A Preliminary Comparison between a Multigrid Method and a Preconditioned Conjugate Gradient Method.

Ivar Gustafsson
Department of Mathematics
Chalmers University of Technology
S-41296 Göteborg, Sweden
ivar@math.chalmers.se

March 8, 2002

Abstract

Multigrid methods are among the most popular methods for solving discretized partial differential equations (pde-systems). One reason for this is the optimal behaviour of the solution time with respect to the degree of discretization i.e. the number of arithmetic operations required for the solution is proportional to the number of unknowns in the system of equations to be solved.

Another well-known class of methods is preconditioned conjugate gradient (pcg) methods. Since these methods are not optimal, like multigrid methods, they are regarded as less efficient in general.

For a fair competition, however, one should also consider the constant in the asymptotic behaviour. Also, the most efficient and robust pcg methods have to be considered. This author believes that often (unmodified) incomplete factorizations are used for the preconditioning while *modified* incomplete factorizations are often much more efficient. In fact, the modified versions are of higher order than the unmodified ones, with respect to a definition of order of convergence rate given in the introduction.

In this paper we make some very preliminar comparisions between a standard multigrid technique and a modified incomplete conjugate gradient technique for a couple of quite simple testproblems.

Key Words: multigrid method, preconditioned conjugate gradient method, incomplete factorization, finite difference method, finite element method.

1 Introduction.

We consider finite difference or finite element approximations of partial differential equations. The derived so called pde-systems are solved by multigrid methods or preconditioned conjugate gradient methods.

We use an implementation of the standard multigrid method as described in [1]. We use the V-cycle version, including interpolation-, restriction- and smoothing operations. The restriction operator is defined in [1] and we usually use linear interpolation. In some cases, however, we have also tested cubic interpolation. We use weighted Jacobi iterations with weight $w = 2/3$ as smoother.

For the preconditioned conjugate gradient method we use modified incomplete (MIC) factorization for construction of the preconditioning matrix, see [2] and [3]. This method is a modification of an incomplete factorization (IC) technique first presented in [8]. In order to become robust, in the sence of applicability to arbitrary symmetric positive definite pde-systems, the incomplete factorization methods, IC as well as MIC, are based on modified element matrices as described in [4]. When this modification on element level is actually performed, the methods are denoted EIC and EMIC, respectively.

Let h be a mesh-size parameter in the finite difference or finite element discretization of a second order elliptic differential equation. Then it is well known, see eg. [3], that the number of iterations for the (E)ICCG method is of order $O(h^{-1})$, $h \to 0$. This is called *first order convergence rate*, since it is actually of the same order as for the unpreconditioned system itself. For the (E)MICCG method, however, the number of iterations is only of order $O(h^{-1/2})$, $h \to 0$, and this is called *second order convergence rate*.

In the section to follow, we present comparisions between the multigrid method and the pcg method for some fairly simple modelproblems. In section 3, we draw some preliminary conclusions.

2 Comparision between multigrid and pcg techniques.

In this section we study a couple of simple testproblems. We compare multigrid methods with pcg-methods with respect to the number of floating point operations required for a small residual. Theoretically, we are interested in the asymptotic behaviour of the two classes of methods when the mesh size parameter $h \to 0$, i.e. when the pde-system becomes large. For practical reasons, however, it is more interesting to find out which class of methods is the better for a given problem of a certain size (a certain mesh parameter h) and to find breakingpoints between the different methods.

Our testproblems are based on the simple partial differential equation in two space dimensions and homogeneouos Dirichlet boundary conditions:

$$\begin{cases} -\nabla(a\ \nabla u) = f,\ (x,y) \in \Omega \\ u(x,y) = 0,\ (x,y) \in \partial\Omega \end{cases} \tag{1}$$

for different righthand sides f and coefficient a. The domain Ω is taken as a unit square (of different size).

Discretization of (1) gives a symmetric positive definite system of equations denoted

$$Ax = b. \tag{2}$$

By our iterative methods we seek an approximate solution of (2) satisfying the breaking-criterion

$$\| r \|_2 \leq 10^{-6} \tag{3}$$

h	1/16	1/32	1/64	1/128
IC(0)	19	33	62	120
IC(2)	12	20	36	60
MIC(0)	14	19	27	38
MIC(2)	9	13	18	26
multigrid	3	3	2	2

Table 1: The number of iterations for multigrid and pcg methods for the Laplace equation on a unit square with different stepsize h in the five-point difference approximation.

where $r = Ax - b$ is the residual and $\| \cdot \|_2$ is the usual Euklidean norm of a vector.

