Inner and Outer Iterations for the Chebyshev Algorithm

Eldar Giladi
Gene Golub
Joseph Keller

CRPC-TR95593
September 1995

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Revised: January, 1996. Also available as *CRPC-95-5* from the Center for Research on Parallel Computation at the California Institute of Technology.

Inner and outer iterations for the Chebyshev algorithm

Eldar Giladi *

Applied Mathematics 217-50

Caltech, Pasadena CA 91125.

Gene H. Golub

Department of Computer Science

Stanford University, Stanford CA 94305.

Joseph B. Keller

Departments of Mathematics and Mechanical Engineering

Stanford University, Stanford CA 94305.

January 17, 1996

Abstract

We analyze the Chebyshev iteration in which the linear system involving the splitting matrix is solved inexactly by an inner iteration. We assume that the tolerance for the inner iteration may change from one outer iteration to the other. When the tolerance converges to zero, the asymptotic convergence rate is unaffected. Motivated by this result, we seek the sequence of tolerance values that yields the lowest cost. We find that among all sequences of slowly varying tolerances, a constant one is optimal. Numerical calculations that verify our results are shown. Our analysis is based on asymptotic methods, such as the W.K.B method, for linear recurrence equations and an estimate of the accuracy of the resulting asymptotic result.

1 Introduction

The Chebyshev iterative algorithm [1] for solving linear systems of equations requires at each step the solution of a subproblem i.e. the solution of another linear system. We

*Phone (o): (818)-395-4550, FAX: (818)-578-0124, e-mail: giladi@ama.caltech.edu

assume that the subproblem is also solved iteratively by an "inner iteration". The term "outer iteration" refers to a step of the basic algorithm. The cost of performing an outer iteration is dominated by the cost of solving the subproblem, and it can be measured by the number of inner iterations. A good measure of the total amount of work needed to solve the original problem to some accuracy ϵ is then, the total number of inner iterations. To reduce the amount of work, one can consider solving the subproblems "inexactly" i.e. not to full accuracy. Although this diminishes the cost of solving each subproblem, it usually slows down the convergence of the outer iteration.

It is therefore interesting to study the effect of solving each subproblem inexactly on the performance of the algorithm. We consider two measures of performance: the asymptotic convergence rate and the total amount of work required to achieve a given accuracy ϵ. The accuracy to which the inner problem is solved may change from one outer iteration to the next. First, we evaluate the asymptotic convergence rate when the tolerance values converge to 0. Then, we seek the "optimal strategy", that is, the sequence of tolerance values that yields the lowest possible cost for a given ϵ.

The present results, extend those in Giladi [2], [3]. The asymptotic convergence rate of the inexact Chebyshev iteration, with a fixed tolerance for the inner iteration, was derived in Golub and Overton [4] (see also [5], [6], [7], [8], [9], [10]). Previous work has mainly concentrated on the convergence rate, whereas we emphasize the cost of the algorithm.

In section 2, we review the Chebyshev method and present the basic error bound for the inexact algorithm. Then, in section 3 we evaluate the asymptotic convergence rate when the sequence of tolerance values gradually decreases to $\eta \geq 0$. In section 4 we seek the "best strategy" i.e the one that yields the lowest possible cost. In section 5, we obtain an asymptotic approximation for the error bound when the sequence of tolerance values is slowly varying. In section 6 we analyze the error in this asymptotic approximation and present a few numerical calculations that demonstrate it's accuracy. In section 7 we use the analysis of section 5, to show that for the Chebyshev iteration, the optimal strategy is constant tolerance. We also estimate the optimal constant. Then, in section 8 we present a few numerical calculations that demonstrate the accuracy of the analysis of section 7. In Section 9, we generalize this result to other iterative schemes.

2 Chebyshev iteration

Chebyshev iteration (see Manteuffel [11]) to solve the real $n \times n$ system of linear equations

$$Ax = b \tag{1}$$

uses the splitting

$$A = M - (M - A). \tag{2}$$

It requires that the spectrum of $M^{-1}A$ be contained in an ellipse, symmetric about the real axis, in the open right half of the complex plane. We denote the foci of such an ellipse by l and u. Furthermore, we assume that $M^{-1}A$ is diagonalizable.

The exact Chebyshev method is defined by

$$x_1 = x_0 + \alpha z_0, \tag{3}$$

$$x_{k+1} = x_{k-1} + \omega_{k+1}(\alpha z_k + x_k - x_{k-1}), \qquad k = 1, 2, ... \tag{4}$$

where

$$M z_k = r_k, \qquad r_k = b - A x_k, \tag{5}$$

$$\alpha = \frac{2}{l+u}, \qquad \mu = \frac{u+l}{u-l} \tag{6}$$

$$\omega_{k+1} = 2\mu \frac{c_k(\mu)}{c_{k+1}(\mu)}. \tag{7}$$

In (3), the initial iterate x_0 is given, and in (7), c_k denotes the Chebyshev polynomial of degree k.

The inexact Chebyshev method is obtained by solving (5) iteratively for z_k. This results in replacing (5) by

$$M z_k = r_k + q_k, \qquad \| q_k \| \leq \delta_k \| r_k \|, \qquad \delta_k \in (0,1). \tag{8}$$

In the variable strategy scheme the tolerance δ_k tends to $\eta \geq 0$ as k increases, while in the constant strategy scheme $\delta_k = \hat{\delta}$ is constant.

We denote the error at step k by

$$e_k = x - x_k. \tag{9}$$

We also define K, V, Σ and σ_j by

$$K = I - \alpha M^{-1} A, \qquad K = V \Sigma V^{-1}, \qquad \Sigma = \text{Diag}(\sigma_j). \tag{10}$$

We use the same derivation as in [4] to show that when $\mu\sigma_j \neq \pm 1$

$$\frac{\| V^{-1} e_k \|}{\| V^{-1} e_0 \|} \leq \frac{\rho^k \tau(k,\delta)}{|\cosh(k \cosh^{-1} \mu)|}, \tag{11}$$

where δ represents a sequence of tolerance values $\{\delta_k\}_{k=1}^{\infty}$. In equation (11), ρ is defined by

$$\rho = \max_j |e^{\theta_j}| \quad \text{where} \quad \theta_j = \cosh^{-1} \mu\sigma_j. \tag{12}$$

The function $\tau(k,\delta)$ satisfies the recurrence equation

$$\tau(k+1,\delta) - 2(1 + \Delta\delta_k)\tau(k,\delta) + \tau(k-1,\delta) = 0 \tag{13}$$

with initial conditions

$$\tau(0,\delta) = 1, \qquad \tau(1,\delta) = 1 + 2\Delta\delta_0. \tag{14}$$

The constant Δ in (13) is given by

$$\Delta = \frac{\alpha|\mu| \parallel V^{-1}M^{-1} \parallel\parallel AV \parallel}{\rho}. \tag{15}$$

The bound (11) is the product of two terms: $\frac{\rho^k}{|\cosh(k\cosh^{-1}(\mu))|}$ and $\tau(k,\delta)$. The former is the bound for the exact algorithm and it is exponentially decaying. The latter is a monotonically increasing term which accounts for the accumulation of errors introduced by solving the inner problem inexactly. We shall obtain asymptotic approximations to $\tau(k,\delta)$ under various assumptions on the sequence δ_k in order to analyze the performance of the inexact algorithm.

3 Asymptotic convergence rate

We shall now estimate the asymptotic convergence rate of the inexact Chebyshev algorithm when the sequence of tolerance values for the inner iteration gradually decreases to 0. Our goal is to show that then, the asymptotic convergence rate of the inexact algorithm is the same as that of the exact scheme. This is in contrast to the case of constant tolerance for which the asymptotic convergence rate of the inexact algorithm is lower than that of the exact algorithm [4].

We base our analysis on the bound (11). Therefore, we wish to compute

$$\overline{\lim}_{k\to\infty}\left(\frac{\rho^k\tau(k,\delta)}{|\cosh(k\cosh^{-1}(\mu))|}\right)^{1/k}. \tag{16}$$

In order to do so, we need to estimate the asymptotic behavior for large k of $\tau(k,\delta)$. By making mild assumptions on the rate at which $\delta_k \to 0$, we will show that

$$\lim_{k\to\infty} \tau(k,\delta)^{1/k} = 1. \tag{17}$$

Upon using (17) in (16), we find that the asymptotic convergence rate of the algorithm is

$$\lim_{k\to\infty} \frac{\rho}{|\cosh(k\cosh^{-1}(\mu))^{1/k}|} = \rho_e, \tag{18}$$

where ρ_e is the asymptotic convergence rate of the exact algorithm.

Equation (17) holds for many sequences δ_k of tolerance values. In order to obtain a general result, we shall assume only that

$$\delta_k \le \frac{C}{k}. \tag{19}$$

The positive constant C in (19) is arbitrary. Hence, if $C \gg 1$ the sequence of tolerance values can decay quite slowly.

We show that (17) holds under assumption (19) in two steps. First we show that $\tau(k,\delta)$ in (13), is bounded by the function $\sigma(k,\hat{\delta})$, where $\sigma(k,\hat{\delta})$ satisfies (13) with δ_k replaced by $\hat{\delta}_k = \frac{C}{k}$. Then, we show that $\lim_{k\to\infty} \sigma(k,\hat{\delta})^{1/k} = 1$. As a first step, we prove the following proposition

Proposition 1 *Let $\tau(k,\delta)$ be a solution to (13) and (14) and let $\sigma(k,\hat{\delta})$ be the solution to the same equation with δ_k replaced by $\hat{\delta}_k$. Assume that $\hat{\delta}_k \geq \delta_k$ and that $\sigma(0,\hat{\delta}) = \tau(0,\delta)$. Then,*

$$\sigma(k,\hat{\delta}) - \sigma(k-1,\hat{\delta}) \geq \tau(k,\hat{\delta}) - \tau(k-1,\hat{\delta}) \tag{20}$$

for $k = 1, 2, \ldots$ and

$$\sigma(k,\hat{\delta}) \geq \tau(k,\delta) \tag{21}$$

for all k.

