Proceedings of ALGORITMY 2000
Conference on Scientific Computing, pp. 264–272

STABILITY OF POLYNOMIAL PRECONDITIONING

Y. LIANG, J. WESTON, M. SZULARZ *

Abstract. Polynomial preconditioners are frequently used in a parallel environment for the computation of the solution of large-scale sparse linear equations ($Ax = b$) because of their easy implementation and trivial parallelization. It is well known that the construction of the preconditioning polynomial can be transformed into a *constrained optimization problem*, namely that of finding an m-degree polynomial in the matrix $A, P_m(A)$, such that $P_m(A) \approx A^{-1}$. Three typical polynomial preconditioners arise, *Neumann-series, Least-squares* and *Chebyshev*. Theoretically, the higher the degree of the polynomial preconditioner used, the closer is the resemblance to A^{-1}. However, high-degree polynomial preconditioning may lead to rapidly propagating rounding errors and consequently worsen the stability of the preconditioned system. As a result, a trade-off is required. This paper is concerned with the stability of the preconditioning operation. Experimental results using a highly parallel machine environment (MPI on IBM SP2) are presented and a related analysis is also included.

Key words. Polynomial preconditioner, linear equations, parallel, perturbation, GMRES.

1. Introduction. The interest in *polynomial preconditioners* [1, 3] is motivated by the need for simple, yet efficient, methods for the speedup of iterative solvers [5, 10] on vector and parallel processors.

Generally the objectives of preconditioning are to improve the stability of the systems of linear equations and to provide efficient solution procedures. It is well known that the *condition number* [5, 10, 12] is an important parameter in the evaluation of linear systems. Consider the polynomial preconditioned linear equations with form

$$P_m(A)Ax = P_m(A)b \tag{1}$$

where $P_m(A)$ is a polynomial in A with degree no more than m (It should be noted that there is no difference between left and right strategies with respect to polynomial preconditioning). $P_m(\lambda)$ may be constructed so that the condition number $\kappa(P_m(A)A)$ is minimised. When constructing a suitable $P_m(A)$, several issues must be considered:

1. Estimation of the spectrum of A and the degree of $P_m(A)$.
2. The additional cost (computation, communication and storage) brought about by the preconditioning operation.
3. The perturbation of the preconditioning operation. In this paper the degree and coefficients of $P_m(A)$, two important factors which determine the stability of preconditioning, will be discussed.

The main topic of this paper is the stability of the preconditioning operation itself. The main focus of attention is the rounding error analysis associated with preconditioning.

*School of Information and Software Engineering, University of Ulster at Coleraine, BT52 1SA, U.K., e-mails¿ y.liang@ulst.ac.uk, jsc.weston@ulst.ac.uk, mszularz@ulst.ac.uk

Let $\sigma(A)$ be the spectrum of A, i.e., the set of all eigenvalues [9] of A. When A is *positive definite*, $\sigma(A)$ is contained by a single interval $\Omega = [a, b]$ where $0 \leq a < b$; when A is *indefinite*, $\sigma(A)$ is contained by two disjoint intervals $\Omega = [a, b] \cup [c, d]$ where $a < b \leq 0 \leq c < d$. For the sake of simplicity and practical application, this paper is concerned with polynomial preconditioning for SPD (symmetric positive definite) linear systems $Ax = b$. Further, because of its broader application, the GMRES (Generalised Minimised RESidual) method is used instead of the more efficient CG (Conjugate Gradient) method which may be used in the case of SPD linear systems.

Experimental results are based on benchmark linear systems taken from the Harwell-Boeing collection [7].