2.1 The standard five-point approximation of Laplace equation.

As a first modelproblem we take (1) with constant coefficient $a = 1$ and constant righthand side $b = 1$ on the square $S_1 = \{-1 \leq x \leq 1, \ -1 \leq y \leq 1\}$. We use standard five-point finite difference approximations to get the dicsretized pde-system (2).

This is a perfectly suited problem for multigrid technique, see [1]. At first we perform one full multigrid (fmg) cycle and if the stopping ctiterion (3) is not fullfilled then, we perform a number of additional multigrid V (mgv) cycles until (3) is fulfilled. The coarsest grid is due to $h = 1/2$ on which the system of equations of size 3×3 is solved by a direct method (Gaussian elimination). As smoother we have used four weighted Jacobi iterations.

For the pcg-method we use incomplete and modified incomplete factorizations with different degree of fill-in; (M)IC(0) indicates no fill-in and (M)IC(2) indicates fill-in in two subdiagonals of the lower factor L in the incomplete Cholesky factorization $A \approx LL^T$, for details see [2] and [3]. As starting approximation x_0 in the pcg-method we use the trivial vector $x_0 = 0$ or we use interpolation from a coarser mesh with double mesh size parameter h. In most cases we have used linear interpolation but later on sometimes also cubic interpolation will be used. For this test example we have used (3) with lower accuracy, $\| r \|_2 \leq 10^{-4}$ for the solution on the coarser mesh.

In table 1 we present the number of iterations needed for the methods to reach the criterion (3) with the trivial starting approximation for the pcg-methods. The number of iterations in the multigrid method refers to the number of additional mgv-cycles. We observe a good agreement with the theory; the number of iterations in the multigrid method is (almost) independent of h, the number of iterations for the modified MIC-CG methods is of order $O(h^{-1/2})$, $h \to 0$, i.e. second order convergence rate, while the number of iterations for the (unmodified) IC-CG methods rather behaves like $O(h^{-1})$, $h \to 0$, i.e. first order convergence rate.

In figure 1 we present the number of floating point operations per unknown required for the different methods for various values of problem size $n = 2/h - 1$, i.e. the system of equations is of order $n^2 \times n^2$. We see that the breakeven point between the multigrid method and the best pcg method, i.e. MIC(2) with starting approximation given by the interpolated solution from a coarser mesh, is $n \approx 230$, i.e. only for very large problems

