



Improving Multigrid and Conventional Relaxation Algorithms for Propagators

T. Kalkreuter

II. Institut für Theoretische Physik, Universität Hamburg

ISSN 0418-9833

NOTKESTRASSE 85 · D - 2000 HAMBURG 52

DESY behält sich alle Rechte für den Fall der Schutzrechtserteilung und für die wirtschaftliche Verwertung der in diesem Bericht enthaltenen Informationen vor.

DESY reserves all rights for commercial use of information included in this report, especially in case of filing application for or grant of patents.

To be sure that your preprints are promptly included in the
HIGH ENERGY PHYSICS INDEX,
send them to (if possible by air mail):

DESY Bibliothek Notkestraße 85 W-2000 Hamburg 52 Germany	DESY-IfH Bibliothek Platanenallee 6 O-1615 Zeuthen Germany
---	---

Improving Multigrid and Conventional Relaxation Algorithms for Propagators *

Thomas Kalkreuter **

II. Institut für Theoretische Physik der Universität Hamburg,
Luruper Chaussee 149, W-2000 Hamburg 50, Germany

July 16, 1992

Abstract

Practical modifications of deterministic multigrid and conventional relaxation algorithms are discussed. One of these modifications can be thought of as “updating on a last layer consisting of a single site”. It eliminates critical slowing down in computations of bosonic and fermionic propagators in a fixed volume. Here critical slowing down means divergence of asymptotic relaxation times as the propagators approach criticality. A remaining volume dependence is weak enough in case of bosons so that conjugate gradient can be outperformed. However, no answer can be given yet if the same is true for staggered fermions on lattices of realizable sizes. Numerical results are presented for propagators of bosons and of staggered fermions in 4-dimensional $SU(2)$ gauge fields.

*Work supported by Deutsche Forschungsgemeinschaft.

**E-mail: I02KAL@DSYIBM.DESY.DE

1. Introduction

In Monte Carlo simulations of lattice gauge theories with fermions the most time-consuming part is the computation of the gauge field dependent fermion propagators. Conjugate gradient (CG) or minimal residual (MR) algorithms are state of the art [1]. Great hopes to do better are attached to multigrid (MG) methods [2-10]. In Ref. [10] the first MG computations without critical slowing down (CSD) in non-Abelian gauge fields (4-d $SU(2)$) were presented. They prove that *the MG method can cope with the frustration which is inherent in non-Abelian gauge fields*. However, elimination of CSD succeeded only when an “optimal” interpolation kernel [10,11] was used. The use of this optimal kernel for production runs is impractical because of computational complexity and storage space requirements.

In this letter practical modifications of MG and conventional relaxation algorithms for propagators are discussed. A propagator ϕ is the solution of a linear equation

$$D\phi = f \quad (1)$$

on a d dimensional lattice Λ of sites z , for given f . In our case, $D = -\Delta + m^2$ for bosons, and $D = -\not{D} + m^2$ for fermions, where Δ and \not{D} are the gauge covariant Laplace or Dirac operators (with periodic boundary conditions) respectively. Color indices are always suppressed, and $\phi(z)$ is an $N_c \times N_c$ matrix where N_c is the number of colors.

CSD in computations in a fixed volume can be eliminated by “updating on a last layer consisting of a single site”. Given an approximation $\phi^{(n)}$ to ϕ , this updating amounts to rescaling $\phi^{(n)}$ by an $N_c \times N_c$ matrix Ω (in case of bosons or Wilson fermions):

$$\phi^{(n)}(z) \mapsto \phi^{(n)}(z)\Omega \quad , \quad \Omega = (\phi^{(n)}, D\phi^{(n)})^{-1}(\phi^{(n)}, f) \quad , \quad (2)$$

where

$$(\varphi, \psi) \equiv \frac{1}{|\Lambda|} \sum_{z \in \Lambda} \varphi(z)^\dagger \psi(z) \quad . \quad (3)$$

We will discuss Eq. (2) and its generalization for staggered fermions, and related modifications in Sec. 2. Numerical results for propagators of bosons and of staggered fermions in 4-dimensional $SU(2)$ gauge fields will be presented in Sec. 3.

