13], Theorem 45.3 and Theorem 46.4, respectively. The convergence of the discrete solution to the continuous one is presented in [13], Theorem 46.4 and 47.1.

3 An Additive Schwarz-Richardson Algorithm

Given a finite element basis $\{\phi_j,\ j=1,\dots,n\}$ of V_h, for simplicity, we will drop the indexes h and m and still denote by u both the finite element approximation $u = \sum_{j=1}^{n} u_j\phi_j$ of the continuous solution and its vector representation $u=(u_1,\dots,u_n)^T$.
Problem (2) is equivalent to the nonlinear algebraic system

$$B(u) = \hat{g}, \tag{3}$$

where $B(u) = (b_1,\dots,b_n)^T,\ b_j = (u,\phi_j) + \tau b(u,\phi_j),\ \hat{g} = (g_1,\dots,g_n)^T,\ g_j = \tau \cdot < f^m, \phi_j > + (u_h^{m-1}, \phi_j)$. We consider a family of subspaces $V_i \subset V_h,\ i=0,\dots,N$ and the interpolation operators $R_i^T : V_i \longrightarrow V_h$.
We assume that V_h admits the following decomposition:

$$V_h = \sum_{i=0}^{N} R_i^T V_i.$$

In addition to the previous properties (1-6), we also assume the following property (verified in most reaction-diffusion problems in applications):

7. b can be written as a sum $b(u,v) = a(u,v) + \tilde{b}(u,v)$ of a bilinear, continuous and coercive form $a : V \times V \longrightarrow \mathbb{R}$ and a nonlinear form $\tilde{b}$ (that is monotone and Lipschitz continuous with constant $\tilde{L}$ due to 1. and 4.).

The bilinear form $a_\tau(u,v) = (u,v) + \tau a(u,v)$ defines a scalar product on V. We introduce the local symmetric, positive definite bilinear forms $\tilde{a}_{\tau,\ i} : V_i \times V_i \longrightarrow \mathbb{R}$ and, as in the abstract Schwarz theory [12], we make the following assumptions:

- *Stable Decomposition.* There exist a constant C_0, such that every $u \in V_h$ admits a decomposition $u = \sum_{i=0}^{N} R_i^T u_i,\ u_i \in V_i,\ i=0,\dots,N$ that satisfies
$$\sum_{i=0}^{N} \tilde{a}_{\tau,i}(u_i,u_i) \le C_0^2 a_\tau(u,u);$$
- *Strengthened Cauchy-Schwarz inequality.* $\exists\ \epsilon_{ij} \in [0,1]\ i,j=1,\dots,N$, s.t.
$$|a_\tau(R_i^T u_i, R_j^T u_j)| \le \epsilon_{ij} a_\tau(R_i^T u_i, R_i^T u_i)^{1/2} a_\tau(R_j^T u_j, R_j^T u_j)^{1/2},\ \forall u_i \in V_i,\ u_j \in V_j;$$
- *Local Stability.* There is $\omega > 0$, such that
$$a_\tau(R_i^T u_i, R_i^T u_i) \le \omega \tilde{a}_{\tau,i}(u_i,u_i),\ \forall u_i \in V_i,\ 0 \le i \le N.$$

We define the "projection"-like operators $\tilde{Q}_i : V_h \longrightarrow V_i$ by $\tilde{a}_{\tau,i}(\tilde{Q}_i(u), v_i) = (u, R_i^T v_i) + \tau b(u, R_i^T v_i),\ \forall v_i \in V_i, u \in V_h$, their extensions $Q_i : V_h \longrightarrow R_i^T V_i \subset V_h$ by $Q_i(u) = R_i^T \tilde{Q}_i(u)$ and $Q(u) = \sum_{i=0}^{N} Q_i(u)$.
Let $\tilde{A}_{\tau,i} \equiv (\tilde{a}_{\tau,i}(\phi_j,\phi_l))_{j,l}$ be the matrix representation of the local bilinear form $\tilde{a}_{\tau,i}$. The matrix form of $Q(u)$ is