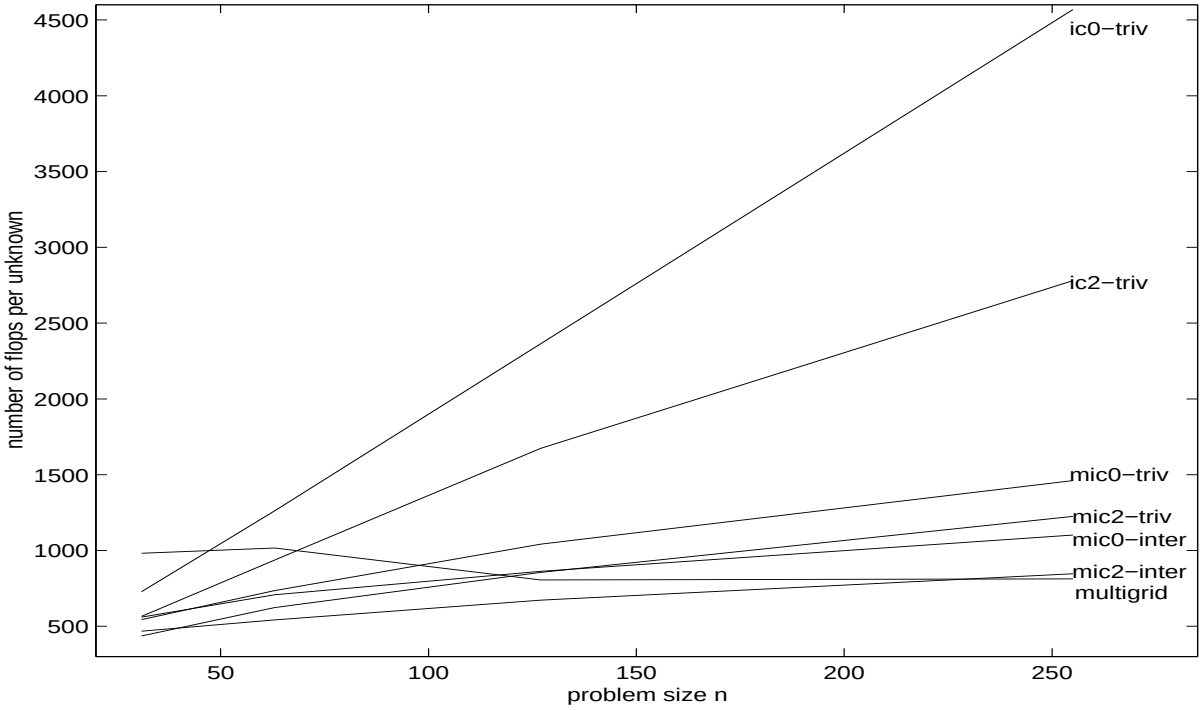

Figure 1: The number of floating point operations per unknown for various methods for the solution of the discretized Laplace equation on a unit square with various values of $n = 2/h - 1$.

with more than about 50 000 unknowns the multigrid method is faster.

2.2 An anisotropic problem.

As a testproblem, which in not so well suited for multigrid methods, we use the anisotropic variant of (1) corresponding to different coefficients in different space directions;

$$\begin{cases} -a\ u''_{xx} - u''_{yy} = f,\ (x,y) \in \Omega \\ u(x,y) = 0,\ (x,y) \in \partial\Omega. \end{cases} \tag{4}$$

With $a = 0.1$ we need more smoothing iterations in the multigrid method than for the corresponding isotropic problem with $a = 1$. If we just use four Jacobi-iterations like in the last subsection the number of additional mgv-cycles increases. If we increase the number of smoothing iterations, however, the number of additional mgv-cycles can still be kept small. To be more concrete, for $n = 127$ and four Jacobi-iterations the number of mgv-cycles, needed for (3) to be fullfilled, is seven. With ten Jacobi-iterations as smoother still only two mgv-cycles are required. Because of the larger number of Jacobi-iterations the total work for the multigrid method increases compared to the isotropic case. Indeed, the work required for the multigrid method increases rapidly with decreasing value of a.

In order to explain the behaviour of Jacobi as a smoother we use discrete Fourier transform, see [5]. After each Jacobi-iteration we compute a two-dimensional discrete Fourier transform of the approximate solution. Then we compute the power spectrum of

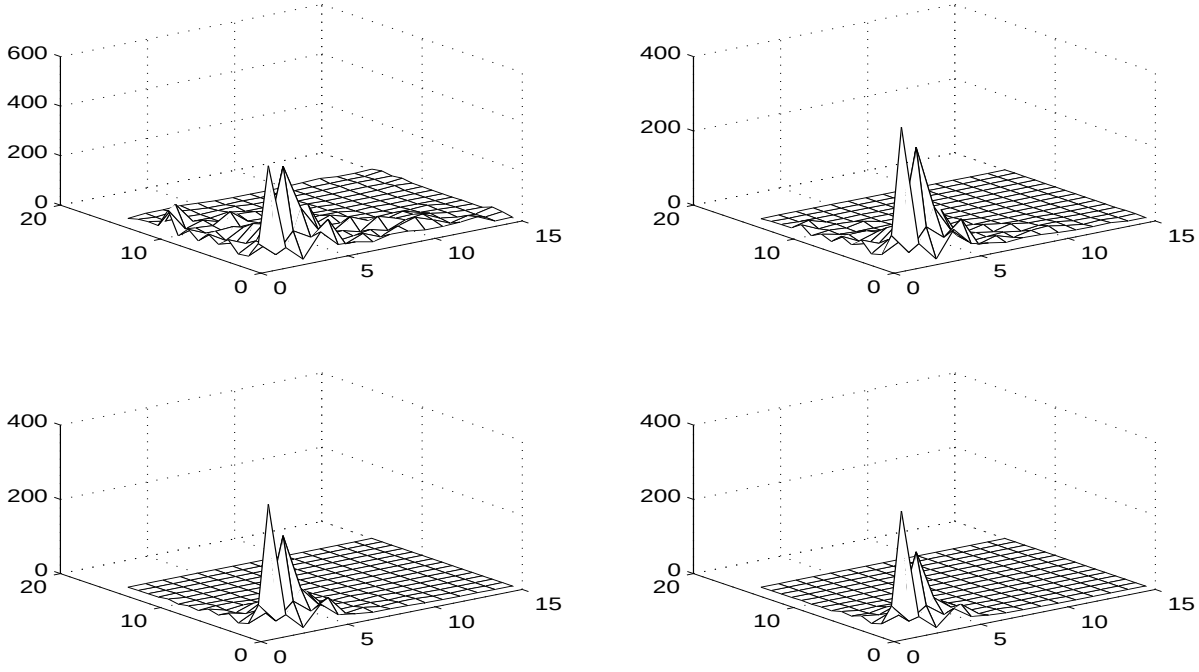

Figure 2: The smoothing effect of Jacobi-iteration on the isotropic problem.

the discrete Fourier transform. The first quarter of this power spectrum matrix (i.e. up to the so called Nyquist frequency in both space directions) shows the amout of energy present at each frequency. In figure 2 we have plotted the first quarter of the power spectrum matrix after each smoothing iteration with weighted Jacobi for the isotropic problem $a = 1$ and $n = 32$. We see that the high frequences are damped out rapidly and that their corresponding energies are very small already after four iterations.