We prove this proposition by induction. For $k = 1$ we obtain from (14) that

$$\sigma(1,\hat{\delta}) = (1 + 2\Delta\hat{\delta}_0) \geq (1 + 2\Delta\delta_0) = \tau(1,\delta). \tag{22}$$

Then, we assume that assertions (20) and (21) are true for all $k = 1, 2, \ldots, N$. In view of (13)

$$\sigma(N+1,\hat{\delta}) - \sigma(N,\hat{\delta}) = \sigma(N,\hat{\delta}) - \sigma(N-1,\hat{\delta}) + 2\Delta\hat{\delta}_N\sigma(N,\hat{\delta}) \tag{23}$$

and

$$\tau(N+1,\delta) - \tau(N,\delta) = \tau(N,\delta) - \tau(N-1,\delta) + 2\Delta\delta_N\tau(N,\delta). \tag{24}$$

By the induction hypothesis, $\sigma(N,\hat{\delta}) - \sigma(N-1,\hat{\delta}) \geq \tau(N,\delta) - \tau(N-1,\delta)$ and $\sigma(N,\hat{\delta}) \geq \tau(N,\delta)$. Furthermore, $\hat{\delta}_k \geq \delta_k$ so the right side of (23) is greater than or equal to the right side of (24). We conclude that

$$\sigma(N+1,\hat{\delta}) - \sigma(N,\hat{\delta}) \geq \tau(N+1,\delta) - \tau(N,\delta) \tag{25}$$

and that $\sigma(N+1,\hat{\delta}) \geq \tau(N+1,\delta)$.

We shall now obtain the asymptotic behavior of $\sigma(k,\delta)$ for large k from (13) with $\delta_k = \frac{C}{k}$. We use the method of [12].

We first replace $\tau(k,\delta)$ by $\sigma(k,\hat{\delta})$ in (13) and set $\hat{\delta}_k = C/k$. Then we introduce the stretched variables

$$x = \epsilon k, \quad \sigma(k,\hat{\delta}) \equiv R(x), \tag{26}$$

to obtain

$$R(x+2\epsilon) - 2\left(1 + \frac{C\Delta\epsilon}{x+\epsilon}\right) R(x+\epsilon) + R(x) = 0. \tag{27}$$

We seek for $R(x)$ an asymptotic approximation valid for $\epsilon \ll 1$ of the form

$$R(x) \sim c(\epsilon) e^{\frac{\psi(x)}{\sqrt{\epsilon}}} (K_0(x) + \epsilon^{1/2} K_1(x) + \epsilon K_2(x) + \epsilon^{3/2} K_3(x) + O(\epsilon^2)). \tag{28}$$

The functions $\psi(x)$, $K_0(x), \ldots, K_3(x)$ are to be determined so that $R(x)$ satisfies equation (27). The constant $c(\epsilon)$ is to be determined so that $R(x)$ is independent of ϵ. After substituting (28) into (27), we express each side of the resulting expression in power series in $\epsilon^{1/2}$ assuming that $\psi(x+\epsilon)$, $K_0(x+2\epsilon)$, $K_0(x+\epsilon)$, $K_1(x+2\epsilon)$.... can be expanded in Taylor series in powers of ϵ. Then, we equate the coefficients of each power of $\epsilon^{1/2}$ on the left side of the resulting expression, to the same power of $\epsilon^{1/2}$ on the right side. The coefficients of ϵ and of $\epsilon^{3/2}$, yield the following equations for $\psi(x, \delta)$ and $K_0(x, \delta)$ respectively:

$$\frac{-2C\Delta K_0(x)}{x} + K_0(x)\psi'(x)^2 = 0, \tag{29}$$

$$-\frac{2C\Delta K_0(x)\psi'(x)}{x} + K_0(x)\psi'(x)^3 + 2\psi'(x)K_0'(x) + K_0(x)\psi''(x) = 0. \tag{30}$$

Upon solving (29) for ψ we find

$$\psi(x) = 2\sqrt{2C\Delta x}. \tag{31}$$

Introducing the right side of (31) into (30) and solving the resulting equation for K_0 we obtain

$$K_0 = D\left(\frac{x}{2C\Delta}\right)^{1/4}. \tag{32}$$

To find the constant D in (32) we could match (28) to another expansion which satisfies the initial conditions (14). However, the value of D is unimportant for our purposes since $\lim_{k\to\infty} D^{1/k} = 1$.

We substitute (31) for ψ and (32) for K_0 into (28) for R. Then, we use the change of variables (26) to obtain

$$\sigma(k, \hat{\delta}) \sim c(\epsilon) D \epsilon^{1/4} \left(\frac{k}{2\Delta C}\right)^{1/4} e^{2\sqrt{2C\Delta k}}. \tag{33}$$

To make the right side of (33) independent of ϵ, we require that $c(\epsilon) = \epsilon^{-1/4}$, and we obtain

$$\sigma(k, \hat{\delta}) \sim D\left(\frac{k}{2\Delta C}\right)^{1/4} e^{2\sqrt{2C\Delta k}}. \tag{34}$$

Therefore,

$$\lim_{k\to\infty} \sigma(k, \hat{\delta})^{1/k} = 1. \tag{35}$$

In a realistic numerical computation δ_k is bounded below by the machine precision η_0. Moreover, the analysis of the iteration with $\delta_k = \hat{\delta}$ reveals that if $\hat{\delta} \gg \eta_0$ is sufficiently small,

the performance of the inexact algorithm is for all practical purposes indistinguishable from that of the exact algorithm. Indeed, solving (13) with $\delta_k = \hat{\delta}$ yields

$$\tau(k, \delta) = \frac{2}{1 + e^{-\Phi(\hat{\delta})}} \sinh(k\Phi(\hat{\delta})) + e^{-k\Phi(\hat{\delta})}, \tag{36}$$

where

$$\Phi(\delta) \equiv \log\left(1 + \Delta\delta + \sqrt{(1 + \delta\Delta)^2 - 1}\right). \tag{37}$$

It follows from (16), (18) and (36) that the asymptotic convergence rate is

$$\rho_e e^{\phi(\hat{\delta})} \sim \rho_e(1 + \sqrt{2\Delta\hat{\delta}}), \quad \hat{\delta} \ll 1. \tag{38}$$

The number $N(\epsilon, \hat{\delta})$ of outer iterations required to achieve an accuracy ϵ with tolerance $\hat{\delta}$ is approximately

$$N(\epsilon, \hat{\delta}) \approx \lceil \frac{\log \epsilon}{\log\left(\rho_e(1 + \sqrt{2\Delta\hat{\delta}})\right)} \rceil. \tag{39}$$

Hence, if $\sqrt{2\Delta\hat{\delta}} \leq 10^{-3}\,|\log \rho_e|$, the inexact scheme requires no more than one more iteration per thousand than the exact scheme. The difference is undetectable when $N(\epsilon, 0) = O(100)$.

This leads us to evaluate the asymptotic convergence rate when $\delta_k \to \eta$ and $\eta > 0$. To obtain the behavior of $\sigma(k, \delta)$ in (13) for large k, when $\eta \leq \delta_k \leq \frac{C}{k} + \eta$, we introduce (26) into (13) to obtain (27) and seek an expansion for $R(x)$ of the form

$$R(x) \sim c(\epsilon) e^{\frac{\psi(x)}{\epsilon}} (K_0(x) + \epsilon K_1(x) + O(\epsilon^2)). \tag{40}$$

We introduce (40) into (27) to obtain, after some manipulation, equations for ψ and K_0:

$$e^{\psi_x} - 2(1 + \Delta\eta) + e^{-\psi_x} = 0, \tag{41}$$

$$2\sinh(\psi_x) K_{0,x} + \left(\psi_{xx}\cosh(\psi_x) - \frac{2\Delta C}{x}\right) K_0 = 0. \tag{42}$$

Then, we solve (41) and (42) and substitute the results into (40) to obtain, with $\Phi(\eta)$ defined in (37), $c(\epsilon) = \epsilon^{-\Delta C/\sinh\Phi(\eta)}$ and D a constant

$$\sigma(k, \delta) \sim D k^{\Delta C/\sinh\Phi(\eta)} e^{k\Phi(\eta)}. \tag{43}$$

Hence,

$$\lim_{k\to\infty} \sigma(k, \delta)^{1/k} = e^{\Phi(\eta)}. \tag{44}$$

In view of (38) and (44) the asymptotic convergence rate is the same as that with $\delta_k = \eta$.

The results (34) and (43) of this formal analysis can be made rigorous. We summarize the above analysis in the following theorem:

Theorem 1 *Assume that a linear system of equations is solved to accuracy ϵ, using the Chebyshev iteration, with a variable strategy $\{\delta_k\}$. Assume that $\delta_k \to \eta$ with $\eta \geq 0$ and that $\eta \leq \delta_k \leq C/k + \eta$ for some positive constant C. Then, the asymptotic convergence rate of the Chebyshev iteration with the variable tolerance is the same as the asymptotic convergence rate of the scheme with the fixed tolerance η.*

4 The optimal strategy problem

Motivated by the result of section 3, we now wish to find the "best" sequence of tolerance values for the inner iterations. More precisely, we seek the sequence of tolerances that yields the lowest possible cost for the algorithm.