2. Typical Polynomial Preconditioners.

2.1. Derivation of Polynomial Preconditioners. Assume that $\hat{P}_m(A)$ is the appropriate preconditioning polynomial in A. It should satisfy

$$\min_{\hat{P}_m \in \wp_m} \kappa(\hat{P}_m(A)A) \tag{2}$$

where $\wp_m$ denotes the set of all polynomials of degree not exceeding m and $\kappa(A)$ denotes the *condition number* [10, 5] of the matrix A. Thus, $\hat{P}_m(\lambda)$ should satisfy

$$\lambda \hat{P}_m(\lambda) \approx \delta, \quad \hat{P}_m \in \wp_m \tag{3}$$

where $\delta \in \Re$ [3]. By assuming $P_m(\lambda) = \hat{P}_m(\lambda)/\delta$ (i.e., scaling operation), the canonical form for the construction of the preconditioning polynomial is given by

$$\min_{P_m \in \wp_m} \|1 - \lambda P_m(\lambda)\| \quad (\lambda \in \Omega) \tag{4}$$

where $\|.\|$ represents a specific norm (i.e., uniform norm) and Ω will usually be estimated. The accuracy of estimation determines the convergence speed of the procedure used to solve preconditioned system. In [6] it is shown that, if $\epsilon_{m+1} = \min_{\lambda \in \Omega, P_m \in \wp_m} \|1 - \lambda P_m(\lambda)\|$, $\kappa(P_m(A)A) \leq \frac{1+\epsilon_{m+1}}{1-\epsilon_{m+1}}$.

Clearly, the construction of a preconditioning polynomial for a system of linear equations is just a polynomial approximation problem. Neumann-series [3, 10] is the simplest method, while Least-squares [3, 10] and Chebyshev [3, 10] are optimum approximation methods with respect to the *quadratic norm* [4] and the *uniform norm* (or *minmax norm*) [6], respectively (Note:optimum approximation methods do not necessarily guarantee the fastest rate of convergence).

2.2. Neumann Series Preconditioning. The *Neumann polynomial preconditioner* [3, 10] is the simplest polynomial preconditioner and it originates from the basic algebraic relation

$$\forall \mid \lambda \mid < 1 (\lambda \in \Re), \frac{1}{1-\lambda} = \sum_{i=0}^{\infty} \lambda^i. \tag{5}$$

The following theorem may be used in the construction of a polynomial approximation for A^{-1}.

THEOREM 2.1. *Let $G \in \Re^{n,n}$, then $\sum_{k=0}^{\infty} G^k$ converges if and only if $\rho(G) < 1$. In addition, $\sum_{k=0}^{\infty} G^k = (I - G)^{-1}$.*

Thus, given any coefficient matrix A, trivially $\omega A \equiv I - (I - \omega A)$ (ω is scaling scalar). Let $G = (I - \omega A)$, then $A^{-1} = \omega(I - G)^{-1}$ where ω can be adjusted so that $\rho(G) < 1$ (ρ indicates the spectral radius of the matrix). It follows that

$$A^{-1} = \omega(I - G)^{-1} = \omega(\sum_{i=0}^{\infty} G^i) \approx \omega(I + G + G^2 + \cdots + G^{m-1}). \tag{6}$$

Consequently $P(A) = \omega(I+G+G^2+\cdots+G^{m-1})$ may be regarded as the required preconditioner. In practical implementations, ω should be determined according to the characteristics of A. For example, if A is *SPD*, we can select $\omega = \lambda_{max}$, where λ_{max} is an estimation for the largest eigenvalue.

2.3. Least Squares Preconditioning Polynomial. The lease-squares preconditioning polynomial [3, 10] is derived from an optimum approximation method based on the *quadratic norm* [4]

$$\min_{P_m \in \wp_m, \lambda \in \Omega} \|1 - \lambda P_m(\lambda)\|_w. \tag{7}$$

According to the definition of the quadratic norm, (7) is transformed into

$$\min_{P_m \in \wp_m} \int_{\Omega} w(\lambda)[1 - \lambda P_m(\lambda)]^2 d\lambda. \tag{8}$$

where $w(\lambda)$ is a non-negative weight function on the interval Ω.

With respect to an appropriate weight function, the least-squares preconditioning polynomial series can be generated via a *three-term recurrence relationship* of orthonormal polynomials. Table 1 lists the low-degree least-squares preconditioning polynomial series which is induced from the *Jacobi weight function*

$$w(\lambda) = \lambda^{\mu-1}(1-\lambda)^{\nu}, \mu > 0 \ \mathit{and} \ \nu \geq \frac{1}{2}. \tag{9}$$

with $\mu = \frac{1}{2}$ and $\nu = -\frac{1}{2}$. The polynomials listed are for the interval $[0, 4]$ as this yields integer coefficients which are not too large. For a general interval $[0, \beta]$, the best polynomial of degree k is $P_k(4\lambda/\beta)$.