For the anisotropic problem with $a = 0.1$ and $n = 32$ we present the corresponding result in figure 3. We see that more smoothing iterations are needed to damp out the high frequences in both space directions.

For the pcg method with incomplete factorizations, on the other hand, the number of iterations decreases with decreasing value of a. This is so because the incomplete factorizations tend to become exact when $a \to 0$. In particular, the (M)IC(2) factorizations give an error matrix $R = A - LL^T$ with elements of order $O(a^2),\ a \to 0$ if a standard rowwise numbering of the meshpoints is used. For the (M)IC(0) factorization the error elements are of order $O(a),\ a \to 0$.

In figure 4 we present the number of floating point operations per unknown for the anisotropic problem (4) with $a = 0.1$ and $f = (1 - x^2)(1 - y^2)$ on S_1 and various methods and problem size n. Since the unmodified IC-CG methods are of lower order of convergence rate than the modified versions we do not consider them any longer. For the MIC(2)-CG method we use different starting approximations, the trivial one, i.e. $x_0 = 0$, and linear interpolation of the solution from a coarser mesh. For this problem it is worth while to solve the problem on the coarser mesh to full accuracy, that is to use the breaking criterion (3) with tolerance 10^{-6}, compared to tolerance 10^{-4}.

It can be observed in the figure that the multigrid method is not competative unless

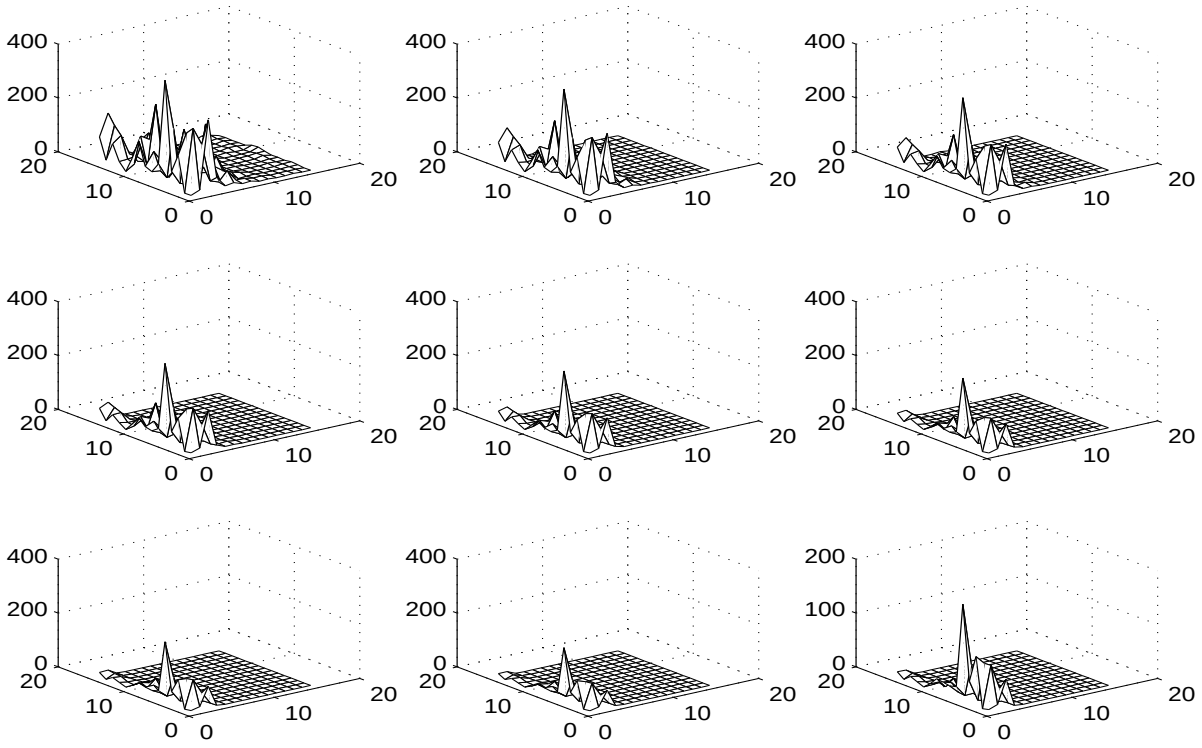

Figure 3: The smoothing effect of Jacobi-iteration on the anisotropic problem with $a = 0.1$.