2. Improving algorithms from a variational point of view

2.1. Updating on a 1^d MG layer

It is well known that the CSD of conventional iterative algorithms for solving Eq. (1) depends only on m^2 and *not* on the lattice size $|\Lambda|$. There is only an implicit dependence on $|\Lambda|$ (and β) through the value of m_{cr}^2 , where $-m_{cr}^2$ denotes the lowest eigenvalue of $-\Delta$ or $-\not{D}^2$. The dimension d enters in the scaling relation for relaxation times only through the constant of proportionality. Therefore one continues to have CSD on a lattice of only 2^d sites, and it seems necessary to go to a 1^d lattice in order to eliminate the appearance of CSD.

When we update on a 1^d sublattice, we make the replacement

$$\phi^{(n)}(z) \mapsto \phi^{(n)}(z) + \mathcal{A}(z)(\Omega - \mathbb{1}) \quad , \quad (4)$$

where \mathcal{A} denotes an interpolation kernel, and $\Omega - \mathbb{1}$ is the error of $\phi^{(n)}$ represented at the last site. In the MG context, $\Omega - \mathbb{1} = (C^*, D\mathcal{A})^{-1}(C^*, f - D\phi^{(n)})$, where C^* is the adjoint of the restriction operator which averages to a 1^d layer. From the variational point of view

(VPV) an algorithm is set up in such a way that the functional $K[\phi] = \frac{1}{2} \langle \phi, D\phi \rangle - \langle \phi, f \rangle \equiv \frac{1}{N} \text{Tr} \left[\frac{1}{2}(\phi, D\phi) - (\phi, f) \right]$ is lowered as far as possible in every iteration. This leads to $C^* = \mathcal{A}$. From the VPV the optimal \mathcal{A} in (4) equals $\phi^{(n)}$. Thus, we obtain Eq. (2). Considered from the VPV alone without thinking of MG, one obtains (2) by rescaling $\phi^{(n)}$ with a matrix Ω that is determined such that $K[\phi^{(n)}\Omega]$ is as low as possible.

In case of staggered fermions we have to consider that there are 2^d different pseudoflavors [7]. In the limiting case of a pure gauge the fermionic problem amounts to computing 2^d decoupled bosonic propagators. Hence for staggered fermions we replace (2) by

$$\phi^{(n)}(z) \mapsto \phi^{(n)}(z) \Omega(H(z)) , \quad (5)$$

where $H(z)$ denotes the pseudoflavor of z . Now the expression for $\Omega(H)$ is more complicated than that given in (2). In practice, we determine the $\Omega(H)$'s by solving one linear $N_c 2^d \times N_c 2^d$ system, or – making use of the independence of the even and odd sublattices – by solving two systems of a quarter of that size.

2.2. Related improvements from a VPV

Some related modifications of (MG) relaxation algorithms will be discussed now.

2.2.1. Modified MG correction updating step. In conventional MG approaches one considers updates of the form $\phi^{(n)} \mapsto \phi^{(n)} + \varphi^{(n)}$ where $\varphi^{(n)}$ is obtained by interpolation of an approximate solution of a residual equation on a coarser lattice. We propose to generalize this to

$$\phi^{(n)}(z) \mapsto \tilde{\phi}^{(n)} \equiv \phi^{(n)}(z) \Omega + \varphi^{(n)}(z) \Theta . \quad (6)$$

The two matrices Ω and Θ are chosen such that $K[\tilde{\phi}^{(n)}]$ is minimized. In particular, this proposal may be an improvement in algorithms where the residual equation is only solved approximately, or in algorithms where coarse grid operators are not defined through the Galerkin prescription. For staggered fermions Ω and Θ in (6) become pseudoflavor dependent.

2.2.2. Modified checkerboard SOR. Consider for illustration the bosonic problem. When we update at the even sites, we propose to modify SOR according to

$$\begin{aligned} \phi^{(n)}(z) &\mapsto \phi^{(n)}(z) \Omega + \varphi^{(n)}(z) \Theta & \text{if } z \text{ is even} , \\ \phi^{(n)}(z) &\mapsto \phi^{(n)}(z) \Xi & \text{if } z \text{ is odd} , \end{aligned} \quad (7)$$

where $\varphi^{(n)}(z) = (2d + m^2)^{-1} [f(z) + \sum_{z' \sim z} U(z, z') \phi^{(n)}(z')]$, and $U(z, z')$ is the gauge field on the link (z, z') . Again, the matrices Ω , Θ and Ξ should be chosen such that the functional K gets minimized. The proposal (7) expresses the view that in gauge theories one should have relaxation matrices rather than relaxation parameters that are real numbers.