To formalize this problem, we let $\boldsymbol{\delta} = \{\delta_j\}_{j=0}^{\infty}$, be a sequence of tolerance values. The jth component of $\boldsymbol{\delta}$, δ_j, is the tolerance, required in the solution of the subproblem at outer iteration j. Therefore $\delta_j \in (0,1)$ and the number of inner iterations at step j is $\lceil \frac{-\log \delta_j}{-\log \rho} \rceil$. In this estimate, ρ is the convergence factor of the method which is used in the solution of the subproblem. Then, we define $N(\epsilon,\boldsymbol{\delta})$ to be the number of outer iterations needed to reduce the initial error by a factor ϵ when the problem is solved with strategy $\boldsymbol{\delta}$. It follows that the total number of inner iterations required to achieve this accuracy ϵ is proportional to

$$\mathbf{C}(\epsilon, \boldsymbol{\delta}\ ) = \sum_{j=0}^{N(\epsilon,\delta)-1} -\log \delta_j. \tag{45}$$

Our objective is to minimize $\mathbf{C}(\epsilon, \boldsymbol{\delta}\ )$ with respect to $\boldsymbol{\delta}$.

We consider the set $\mathcal{S}$ of slowly varying strategies

$$\mathcal{S} = \left\{ \boldsymbol{\delta}\ \mid \delta_k = \delta(\beta k),\ \ 0 < \beta << 1,\ \ \forall x \geq 0 :\ \ \frac{\delta'}{\delta}, \frac{\delta''}{\delta} = O(1)\ \text{ and }\ \delta(0) \geq \delta(x) \geq \eta \right\}. \tag{46}$$

In (46), the function $\delta(x)$ is assumed to be twice continuously differentiable and δ' denotes it's derivative. The condition $\frac{\delta'}{\delta} = O(1)$ ensures that $\delta(\beta k)$ varies slowly as a function of k if $\beta \ll 1$.

In order to simplify the analysis, we use the fact that

$$\sum_{j=0}^{N(\epsilon,\delta)-1} (-\log \delta_j) \approx -\int_0^{N(\epsilon,\delta)} \log \delta(\beta t) dt,$$

and redefine the cost as

$$\mathcal{C}(\epsilon, \delta) = -\int_0^{N(\epsilon,\delta)} \log \delta(\beta t) dt. \tag{47}$$

We can now restate the problem as follows. Find $\delta^* \in \mathcal{S}$ such that

$$\mathcal{C}(\epsilon, \delta^*) = \min_{\delta \in \mathcal{S}} \mathcal{C}(\epsilon, \delta). \tag{48}$$

5 Error bound for slowly varying strategies

Now we shall approximate the error bound (11), under the assumption that $\delta \in \mathcal{S}$. First, we obtain an asymptotic approximation for $\tau(k, \delta)$, valid for $\beta \ll 1$. To emphasize the fact that $\tau(k, \delta)$ depends on β, we denote it $\tau(k, \delta, \beta)$.

To simplify the analysis we assume that the function $\delta(x)$ is constant on $[0, \beta]$. This assumption is not very restrictive since it requires only that we change the value of δ_0 to equal δ_1. Moreover, since δ_k is slowly varying the impact of this change on the cost is negligible.

The method we use is similar to the W.K.B method [13] for linear ordinary differential equations with a small parameter, and the ray method Keller [14] for linear partial differential equations with a small parameter. These methods have recently been adapted to linear difference equations with small parameters [12], [15].

We now obtain an approximate solution to equation (13) when $\delta_k = \delta(\beta k)$ belongs to $\mathcal{S}$. Since we are looking for an asymptotic expansion of $\tau(k, \delta, \beta)$ for small β, we introduce the new scaled variables

$$x \equiv \beta k, \quad R(x, \delta, \beta) = \tau(k, \delta, \beta). \tag{49}$$

Upon performing the change of variables (49) in (13), we obtain

$$R(x + \beta, \delta, \beta) = 2(1 + \Delta\delta(x))R(x, \delta, \beta) - R(x - \beta, \delta, \beta). \tag{50}$$

We seek an asymptotic expression for $R(x, \delta, \beta)$ for small β, in the form

$$R(x, \delta, \beta) \sim e^{\psi(x,\delta)/\beta}[K(x, \delta) + \beta K_1(x, \delta) + \beta^2 K_2(x, \delta) + \ldots]. \tag{51}$$

The functions $\psi(x, \delta)$, $K(x, \delta)$, $K_1(x, \delta)\ldots$ are to be determined to make R satisfy (50).

Substitution of (51) into (50), and multiplication by $e^{-\psi/\beta}$ yields

$$\begin{aligned} e^{(\psi(x+\beta,\delta)-\psi(x,\delta))/\beta}(K(x + \beta, \delta) + \beta K_1(x + \beta, \delta) + \ldots) = \\ 2(1 + \Delta\delta(x))[K(x, \delta) + \beta K_1(x, \delta) + \ldots] - \\ e^{-(\psi(x)-\psi(x-\beta))/\beta}(K(x - \beta, \delta) + \beta K_1(x - \beta, \delta) + \ldots). \end{aligned} \tag{52}$$

We now express each side of (52) in powers of β, assuming that $\psi(x + \beta, \delta)$, $\psi(x - \beta, \delta)$, $K(x + \beta, \delta)$, etc.. can be expanded in Taylor series in powers of β. Then, we equate coefficients of powers of β. The coefficients of β^0 and of β^1 yield

$$e^{2\psi_x} - 2(1 + \Delta\delta(x))e^{\psi_x} + 1 = 0, \tag{53}$$

$$\tanh \psi_x K_x + \frac{\psi_{xx}}{2} K = 0. \tag{54}$$

Solving (53) for ψ_x yields

$$\psi_x(x, \delta) = \Phi(\delta), \tag{55}$$

with $\Phi(\delta)$ given by (37). Integrating (55) yields, with a a constant of integration

$$\psi(x, \delta) = \pm \int_0^x \Phi(\delta(t))dt + a. \tag{56}$$

We now rewrite (54) as

$$\frac{K_x}{K} = -\frac{\cosh \psi_x}{\sinh \psi_x} \frac{\psi_{xx}}{2}. \tag{57}$$

Integrating (57), with b a constant of integration, gives

$$K(x, \delta) = \frac{b}{\sqrt{|\sinh(\psi_x(x, \delta))|}}. \tag{58}$$

Now, we use expression (55) for ψ_x in (58) to obtain

$$K(x, \delta) = \frac{b}{((1 + \Delta\delta(x))^2 - 1)^{1/4}}. \tag{59}$$

To obtain the leading order term in $\tau(k, \delta, \beta)$, we substitute the two values (56) for ψ into (51) for R and add the two terms. Then, we use the result in (49) and set $x \equiv \beta k$ to find

$$\tau(k, \delta, \beta) \sim K(\beta k, \delta)(A e^{\frac{1}{\beta}\int_0^{\beta k} \Phi(\delta(t))dt} + B e^{-\frac{1}{\beta}\int_0^{\beta k} \Phi(\delta(t))dt}). \tag{60}$$

Here $\Phi(\delta)$ is defined in (37) and $K(x, \delta)$ is given by (59). The constants A and B are determined to make (60) satisfy the initial conditions (14):

$$A = \frac{1}{2\sinh[\frac{1}{\beta}\int_0^\beta \Phi(\delta(t))dt]} \left(\frac{1 + 2\Delta\delta(0)}{K(\beta, \delta)} - \frac{e^{-\frac{1}{\beta}\int_0^{\beta k} \Phi(\delta(t))dt}}{K(0, \delta)} \right), \qquad B = \frac{1}{K(0, \delta)} - A. \tag{61}$$

Since $\delta(x)$ is constant on $[0, \beta]$, (59) shows that $K(0, \delta) = K(\beta, \delta)$ and $\frac{1}{\beta}\int_0^\beta \Phi(\delta(t))dt = \Phi(\delta(0))$. We substitute (61) into (60) to obtain, after some manipulation,

$$\tau(k, \delta, \beta) \sim \frac{K(\beta k, \delta)}{K(0, \delta)} \left[\frac{2}{1 + e^{-\Phi(\delta(0))}} \sinh\left(\frac{1}{\beta} \int_0^{\beta k} \Phi(\delta(t))dt \right) + e^{-\frac{1}{\beta}\int_0^{\beta k} \Phi(\delta(t))dt} \right]. \tag{62}$$

When $\delta(x) \equiv \hat{\delta}$ is constant, (55) implies that $\psi_{xx} = 0$ and (54) shows that K is also constant. Hence, (62) simplifies to the exact solution (36) of (13) and (14) when $\delta_k \equiv \hat{\delta}$ is a constant.

The exponentially decaying term in (62) can be neglected after a few outer iterations. Then we set $s = t/\beta$ in (62) and introduce the function

$$\sigma(k, \delta) = \frac{K(\beta k, \delta)}{K(0, \delta)} \frac{2}{1 + e^{-\Phi(\delta(0))}} \sinh\left(\int_0^k \Phi(\delta(\beta s))ds \right). \tag{63}$$

Now, we approximate $\tau(k, \delta)$ by $\sigma(k, \delta)$, and the bound for the error in the right hand side of (11) becomes

$$B(k, \delta) = \frac{\sigma(k, \delta)\rho^k}{|\cosh(k \cosh^{-1}(\mu))|}. \tag{64}$$

In the next section, we shall analyze the validity of the approximation (64).