Degree	Least-square polynomial
$P_1(\lambda)$	$5 - \lambda$
$P_2(\lambda)$	$14 - 7\lambda + \lambda^2$
$P_3(\lambda)$	$30 - 27\lambda + 9\lambda^2 - \lambda^3$
$P_4(\lambda)$	$55 - 77\lambda + 44\lambda^2 - 11\lambda^3 + \lambda^4$

TABLE 1
Least-square Preconditioning Polynomials

2.4. Chebyshev Preconditioning Polynomial. The Chebyshev preconditioning polynomial [10] is derived using a minmax polynomial approximation method based on the *uniform norm*

$$\min_{P_m \in \wp_m} \max_{\lambda \in \Omega} \mid 1 - \lambda P_m(\lambda) \mid . \tag{10}$$

ALGORITHM 2.1. The construction of the Chebyshev Preconditioning Polynomial (for SPD matrix)

$$\begin{cases} \sigma_0 = 1, \sigma_1 = \frac{\theta}{\delta}, \sigma_{k+1} = 2\frac{\theta}{\delta}\sigma_k - \sigma_{k-1}; \\ P_0(\lambda) = \frac{1}{\theta}, P_1(\lambda) = \frac{4\theta-2\lambda}{2\theta^2-\delta^2}, \\ P_k(\lambda) = \frac{2\sigma_k}{\delta\sigma_{k+1}} + \frac{2\sigma_k(\theta-\lambda)}{\sigma_{k+1}\delta}P_{k-1}(\lambda) - \frac{\sigma_{k-1}}{\sigma_{k+1}}P_{k-2}(\lambda). \end{cases} \tag{11}$$

where δ and θ are the centre and mid-width of the estimation of the $\sigma(A)$ (A is SPD).

Algorithm 2.1 provides the formula for Chebyshev preconditioning polynomials. It is derived using a 3-term recurrence relationship. Low-degree Chebyshev preconditioning polynomials may be easily constructed using Algorithm 2.1. Table 2 illustrates the Chebyshev preconditioning polynomials with degree 0 to 4. When using predetermined preconditioning polynomials, the cost of construction may be neglected.

$P_0(\lambda)$	$\frac{1}{\theta}$
$P_1(\lambda)$	$\frac{4\theta-2\lambda}{2\theta^2-\delta^2}$
$P_2(\lambda)$	$\frac{1}{\theta} + \frac{4(\lambda-\theta)(\lambda-2\theta)}{4\theta^3-3\theta\delta^2}$
$P_3(\lambda)$	$\frac{8(2\theta-\lambda)[\theta^2-\delta^2+(\theta-\lambda)^2]}{8\theta^4-8\theta^2\delta^2+\delta^4}$
$P_4(\lambda)$	$\frac{[16\lambda^4-80\theta\lambda^3+(160\theta^2-20\delta^2)\lambda^2+(60\theta\delta^2-160\theta^3)\lambda+(80\theta^4-60\theta^2\delta^2+5\delta^4)]}{16\theta^5-20\theta^3\delta^2+5\theta\delta^4}$

TABLE 2

Chebyshev Preconditioning Polynomials ($P_0(\lambda)$ through $P_4(\lambda)$)

2.5. Polynomial Preconditioned GMRES. The GMRES (Generalised Minimal REsidual method) [10, 2, 3], is applicable to general nonsymmetric matrices. However, in order to avoid the difficulties associated with estimating the spectum of such matrices [9], this paper is concerned with SPD systems only. GMRES is finally attributed to the following least-squares problem [4]

$$\min_{x_s=x_0+\mathcal{K}(A,r_0)} \|b - Ax_s\|_2 \tag{12}$$

where

$$\mathcal{K}_s(A, r_0) \equiv span\{r_0, Ar_0, A^2r_0, \cdots, A^{s-1}r_0\}. \tag{13}$$

Here $\mathcal{K}_s(A, r_0)$ ($r_0 = b - Ax_0$ and x_0 is the initial guess of solution) is called the s-dimensional *Krylov subspace* [10]. GMRES [2] solves (12) by constructing an orthonormal basis $\{v_1, v_2, \cdots, v_s\}$ for $\mathcal{K}_s(A, r_0)$ using the *Arnoldi algorithm* [5, 10] and then finding the optimum solution based on the above basis.

