Journal of Computational Mathematics
Vol.32, No.2, 2014, 183–194.
http://www.global-sci.org/jcm
doi:10.4208/jcm.1310-FE1

TIME-EXTRAPOLATION ALGORITHM (TEA) FOR LINEAR PARABOLIC PROBLEMS *

Hongling Hu Chuanmiao Chen
College of Mathematics and Computer Science, Key Laboratory of High Performance Computing and Stochastic Information Processing (HPCSIP) (Ministry of Education of China), Hunan Normal University, Changsha 410081, China
Email: hhling625@163.com cmchen@hunnu.edu.cn
Kejia Pan
School of Mathematics and Statistics, Central South University, Changsha 410083, China
Email: pankejia@hotmail.com

Abstract

The fast solutions of Crank-Nicolson scheme on quasi-uniform mesh for parabolic problems are discussed. First, to decrease regularity requirements of solutions, some new error estimates are proved. Second, we analyze the two characteristics of parabolic discrete scheme, and find that the efficiency of Multigrid Method (MG) is greatly reduced. Numerical experiments compare the efficiency of Direct Conjugate Gradient Method (DCG) and Extrapolation Cascadic Multigrid Method (EXCMG). Last, we propose a Time-Extrapolation Algorithm (TEA), which takes a linear combination of previous several level solutions as good initial values to accelerate the rate of convergence. Some typical extrapolation formulas are compared numerically. And we find that under certain accuracy requirement, the CG iteration count for the 3-order and 7-level extrapolation formula is about 1/3 of that of DCG's. Since the TEA algorithm is independent of the space dimension, it is still valid for quasi-uniform meshes. As only the finest grid is needed, the proposed method is regarded very effective for nonlinear parabolic problems.

Mathematics subject classification: 65B05, 65M06, 65M12, 65N22.
Key words: Parabolic problem, Crank-Nicolson scheme, Error estimates, Time-extrapolation algorithm, CG-iteration.

1. Introduction

In modern science and technique, the high-dimensional parabolic problems, such as high heat transmission, superconductor, semi-conductor, nuclear-fusion and so on, are more and more important. However their computation is still very difficult.

As we all know, for the linear systems derived from elliptic problems, solving by direct methods is very difficult when the number of unknowns is more than tens of thousands. Therefore various iterative methods have emerged, such as

- Conjugate Gradient Method (CG) is efficient for solving symmetric positive definite systems, but the efficiency of CG reduces significantly when the condition number of coefficient matrix is greater than 10^3. The use of precondition techniques may be appropriate to improve the convergence rate, but the computational complexity increases.

* Received December 28, 2012 / Revised version received July 15, 2013 / Accepted October 9, 2013 / Published online March 31, 2014 /

- Multigrid Method (MG) was presented by Fedorenko (1961) and Brandt (1977) [1]. For the linear systems derived from elliptic problems, the computational work $W = O(N)$ of MG is proportional to the number of unknowns N. So MG, widely used in scientific and engineering computing, has become one of the most effective algorithms to solve large scale problem. It should be noted that grids generated by MG method just satisfy the requirements for superconvergence and extrapolation.

- Cascadic Multigrid Method (CMG) presented by Bornemann [2], Deuflhard [8] and Shaidurov [17] is an one-way multigrid method which may be viewed as a multilevel method without the coarse mesh correction. Since 1998, Shi et al. made a lot of theoretical analysis [19, 20]. Because of its high efficiency, CMG has been quickly applied to a series of problems [10, 15, 16, 19, 24].

In recent years, we have proposed an extrapolation cascadic multi-grid method (EXCMG) [3, 6] and a new extrapolation formula, which use a linear combination of the solutions on previous 2 level coarse meshes to provide a good initial value of finite element solution on next-level fine mesh. This method is of high accuracy and converges for both the function and its derivative. The numerical experiments show that the work of EXCMG is close to that of MG algorithm for simple linear problems. However, EXCMG requires less iterations on the finest mesh and converges more quickly, see [7] for details.

Of course, the above methods can be used to solve parabolic problems, direct methods or indirect methods (i.e. the combination and iteration on multi-levels in time). Brandt studied early the indirect method, Hackbush [12] also suggested a time parallel MG algorithm. Later, Horton [13](1992), Horton and Vandewalle [14](1995), Gander and Vandewalle [11](2007) presented further developments. Shi and Xu [19](2000), Du and Ming [10](2008) directly used the CMG to solve the resulting elliptic problems with a discrete in time formulation for parabolic problems. In a word, all these methods are effective.