6 Validity of the asymptotic expansion

Now we shall show that the leading order expression for $\tau(k,\delta)$, given by (62), is indeed asymptotic to $\tau(k,\delta)$ as $\beta \to 0$. We denote this expression by $\bar{\tau}(k,\delta)$ and define the residual associated with it by $r(k,\delta)$:

$$r(k,\delta) = \bar{\tau}(k+2,\delta) - 2(1+\Delta\delta(\beta(k+1)))\bar{\tau}(k+1,\delta) + \bar{\tau}(k,\delta). \tag{65}$$

To evaluate $r(k,\delta)$ we substitute (60) for $\bar{\tau}(k,\delta)$ into (65) and then expand ψ and K in Taylor series, with remainders up to order β^3 and β^2, respectively. We use (59) and (56) in the resulting expression to obtain, after some manipulation,

$$|r(k,\delta)| \le \beta^2 M e^{\int_0^{k+1} \Phi(\delta(\beta t))dt}, \quad k \ge 0. \tag{66}$$

Here $M = O(\eta^{-1/4})$ for all δ in $\mathcal{S}$ and is independent of k and β.

The error in the asymptotic approximation, $e(k,\delta) = \bar{\tau}(k,\delta) - \tau(k,\delta)$, satisfies

$$e(k+2,\delta) - 2(1+\Delta\delta(\beta(k+1)))e(k+1,\delta) + e(k,\delta) = r(k,\delta). \tag{67}$$

This equation is obtained by subtracting (13) from (65). The initial conditions for $e(k,\delta)$ are

$$e(0,\delta) = e(1,\delta) = 0. \tag{68}$$

Our goal is to show that for any constant C and all $k \le C$

$$\left|\frac{e(k,\delta)}{\tau(k,\delta)}\right| = O(\beta^2) \quad \text{as} \quad \beta \to 0. \tag{69}$$

To estimate the left side of (69), we obtain an explicit formula for $e(k,\delta)$, by solving (67) and (68). We use the method of reduction of order [13]. Specifically, we seek a solution of the form

$$e(k,\delta) = x_k \tau(k,\delta), \tag{70}$$

where $\tau(k,\delta)$ is the solution to equation (13), (14) and x_k is to be determined. Upon substituting (70) into (69) we find that (69) will hold if

$$|x_k| = O(\beta^2), \quad \text{as} \quad \beta \to 0. \tag{71}$$

We obtain an expression for x_k by substituting (70) for $e(k,\delta)$ into (67). Then, we eliminate $\tau(k+1,\delta)$ from the resulting expression by using (13) and we find that

$$\tau(k+2,\delta)(x_{k+2} - x_{k+1}) - \tau(k,\delta)(x_{k+1} - x_k) = r(k,\delta). \tag{72}$$

Now, we introduce

$$X_k = x_{k+1} - x_k \tag{73}$$

into (72) to obtain a linear first order equation for X_k. The initial conditions (68) yield

$$x_0 = x_1 = 0, \quad \Longrightarrow \quad X_0 = 0. \tag{74}$$

The solution of (72) and (74) is

$$X_k = \frac{1}{\tau(k,\delta)\tau(k+1,\delta)} \sum_{l=0}^{k-1} r(l,\delta)\tau(l+1,\delta). \tag{75}$$

We take the absolute value of each side of (75) and use (66) to obtain

$$|X_k| \leq \beta^2 M \frac{e^{\int_0^k \Phi(\delta(\beta t))dt}}{\tau(k,\delta)} \sum_{l=1}^{k} \frac{\tau(l,\delta)}{\tau(k+1,\delta)}. \tag{76}$$

Here $\Phi(\delta)$ is defined in (37).

In lemma 1 we shall show that $\sum_{l=1}^{k} \frac{\tau(l,\delta)}{\tau(k+1,\delta)}$ is bounded by a constant independent on k and β. In lemma 2 we shall show that for a non-increasing strategy $\delta(x)$ in $\mathcal{S}$

$$\frac{e^{\int_0^k \Phi(\delta(\beta t))dt}}{\tau(k,\delta)} \leq P, \tag{77}$$

where the constant P is independent of k and β. We now use these bounds in the right side of (76) and conclude that for all $k \geq 1$

$$|X_k| \leq C\beta^2, \tag{78}$$

where the constant C is independent of k and β.

Equation (73) and the condition for x_1 in (74) determine x_k through

$$x_k = \sum_{j=1}^{k-1} X_j. \tag{79}$$

To derive the bound (71) for $|x_k|$, we take the absolute value of each side of (79) and use (78) to obtain

$$|x_k| \leq kC\beta^2. \tag{80}$$

We summarize the above analysis in the following theorem:

Theorem 2 *Let $\tau(k,\delta)$ satisfy (13) and (14). Let $\bar{\tau}(k,\delta,\beta)$ be the expression on the right side of (62). Assume that $\delta(x)$ is a non-increasing strategy and that $\delta(x) \in \mathcal{S}$ with $\mathcal{S}$ defined in (46). Then,*

$$\left| \frac{\tau(k,\delta) - \bar{\tau}(k,\delta,\beta)}{\tau(k,\delta)} \right| = O(\beta^2) \quad \text{as} \quad \beta \to 0. \tag{81}$$

Furthermore, the coefficient of β^2 in (81) is bounded by a linear function of k.

We now briefly discuss the validity of the approximation (63). When $\delta \equiv \hat{\delta}$ is constant, (63) is exact up to an exponentially decaying term, and it is very accurate after a few iterations. When δ is not a constant, the approximation is based on (62), which is valid for $\beta \ll 1$ and $k = o(\beta^2)$. Therefore, the accuracy decreases as the number of outer iterations $k \to \infty$, and for a fixed k, increases as $\beta \to 0$. At the end of this section we present a few numerical calculations that demonstrate the accuracy of the expansion for a few variable strategies in $\mathcal{S}$. As we shall see, even for large values of k, it is very accurate.

Lemma 1 *Let $\tau(k, \delta)$ satisfy (13) and (14). Then*

$$\frac{\tau(j+1, \delta)}{\tau(j, \delta)} \geq 1 + 2\Delta\delta(\beta j), \quad j \geq 0, \tag{82}$$

and

$$\sum_{j=1}^{k} \frac{\tau(j, \delta)}{\tau(k+1, \delta)} \leq \frac{1}{2\eta\Delta} \quad , k \geq 1. \tag{83}$$

Proof: Inequality (82) is shown by induction. For $j = 0$ it follows from initial conditions (14). Now assume by induction that (82) holds for all $j = 0, \dots, N-1$. Then from (13)

$$\frac{\tau(N+1, \delta)}{\tau(N, \delta)} = 2(1 + \Delta\delta(\beta N)) - \frac{\tau(N-1, \delta)}{\tau(N, \delta)}. \tag{84}$$

By the induction hypothesis

$$-\frac{\tau(N-1, \delta)}{\tau(N, \delta)} \geq \frac{-1}{1 + 2\Delta\delta(\beta(N-1))} \geq -1. \tag{85}$$

We use (85) in (84) to complete the induction.

In order to prove (83), we recall from (46) that $\delta_k \geq \eta$ and we use this bound in (82) to obtain

$$\frac{\tau(j+1, \delta)}{\tau(j, \delta)} \geq 1 + 2\Delta\eta \quad j \geq 0. \tag{86}$$

Furthermore, we note that

$$\prod_{l=j}^{k} \frac{\tau(l, \delta)}{\tau(l+1, \delta)} = \frac{\tau(j, \delta)}{\tau(k+1, \delta)}. \tag{87}$$

We use (86) in (87) to obtain

$$\frac{\tau(j, \delta)}{\tau(k+1, \delta)} \leq \left(\frac{1}{1 + 2\Delta\eta}\right)^{k+1-j}. \tag{88}$$

It follows that

$$\sum_{j=1}^{k} \frac{\tau(j, \delta)}{\tau(k, \delta)} \leq \sum_{j=1}^{k} \frac{1}{(1 + 2\Delta\eta)^j}. \tag{89}$$

Inequality (83) follows from inequality (89).

Lemma 2 *Let $\delta(x)$ be a non-increasing strategy such that $\delta(x) \in \mathcal{S}$, with $\mathcal{S}$ defined in (46). Then*

$$\frac{e^{\int_0^k \Phi(\delta(\beta t))dt}}{\tau(k,\delta)} \leq P \tag{90}$$

where the constant P is independent of k and β.