$$Q(u) = \mathcal{M}^{-1} B(u), \tag{4}$$

where $\mathcal{M} = \left(\sum_{i=0}^{N} R_i^T \tilde{A}_{\tau,i}^{-1} R_i \right)^{-1}$. The matrix $\mathcal{M}$ is symmetric and positive definite and consequently it defines a norm, $||u||_{\mathcal{M}}^2 = u^T \mathcal{M} u$. Denoting $\breve{g} = \mathcal{M}^{-1}\hat{g}$, and using the matrix form of the nonlinear operator Q, it is straightforward to prove that the nonlinear system (3) is equivalent to the system

$$Q(u) = \breve{g}. \tag{5}$$

Additive Schwarz-Richardson (ASR) algorithm: for a fixed time t and for a properly chosen parameter λ, iterate for $k = 0, 1, \ldots$ until convergence

$$u^{k+1} = u^k + \lambda s^k, \tag{6}$$

where $s^k = -\mathcal{M}^{-1}\left(B(u^k) - \hat{g}\right) \Longleftrightarrow s^k = -\left(Q(u^k) - \breve{g}\right)$.
The operator Q satisfies the following Lemmas (for complete proofs see [8]).

Lemma 1. *There exists a positive constant $\delta_0 = \frac{1}{C_0^2}$ such that*

$$(Q(u+z) - Q(u), z)_{\mathcal{M}} \geq \delta_0 ||z||_{\mathcal{M}}^2 \ \forall u, v \in V_h.$$

Lemma 2. *There exists a positive constant $\delta_1 = C\sqrt{\omega^3(1+\rho(\epsilon))^3(1+\tilde{L})^2 C_0^2}$, where C is a positive constant independent of mesh size or time-step, such that*

$$||Q(u+z) - Q(u)||_{\mathcal{M}} \leq \delta_1 ||z||_{\mathcal{M}} \ \forall u, v \in V_h.$$

Using this lemmas, we can prove the following convergence result.

Theorem 1. *If we choose $0 < \lambda < 2\delta_0/\delta_1^2$ then* **ASR** *converges in the $\mathcal{M}$ norm to the solution u^* of* (5), *i.e.*

$$||u^k - u^*||_{\mathcal{M}}^2 \leq P(\lambda)^k ||u^0 - u^*||_{\mathcal{M}}^2,$$

where $P(\lambda) = 1 - 2\lambda\delta_0 + \lambda^2\delta_1^2$.

Proof. We define the error $e^k = u^k - u^*$ and the residual $r^k = Q(u^k) - Q(u^*)$. The error of the $k+1$ step of the ASR-iteration can be expressed in terms of the error and residual at the k step:

$$e^{k+1} = u^{k+1} - u^* = u^k - \lambda r^k - u^* = e^k - \lambda r^k.$$

Using the linearity of $(\cdot,\cdot)_{\mathcal{M}}$:

$$\begin{aligned} ||e^{k+1}||_{\mathcal{M}}^2 &= (e^{k+1}, e^{k+1})_{\mathcal{M}} = (e^k - \lambda r^k, e^k - \lambda r^k)_{\mathcal{M}} \\ &= ||e^k||_{\mathcal{M}}^2 - 2\lambda(e^k, r^k)_{\mathcal{M}} + \lambda^2 ||r^k||_{\mathcal{M}}^2. \end{aligned}$$

Lemma 1 implies:

$$\begin{aligned} -(e^k, r^k)_{\mathcal{M}} &= -(u^k - u^*, Q(u^k) - Q(u^*))_{\mathcal{M}} \\ &= -(u^k - u^*, Q(u^k - u^* + u^*) - Q(u^*))_{\mathcal{M}} \\ &\leq -\delta_0 ||e^k||_{\mathcal{M}}^2. \end{aligned}$$