However the linear systems derived from parabolic problems have the own characteristics. For example, the condition number, $Cond(A) \approx 4r, \quad r = a^2k/(2h^2)$, is much smaller than the condition number of corresponding elliptic problems ($O(h^{-2})$). And the solution U^{j-1} obtained by previous level provides a good initial value of the solution U^j of the present level ($U^j = U^{j-1} + O(k)$). Numerical experiments show that for elliptic problems MG and EXCMG have the absolute advantage, but for parabolic problems things change. Especially when the condition number is not large(such as $r < 1000$), the efficiency of the two algorithms is greatly reduced, and they have lost the absolute advantage even though still better than the direct CG-iteration. Therefore we should develop other efficient algorithms for parabolic problems.

C.C. Douglas [9](1996) predicted that "some excellent time-extrapolation methods exist which can be coupled to conjugate gradient- like methods. Once the first few time steps are solved (slowly), the extrapolation method provides such a good initial value to the solution on the next time step that only a few (say, one or two) iterations of the conjugate gradient-like method are needed before moving on. In this situation multigrid cannot compete." However the numerical experiments and our research show that this problem is quite complicated, far from being so simple as he said.

In this paper we should derive some new estimates of Crank-Nicolson scheme at first , then we will discuss the following algorithms used to solve parabolic problems:

- Direct CG-iteration (DCG) is acceptable when $r < 700$.

- For the model problem EXCMG needs about 16 iterations of CG (convert into the finest grid), independent of r.
- For TEA, under the same control accuracy, CG iteration count of the extrapolation formula 3D7L of 3-order and 7-level is about 1/3 of DCG's. It seems that TEA will be more favorable for nonlinear parabolic problems.

The rest of our paper is organized as follows: In Section 2, we derive some new estimates of Crank-Nicolson scheme. In Section 3, we compare DCG and EXCMG. A new Time-Extrapolation Algorithm(TEA) should be proposed in Section 4. In the last section, the numerical experiments are reported to support our theory.

In the following, the symbol C is used for a positive constant which may vary with the context but is independent of the mesh size h.

2. New Estimates of Crank-Nicolson Scheme

In a cylindric domain $Q = (0, T) \times \Omega$ we discuss a linear parabolic problem

$$u_t = a^2 \Delta u + f(t, x), \quad (t, x) \in Q, \quad u(0, x) = \psi(x), \quad u(t, x) = 0 \ on \ \partial\Omega, \tag{2.1}$$

and its weak formulation

$$\int_0^t [(u_t, v) + A(u, v) - (f, v)]dt = 0, \ v \in H_0^1(\Omega), \tag{2.2}$$

where the bilinear form

$$A(u, v) = \int_\Omega a^2 DuDvdx,$$

is assumed to be bounded and H_0^1-coercive.

Take the time nodes $t_j = jk$ and elements $I_j = (t_j, t_{j+1}), \quad j = 0, 1, 2, ..., N, \quad T = Nk$, The domain Ω is partitioned into quasi-uniform mesh $\Omega_h = \cup \ \tau$, denote linear finite element subspace by

$$S^h = \{v \in C(\Omega), \quad v|_\tau = linear, \quad v|_\Gamma = 0\}.$$

Define the finite element solution $U^j(x) \in S^h$ satisfying Crank-Nicolson scheme

$$\left(\frac{U^{j+1} - U^j}{k}, v\right) + A\left(\frac{U^{j+1} + U^j}{2}, v\right) = (f^{j+1/2}, v), \ v \in S^h, \ U^0 = \psi_h. \tag{2.3}$$

Following M.Wheeler[17](1973), we take the elliptic projection $R_h u$ of u as a comparison function,

$$A(u - R_h u, v) = 0, v \in S^h, \quad ||u - R_h u|| \le Ch^2 ||u(t)||_2,$$

and decompose the error of the discrete solution U into

$$u - U = E_h u - \theta, \quad E_h u = u - R_h u, \quad \theta = U - R_h u.$$

It is easy to prove the following lemma.

Lemma 1. *In the case of the quasi-uniform subdivision of Ω, the elliptic projection error $\rho = u - R_h u$ in a time interval $I_j = (t_j, t_{j+1})$ is estimated as follows,*

$$||E_h u^{j+1} - E_h u^j|| = ||\int_{I_j} \rho_t dt|| \le Ch^2 \int_{I_j} ||u_t||_2 dt \le Ckh^2 \max_{t_j \le t \le t_{j+1}} ||u_t||_2.$$

In [22] (P.17) Thomee summarized some standard results. For example, the C-N scheme has the error estimate

$$||u(t_j)-U^j|| \le Ch^2\Big(||\psi||_2+\int_0^{t_n}||u_t||_2 dt\Big)+Ck^2\int_0^{t_n}\Big(||\Delta u_{tt}||+||u_{ttt}||\Big)dt. \tag{2.4}$$

in which the regularity requirement of u for space-discretization is optimal, but the regularity of u for time-discretization is not optimal.