Proof: We note that when $\delta(x)$ is a non increasing function of x, it follows from the monotonicity of $\Phi(\delta)$ in (37) that

$$\int_0^k \Phi(\delta(\beta t))dt \leq \sum_{j=0}^{k-1} \Phi(\delta(\beta j)). \tag{91}$$

We introduce the right side of (91) into (90) and use (37) for $\Phi(x)$, to obtain

$$\frac{e^{\int_0^k \Phi(\delta(\beta t))dt}}{\tau(k,\delta)} \leq \frac{\prod_{j=0}^{k-1}\left(1+\Delta\delta(\beta j)+\sqrt{(1+\Delta\delta(\beta j))^2-1}\right)}{\tau(k,\delta)}. \tag{92}$$

We now seek a lower bound on $\tau(k,\delta)$. In view of the left condition in (14), we can write write $\tau(k,\delta)$ as the product

$$\tau(k,\delta) = \rho_0\rho_1\cdots\rho_{k-1}, \quad k \geq 1, \tag{93}$$

where

$$\rho_j = \frac{\tau(j+1,\delta)}{\tau(j,\delta)}, \quad j \geq 0. \tag{94}$$

It follows from (13) and (14) that ρ_j satisfies the equation

$$\rho_j = 2(1+\Delta\delta(\beta j)) - \frac{1}{\rho_{j-1}}, \quad j \geq 1, \tag{95}$$

with

$$\rho_0 = 1 + 2\Delta\delta(0). \tag{96}$$

To obtain a lower bound for the product in (93), we introduce the sequence ρ_k^*, $k = 0, 1, 2, \ldots$ which satisfies

$$\rho_k^* \leq \rho_k, \quad k \geq 0. \tag{97}$$

The number ρ_k^* is computed with the aid of the intermediate quantities $\rho_{k,j}^*$, $j = 0, \ldots, k$ as follows:

$$\rho_{k,0}^* = 1 + 2\Delta\delta(\beta k), \tag{98}$$

$$\rho_{k,j}^* = 2(1+\Delta\delta(\beta k)) - \frac{1}{\rho_{k,j-1}^*}, \quad 1 \leq j \leq k. \tag{99}$$

We define

$$\rho_k^* = \rho_{k,k}^*. \tag{100}$$

In order to demonstrate (97), we show by induction on j that for all $0 \leq j \leq k$

$$\rho_{k,j}^* \leq \rho_j. \tag{101}$$

For $j = 0$ it follows from (96), (98) and the fact that $\delta(x)$ is non-increasing that

$$\rho_{k,0}^* = 1 + 2\Delta\delta(\beta k) \leq 1 + 2\Delta\delta(0) = \rho_0. \tag{102}$$

Now, we assume that (101) is true for all $j = 0, \ldots, l-1$. Then, it follows from (95), (99) and the fact that $\delta(x)$ is non-increasing that

$$\rho_{k,l}^* = 2(1 + \Delta\delta(\beta k)) - \frac{1}{\rho_{k,l-1}^*} \;\leq\; 2(1 + \Delta\delta(\beta l)) - \frac{1}{\rho_{l-1}} = \rho_l. \tag{103}$$

The next step in the proof is to evaluate ρ_k^* explicitly and obtain a lower bound for it. This is done by solving the non-linear recurrence equation (99) for $\rho_{k,j}^*$, subject to the initial condition (98). We solve this equation with a method analogous to the one described in section 16.7 of [16] and obtain

$$\rho_{k,j}^* = \frac{2S}{\frac{S - \Delta\delta(\beta k)}{S + \Delta\delta(\beta k)} \frac{1}{(1 + \Delta\delta(\beta k) + S)^{2j}} + 1} + 1 + \Delta\delta(\beta k) - S, \quad j \geq 1, \tag{104}$$

where

$$S = \sqrt{(1 + \Delta\delta(\beta k))^2 - 1}. \tag{105}$$

From equation (105) $S > \Delta\delta(\beta k)$ so that

$$0 < \frac{S - \Delta\delta(\beta k)}{S + \Delta\delta(\beta k)} < 1. \tag{106}$$

Furthermore, it follows from (105) and the definition of η in (46) that

$$\frac{1}{(1 + \Delta\delta(\beta k) + S)^2} \leq \left(\frac{1}{1 + \Delta\eta + \sqrt{2\Delta\eta + (\Delta\eta)^2}} \right)^2 = \mu, \tag{107}$$

where the equality on the right defines the constant μ. We use (107) and (106) in (104) and obtain, in view of (100),

$$\rho_k^* \;\geq\; \frac{2S}{1 + \mu^k} + 1 + \Delta\delta(\beta k) - S, \quad k \geq 1. \tag{108}$$

Further manipulation of (108) yields

$$\rho_k^* \;\geq\; (1 + \Delta\delta(\beta k) + S) \left(1 - \frac{\mu^k}{(1 + \mu^k)} \frac{2S}{1 + \Delta\delta(\beta k) + S} \right), \quad k \geq 1. \tag{109}$$

Finally, we note that $1/(1+\mu^k) < 1$ and $2S/(1+\Delta\delta(\beta k)+S) < 1$, where the latter inequality follows from (105). We use these bounds in (109) and use (97) to get

$$\rho_k \geq \left(1+\Delta\delta(\beta k)+\sqrt{(1+\Delta\delta(\beta k))^2-1}\right)\left(1-\mu^k\right), \quad k \geq 1. \tag{110}$$

We are now ready to prove the lemma. First, we substitute the right side of (93) for $\tau(k,\delta)$ in (92). Then, we use (110) and (96) to find

$$\frac{e^{\int_0^k \Phi(\delta(\beta t))dt}}{\tau(k,\delta)} \leq \frac{1+\Delta\delta(0)+\sqrt{(1+\Delta\delta(0))^2-1}}{1+2\Delta\delta(0)} \prod_{j=1}^{k-1} \frac{1}{1-\mu^j}. \tag{111}$$

The infinite product $\prod_{j=1}^{\infty} \frac{1}{1-\mu^j}$ is convergent because $\sum_{j=1}^{\infty} \mu^j < \infty$. Hence, the right side of (111) is bounded by a number P which is independent of β and k.

We now present a few numerical calculations that demonstrate the accuracy of the expansion derived in section 5. First, we solve (13) for $\tau(k,\delta,\beta)$ by iteration and then we compute the approximate solution $\sigma(k,\delta)$ given by (63), for all $2 \leq k \leq 2000$. We present the relative error in this approximation.

We use strategies from the three parameter family

$$\delta_k = \frac{A}{B(1+(\beta k)^\gamma)} + \eta, \quad k \geq 1, \quad \delta_0 = \delta_1. \tag{112}$$

The minimal tolerance in (112) is $\eta = 10^{-12}$. In all our calculations $\delta_k \gg \eta$ and for all practical purposes η can be neglected. The value of parameters A and γ is fixed at 1. The parameter B and the value of β vary from one calculation to the other. The value of Δ in (13) is set to 37. We performed analogous calculations with larger values of Δ and with $\gamma = 1.5$ and obtained similar results.

In table 1, we present the maximum with respect to k, of the absolute value of the relative error in percent. Each entry in this table corresponds to a calculation with a different strategy. The strategy is determined by the parameters B and β. Figure 1 depicts the relative error in percent between $\sigma(k,\delta)$ and $\tau(k,\delta,\beta)$, for all $2 \leq k \leq 2000$. Each graph corresponds to different values of B and β. We note that the approximation is accurate even for large values of k.

7 Constant strategy is optimal

Now, using (47) and (64) we seek the optimal strategy for the Chebyshev iteration. The numbers $N(\epsilon,\delta)$ and $\mathcal{C}(\epsilon,\delta)$ in (47), are hard to determine precisely. Therefore, we introduce the quantities $N_B(\epsilon,\delta)$ and $\mathcal{C}_B(\epsilon,\delta)$, which are the number of outer iterations required to reduce the error bound (64) to ϵ and the associated cost, respectively. The following theorem shows that a constant strategy is optimal.

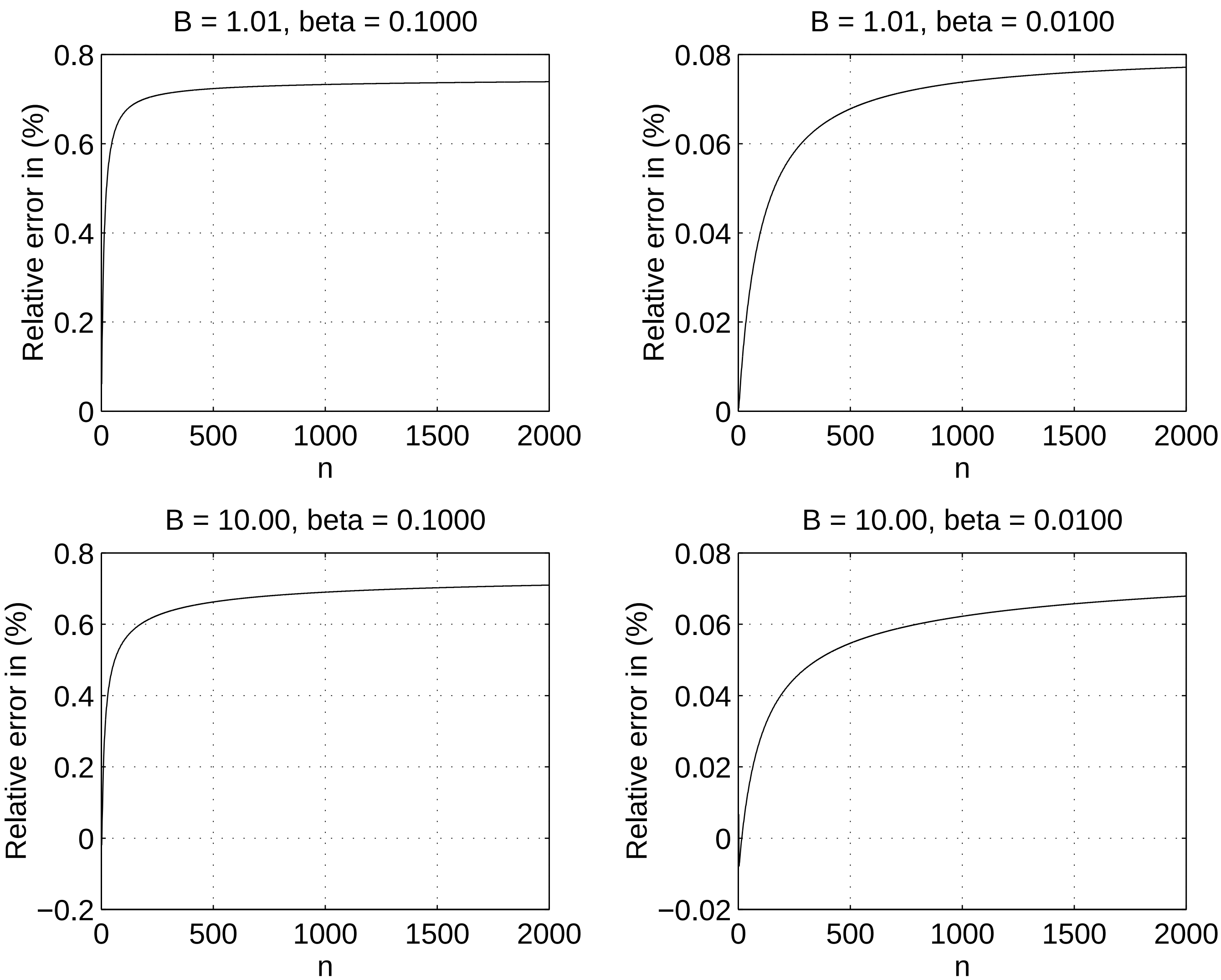

Figure 1: The relative error $|\tau(n, \delta, \beta) - \sigma(n, \delta)| \, / \, |\tau(n, \delta, \beta)|$ in (%) for $2 \leq n \leq 2000$. Each graph corresponds to different values of B and β.