2.2.3. Modified damped Jacobi relaxation. Damped Jacobi relaxation can be generalized according to (6) with $\varphi^{(n)} = (2d + m^2)^{-1} \tau^{(n)}$. If one fixes $\Omega = \mathbb{1}$, one recovers the MR algorithm that was used by Hulsebos et al. [5,8,9].

Finally we note that the proposals (2), (5), (6) and (7) respect gauge covariance. Iterative algorithms are gauge covariant in the sense that all $\phi^{(n)}$ are gauge transformed by g if g is applied before relaxation is started. Ω is gauge invariant (or transforms like a matter field sitting at site w in the adjoint representation, i. e. $\Omega \mapsto g_w \Omega g_w^{-1}$, when $f \mapsto \delta_{z,w}$) etc.

3. Results for propagators in 4-dimensional $SU(2)$ gauge fields

It is well known [1] that the convergence rate of iterative algorithms for propagators is governed by the condition number¹⁾ of $(-\Delta + m^2)$ or $(-\not{D}^2 + m^2)$. Therefore the CSD scaling relation for relaxation times τ reads

$$\tau \propto (\Delta m^2)^{-z/2} \text{ for small } \Delta m^2 = m^2 - m_{c*}^2 , \quad (8)$$

where z denotes the critical exponent. τ is defined by the *asymptotic* exponential decay of the norm of the residual. We measure all τ 's and number of iterations in units of cycles which involve only one sweep through the finest lattice. The variational MG method used for solving Eq. (1) is described in some detail in Refs. [7,10,11]. Its implementation is actually a twogrid algorithm where the residual equation is solved exactly by CG. A scale factor of 3 is chosen in blocking. We use the gauge covariant ground-state projection MG method. An efficient algorithm for computing averaging kernels C was described in Ref. [12], and was used in this work. The (non-optimized) implementations of the MG programs require a factor of 3.2/7.8 (bosons/staggered fermions) more arithmetic operations than CG when updating on a 1^d sublattice is included. Without that inclusion the factor is 2.1/4.5. The amount of work for C on the whole lattice is equivalent to less than 20 CG iterations.

3.1. Propagators in pure gauges

In pure gauges the bosonic and fermionic problems are equivalent. There MG does a perfect job, CSD is completely eliminated with short τ 's and $z = 0$, Ref. [10]. Rescaling (2) does not improve the performance any further. The improved correction scheme (6), however, is able to halve τ and to bring it close to 1. Combining (2) with conventional 1-grid relaxation causes CSD (in the sense stated above) to disappear. This is shown in Fig. 1. The norm of the residual is not monotonically decreased for small m^2 . This feature also shows up in CG which minimizes K (i. e. the residual r in the norm induced by the scalar product $\langle \cdot, D^{-1} \cdot \rangle$) rather than $\|r\|$ itself. This is an interesting point: Instead of (2), one could think of rescaling $\phi^{(n)}$ by another matrix Ω' which is chosen such that $\|r^{(n)}\|$ is minimized. However, in this case there is no difference to 1-grid relaxation. This remains true in nontrivial gauge fields. Finally we note that practically $\Omega = \mathbb{1}$ as soon as $\|r\|$ decays exponentially. Then the step (2) could be switched off. This statement holds also in nontrivial gauge fields and for MG.

3.2. Bosonic propagators in nontrivial gauge fields

The remarks made above about 1-grid relaxation plus (2) apply in nontrivial gauge fields as well. Fig. 1 looks the same when m^2 is replaced by Δm^2 . MG plus (2) beats CSD, too. Compared with CG and 1-grid plus (2), this modified MG performs better the more critical the system is. Convergence for $\Delta m^2 = 10^{-6}$ on 12^4 and 18^4 lattices is shown in Fig. 2.²⁾ It must be emphasized that Ω really needs to be a matrix. Simple rescaling of $\phi^{(n)}$ with a real number does not succeed in eliminating CSD.³⁾ The improved MG correction scheme (6) also beats CSD. However, at finite β it is not able to halve τ 's. Using the modified SOR version (7) does not pay. If one fixes $\Xi = \mathbb{1}$, one has nothing else but conventional Gauss-Seidel relaxation. Retaining Ξ yields little difference in performance.