To decrease the regularity, we shall improve (2.4) in the following theorem.

Theorem 2.1. *For $\theta = U - R_h u$, there is the error estimate*

$$||\theta^n|| \le Ch^2\int_0^{t_n}||u_t||_2 dt + Ck^2\Big(\big(\int_0^{t_n}||u_{tt}||_1^2 dt\big)^{1/2}+\int_0^{t_n}||f_{tt}||dt\Big), \tag{2.5}$$

in comparison with (2.4)*, where $||\Delta u_{tt}||$ and $||u_{ttt}||$ are replaced by $||u_{tt}||_1$ and $||f_{tt}||$, respectively.*

Proof. Firstly, for $v \in S^h$, integrating (2.2) in I_j by trapezoid and mid-point formulas respectively, we have

$$(u^{j+1}-u^j,v)+\frac{k}{2}A(u^{j+1}+u^j,v)=k(f^{j+1/2},v)+r_1^j(v)+r_2^j(v), \tag{2.6}$$

where the residues can be expressed in the form (or asymptotic expansion)

$$r_1^j(v)=O(k^2)\int_{I_j}||f_{tt}||\,||v||dt = c_1k^2\int_{I_j}(f_{tt},v)dt+O(k^3)\int_{I_j}||f_{ttt}||\,||v||dt,$$
$$r_2^j(v)=O(k^2)\int_{I_j}||u_{tt}||_1||v||_1 dt = c_2k^2\int_{I_j}A(u_{tt},v)dt+O(k^3)\int_{I_j}||u_{ttt}||_1||v||_1 dt.$$

Note that here we do not use u_{ttt} and Au_{tt} to denote $f_{tt}=u_{ttt}+Au_{tt}$, as done in [22](p.17). In the theory of partial differential equations, the regularity $u_{ttt}, Au_{tt}\in L^2(Q)$ requires higher consistence of the data, whereas $f_{tt}\in L^2(Q)$ is only differentiability of f.

Comparing (2.3) with (2.6), the error $e=u-U$ satisfies the equation

$$(e^{j+1}-e^j,v)+\frac{k}{2}A(e^{j+1}+e^j,v)=r_1^j(v)+r_2^j(v).$$

Suppose $e=u-U=E_hu-\theta$, and notice $A(E_hu,v)=0$. Then θ satisfies

$$(\theta^{j+1}-\theta^j,v)+\frac{k}{2}A(\theta^{j+1}+\theta^j,v)=r_h^j(v), \tag{2.7}$$

where $r_h^j(v)=r_1^j(v)+r_2^j(v)+r_3^j(v)$,

$$r_3^j(v)=(E_h(u^{j+1}-u^j),v)=\int_{I_j}(E_hu_t,v)dt=O(h^2)\int_{I_j}||u_t||_2||v||dt.$$

On the quasi-uniform mesh, r_3^j has the following high order estimate in time

$$r_3^j(v)-r_3^{j-1}(v)=O(kh^2)\int_{I_j+I_{j-1}}||u_{tt}||_2||v||dt.$$

Taking $v^j = \theta^{j+1} + \theta^j$ (notice $\theta^0 = 0$), and summing over j, we have

$$||\theta^n||^2 + \frac{1}{2}|||v|||_1^2 = r(v), \quad |||v|||_1 = \Big(\sum_{j=0}^{N} ||v^j||_1^2 k\Big)^{\frac{1}{2}}, \tag{2.8}$$

where

$$|r(v)| = \Big|\sum_j r_h^j(v)\Big| \le Ck^2 \int_J \Big(||f_{tt}||\ ||v|| + ||u_{tt}||_1||v||_1\Big)dt + Ch^2 \int_J ||u_t||_2||v||dt,$$
$$\le Ck^2|||u_{tt}|||_1|||v|||_1 + \Big(Ck^2 \int_J ||f_{tt}||dt + Ch^2 \int_J ||u_t||_2 dt\Big) \max_{j\le n} ||\theta^j||.$$

Using Young inequality, we can eliminate the right term $|||v|||_1$. To eliminate the term $||\theta^j||$, we follow the method in [22]. Suppose $||\theta^m|| = \max_{j\le n} ||\theta^j||$, n in Eq. (2.8) will be substituted by m. Then the right term $||\theta^m||$ can be eliminated by using Young inequality. Finally owing to $||\theta^n|| \le ||\theta^m||$, for any n we obtain the estimate (2.5). Hence Theorem 2.1 is established. □