$B\backslash\beta$	.1	.01
1.01	0.74	0.05
1.10	0.74	0.05
1.50	0.74	0.05
2.00	0.73	0.05
5.00	0.72	0.07
10.00	0.71	0.07
100.00	0.70	0.11

Table 1: The maximum over $2 \leq n \leq 2000$ of $|\tau(n,\delta,\beta) - \sigma(n,\delta)| \,/\, |\tau(n,\delta,\beta)|$ in (%).

Theorem 3 *Suppose that a linear system of equations is solved to accuracy ϵ by the Chebyshev iteration using inner iterations with a sequence of tolerances $\{\delta_k\}$ in $\mathcal{S}$. There exists a constant strategy $\hat{\delta}(\delta,\epsilon)$, for which the cost is smaller, i.e.*

$$\mathcal{C}_B(\epsilon,\hat{\delta}) \leq \mathcal{C}_B(\epsilon,\delta).$$

Proof: Given the variable strategy δ and the accuracy ϵ used in the solution of the linear system, we define the associated constant strategy $\hat{\delta}(\delta,\epsilon)$

$$\hat{\delta}(\delta,\epsilon) = \Phi^{-1}\left(\frac{\int_0^{N_B(\epsilon,\delta)} \Phi(\delta(\beta t))dt}{N_B(\epsilon,\delta)}\right). \tag{113}$$

In Lemma 3, we show that $N_B(\epsilon,\hat{\delta}) \leq N_B(\epsilon,\delta)$. Therefore,

$$\mathcal{C}_B(\epsilon,\hat{\delta}) = -\int_0^{N_B(\epsilon,\hat{\delta})} \log\hat{\delta}dt = -N_B(\epsilon,\hat{\delta})\,\log\hat{\delta} \leq -N_B(\epsilon,\delta)\,\log\hat{\delta}. \tag{114}$$

In Lemma 4, we show that

$$-\log\hat{\delta}N_B(\epsilon,\delta) \leq \mathcal{C}_B(\epsilon,\delta). \tag{115}$$

Using (115) in the right hand side of (114), proves the theorem.

Lemma 3

$$N_B(\epsilon,\hat{\delta}) \leq N_B(\epsilon,\delta) \tag{116}$$

Proof: By definition of $N_B(\epsilon,\delta)$ the bound for the error $B(k,\delta)$ in (64) satisfies

$$B(N_B(\epsilon,\delta),\delta) \leq \epsilon.$$

Therefore, to prove (116) it is sufficient to show that after $N_B(\epsilon, \delta)$ outer iterations, the bound for the error associated with the variable strategy is greater than the one associated with the constant strategy. Hence we need to show

$$B(N_B(\epsilon,\delta),\delta) \geq B(N_B(\epsilon,\delta),\hat{\delta}). \tag{117}$$

We see from (64) that (117) is equivalent to the inequality

$$\sigma(N_B(\epsilon,\delta),\delta) \geq \sigma(N_B(\epsilon,\delta),\hat{\delta}), \tag{118}$$

where σ is defined in (63). To prove (118) we begin by rewriting expression (63) for $\sigma(k,\delta)$ with $k = N_B(\epsilon,\delta)$,

$$\sigma(N_B(\epsilon,\delta),\delta) = \frac{2}{1+e^{-\Phi(\delta(0))}} \frac{K(\beta N_B(\epsilon,\delta),\delta)}{K(0,\delta)} \sinh\left(\int_0^{N_B(\epsilon,\delta)} \Phi(\delta(\beta t))dt\right). \tag{119}$$

Then, we note from (37) that Φ is monotonically increasing and that for all non-negative x, $\delta(0) \geq \delta(x)$ from (46). Therefore,

$$\Phi(\hat{\delta}) = \frac{\int_0^{N_B(\epsilon,\delta)} \Phi(\delta(\beta t))dt}{N_B(\epsilon,\delta)} \leq \Phi(\delta(0)), \tag{120}$$

$$\frac{2}{1+e^{-\Phi(\delta(0))}} \geq \frac{2}{1+e^{-\Phi(\hat{\delta})}}. \tag{121}$$

Furthermore, we see that $K(\beta N_B(\epsilon,\delta),\delta)/K(0,\delta) \geq 1$ from equation (59). Using this and (121) in the right hand side of (119) we obtain

$$\sigma(N_B(\epsilon,\delta),\delta) \geq$$

$$\frac{2}{1+e^{-\Phi(\hat{\delta})}} \sinh\left(N_B(\epsilon,\delta)\Phi\Phi^{-1}\left(\frac{\int_0^{N_B(\epsilon,\delta)} \Phi(\delta(\beta t))dt}{N_B(\epsilon,\delta)}\right)\right) = \sigma(N_B(\epsilon,\delta),\hat{\delta}). \tag{122}$$

Lemma 4

$$-\log \hat{\delta} N_B(\epsilon,\delta) \leq \mathcal{C}_B(\epsilon,\delta)$$

Proof: The definition (37) of Φ shows that $\Phi^{-1}(x) = \frac{\cosh(x)-1}{\Delta}$. Therefore,

$$\frac{d^2}{dx^2}\left(-\log\Phi^{-1}(x)\right) = (\cosh(x)-1)^{-1} > 0$$

and $-\log\Phi^{-1}(x)$ is strictly convex on the interval $\{\Phi(\delta(x)) \mid 0 \leq x \leq N_B(\epsilon,\delta)\}$. It follows from Jensen's inequality that

$$-\log\hat{\delta} =$$

$$-\log\Phi^{-1}\left(\frac{\int_0^{N_B(\epsilon,\delta)} \Phi(\delta(\beta t))dt}{N_B(\epsilon,\delta)}\right) \leq \frac{-\int_0^{N_B(\epsilon,\delta)} \log\Phi^{-1}\Phi(\delta(\beta t))dt}{N_B(\epsilon,\delta)} = \frac{\mathcal{C}_B(\epsilon,\delta)}{N_B(\epsilon,\delta)}. \tag{123}$$

Multiplying (123) by $N_B(\epsilon, \delta)$ proves the lemma.

We now show how to estimate the optimal constant $\hat{\delta}$. We note from (64) that for any iteration N

$$B(N, \hat{\delta}) \leq 2\frac{\rho^N \sinh(N\Phi(\hat{\delta}))}{|\cosh(N\cosh^{-1}(\mu))|} \approx 2e^{N[\log\rho + \Phi(\hat{\delta}) - \mathrm{Re}(\cosh^{-1}(\mu))]}. \tag{124}$$

Then, by equating the right side of (124) to ϵ and using (37), we obtain

$$N_B(\epsilon, \hat{\delta}) \approx \frac{\log 2 - \log\epsilon}{\mathrm{Re}(\cosh^{-1}(\mu)) - \log\rho - \cosh^{-1}(1 + \Delta\hat{\delta})}. \tag{125}$$

An estimate of the cost is then

$$\mathcal{C}_B(\epsilon, \hat{\delta}) \approx \frac{-\log\hat{\delta}(\log 2 - \log\epsilon)}{\mathrm{Re}(\cosh^{-1}(\mu)) - \log\rho - \cosh^{-1}(1 + \Delta\hat{\delta})}. \tag{126}$$

The right side of (126) can be minimized easily with respect to $\hat{\delta}$ using a standard minimization technique. The original variational problem (48) is thus reduced to a simple optimization problem. Since $B(N, \hat{\delta})$ approximates a bound for the error, the tolerance obtained by this method will be a lower bound for the optimal tolerance. The estimation of the optimal constant depends on the parameters μ and ρ in expression (126). These are often determined adaptively while solving the system [17].

8 Numerical calculations

We now present a few numerical calculations that verify the analysis of section 7. In each experiment, we solve a linear system with Chebyshev iteration to accuracy ϵ, using a variable strategy δ. Then, we solve the same system with the associated constant strategy $\hat{\delta}(\delta, \epsilon)$, where $\hat{\delta}$ is defined in (113) with $N_B(\epsilon, \delta)$ replaced by $N(\epsilon, \delta)$. We recall that $N(\epsilon, \delta)$ is the exact number of outer iterations required to achieve an accuracy ϵ, when solving the problem with strategy δ. This number is obtained from our numerical experiment. Our goal is to verify that the predictions of lemma 3 and theorem 3 hold in practice.