From Lemma 2 we have

$$||r_k||_{\mathcal{M}}^2 = ||Q(u^k) - Q(u^*)||_{\mathcal{M}}^2 = ||Q(u^k - u^* + u^*) - Q(u^*)||_{\mathcal{M}}^2 \le \delta_1^2 ||e^k||_{\mathcal{M}}^2,$$

hence

$$||e^{k+1}||_{\mathcal{M}}^2 \le (1 - 2\lambda\delta_0 + \lambda^2\delta_1^2)||e_k||_{\mathcal{M}}^2.$$

We define $P(\lambda) = 1 - 2\lambda\delta_0 + \lambda^2\delta_1^2$. If we choose $0 < \lambda < \frac{2\delta_0}{\delta_1^2}$ then $P(\lambda) < 1$ and the convergence holds. We remark that $P(\lambda)$ has its minimum in $\lambda_{min} = \frac{\delta_0}{\delta_1^2}$ and $P(\lambda_{min}) = 1 - \frac{\delta_0^2}{\delta_1^2} < 1$.

Remark 1. If we drop the coarse space V_0 and we define $Q(u) = \sum_{i=1}^{N} Q_i(u)$, the ASR-algorithm is convergent. In this case, it is possible to prove that $\delta_0 = \frac{1}{C_0^2}$ and $\delta_1 = C_0\rho(\epsilon)\omega(1 + \tilde{L})$.

Remark 2. The algorithm depends on the choice of the parameter λ. Numerical tests have shown that the step-length selection described in [10] performs well.

4 Numerical Results

We consider the variational nonlinear parabolic problem: given $u(t_0, x) = u_0(x)$ and $T > t_0$, for all $t \le T$ find $u(t) \in H_0^1(\Omega)$ such that

$$\left(\frac{\partial u(t)}{\partial t}, v\right) + a(u(t), v) + (f(u(t)), v) = (g, v) \quad \forall v \in H_0^1(\Omega),$$

where

$$a(u, v) = \int_\Omega \sum_{i,j} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j}\, dx,$$

with $a_{ij} \in C^1(\Omega)$ such that $a_{ij}(x) = a_{ji}(x), \quad \forall x \in \Omega,\ \forall i, j$ and f monotone.

Table 1. Scalability of 1-level and 2-level ASR method for fixed overlap size $\delta = h$, subdomain size $H/h = 4$ and increasing number of subdomains (and nodes).

$\lambda = 0.4$

N	1-level		2-level	
	iter	err	iter	err
2×2	40	6.75e-3	44	6.75e-3
4×4	70	1.67e-3	37	1.67e-3
6×6	123	7.44e-4	37	7.44e-4
8×8	197	4.18e-4	38	4.18e-4
10×10	293	2.67e-4	38	2.67e-4
12×12	410	1.85e-4	39	1.85e-4
14×14	-	-	39	1.36e-4

λ : step-length selection

N	1-level		2-level	
	iter	err	iter	err
2×2	25	6.75e-3	24	6.75e-3
4×4	37	1.67e-3	25	1.67e-3
6×6	69	7.44e-4	24	7.44e-4
8×8	117	4.18e-4	24	4.18e-4
10×10	157	2.67e-4	27	2.67e-4
12×12	223	1.85e-4	22	1.85e-4
14×14	-	-	25	1.36e-4

Table 2. Iteration counts and relative errors for fixed overlap size $\delta = h$, mesh size $h = 1/48$ and increasing number of subdomains.

$\lambda = 0.4$

N	1-level iter	1-level err	2-level iter	2-level err
2×2	155	1.74e-4	70	1.74e-4
4×4	192	1.74e-4	58	1.74e-4
6×6	245	1.74e-4	47	1.74e-4
8×8	301	1.74e-4	41	1.74e-4

λ : step-length selection

N	1-level iter	1-level err	2-level iter	2-level err
2×2	86	1.74e-4	43	1.74e-4
4×4	114	1.74e-4	32	1.74e-4
6×6	149	1.74e-4	27	1.74e-4
8×8	-	-	23	1.74e-4