the problem is very large. Following the asymptotic behaviour we estimate the breakeven piont between multigrid and the less efficient MIC method, i.e. MIC(2)-CG with trivial starting approximation, to be about $n \approx 1000$, i.e. a problem with about one million unknowns.

2.3 A discontinuous problem.

In this subsection we consider the testproblem (1) on the unit square $S_0 = \{0 \leq x \leq 1,\ 0 \leq y \leq 1\}$ with righthand side $f = 1$ and discontinuous coefficient a such that $a = d > 1$ inside the four subdomains $\Omega_1 = \{0.125 \leq x \leq 0.25,\ 0.125 \leq y \leq 0.375\}$, $\Omega_2 = \{0.125 \leq x \leq 0.25,\ 0.5 \leq y \leq 0.875\}$, $\Omega_3 = \{0.375 \leq x \leq 0.875,\ 0.125 \leq y \leq 0.375\}$ and $\Omega_4 = \{0.375 \leq x \leq 0.875,\ 0.5 \leq y \leq 0.875\}$ and $a = 1$ outside these subdomains.

This problem is used in [6] as a testproblem for a recent incomplete factorization method of drop tolerance type. In [7] this newer incomplete factorization technique is compared with the former (M)IC factorization methods based on fill-in in positions, using this testproblem among others. Although this testproblem for large degree of discontinuity, i.e. $d >> 1$, is not well suited for incomplete factorization based on certain diagonal positions like our (M)IC-methods, the comparision in [7] is in favour of these methods.

As we will see, however, for this problem the multigrid method is preferable compare to MIC-CG already for moderate zise of the problem. Here we just perform four weighted Jacobi-iterations in the smoothing operation of the multigrid method. In figure 5 and