¹⁾ i. e. the ratio of the largest to the smallest eigenvalue

²⁾ Using SOR as a smoother contradicts the conventional MG wisdom. However, in nontrivial gauge fields any over-relaxation yields better performance than Gauss-Seidel and much better performance than damped Jacobi or MR. A similar result was reported in a recent paper on MG gauge fixing [13].

³⁾ A remark about 1-grid plus (2) is in order here: It can be proved by induction that in $SU(2)$ gauge fields $\Omega \propto \mathbb{1}$ always holds if one starts relaxation with $\phi^{(0)} = 0$. However, $\Omega \propto \mathbb{1}$ is not true in general.

3.3. Propagators of staggered fermions in nontrivial gauge fields

In case of staggered fermions m_c^2 is much closer to zero than in case of bosons. Actually, it is often assumed that the finite value of m_c^2 can be neglected completely in (8) so that $\tau \propto m^{-2}$. However, this neglect of m_c^2 is not justified on lattices amenable in size to date. For bosons, scaling (8) without violations is observed for $\Delta m^2 \lesssim 0.01$, Ref. [10]. For staggered fermions it was verified that the scaling relation (8) also holds. The critical exponent z equals 2 for conventional relaxation (fixed ω) and for MG without (5). It was also verified that the constant of proportionality in (8) is independent of the lattice size. To obtain these results requires a careful analysis of data. The subtle point is that for fermions (8) is obeyed only for $\Delta m^2 \lesssim 0.001$, and asymptotic decay in the sense that only the slowest mode governs convergence does not set in before 400 – 500 iterations in 4-dimensional $SU(2)$ gauge fields.

The lesson is that m_c^2 must not be neglected, at least for lattices up to 18^4 . In practice this might mean that on such relatively small lattices m^2 must also be given small negative values in order to study effects of CSD. This is of course artificial, but it is necessary when one wants to obtain results for z which are reliable for predictions how algorithms perform on large lattices. If one does not investigate systems close enough to criticality and if one does not take care that decay rates are really asymptotic, there is the danger of extracting wrong values for z , even in case of pure gauges.

If one is interested in the question how well algorithms perform on lattices of sizes that are used in present day simulations, the foregoing discussion might appear too academic. It might appear more natural to ask how many iterations are needed to obtain a given accuracy. From this viewpoint it turned out that MG is competitive or even superior to conventional algorithms in 2-d models [3,6]. However, it is more difficult to reach decisive conclusions in $d = 4$. Preliminary results on 16^4 [9] and 18^4 [14] lattices (both at $\beta = 2.7$) indicated that the MG methods tested so far will not be able to outperform CG. But we expect that the situation will be different on larger lattices. Details will be reported elsewhere [15].

We carry on with results of the modifications proposed in this article. Table 1 gives a survey of convergence on 12^4 and 18^4 lattices. Rescaling (5) does not pay for positive m^2 on small lattices. But including (5) in algorithms brings z down to zero, i. e. CSD is eliminated. An analogue of Fig. 2 for staggered fermions is given with Fig. 3. Mind, however, that relaxation algorithms for fermions are used with lexicographic, not checkerboard, updating.

We conclude this section with remarks on MR. Since MR can be viewed as an optimized Jacobi relaxation, it comes as no surprise that its convergence properties are worse than those of SOR. Working with pseudoflavor dependent matrices leads to no practical improvement. Over-relaxed MR versions have not been investigated yet.

4. Conclusions

Updating on a last 1^d MG layer provides an astonishingly simple modification which eliminates CSD of asymptotic relaxation times in MG and even in 1-grid relaxation algorithms for propagators. As soon as the decay of the error is exponential, updating on the 1^d sublattice can be switched off. Therefore additional work must only be invested in the initial and in an intermediate stage of computations. Since the modified algorithms have $z = 0$, we expect that CG will eventually be outperformed. What remains, however, is a volume dependence on how fast the asymptotic regime is reached. For bosons it was shown that CG is outperformed. But we feel unable to predict whether the same methods will pay for staggered fermions on lattices of realizable sizes.

ACKNOWLEDGMENTS

The investigation of “updating on a last 1^d MG layer” was motivated by the observation of Gerhard Mack that this eliminates CSD in a simple toy model. I am indebted to him for stimulating discussions. I would also like to thank S. Meyer for his interest in this work and for discussions. Financial support by Deutsche Forschungsgemeinschaft is gratefully acknowledged. For providing resources, advice and help I wish to thank HLRZ Jülich and its staff.