The main drawback to GMRES is the increasing computational/memory cost as the number of iterations increases [2, 10]. In this paper two remedies will be employed. The first is to *restart* algorithm periodically [10]. For example, Algorithm 2.2 restart every s steps. The second is preconditioning. Note that the use of the Arnoldi algorithm to construct orthonormal bases for iterative methods for the solution of the eigenvalue problem is also the most costly operation in such methods. Observe also that, with little extra cost, an estimation for $\sigma(A)$ may be obtained using the GMRES algorithm.

ALGORITHM 2.2. *Adaptively Polynomial Preconditioned ($P_m(A)$) Restarted GMRES*

```
(1)  Start: choose x_0, s (dimension of the Krylov subspace),
            m (degree of P_m) and Ω(estimation about σ(A));
(2)  Arnoldi process:
(3)       r_0 = b - A x_0; β = ||r_0||_2; v_1 = r_0/β;
(4)       FOR j = 1, ···, s DO
(5)            z_j = P_m(A) v_j where P_m is constructed according to Ω;
(6)            w = A z_j
(7)            FOR i = 1, ..., j DO
(8)                 h_{i,j} = (w, v_i);
(9)                 w = w - h_{i,j} v_i;
(10)           ENDFOR
(11)           h_{j+1,j} = ||w||_2; v_{j+1} = w/h_{j+1,j};
(12)      ENDFOR
(13)      Define Z_s = [z_i, ···, z_s] and Ĥ_m = {h_{i,j}} (1 ≤ i ≤ j+1, 1 ≤ j ≤ s),
          Compute the extremal Ritz value so as to adjust Ω;
(14) Achieve the approximate solution:
(15)      x_s = x_0 + Z_s y_s such that y_s = argmin_y ||β e_1 - Ĥ_s y||_2;
(16) Restart: IF (not convergent) x_0 = x_s and then GOTO (2)
```

In this paper GMRES-Neum, GMRES-LS and GMRES-Cheby represent Neumann series, Least-squares and Chebyshev polynomial preconditioning GMRES respectively. The experimental results show that, in the case of low degree (≤ 5), the convergence performance of GMRES-Cheby is generally better than that of GMRES-Neum and GMRES-LS.

3. Perturbation of Polynomial Preconditioning. *Perturbation analysis* [11] is an important issue in scientific and engineering computation. Firstly, there may be errors in the input data, caused by prior calculations or perhaps measurement errors. Secondly, there are errors caused by the algorithm itself, or the approximations made within the algorithm. In order to estimate the errors in the computation from both these sources, we need to understand how much the solution of a problem is changed if the input data is slightly perturbed [11].

Algorithm 2.2 shows that preconditioning operation is

$$P_m(A)v = (\sum_{k=0}^{m} \alpha_k A^k)v, \tag{14}$$

which is actually considered as a *polynomial evaluation problem* [5]. With respect to the characteristics of matrix computation, (14) is implemented using *Horner's rule* [5]. The corresponding algorithm is given below.

ALGORITHM 3.1. $z \leftarrow P_m(A)v$ *($P_m = \sum_{k=0}^{m} \alpha_k A^k$)*

```
z^{(m)} = α_m v
for k = m - 1 downto 0
     z^{(k)} = A z^{(k+1)} + α_k v;
endfor
```

In Algorithm 3.1, $z^{(k)}$ represents the value of z at $(m-k)$th iteration step, thus $z = z^{(0)} = P_m(A)v$.