The numerical tests were performed for the bilinear form $a(u,v) = (\nabla u, \nabla v)$ and the nonlinear function $f(u) = 0.5u + u^3$. The domain is the unit square $\Omega = (0,1) \times (0,1)$ and g is chosen so that $u^*(t,x) = t\sin(\pi x)\sin(\pi y)$ is the exact solution. We consider $t_0 = 0,\ u_0(x) = 0$ and we compute the solution for $t = \tau = 0.01$. The iterations process is stopped when $||r_k||_{\mathcal{M}}/||r_0||_{\mathcal{M}} \leq 1e-8$ and we denote the relative error by $err = ||u - u^*||_{l^2(\Omega)}/||u^*||_{l^2(\Omega)}$.
Our additive Schwarz preconditioner is build as in the linear case. We partition the domain Ω into shape regular nonoverlapping subdomains $\{\Omega_i,\ 1 \leq i \leq N\}$ of diameter H defining a shape-regular coarse mesh $\mathcal{T}_H$. Each subregion Ω_i is extended to a larger one, Ω_i' such that the fine mesh $\mathcal{T}_h$ gives rise to N local meshes $\mathcal{T}_{h,i}$, and the partition $\{\Omega_i'\}$ satisfies the finite covering assumption [12]. Using the above decomposition, a 1-level method is defined by the local spaces $V_i = \{v \in H_0^1(\Omega_i') |\ v|_T \text{ is linear, } \forall T \in T_{h,i}\},\ 1 \leq i \leq N$, and the local bilinear forms $\tilde{a}_{\tau,i}(u_i, v_i) = a_\tau(R_i^T u_i, R_i^T v_i),\ \forall u_i,\ v_i \in V_i$, with zero extension interpolation operators $R_i^T : V_i \longrightarrow V,\ 1 \leq i \leq N$. We then build a 2-level algorithm by defining the coarse finite element space $V_0 = \{v \in H_0^1(\Omega) |\ v \text{ is continuous and } v|_T \text{ is linear, } \forall T \in T_H\}$ and the operator R_0^T which interpolates the coarse functions onto the fine mesh. It can be proved that the stable decomposition constant is

$$C_0^2 = C\max\{1 + \frac{H}{\delta},\ 1 + \frac{\tau}{H\delta}\}\ \text{(1-level)}, \qquad C_0^2 = C(1 + \frac{H}{\delta}),\ \text{(2-level)}, \quad (7)$$

where δ measures the width of region $\Omega_i' \backslash \Omega_i$, i.e. the overlap size.

Table 3. Iteration counts and relative errors for fixed mesh size $h = 1/48$, number of subdomain $N = 2 \times 2$, $\lambda = 0.4$ and increasing the overlap size δ

overlap	1-level iter	1-level err	2-level iter	2-level err
h	155	1.74e-4	70	1.74e-4
2h	82	1.74e-4	46	1.74e-4
3h	59	1.74e-4	37	1.74e-4
4h	49	1.74e-4	37	1.74e-4

Table 1 reports the iteration counts and relative errors of our ASR method with fixed overlap $\delta = h$, increasing the number of nodes and subdomains so that $H/h = 4$ is kept fixed (scaled speedup). The parameter λ is fixed at 0.4 (left table) or chosen by

Table 4. Same as Table 1 but with random right-hand side; $\lambda = 0.4$.

N	1-level iter	2-level iter
2×2	40	42
4×4	70	38
6×6	122	38
8×8	196	38
10×10	291	41
12×12	407	42