In section 4 we define the cost at outer iteration j by using

$$\frac{-\log\delta_j}{-\log\rho} \tag{127}$$

for the number of inner iterations required to achieve accuracy δ_j instead of $\lceil\frac{-\log\delta_j}{-\log\rho}\rceil$. Here, ρ is the convergence factor for the inner iteration. If ρ is close to 1, then the relative error in using (127) is usually small and the cost (45) is truly proportional to the total number of inner iterations. In this case, we expect good agreement between the analysis and

the numerical calculations. Moreover, we expect some fluctuations around the predicted behavior when $\rho \ll 1$. We covered both cases in our experiments.

We solve the symmetric system

$$Ax = b, \tag{128}$$

arising from the central difference discretization of the operator

$$-\frac{d^2}{dx^2} + (.8\sin(10x) + 1)C, \tag{129}$$

in the interval $[0,1]$ with homogeneous Dirichlet boundary conditions. The right side b in (128) is chosen at random. The splitting matrix M is obtained from the discretization of the operator

$$-\frac{d^2}{dx^2} + C, \tag{130}$$

with homogeneous Dirichlet boundary conditions. The mesh parameter in this discretization is $h = 1/100$. The tolerance for the outer iteration is $\epsilon = 10^{-12}$. The initial iterates for both the inner and outer iterations are 0.

In all our experiments, we use strategies from the family (112). The values of γ and A are fixed at 1. The parameter B and the value of β vary from one experiment to the other. For each variable strategy δ, the associated constant strategy $\hat{\delta}$ is computed using (113) with $N_B(\epsilon, \delta)$ replaced by $N(\epsilon, \delta)$. We note from (113) that Φ depends on Δ. We evaluate Δ exactly but find that $\hat{\delta}$ is not very sensitive to the value of Δ. We performed calculations with various values of C in (129) and (130) and we shall report on a representative sample obtained with $C = 30$.

We use two methods for the inner iteration. The symmetric Gauss Seidel, with the convergence factor 0.993, close to 1, and the symmetric successive over relaxation method [18] (S.S.O.R) with the smaller convergence factor 0.925. In the S.S.O.R iteration, the relaxation parameter ω is the optimal parameter ω^* of S.O.R. In each experiment, we record the number of outer iterations and the total number of inner iterations for the variable and constant strategy cases.

Tables 2-5 correspond to the case where the inner iteration is symmetric Gauss Seidel. In table 2 we report the difference in the total number of inner iterations between the variable strategy case and the associated constant strategy case. All entries in the table are in (%) and are computed from

$$\frac{N_{in}(\epsilon, \delta) - N_{in}(\epsilon, \hat{\delta})}{N_{in}(\epsilon, \hat{\delta})} * 100. \tag{131}$$

Here $N_{in}(\epsilon, \delta)$ is the total number of inner iterations performed when solving the system to accuracy ϵ with strategy δ.

Each entry in table 2 corresponds to a different strategy. The strategy is determined by the parameters B and β. Note that not all strategies are slowly varying since $\beta \not\ll 1$

in the two rightmost columns of that table. The important thing to note in table 2 is that all entries are positive. Therefore, the number of inner iterations associated with the variable strategy is greater than or equal to the number of inner iterations with the constant strategy. Hence, there is agreement with Theorem 3.

In table 3, we present the difference in the number of outer iterations between the variable strategy case and the constant strategy case i.e. $N(\epsilon,\delta) - N(\epsilon,\hat{\delta})$. We see that all entries are non-negative and there is very good agreement with Lemma 3.

In table 4 we present the total number of inner iterations with the associated constant strategy. The lowest number of inner iterations is found at the top left entry. This entry corresponds to the lowest tolerance for the inner iteration. Table 5 presents the total number of outer iterations. We see that the top left entry maximizes the number of outer iterations. Hence, among all strategies considered in this table, the strategy which yields the lowest convergence rate also yields the lowest cost.

Tables 6 and 7 present the difference in number of inner and outer iterations, respectively, when the inner iteration is S.S.O.R. Since the convergence factor is not close to 1 some fluctuations from the predicted behavior are expected. Indeed, two entries in table 6 are negative. However, the fluctuations are small and the constant strategy performs essentially as well as the variable one.

In our numerical calculations we have used both slowly varying strategies, $\beta = .1$, and "rapidly" varying ones, $\beta = 2 \not\ll 1$. Although our theory was developed for slowly varying strategies, the conclusion of theorem 3 is found to hold for all the strategies considered.

9 Generalization to other iterative procedures

We now consider a general iterative algorithm in which, at iteration k, a subproblem is solved by an inner iteration to accuracy δ_k. The norm of the error at step k, e_k, satisfies the relation

$$e_{k+1} = \rho(k, x_0, \boldsymbol{\delta}\ )e_k. \tag{132}$$

In (132), $\rho(k, x_0, \boldsymbol{\delta}\ )$, the convergence factor at step k, depends on the initial iterate x_0 and on the sequence of tolerance values $\boldsymbol{\delta}$. We assume that $\rho(k, x_0, \boldsymbol{\delta}\ )$ is a product

$$\rho(k, x_0, \boldsymbol{\delta}\ ) = e^{f(k,x_0)} e^{g(\phi_k)}, \tag{133}$$

with

$$\delta_k = e^{\phi_k}. \tag{134}$$

Hence, the only tolerance upon which $\rho(k, x_0, \boldsymbol{\delta}\ )$ depends is δ_k, the tolerance at outer iteration k. Furthermore, the dependence of $\rho(k, x_0, \boldsymbol{\delta}\ )$ on δ_k is the same at each iteration of the algorithm. We can prove a result similar to the one of section 7 for an iteration satisfying (133).

$B\backslash\beta$	.1	.5	1	2
1.01	4.18	4.07	5.23	6.00
1.10	2.11	8.54	10.23	11.66
1.50	2.92	5.02	6.07	6.51
2.00	0.02	5.93	6.77	7.18
5.00	0.85	1.85	8.46	2.38
10.00	0.81	8.30	9.29	3.34
100.00	1.00	2.57	3.29	3.85

Table 2: The difference in number of inner iterations $(N_{in}(\epsilon,\delta) - N_{in}(\epsilon,\hat{\delta}))/N_{in}(\epsilon,\hat{\delta})$ in (%). The tolerances δ_k and $\hat{\delta}$ are defined by (112) and (113), respectively. Inner iteration is symmetric Gauss Seidel.

$B\backslash\beta$	.1	.5	1	2
1.01	2	1	1	1
1.10	1	2	2	2
1.50	1	1	1	1
2.00	0	1	1	1
5.00	0	0	1	0
10.00	0	1	1	0
100.00	0	0	0	0

Table 3: The difference in number of outer iterations $N(\epsilon,\delta) - N(\epsilon,\hat{\delta})$. The tolerances δ_k and $\hat{\delta}$ are defined by (112) and (113), respectively. Inner iteration is symmetric Gauss Seidel.

$B\backslash\beta$	.1	.5	1	2
1.01	4232	4786	5219	5814
1.10	4310	4792	5159	5685
1.50	4458	5144	5369	6203
2.00	4666	5143	5571	6324
5.00	5199	6208	6444	7447
10.00	5697	6722	7167	8093
100.00	8660	9423	10228	11137

Table 4: The number of inner iterations $N_{in}(\epsilon,\hat{\delta})$ with $\hat{\delta}$ given in (113). Inner iteration is symmetric Gauss Seidel.

$B\backslash\beta$	.1	.5	1	2
1.01	37	24	21	19
1.10	36	23	20	18
1.50	31	22	19	18
2.00	28	20	18	17
5.00	21	18	16	16
10.00	18	16	15	15
100.00	15	14	14	14

Table 5: Number of outer iterations $N(\epsilon, \hat{\delta})$ with $\hat{\delta}$ given in (113). Inner iteration is symmetric Gauss Seidel.

$B\backslash\beta$	.1	.5	1	2
1.01	2.36	8.60	4.29	8.62
1.10	5.48	6.21	8.15	3.77
1.50	2.26	3.66	4.34	3.42
2.00	0.98	3.43	4.79	4.76
5.00	0.64	5.87	4.64	5.44
10.00	-0.77	-0.67	5.93	0.14
100.00	0.42	0.60	0.78	1.03

Table 6: The difference in number of inner iterations $(N_{in}(\epsilon, \delta) - N_{in}(\epsilon, \hat{\delta}))/N_{in}(\epsilon, \hat{\delta})$ in (%). The tolerances δ_k and $\hat{\delta}$ are defined by (112) and (113), respectively. Inner iteration is S.S.O.R.

$B\backslash\beta$	.1	.5	1	2
1.01	2	2	1	2
1.10	1	2	2	1
1.50	0	1	1	1
2.00	0	1	1	1
5.00	0	1	1	1
10.00	0	0	1	0
100.00	0	0	0	0

Table 7: The difference in number of outer iterations $N(\epsilon, \delta) - N(\epsilon, \hat{\delta})$. The tolerances δ_k and $\hat{\delta}$ are defined by (112) and (113), respectively. Inner iteration is S.S.O.R.

Theorem 4 *Consider an iterative algorithm in which at step k, a subproblem is solved by inner iteration to accuracy δ_k. Assume that the norm of the error satisfies (132), with $\rho(k, x_0, \boldsymbol{\delta})$ of the form (133). Assume that $g(\phi)$ is a convex non decreasing function. Let $\epsilon(N, \boldsymbol{\delta})$ be the reduction of the error after N outer iterations. Then, for any variable strategy $\boldsymbol{\delta}$ and any number of outer iterations N, there exists a constant $\hat{\delta}(N, \boldsymbol{\delta}) = e^{\hat{\phi}(N, \boldsymbol{\delta})}$ with the following properties.*

1. *After N outer iterations with the constant tolerance $\hat{\delta}(N, \boldsymbol{\delta})$ for the inner iteration, the error is reduced by exactly $\epsilon(N, \boldsymbol{\delta})$.*
2. *The cost (45) of performing N outer iterations with the constant tolerance $\hat{\delta}(N, \boldsymbol{\delta})$ is lower than the cost of performing N outer iterations with the variable tolerance $\boldsymbol{\delta}$.*

In other words, for such an iteration a constant strategy is optimal.