Only the rounding errors [11] associated with matrix-vector products need be considered, since the errors associated with perturbations in v and A and the rounding errors of other operations are relatively insignificant. Let $(Av)_{fl}$ be the output of the matrix-vector product Av, then

$$(Av)_{fl} = Av + \delta(Av). \tag{15}$$

Define the diagonal *modification matrix* $\Gamma \in R^{n,n}$, whose entries satisfy $0 \leq \gamma_{l,l} \leq \varepsilon$ where ε is the machine precision (roundoff unit). Then (15) yields:

$$(Av)_{fl} = (I + \Gamma)Av, \tag{16}$$

which may be used to obtain the accumulated error of Algorithm 3.1. This approach is simpler and more optimistic than that of Wilkinson [11].

For the first iteration:

$$z_{fl}^{(m)} = \alpha_m v.$$

Due to the influence of rounding errors,

$$z_{fl}^{(m-1)} = \alpha_m (Av)_{fl} + \alpha_{m-1} v = (I + \Gamma_{m-1}^{(1)})\alpha_m Av + \alpha_{m-1} v$$

Similarly,

$$z_{fl}^{(m-2)} = (I + \Gamma_{m-2}^{(1)})(I + \Gamma_{m-1}^{(1)})\alpha_m A^2 v + (I + \Gamma_{m-2}^{(2)})\alpha_{m-1} Av + \alpha_{m-2} v. \tag{17}$$

By induction, it follows that

$$z_{fl}^{(0)} = \sum_{i=0}^{m} [\prod_{k=1}^{i} (I + \Gamma_k^{(m-i)})]\alpha_i A^i v. \tag{18}$$

then

$$\begin{aligned} \|z_{fl}^{(0)} - z^{(0)}\|_2 &= \| \textstyle\sum_{i=0}^{m} [\prod_{k=0}^{i} (I + \Gamma_k^{(m-i)})]\alpha_i A^i v - \alpha_i A^i v\|_2 \\ &\leq \textstyle\sum_{i=0}^{m} \|[\prod_{k=0}^{i} (I + \Gamma_k^{(m-i)})]\alpha_i A^i v - \alpha_i A^i v\|_2. \end{aligned} \tag{19}$$

It follows that

$$\begin{aligned} \|z_{fl}^{(0)} - z^{(0)}\|_2 &\leq \textstyle\sum_{i=0}^{m} \|(1+\varepsilon)^i \alpha_i A^i v - \alpha_i A^i v\|_2 \\ &\leq \textstyle\sum_{i=0}^{m} \|[(1+\varepsilon)^i - 1]\alpha_i A^i v\|_2 \end{aligned} \tag{20}$$

Note $\varepsilon \ll 1$, thus $(1+\varepsilon)^k = 1 + k\varepsilon + O(\varepsilon^2) \approx 1 + k\varepsilon$. Therefore,

$$\begin{aligned} \|z_{fl}^{(0)} - z^{(0)}\|_2 &\leq \textstyle\sum_{i=0}^{m} \|i\varepsilon\alpha_i A^i v\|_2 \\ &\leq m\varepsilon \textstyle\sum_{i=0}^{m} (\|\alpha_i A^i v\|_2) \\ &\leq m\varepsilon \textstyle\sum_{i=0}^{m} \mid \alpha_i \mid \|A\|_2^i \|v\|_2. \end{aligned} \tag{21}$$

Since polynomial preconditioners are commonly used in GMRES (e.g., Algo/ rithm 2.2) and other *Krylov subspace solvers* [10], $\|v\|_2$ is always 1. In addition, without loss of generality, the matrix A can be *scaled* [1] such that $\|A\|_2 = 1$. As a result, from (21) it follows that

$$\|z_{fl}^{(0)} - z^{(0)}\|_2 \leq m\varepsilon \textstyle\sum_{i=0}^{m} \mid \alpha_i \mid . \tag{22}$$

Clearly, the bound of $\|z_{fl}^{(0)} - z^{(0)}\|_2$ is dependent upon:

[1] Scaling may change $\sigma(A)$ and eventually influence the coefficient of $P_m(A)$.