Theorem 2.2 *Assume that the partition of Ω is quasi-uniform and U^n is the discrete solution of C-N scheme. Then the time-difference of error $e = u - U$ has high order error:*

$$||(u-U)^n - (u-U)^{n-1}|| \le C(u)k(h+k)^2, \tag{2.9}$$

where

$$C(u) \le C\Big(\max_{t\le k}(||f_{tt}|| + ||u_t||_2 + ||u_{tt}||_2\Big) + \Big(\int_J ||u_{ttt}||_1^2 dt)^{1/2} + \int_J ||f_{ttt}||dt\Big). \tag{2.10}$$

Proof. In order to prove (2.9), the regularity requirement of (2.10) is higher than (2.5). By (2.6), Eq. (2.7) can be written in the form of asymptotic expansion

$$(\theta^{j+1} - \theta^j, v) + \frac{k}{2}A(\theta^{j+1} + \theta^j, v) = g(v) + r_{kh}(v), \quad v \in S^h,$$

where the main part

$$g(v) = C_1k^2 \int_{I_j} ((f_{tt}, v) + A(u_{tt}, v))dt + \int_{I_j} (E_h u_t, v)dt,$$

and the residual of three order

$$|r_{kh}(v)| \le Ck^3 \int_{I_j} \Big(||f_{ttt}||\ ||v|| + ||u_{ttt}||_1||v||_1\Big)dt + Ckh^2 \int_{I_j} ||u_{tt}||_2||v||dt.$$

On the time level $t_1 = k$, we temporarily assume that there holds

$$||\theta^1|| \le C(u)k(k+h)^2. \tag{2.11}$$

Because the space subdivision is not involved in the estimate (2.9), we consider the difference of θ^j in time

$$\eta^{j+1} = \theta^{j+1} - \theta^j, \quad \theta^{j+1} + \theta^j - (\theta^j + \theta^{j-1}) = \eta^{j+1} + \eta^j.$$

Using the two equalities obtained from I_j and I_{j-1} give

$$(\eta^{j+1} - \eta^j, v) + \frac{k}{2}A(\eta^{j+1} + \eta^j, v) = \delta^j(v) + r_{kh}(v), \tag{2.12}$$

where

$$\begin{aligned}\delta^j(v) =& g^{j+1/2}(v) - g^{j-1/2}(v)\\ \le& Ck(k^2+h^2)\int_{I_{j-1}+I_j}\Big((||f_{ttt}||+||u_{tt}||_2)||v||+||u_{ttt}||_1||v||_1\Big)dt.\end{aligned}$$

Notice that the integrand has been included in $r_{kh}(v)$ without additional discussion. Now taking $v^j = \eta^{j+1} + \eta^j$ and summing (2.12) over j, we have

$$\begin{aligned}&||\eta^n||^2 + \frac{\nu}{2}|||v|||_1^2\\ \le& ||\eta^1||^2 + Ck(k^2+h^2)\left(\int_J (||f_{ttt}||+||u_{tt}||_2)dt\max||v^j|| + |||u_{ttt}|||_1|||v|||_1\right),\end{aligned}$$

where $\eta^1 = \theta^1 - \theta^0 = \theta^1$ can be bounded by (2.11). Note $||\eta^n|| = \max\limits_{j\le N}||\eta^j||$, $|||v|||_1$ and $||\eta^n||$ can be eliminated by using Young inequality, we obtain

$$||\theta^n - \theta^{n-1}|| = ||\eta^n|| \le Ck(k+h)^2,$$

We now come back to prove the estimate (2.11). Noting $\theta^0 = 0$, for $j = 1$ and taking $v = \theta^1$ in (2.7), we have

$$||\theta^1||^2 + \frac{k}{2}||\theta^1||_1^2 = r_h(v), \quad v \in S^h, \tag{2.13}$$

where we have to use a stronger norm estimate

$$|r_h(v)| \le C_1 k\Big(k^2\max_{t\le k}(||f_{tt}||+||u_{tt}||_2) + h^2\max_{t\le k}||u_t||_2\Big)||\theta^1||.$$

Finally, using $(u-U)^n = E_h u^n - \theta^n$, (2.13) and Lemma 1, we have (2.9). Consequently, Theorem 2.2 holds. □

3. Direct CG-iteration (DCG)

Assume that $\Omega = (0,1)^2$ is a square. Denote second order central difference $\delta_x^2 u_i^j = u_{i+1}^j - 2u_i^j + u_{i-1}^j$, and consider the Crank-Nicolson difference scheme at (x_i, y_l)

$$U^{j+1} - U^j = \frac{a^2k}{2h^2}\Big(\delta_x^2U^{j+1} + \delta_x^2U^j + \delta_y^2U^{j+1} + \delta_y^2U^j\Big) + kf^{j+1/2}. \tag{3.1}$$