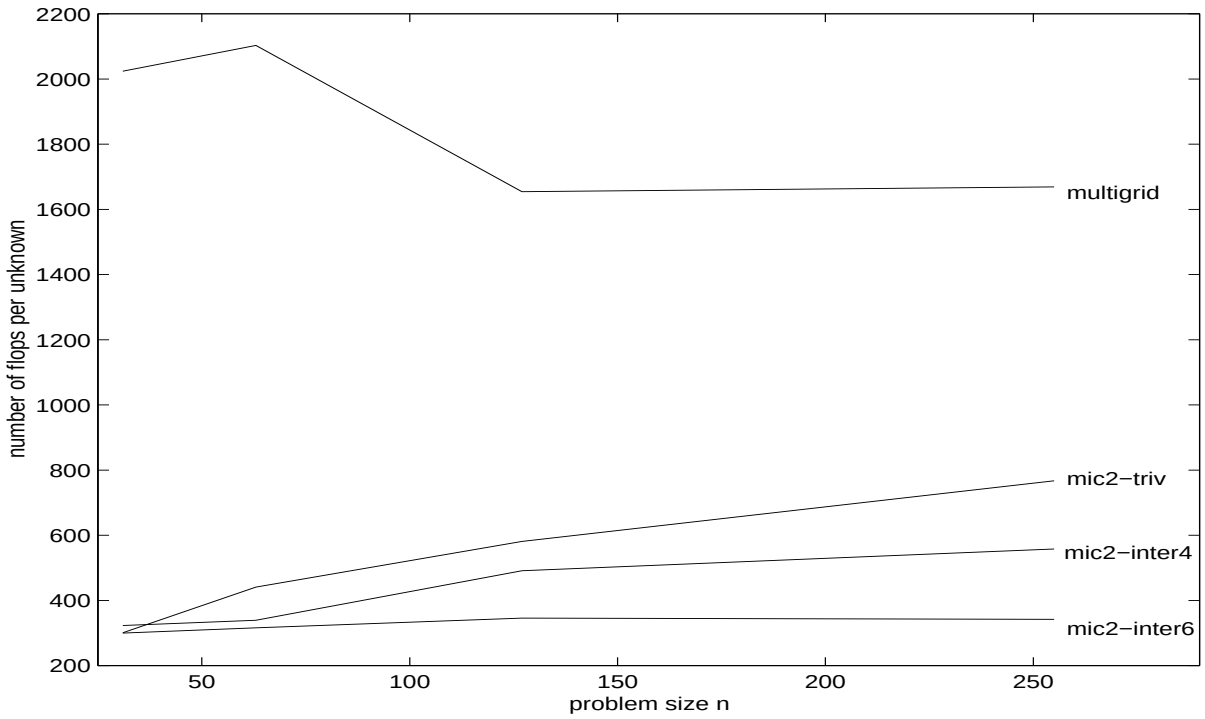

Figure 4: The number of floating point operations per unknown for various methods for the solution af the anisotropic problem (4) with $a = 0.1$ and $f = (1 - x^2)(1 - y^2)$ and various problem size $n = 2/h - 1$.

figure 6 we present the number of floating point operations for $d = 10$ and $d = 1000$, respectively. It is observed that the breakeven point between multigrid and MIC-CG uccurs for smaller value of n when the value of d is increased. We also see that it does not pay off much to use a better starting approximation derived by interpolation from a coarser mesh. This is so because we have to interpolate across the discontinuity lines.

2.4 Discretization by quadratic finite elements.

We solve the modelproblem (1) with constant coefficient $a = 1$ and righthand side $f = 1$ on $S_0 = \{0 \leq x \leq 1,\ 0 \leq y \leq 1\}$ by quadratic finite elements based on a uniform rightangled triangulation. Since the derived finite element matrix in the pde-system (2) is not an M-matrix in this case, we use modification on element level as decribed in [4] in order to ensure stability and fast convergence in the MIC-CG methods, i.e. we use EMIC factorizations.

In the multigrid method as well as when interpolating from a coarser mesh to get a good starting approximation for the EMIC-CG method, it is worth while to use higher order interpolation for this higher order finite element approximation. In the tests we have compared linear and cubic interpolation at this point. As smoother in the multigrid method we have used four weighted Jacobi iterations.

In figure 7 we present the number of floating point oprations for the compared methods and various size of the problem $n = 2/h - 1$. We see that the EMIC(2)-CG method, with

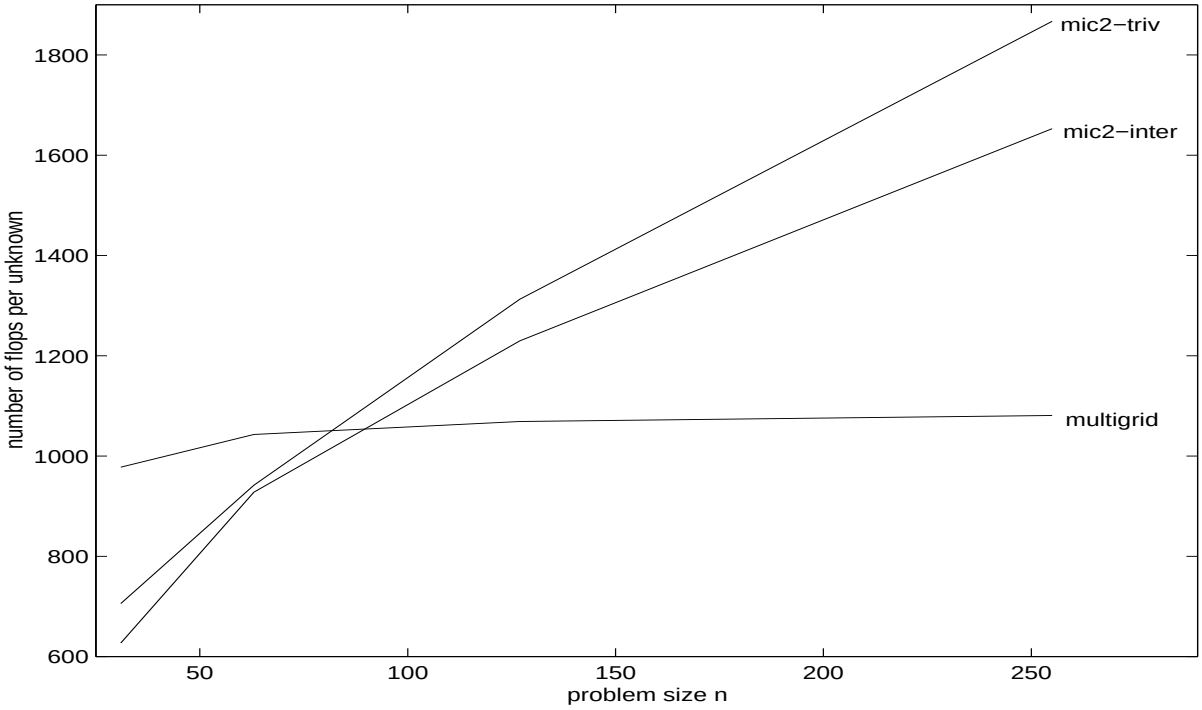

Figure 5: The number of floating point operations per unknown for various methods for the solution of the discontinuous modelproblem with $d = 10$ and with various values of $n = 1/h - 1$.

starting approximation based on cubic interpolation from a coarser mesh, is the most effecient method among the compared ones. Here we use full accuracy when solving the problem on the coarser mesh, i.e. we use (3) with tolerance 10^{-6}.