1. m, the degree of the preconditioning polynomial. Theoretically, the higher the degree of the preconditioning polynomial, the closer is the approximation to A^{-1}. However, due to the limitation of the finite precision of the computer, the construction of a high degree polynomial may incur great accumulated errors and yield poor convergence performance. This conclusion is supported by the experimental results presented in Figure 1((II) and (III)), which record the *convergence* [2] traces of —GMRES-LS— and —GMRES-Cheby— using several benchmark linear systems with dimensions ranging from 729 to 11948. The degrees of the polynomial preconditioners in these figures range from 0 to 10 .
2. $\sum_{i=0}^{m} \mid \alpha_i \mid$, the sum of the absolute values of the coefficient of the preconditioning polynomial. Table 3 shows that the coefficients of least-squares and Chebyshev preconditioning polynomial increase sharply while the Neumann series preconditioning polynomial satisfies ($\sum_{i=0}^{m} \mid \alpha_i \mid = m+1$). The experimental results presented in Figure 1(I) and Table 4 show that the convergence of —GMRES-Neum— is not seriously influenced by m.

$\sum_{i=0}^{m} \mid \alpha_i \mid$	degree of $P_m(A)$				
↘	0	1	2	3	4
Neuman-series	1	2	3	4	5
Least-squares	1	9	58	364	2027
Chebyshev	2	16	98	576	3362

TABLE 3
$\sum_{i=0}^{m} \mid \alpha_i \mid$ *($\sigma(A) \subseteq [0,1]$)*

convergent iteration count		degree of $P_m(A)$							
↘		1	2	4	8	16	32	64	128
GMRES-Neum	nos7	27	16	24	21	7	6	3	3
	plat1919	37	33	28	27	27	27	27	27
	bcsstk18	74	70	61	47	34	27	20	27
GMRES-LS	nos7	29	33	28	29	-	-	-	-
	plat1919	38	28	39	41	-	-	-	-
	bcsstk18	76	62	38	29	-	-	-	-
GMRES-Cheby	nos7	29	27	28	526	∞	∞	∞	∞
	plat1919	49	44	30	293	∞	∞	∞	∞
	bcsstk18	93	84	72	43	∞	∞	∞	∞
NOTE	"-" : The implementation is inapplicable yet; "∞": The solver does not converge.								

TABLE 4
Convergence of GMRES-polynomial

4. Results and Conclusions.

The analysis and supporting experimental results presented in this paper show that the stability of the major operation $P_m(A)v$ is determined by two parameters m and $\sum_{i=0}^{m} \mid \alpha_i \mid$, the degree of the preconditioning polynomial $P_m(A)$ and the sum of the absolute values of the coefficients of the polynomial, respectively. Further, it would appear to be the case that the second parameter is the more significant of the two.

[2]The convergence criterion is $\frac{\|b-Ax_s\|_2}{\|b\|_2} \leq .111 * 10^{-15}$ where x_s is the approximate solution at sth step. This seems a little bit stringent but that level of accuracy is required in practical applications such as eigenvalue computations.