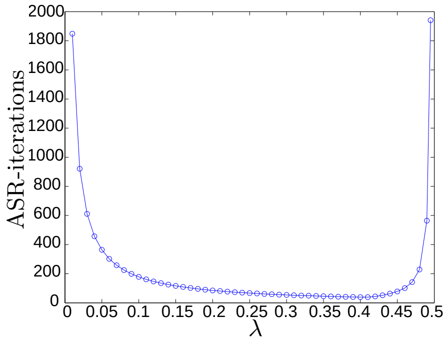

Fig. 1. ASR iterations counts as a function of the parameter λ

the step-length strategy of [10] (right table). According to the theory, in the 1-level case the number of iterations increases, because H decreases to zero while δ and τ are kept constant in (7). On the other hand, the iteration counts of the 2-level method remain bounded, because H/δ is kept fixed in (7). The same quantities are reported in Table 2, keeping now $h = 1/48$ fixed and increasing the number of subdomains (standard speedup). Only in the 2-level case the iteration counts improve as the subdomain size decreases. Table 3 shows that the iteration counts improve with increasing overlap size, as in the linear case and Table 4 is the same as Table 1 with $\lambda = 0.4$ but with random right-hand size. Finally, Fig. 1 confirms the theoretical prediction of Theorem 1, showing the ASR iteration counts as a function of the parameter λ for $h = 1/16, N = 2 \times 2, \delta = h$: the ASR convergence rate attains a minimum inside an interval $(0, \alpha)$, $\alpha > 0$ and degenerates at the interval endpoints.

Acknowledgement. Thanks are due to Maksymilian Dryja for very helpful comments and suggestions.

References

[1] H-B. An. On convergence of the Additive Schwarz Preconditioned Inexact Newton Method. *SIAM J. Numer. Anal.*, 43(5):1850–1871, 2005.

[2] X-C. Cai and M. Dryja. Domain decomposition methods for monotone nonlinear elliptic problems. In J. Xu D.E. Keyes, editor, *Seventh International Conference*

on Domain Decomposition Methods in Scientific and Engineering Computing, pages 21–27, 1994. Contemp. Math. 180, Amer. Math. Soc., Providence, RI.
[3] X-C. Cai, W.D. Gropp, D.E. Keyes, R.G. Melvin, and D.P. Young. Parallel Newton-Krylov-Schwarz algorithms for the transonic full potential equation. *SIAM. J. Sci. Comput.*, 19(1):246–265, 1998.
[4] X-C. Cai, W.D. Gropp, D.E. Keyes, and M.D. Tidriri. Newton-Krylov-Schwarz methods in CFD. In *Proceedings of the International Workshop on Numerical Methods for the Navier-Stokes Equations*, pages 183–200, 1995. Vieweg, Braunschweig.
[5] X-C. Cai and D.E. Keyes. Nonlinearly preconditioned inexact Newton algorithms. *SIAM. J. Sci. Comput.*, 24(1):183–200, 2002.
[6] M. Dryja and W. Hackbush. On the nonlinear domain decomposition method. *BIT*, 37(2):296–311, 1997.
[7] F-N. Hwang and X-C. Cai. A parallel nonlinear additive Schwarz preconditioned inexact Newton algorithm for incompressible Navier-Stokes equations. *J. Comput. Phys.*, 204(2):666–691, 2005.
[8] M. Munteanu. *Overlapping Additive Schwarz Methods for Nonlinear Parabolic Rection-Diffusion Problems*. PhD thesis, Univ. of Milan. Dept. of Math., 2007.
[9] M. Murillo and X-C. Cai. A fully implicit parallel algorithm for simulating the non-linear electrical activity of the heart. *Numer. Linear Algebra Appl.*, 11(2-3):261–277, 2004.
[10] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer-Verlag, New-York, 1999. Springer Series in Operations Research.
[11] S. Ovtchinnikov and X-C. Cai. One-level Newton-Krylov-Schwarz algorithm for unsteady non-linear radiation diffusion problem. *Numer.Linear Algebra Appl.*, 11(10):867–881, 2004.
[12] A. Toselli and O. Widlund. *Domain Decomposition Methods-Algorithms and Theory*. Springer-Verlag, Berlin, 2005. Springer Series in Computationl Mathematics, 34.
[13] A. Ženišek. *Nonlinear Elliptic and Evolution Problems and Their Finite Element Approximtions*. Academic Press, Inc., London, 1990. Computational Mathematics and Applications.