References

- [1] C.B. Chalmers, R.D. Kenway and D. Roweth, J. Comp. Phys. 70 (1987) 500;
P.B. Mackenzie, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 103;
D. Henty, R. Setoodeh and C.T.H. Davies, Nucl. Phys. B337 (1990) 487
- [2] R.C. Brower, K.J.M. Moriarty, E. Myers and C. Rebbi, in: Multigrid Methods, ed. S.F. McCormick (Marcel Dekker, New York, 1988)
- [3] R. Ben-Av, A. Brandt and S. Solomon, Nucl. Phys. B329 (1990) 193;
R. Ben-Av, A. Brandt, M. Haratz, E. Katznelson, P.G. Lauwers, S. Solomon and K. Wolowesky, Phys. Lett. B253 (1991) 185; Nucl. Phys. B (Proc. Suppl.) 20 (1991) 102;
R. Ben-Av, P.G. Lauwers and S. Solomon, Nucl. Phys. B 374 (1992) 249;
P.G. Lauwers and S. Solomon, Int. J. Mod. Phys. C3 (1992) 149
- [4] R.C. Brower, C. Rebbi and E. Vicari, Phys. Rev. D43 (1991) 1965; Phys. Rev. Lett. 66 (1991) 1263;
R.C. Brower, K.J.M. Moriarty, C. Rebbi and E. Vicari, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 89; Phys. Rev. D43 (1991) 1974
- [5] A. Hulsebos, J. Smit and J.C. Vink, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 94; Int. J. Mod. Phys. C3 (1992) 161; Nucl. Phys. B368 (1992) 379
- [6] R.C. Brower, R.G. Edwards, C. Rebbi and E. Vicari, Nucl. Phys. B366 (1991) 689
- [7] T. Kalkreuter, G. Mack and M. Speh, Int. J. Mod. Phys. C3 (1992) 121
- [8] J.C. Vink, Phys. Lett. B272 (1991) 81
- [9] J.C. Vink, Nucl. Phys. B (Proc. Suppl.) 26 (1992) 607
- [10] T. Kalkreuter, Phys. Lett. B276 (1992) 485
- [11] G. Mack, T. Kalkreuter, G. Palma and M. Speh, *Effective Field Theories*, preprint DESY 92-070, to appear in the proceedings of the 31st IUKT, Schladming, February 1992
- [12] T. Kalkreuter, Nucl. Phys. B376 (1992) 637
- [13] A. Hulsebos, M.L. Laursen and J. Smit, *SU(N) Multigrid Landau Gauge Fixing*, KFA Jülich preprint HLRZ-92-15, May 1992
- [14] T. Kalkreuter, *Multigrid for staggered fermions in 4-dimensional SU(2) gauge fields*, talk presented at the DFG-Colloquium “Dynamical Fermions” held in Leipzig, March 1992
- [15] T. Kalkreuter, *Ph.D. thesis*, in preparation

Table

Table 1 Convergence in computations of propagators of staggered fermions with $f(z) = \delta_{z,0}$ in 4- d $SU(2)$ gauge fields. Given is the number of iterations necessary for reducing $\ln||\tau^{(0)}||$ by 10. All propagators are initialized with zero. 1-grid and MG SOR are swept in lexicographic ordering.

pure gauge configurations ($\beta = \infty$)						
algorithm and lattice size	$m^2 =$					
	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
CG on 12^4	17	17	17	17	17	17
CG on 18^4	28	30	33	35	37	39
1-grid SOR plus (5), $\omega = 1.90$, on 12^4	105	125	155	195	230	270
1-grid SOR plus (5), $\omega = 1.90$, on 18^4	110	125	155	195	230	270
MG SOR without (5), $\omega = 1.09$, on 12^4	21	23	23	23	23	23
MG SOR without (5), $\omega = 1.09$, on 18^4	20	23	23	23	23	23
nontrivial gauge fields ($\beta = 2.7$)						
algorithm and lattice size	$\Delta m^2 =$					
	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
CG on 12^4	65	175	265	300	320	340
CG on 18^4	65	180	350	415	455	495
1-grid SOR plus (5), $\omega = 1.90$, on 12^4	100	130	530	560	630	710
1-grid SOR plus (5), $\omega = 1.90$, on 18^4	95	125	710	1090	1320	1540
MG SOR plus (5), $\omega = 1.96$, on 12^4	180	185	385	425	475	535
MG SOR plus (5), $\omega = 1.96$, on 18^4	180	185	530	775	950	1070