The observed decrease in the number of floating points operations for the multigrid method with cubic interpolation is due to the fact that fewer additional mgv-cycles are needed, when n is increased, in this case.

3 Conclusions.

We have compared multigrid methods with preconditioned conjugate gradient methods for the solution of a couple of partial differential equation modelproblems discretized by finite difference or finite element techniques. In order to obtain a stable and fast preconditioning we have considered modified incomplete factorization, possibly applied to a matrix arising from modification of the element matrices, i.e. we have used the EMIC factorisations decribed in [4], In the multigrid method we have simply used the weighted Jacobi method as smoother.

From these preliminary comparisons one may conclude that the pcg methods are preferable for problems with anisotropic material coefficients if the problem is not very large. For problems with discontinuous material coefficients, on the other hand, the multigrid method turned out to be the most efficient. For a finite element discretization

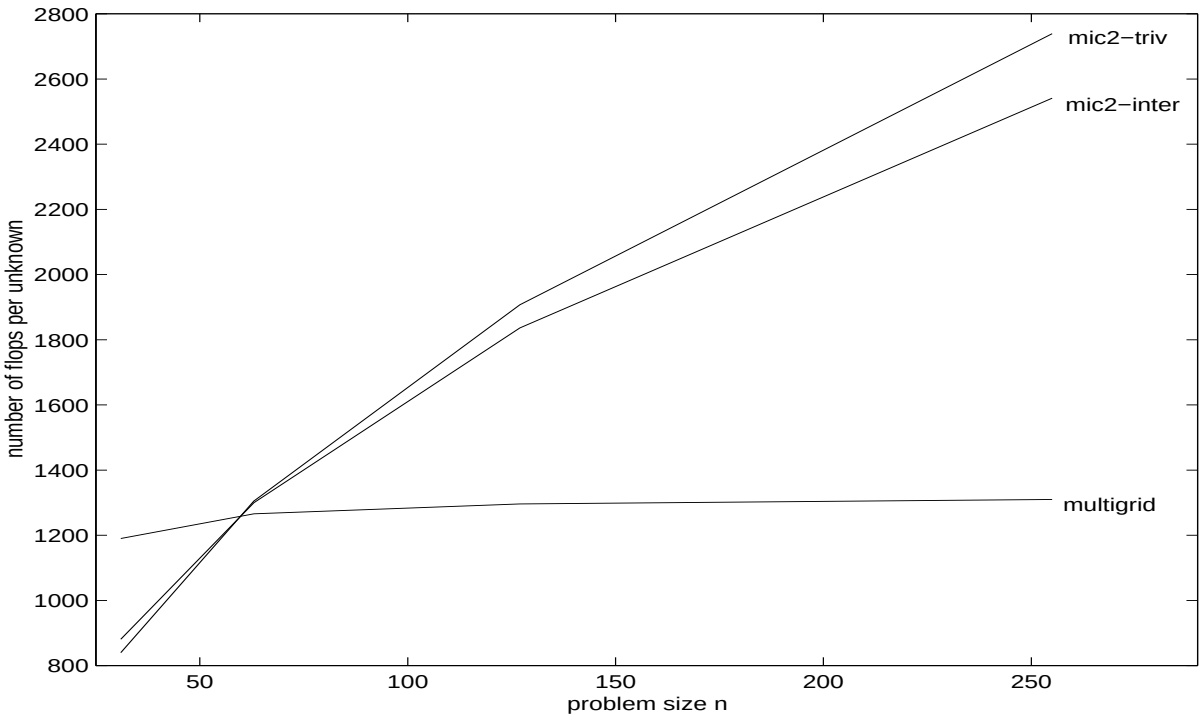

Figure 6: The number of floating point operations per unknown for various methods for the solution of the discontinuous modelproblem with $d = 1000$ and with various values of $n = 1/h - 1$.

with quadratic elements, the EMIC-CG method was faster than the used multigrid method for problems of reasonable size.