Denote the mesh-ratio $r = a^2k/(2h^2)$. Then the scheme (3.1) can be written as linear system

$$AU^{j+1} = F(U^j) = BU^j + kf^{j+1/2},$$

where

$$AU^{j+1} \equiv U^{j+1} - r(\delta_x^2U^{j+1} + \delta_y^2U^{j+1}), \quad BU^j \equiv U^j + r(\delta_x^2U^j + \delta_y^2U^j).$$

To solve a discrete elliptic system $AU = b$, we know that the solution U^n of the DCG has the classical estimate

$$||U^n - U||_A \le 2\rho^n||U^0 - U||_A, \quad \rho = \frac{\sqrt{cond(A)}-1}{\sqrt{cond(A)}+1}.$$

For example, for five-point scheme of Poisson equation, denoting by h and n the step-length in space and iteration count of CG respectively, the condition number of coefficient matrix A

$$cond(A) \approx \frac{1}{2}h^{-2}, \quad \text{so} \quad \rho \approx 1-h, \quad \rho^n \approx e^{-nh},$$

i.e., taking $n = h^{-1}$, the error will be contracted to $1/e$.

However, for parabolic problem, we found numerically, when $r > 4$, the C-N scheme (3.1) has smaller condition number

$$cond(A) \approx 4r, \quad r = \frac{a^2 k}{2h^2}, \quad \text{if} \quad k = O(h),$$

So DCG has a better estimate as follows,

$$||U^n - U||_A \leq 2e^{-n/\sqrt{r}}||U^0 - U||_A, \quad \sqrt{r} = \frac{a}{h}\sqrt{\frac{k}{2}}.$$

When $n = \sqrt{r}$, the error will be contracted to $1/e$, which is better than the case of elliptic problems.

Note that the parabolic problems require often to solve the linear systems on many time levels, which may require more CPU-time. This is the new difficulty.

We take **Direct CG-iteration (DCG)** as a comparing mode:

- $U^{n,0} = U^{n-1}$, *with the error* $U^{n-1} - U^n = O(k)$;
- CG-iterations $U^{n,p} = S^p(U^{n,0})$ are stopped by (control) residues

$$Res = ||F - AU^{n,p}|| \leq \epsilon, \quad U^n = U^{n,p}.$$

For Problem (2.1) with $a = 1$ and the exact solution $u = \sqrt{1+t}\sin(\pi x)\sin(\pi y)$, numerical results obtained are shown in Table 3.1. Because the solution increases with time, we use relative error and relative control residual. Here fix k/h, take $T = nh = 3$, and let the iteration count of EXCMG on $i-th$ level be $4 * 4^{L-i}$, where L is the total number of levels.

Table 3.1: Comparison of DCG and EXCMG.

h	1/128	1/256	1/512	1/1024	remark
Unknowns	$15K$	$65.5K$	$262K$	$1050K$	$K = 10^3$
k	1/100	1/200	1/400	1/800	
n	300	600	1200	2400	
ratio $r = k/(2h^2)$	82	164	328	656	$cond(A) \approx 4r$
EXCMG, relative error	$4.96e-5$	$1.24e-5$	$3.10e-6$	$7.70e-7$	
ratio of error e	—	4.00	3.99	4.03	$e = O(k^2 + h^2)$
EXCMG, CPU T1(s)	1.58	14.94	133.8	1270	$p(r) \approx 16$
DCG, iteration count $p(r)$	26.6	40.1	55.5	81.5	$p(r) \approx O(\sqrt{r})$
DCG, CPU T2(s)	2.64	27.38	440	6070	
T2/T1	1.67	1.83	3.29	4.78	

The following observations can be made from Table 3.1:

- For the model problem, take $L = 6$, the iteration count of CG in EXCMG is about $p(r) \approx 16$ (to be converted into the finest grid), independent of r and control residue.

- Direct CG-iteration (DCG) is acceptable when $r < 700$. The iteration count of DCG $p(r) \approx O(\sqrt{r})$ increases with r slowly .
- For parabolic problems, although EXCMG is four times faster than DCG when $r = 700$, its advantages far less than the case of elliptic problems, so a wide research space in the other algorithms remains.

4. Time-Extrapolation Algorithm (TEA)

Our new idea is to propose Time-Extrapolation Algorithm (TEA):
Step 1. use linear combination of previous l level solutions as an initial value

$$U^{n,0} = I_l(U) \equiv \sum_{j=1}^{l} a_j U^{n-j}, \quad l = 3 \sim 7. \tag{4.1}$$

Step 2. take CG-iterations up to $Res = ||F - AU^{n,p}|| < \epsilon$.