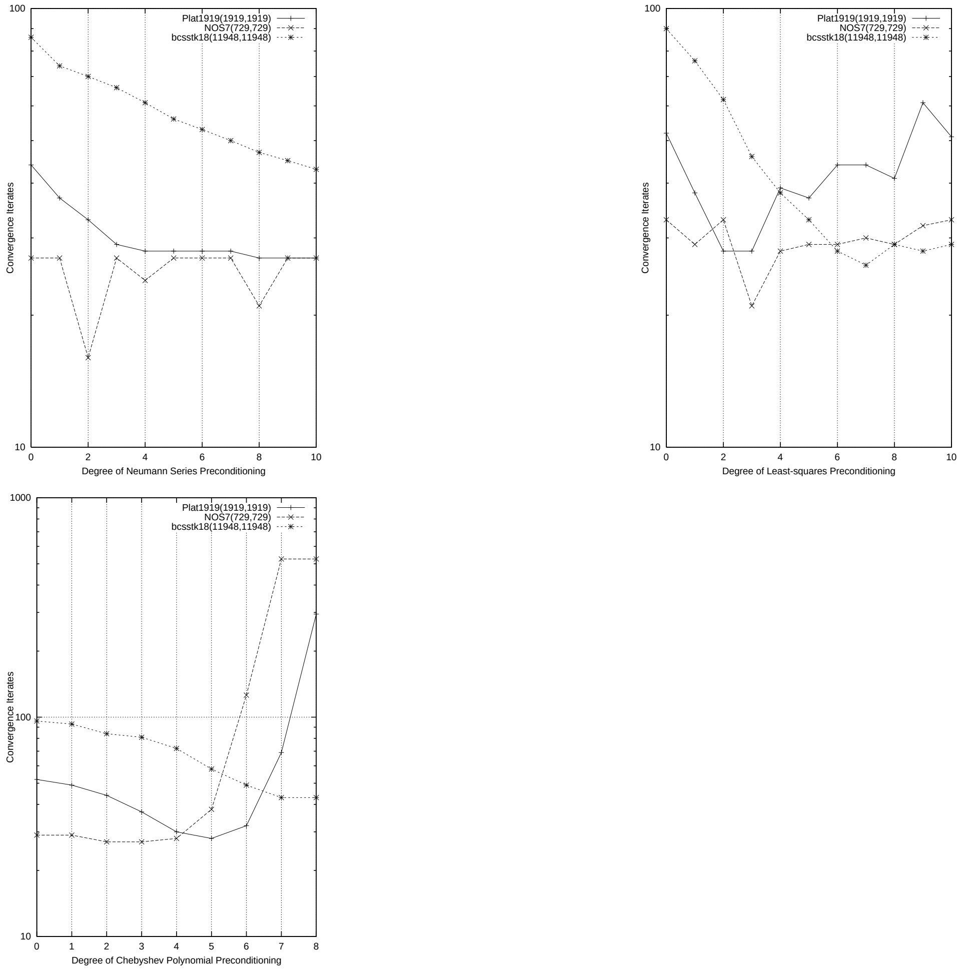

FIG. 1. Convergence of GMRES-polynomial: (I)GMRES-Neum, (II)GMRES-LS, (III)GMRES-Cheby.

Thus, the use of the Neumann series polynomial preconditioner in preference to the Least-Squares and the Chebyshev preconditioners is to be recommended when the degree of the polynomial used is large. This is so despite the fact that the corresponding Least-squares and Chebyshev polynomial preconditioners of the same degree are theoretically better approximations to A^{-1} than the Neumann-series preconditioner.

Undoubtedly, a comprehensive analysis of the stability of preconditioning involves far more than a study of the above parameters. Thus, for example, both the accuracy of the estimation of $\sigma(A)$ and the plot-structure of A may influence the perturbation of preconditioning. Issues such as these will be addressed in future work.

REFERENCES

[1] S. F. Ashby, *Polynomial preconditioning for conjuagte gradient methods*, PhD thesis, Dept. of Computer Science, Univ. of Illinois at Urbana-Champaign, Illinois, 1987.

[2] J.Baglama et al, *Adaptively preconditioned GMRES algorithms*, SIAM J. Sci. Comput. Vol. 20,

No.1, pp.243-269.

[3] Are Magnus Bruaset, *A survey of preconditioned iterative methods*, Longman Scientific & Technical, Copublished in the United States with John Wiley & Songs, Inc., New York.

[4] A. Bjorck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, 1995.

[5] James W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, PA, 1997.

[6] Hong Jiang, *On the orthogonality of residual polynomials of minimax polynomial preconditioning*, Numerische Mathematik, 67:pp345-364, 1994.

[7] *Harwell-Boeing Sparse Matrices Collection*, URL http://math.nist.gov/MatrixMarket.

[8] Yousef Saad, *Iterative solution of indefinite symmetric linear systems by methods using orthogonal polynomial over two disjoint interval*, SIAM J. Numer. Anal., Vol20, Aug. pp362-376, 1983.

[9] Yousef Saad, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press Series in Algorithms and Architecture for Advanced Scientific Computing, 1994.

[10] Yousef Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing, 1996.

[11] JH Wilkinson, *Rounding Errors in Algebraic Processes.* London: H.M. Stationery office, 1963.

[12] JH Wilkinson, et. al., *An Estimate for the Condition Number of a Matrix.* SIAM J. Num. Anal. 1979, Vol.16: pp.368-375.