We have also demonstrated the positive effect of using a good starting approximation for the pcg method. This starting approximation has been chosen as the interpolant of the solution on a coarser mesh. Hence, by just using two levels of mesh, we have in many cases obtained a faster method than the standard multigrid method, for not unnecessarily large problems.

References

[1] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[2] I. Gustafsson, A class of first order factorization methods, *BIT 18*, 142-156, 1978.

[3] I. Gustafsson, Modified incomplete Cholesky (MIC) methods. In *Preconditioning Methods - Theory and Applications,* D. Evans (ed.), Gordon and Breach Publishers, New York-London-Paris, 262-293, 1983.

[4] I. Gustafsson, An incomplete factorization preconditioned method for FEM-discretized partial differential equations based on modification of element matrices, *BIT 36*, 86-100, 1996.

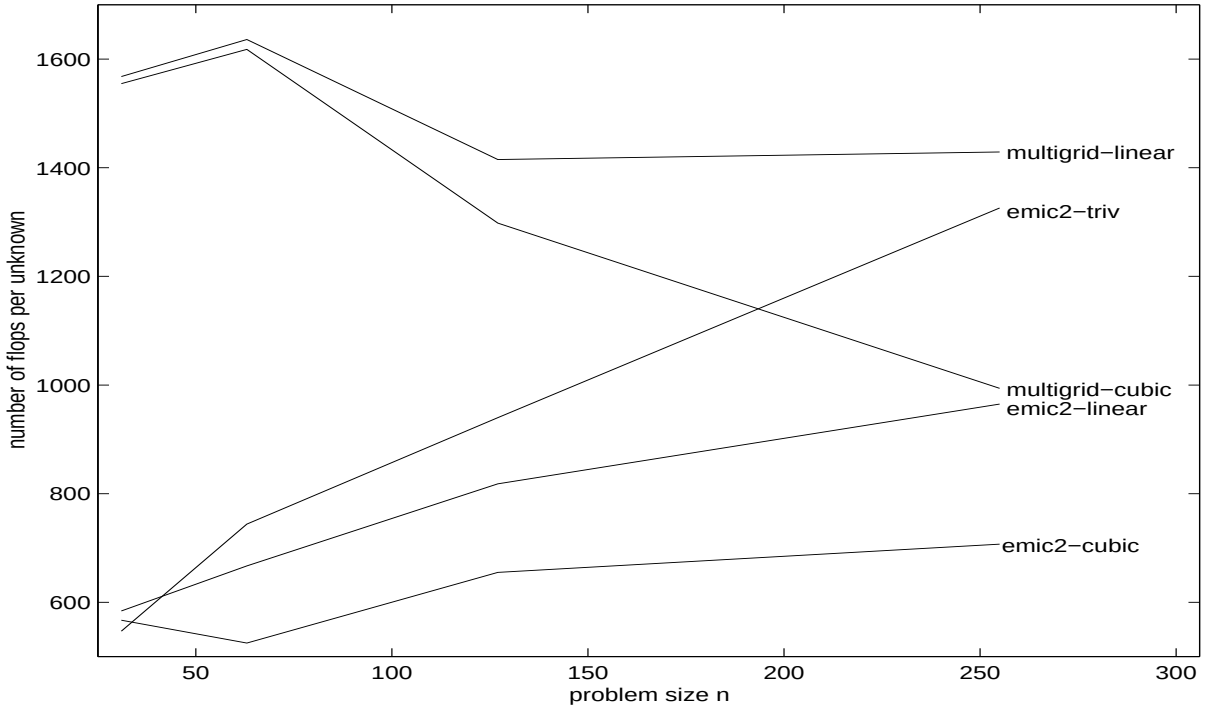

Figure 7: The number of floating point operations per unknown for the compared methods for the solution of the modelproblem, discretized by quadratic finite elements on a uniform rightangled triangulation, for various sizes of $n = 2/h - 1$.

[5] M. T. Heath, *Scientific Computing, An introductory survey*, Mc Graw Hill, 2002.

[6] I. E. Kaporin, High quality preconditioning of a general symmetric positive definite matrix based on its $U^TU + U^TR + R^TU$-decomposition, *Numer. Linear Algebra Appl. 5*, 483-509, 1998.

[7] G. Lindskog and I. Gustafsson, Preconditioning by incomplete factorization; an overview and an application to a linear elasticity problem. Report 2001:78, Department of Mathematics, Chalmers University of Technology, Göteborg, Sweden, 2001.

[8] J. A. Meijerink and H. A. van der Vorst, An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix, *Math. Comp. 31*, 148-162, 1977.