By the function approximation, we take the following formula

$$u^n = 2u^{n-1} - u^{n-2} + O(k^2), \quad u^n = 3u^{n-1} - 3u^{n-2} + u^{n-3} + O(k^3), \cdots .$$

For example, Thomee [22] (p.186) proposed to use the better initial value

$$U^{n,0} = 2U^{n-1} - U^{n-2}, \quad ||U^{n,0} - U^n|| = O(k^2),$$

"as a second order accurate extrapolation approximation to U^n", although he did not discuss the computational efficiency. Numerical experiments show that this approach is not good (see the results of 3D3L in Table 5.1). The reason is that both the combination coefficients and the sum of the coefficients are larger so the rounding errors are continued to increase. Recall that the contraction ratio of CG is only $\rho \approx e^{-1/\sqrt{r}}$.

We shall discuss (4.1) in the general framework. First, the error e^n of the C-N scheme is decomposed into three parts [5]

$$\begin{aligned} &e^n = U^n - u^n = C^n k^2 + \eta^n + O(k(k+h)^2), \\ &\eta^n = (u - R_h u)^n = O(h^2), \text{ on quasiuniform grid,} \end{aligned}$$

and it is proved in Theorems 2.1 and 2.2 that

$$||e^n - e^{n-1}||, \quad ||\eta^n - \eta^{n-1}|| \le Ck(k+h)^2. \tag{4.2}$$

Rewrite the exact solution U^n as $U^n = u^n + e^n$, and denote the iteration solution W^j has error $W^j - U^j = \epsilon^j$. Then we have

$$W^j = u^j + \epsilon^j + e^j, \quad j < n.$$

The extrapolation formula (4.1) provides an initial value

$$U^{n,0} = I_l W = \sum_{j=1}^{l} a_j W^{n-j} = I_l u + I_l \epsilon + I_l e,$$

which has the error

$$U^n - U^{n,0} = (u^n - I_l u) - I_l \epsilon + (e^n - I_l e).$$

By (4.2), we obtain $||e^n - I_l(e)|| \le Ck(k+h)^2$. The main error for $U^{n,0}$ consists of two parts

$$||U^n - U^{n,0}|| \le ||u^n - I_l u|| + ||I_l \epsilon|| + O(k(k+h)^2).$$

Therefore the study of TEA is related to three important issues:

1). Interpolation $I_l u$ should have the accuracy of order m,

$$||u^n - I_l u|| \le Ck^m, \quad m = 2, 3,$$

and for $u = t^m$, the residual of $I_l u$

$$(t_l)^m - I_l(t^m) = \Big(l^m - \sum_{j=1}^{l} a_j (j-1)^m\Big) k^m = \beta_m k^m,$$

should be smaller. In fact, $m = 3$ is better.

2). The errors ϵ^j are interpolated and transferred to the present level,

$$||I_l \epsilon|| = ||\sum_{j=1}^{l} a_j \epsilon^{n+l-1-j}|| \le S \max_{1 \le j \le l} ||\epsilon^{n+l-1-j}||, \quad S = \sum_{j=1}^{l} |a_j|.$$

Interpolation $I_l \epsilon$ is of good stability, if and only if the coefficients $|a_j|$ and sum S should be as smaller as possible. How to chose these coefficients a_j is quite complicated.

3). How to control residues $Res = ||F - AU^{n,p}|| \le \epsilon$ to guarantee that the deviation(accuracy) satisfies

$$Er = ||U^{n,p} - U^n|| < ||U^n - u|| = e_0 \ \ (\textit{basic error}).$$

5. Numerical Examples

In a cylindric domain $Q = (0,1)^2 \times (0,T)$ we consider

$$u_t = a^2 \Delta u + f(x,t), \quad u(x,0) = \psi(x), \quad u(x,t) = 0 \ \text{ on } \ \partial\Omega,$$

with an increasing-type solution

$$u = (1+t)^{1/2} \sin \pi x_1 \sin \pi x_2, \quad a = 1.$$

Taking the step-length

$$k = 0.01, \quad h = 1/N, \ N = 128, 256, 512, \quad \max r = \frac{a^2 k}{2h^2} \approx 1311.$$

we discuss the Crank-Nicolson difference scheme.