Figure captions

Fig. 1 1-grid SOR plus (2) eliminates CSD, shown here for a pure gauge. ($\omega = 1.90$, checkerboard updating) The 6 curves correspond to $m^2 = 0.1, 0.01, \dots, 10^{-6}$ on a 12^4 (non-staggered) lattice with m decreasing from left to right. A small volume effect remains, not for τ (i. e. the asymptotic decay rate) but with respect to how fast the asymptotic regime is reached: e. g. on an 18^4 lattice the number of iterations needed to obtain a given accuracy of $||\tau||$ for $m^2 = 10^{-6}$ is increased by ~ 20 . The figure looks the same for bosons in nontrivial gauge fields when m^2 is replaced by Δm^2 .

Fig. 2 Convergence for bosonic propagators with $\Delta m^2 = 10^{-6}$ in quenched 4- d $SU(2)$ gauge fields equilibrated with Wilson's action at $\beta = 2.7$. The numbers refer to the following algorithms: 1/2: variational MG SOR ($\omega = 1.50$) plus (2) on a $12^4/18^4$ lattice; 3/4: 1-grid SOR ($\omega = 1.91$) plus (2) on a $12^4/18^4$ lattice; 5/6: CG on a $12^4/18^4$ lattice. Relaxation algorithms are swept in checkerboard fashion. The critical masses are $m_{cr}^2 = -0.7726281/-0.7554339$. Without (2) MG SOR and 1-grid SOR have τ 's of $O(10^5)$ [10].

Fig. 3 Convergence for propagators of staggered fermions with $\Delta m^2 = 10^{-6}$ in quenched 4- d $SU(2)$ gauge fields at $\beta = 2.7$. The numbers refer to the following algorithms: 1/2: variational MG SOR ($\omega = 1.96$) plus (5) on a $12^4/18^4$ lattice; 3/4: 1-grid SOR ($\omega = 1.90$) plus (5) on a $12^4/18^4$ lattice; 5/6: CG on a $12^4/18^4$ lattice. Relaxation algorithms are swept in lexicographic ordering. The critical masses are $m_{cr}^2 = -0.0368447/-0.0096640$. Without (5) MG SOR and 1-grid SOR have τ 's of $O(10^5)$.

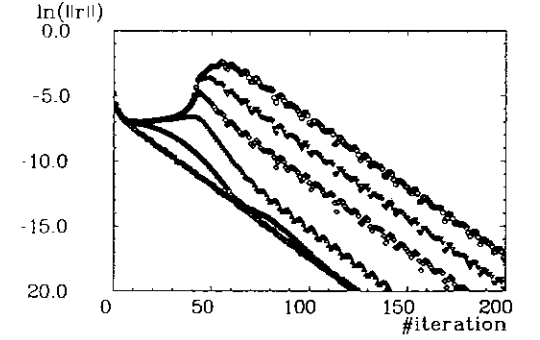


Fig. 1

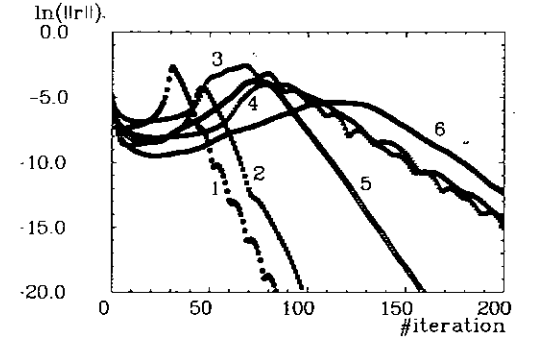


Fig. 2

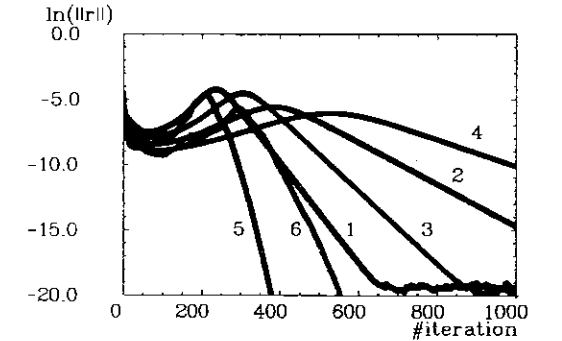


Fig. 3