Below we take DCG as a comparing mode. Compare the extrapolation formulas of order m and level l,

$2D2L:$	$W^n = [2, -1],$	$S = 3, \ \beta_2 = 2;$ *bad*
$2D3L-1:$	$W^n = [1.5, -0.5],$	$S = 2, \ \beta_2 = 3;$ *not good*
$2D3L:$	$W^n = [1, 1, -1],$	$S = 3, \ \beta_2 = 4;$ *not good*
$3D3L:$	$W^n = [3, -3, 1],$	$S = 7, \ \beta_3 = 6;$ *bad*
$3D4L:$	$W^n = [2, 0, -2, 1],$	$S = 5, \ \beta_3 = 12;$
$3D5L:$	$W^n = [1.25, 1.25, -1.25, -1.25, 1],$	$S = 6, \ \beta_3 = 22.5;$
$3D6L:$	$W^n = [1, 1, 0, -1, -1, 1],$	$S = 5, \ \beta_3 = 36;$ *good*
$3D7L:$	$W^n = [1, 1, -0.95, 0.9, -1, -0.9, 0.95],$	$S = 6.7, \ \beta_3 = 47.4.$ *better*

In the calculation for long time, first we obtain the approximate solutions for first 10-20 levels by DCG, then use TEA. For first 10-20 levels, DCG has a faster convergence rate and higher precision. Later with TEA iteration, the precision is reduced and a cyclical oscillations of accuracy are generated, which is disadvantageous to reduce the number of iterations. For this reason, we adopt different control accuracy for first 10-20 levels and later levels. For example, $Res = (e-p, e-(p+1))$ means that DCG iterations use $Res = 10^{-p}$, and TEA iterations $Res = 10^{-(p-1)}$. It should be noted that less iterations are required with $Res = (e-3, e-4)$, but it does not reach the accuracy of the finite difference solution; the accuracy requirement is too high with $Res = (e-5, e-6)$, which increases the iteration count; therefore $Res = (e-4, e-5)$ is more appropriate.

For $0 \le t \le 5$, we have compared the numerical results for several algorithms in Table 5.1.

Table 5.1: Average iteration counts for DCG and TEA algorithms with fixed $k = 0.01$ and $T = 5s$.

$1/h$	Res	DCG	2D2L	2D3L	3D3L	3D4L	3D5L	3D6L	3D7L*
128	$e-3, e-4$	23.8(5.0s)	5.8	5.1	14.8	9.6	10.5	8.3	6.7
	$e-4, e-5^*$	23.7(3.6s)	15.7	17.4	14.9	10.0	11.1	9.1	7.8*
	$e-5, e-6$	23.2(4.7s)	27.9	22.7	16.3	11.6	13.4	11.9	11.9
256	$e-3, e-4$	47.3(33.3s)	13.6	12.1	30.4	19.5	22.4	16.5	13.6
	$e-4, e-5^*$	47.4(31.4s)	36.0	37.8	30.5	20.1	22.7	19.6	15.4*
	$e-5, e-6$	46.6(34.4s)	56.5	45.6	33.2	23.7	27.9	25.9	27.5
512	$e-3, e-4$	95.8(334s)	31.6	31.9	60.4	39.7	43.2	35.9	28.2
	$e-4, e-5^*$	95.1(335s)	79.8	81.7	61.4	40.9	45.5	36.5	34.3*
	$e-5, e-6$	93.0(333s)	114	91.2	67.3	51.5	58.0	53.7	57.0

It is observed from Table 5.1 that

1. Different control residues have little effect on the number of iterations of DCG, but a great influence on the number of iterations of the TEA;

2. The 2-order formulas are generally not good. The 3-order formula 3D3L is also not good, because two largest coefficients ± 3 in the formula seriously enlarge the rounding errors of approximate solution on former three levels, which requires more iterations;

3. Third order formula 3D6L and 3D7L, whose coefficient $|a_j| \le 1$, are relatively good. Especially for 3D7L, its coefficients have some antisymmetry and CG iterations count can be reduced to 1/3 of DCG's. We should find better extrapolation formulas.

6. Conclusions

Summarizing above results, we have the following conclusions:

1. For linear parabolic problems, Direct CG-iteration(DCG) is acceptable when $r < 700$, the iteration count $p(r) \approx O(\sqrt{r})$ increases slowly with r. Although EXCMG is four times faster than DCG, its advantages far less than the elliptic case;

2. The iteration count of EXCMG is almost unchanged when r changes in a large range. EXCMG is significantly better than the DCG only when $r \ge 2000$;

3. TEA is a universal algorithm, very simple, only extrapolation in time needed, independent of space. Therefore, it is effective for the quasi-uniform grid in space (what important is that only the finest grid is needed), independent of the dimension of space;

4. With a good initial value, TEA will be more favorable for nonlinear parabolic problems [4, 5], only one Newton iteration and re-calculating of the stiffness matrix in every time level are needed. Therefore TEA is so simple, equivalent to solve a linear problem with variable coefficients.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (No. 11071067, 41204082, 11301176), the Research Fund for the Doctoral Program of Higher Education of China (No. 20120162120036), the Hunan Provincial Natural Science Foundation of China (No.14JJ3070) and the Construct Program of the Key Discipline in Hunan Province.

References

[1] A. Brandt, Multi-level adaptive solutions to boundary value problems, *Math. Comp.*, **31** (1977) 333–390.

[2] F. Bornemann, P. Deuflhard, The cascadic multigrid method for elliptic problems, *Numer Math.*, **75**:2 (1996), 135–152.

[3] C.M. Chen, Z.Q. Xie, C.L. Li and H.L. Hu, Study of A New Extrapolation Multigrid Method, *Journal of Natural Science of Hunan Normal University*, **30**:2 (2007), 1-5.

[4] C.M. Chen, S. Larsson and N.Y. Zhang, Error estimates of optimal order for finite element methods with interpolated coefficients for the nonlinear heat equation, *IMA J. Numer. Anal.*, **9** (1989), 507-524.

[5] C.M. Chen, Y.Q. Huang, Higher Accuracy Theory in Finite Element Methods (in Chinese). Changsha: Hunan Science Technique Press, 1995.

[6] C.M. Chen, H.L. Hu, Z.Q. Xie and C.L. Li, Analysis of extrapolation cascadic multigrid method Science China, *Series A*, **51**:8 (2008), 1349-1360.

[7] C.M. Chen, Z.C. Shi and H.L. Hu, On Extrapolation Cascadic Multigrid Method, *Journal of Computational Mathematics*, **29**:6 (2011), 684-697.

[8] P. Deuflhard, P. Leinen, H. Yserentant, Concepts of an adaptive hierarchical finite element code, *IMPACT Comput Sci Engrg*, **1** (1989), 3–35.

[9] C.C. Douglas, Multigrid Methods in Science and engineering, *IEEE Comput. Sci. Engrg.*, **3** (1996), 55-68.

[10] Q. Du, P.B. Ming, Cascadic multigrid methods for parabolic problems, *Science in China Series A*, **51**:8 (2008), 1415–1439.

[11] M.J. Gander, S. VandewalleE, Analysis of the parabolic time-parallel time-integration method, *SIAM J. Sci. Comput.*, **29**:2 (2007), 556-578.

[12] W. Hackbusch, Parabolic multigrid methods, in Computing Methods in Applied Sciences and Engineering, VI, R. Glowinski and J.L. Lions, eds.,North C Holland, Amsterdam: 1984, pp. 189-197.

[13] G. Horton, The Time-Parallel Multigrid Method, *Comm. Appl. Numer. Meth.*, **8** (1992), 585-595.

[14] G. Horton, S. Vandewalle, A space-time multigrid method for parabolic partial differential equations, *SIAM J. Sci. Comput.*, **16**:4 (1995), 848-864.

[15] K.J. Pan, H.L. Hu, C.M. Chen, et al., Parallelization of extrapolation cascadic multigrid method using OpenMP(in Chinese), *Mathematica Numerica Sinica*, **34**:4 (2012), 425-436.

[16] K.J. Pan, J.T. Tang, H.L. Hu, et al., Extrapolation cascadic multigrid method for 2.5D direct current resistivity modeling(in Chinese), *Chinese J. Geophys*, **55**:8 (2012), 2769-2778.

[17] V. Shaidurov, Some estimates of the rate of convergence for the cascadic conjugate gradient method, *Comp Math Appl.*, **31**:4-5 (1996), 161–171.

[18] Z.C. Shi, X.J. Xu, Cascadic multigrid method for elliptic problems, *East-West J Numer Math.*, **7** (1999), 199–211.

[19] Z.C. Shi, X.J.Xu, Cascadic multigrid for parabolic problems, *J. Comput. Math.*, **18**:5 (2000), 551-560.

[20] Z.C. Shi, X.J. Xu, A new cascadic multigrid, *Science in China Series A*, **44**:1 (2001), 21-30.

[21] Z.C. Shi, X.J. Xu, Y.Q. Huang, Economaical cascadic multigrid methods(ECMG), *Science in China Series A*, **50**:12 (2007), 1765–1780.

[22] V. Thomee, Galerkin Finite Element Methods for Parabolic Problems, Second edition, New York: Springer-Verlag, 2006.

[23] M. Wheeler, A Priori L_2 Error Estimates for Galerkin Approximations to Parabolic Partial Differential Equations, *SIAM J. Numer. Anal.*, **10**:4 (1973), 723-759.

[24] S.Z. Zhou, H.X. Hu, On the convergence of a cascadic multigrid method for semilinear elliptic problem, *Appl. Math. Comput.*, **159**:2 (2004), 407-417.