Proof: From (132) and (133) we find that after N outer iterations of the algorithm with the variable tolerance $\boldsymbol{\delta}$, the error is reduced by

$$\epsilon(N, \boldsymbol{\delta}) = \frac{e_N}{e_0} = e^{\sum_{k=0}^{N-1} f(k, x_0)} e^{\sum_{k=0}^{N-1} g(\phi_k)}. \tag{135}$$

Let

$$\hat{\phi}(N, \boldsymbol{\delta}) = g^{-1}\left(\frac{\sum_{k=0}^{N-1} g(\phi_k)}{N}\right) \quad \text{and} \quad \hat{\delta}(N, \boldsymbol{\delta}) = e^{\hat{\phi}(N, \boldsymbol{\delta})}. \tag{136}$$

Then, it follows from equations (132) and (133) that after N iterations with the constant strategy $\hat{\delta}(N, \boldsymbol{\delta})$ the error is reduced by

$$\frac{e_N}{e_0} = e^{\sum_{k=0}^{N-1} f(k, x_0)} \exp\left(\sum_{j=0}^{N-1} g \circ g^{-1}\left(\frac{\sum_{k=0}^{N-1} g(\phi_k)}{N}\right)\right) = \epsilon(N, \boldsymbol{\delta}). \tag{137}$$

The right hand side of equation (137) is exactly $\epsilon(N, \boldsymbol{\delta})$.

Using (45) and (136) we find that the cost associated with N steps of the constant tolerance iteration is

$$\mathbf{C}(\epsilon, \hat{\delta}) = -N g^{-1}\left(\frac{\sum_{k=0}^{N-1} g(\phi_k)}{N}\right), \tag{138}$$

while the cost associated with the variable tolerance is

$$\mathbf{C}(\epsilon, \boldsymbol{\delta}) = -\sum_{k=0}^{N-1} \phi_k. \tag{139}$$

Now, the right side of (138) is no greater then the right side of (139) since g is convex.

The error bound (64) for the Chebyshev iteration is analogous to (135) with the sum over $g(\phi)$ replaced by an integral and the term $e^{\sum_{k=0}^{N-1} f(k, x_0)}$ replaced by a function $F(k, x_0)$, independent of δ. Hence, theorem 3 is essentially a continuous version of theorem 4. The proof of the former is complicated by the presence of the amplitude term (59) in (63).

10 Acknowledgements

This work was supported in part by NSF under cooperative agreement no CCR-9120008 and grant CCR-9505393, ONR, and AFOSR.

A

In this appendix, we derive an expression for the residual defined in equation (65) and obtain a bound for it. We find it convenient to work with the stretched variable x of equation (49). Hence, we change variables in equation (65) and also center the differences in that equation around x to obtain

$$r(x-\beta,\delta,\beta)=\bar{\tau}(x+\beta,\delta,\beta)-2(1+\Delta\delta(x))\bar{\tau}(x,\delta,\beta)+\bar{\tau}(x-\beta,\delta,\beta),\quad x\geq\beta. \tag{140}$$

In equation (140) $\bar{\tau}(x,\delta,\beta)=\bar{\tau}(k,\delta)$ so that

$$\bar{\tau}(x,\delta,\beta)=K(x)\left(Ae^{\frac{\psi(x)}{\beta}}+Be^{\frac{-\psi(x)}{\beta}}\right), \tag{141}$$

where $\psi(x)$ and $K(x)$ are defined in equations (58) and (56), respectively. We expand in the resulting expression, $\psi(x+\beta,\delta)$, $\psi(x-\beta,\delta)$, $K(x+\beta)$ and $K(x-\beta)$ in Taylor series with remainder. We expand ψ up to order β^3 and K up to order β^2. Upon collecting coefficients of equal powers in β and using equations (53) and (54) we obtain the following expression for the residual

$$r(x-\beta,\delta,\beta)=\beta^2\left(Ae^{\frac{\psi(x)}{\beta}}(S_1^+ +\beta S_2^+ +\beta^2 S_3^+)+Be^{\frac{-\psi(x)}{\beta}}(S_1^- +\beta S_2^- +\beta^2 S_3^-)\right). \tag{142}$$

In equation (142),

$$\begin{aligned}S_1^+ = {}& e^{\Phi(x)}\left(K(x)R_1(x,\beta)+\frac{\Phi^{'}(x)K^{'}(x)}{2}+\frac{K^{''}(\nu_1)}{2}\right)+\\ & e^{-\Phi(x)}\left(K(x)R_2(x,\beta)-\frac{\Phi^{'}(x)K^{'}(x)}{2}+\frac{K^{''}(\nu_2)}{2}\right),\end{aligned} \tag{143}$$

$$S_2^+ = e^{\Phi(x)}\left(R_1(x,\beta)K^{'}(x)+\frac{\Phi^{'}(x)K^{''}(\nu_1)}{4}\right)+e^{-\Phi(x)}\left(-R_2(x,\beta)K^{'}(x)+\frac{\Phi^{'}(x)K^{''}(\nu_2)}{4}\right), \tag{144}$$

$$S_3^+ = e^{\Phi(x)}\frac{R_1(x)K^{''}(\nu_1)}{2}+e^{-\Phi(x)}\frac{R_2(x)K^{''}(\nu_2)}{2}. \tag{145}$$

In equations (143)-(145) $\Phi(x)$ is defined in equation (37) and the points ν_1, ν_2 satisfy

$$x\leq\nu_1\leq x+\beta,\qquad x-\beta\leq\nu_2\leq x. \tag{146}$$

Furthermore, the functions $R_1(x,\beta)$ and $R_2(x,\beta)$ are defined by

$$R_1(x,\beta)=\frac{\Phi^{''}(\eta_1)}{6}+\left(\frac{\Phi^{'}(x)}{2}+\frac{\beta}{6}\Phi^{''}(\eta_1)\right)^2\frac{e^{\xi_1}}{2}, \tag{147}$$

$$R_2(x,\beta)=\frac{-\Phi^{''}(\eta_2)}{6}+\left(\frac{\Phi(x)^{'}}{2}-\frac{\beta}{6}\Phi^{''}(\eta_2)\right)^2\frac{e^{\xi_2}}{2}. \tag{148}$$

In equations (147) and (148) the points η_1, η_2 satisfy

$$x \le \eta_1 \le x+\beta, \quad x-\beta \le \eta_2 \le x. \tag{149}$$

The point ξ_1 is between 0 and $\beta(\frac{\Phi^{'}(x)}{2}+\frac{\beta}{6}\Phi^{''}(\eta_1))$ and ξ_2 is between 0 and $\beta(\frac{\Phi^{'}(x)}{2}-\frac{\beta}{6}\Phi^{''}(\eta_2))$. The terms S_j^- in equation (142) are obtained from the corresponding terms S_j^+ in equations (143)-(145), by replacing Φ, Φ' and Φ'' by $-\Phi$, $-\Phi'$ and $-\Phi''$ respectively.

Upon introducing expressions (143)-(148) into equation (142) we find after some manipulations that

$$|r(x-\beta,\delta)| \le \beta^2 e^{\psi(x)/\beta}(A+B)(1+\Delta)(2R\bar{K}+\bar{K}'\bar{\Phi}'+\bar{K}''+\beta(2R\bar{K}'+\frac{\bar{K}''\bar{\Phi}'}{2})+\beta^2\bar{K}''R), \tag{150}$$

where

$$R=\frac{\bar{\Phi}''}{6}+\left(\frac{\bar{\Phi}'}{2}+\beta\frac{\bar{\Phi}''}{6}\right)^2\frac{e^{\xi}}{2}, \quad |\xi|=\beta\frac{\bar{\Phi}'}{2}+\beta^2\frac{\bar{\Phi}''}{6}. \tag{151}$$

The terms $\bar{K}', \bar{K}'', \bar{\Phi}', \bar{\Phi}''$ in equations (150) and (151) denote the maximum over all x of the functions $|K^{'}|, |K^{''}|, |\Phi^{'}|, |\Phi^{''}|$ respectively. In order to obtain bounds for these maxima, we evaluate the derivatives of $K(x,\delta)$ and $\Phi(x,\delta)$ using equations (37) and (59) to obtain

$$K^{'}(x)=\frac{-\Delta\ (1+\Delta\,\delta(x))\ \delta'(x)}{2\left((1+\Delta\,\delta(x))^2-1\right)^{\frac{5}{4}}}, \tag{152}$$

$$K^{''}(x)=\frac{5\,\Delta^2\,(1+\Delta\,\delta(x))^2\,\delta'(x)^2}{4\left((1+\Delta\,\delta(x))^2-1\right)^{\frac{9}{4}}}-\frac{(\Delta\delta^{'}(x))^2+\Delta(1+\Delta\delta(x))\delta^{''}(x)}{2\,((1+\Delta\delta(x))^2-1)^{\frac{5}{4}}}, \tag{153}$$

$$\Phi^{'}(x)=\frac{\Delta\delta'(x)}{\sqrt{(1+\Delta\delta(x))^2-1}}, \tag{154}$$

and

$$\Phi^{''}(x)=\frac{\Delta\delta^{''}(x)}{\sqrt{(1+\Delta\delta(x))^2-1}}-\frac{(\Delta\delta'(x))^2(1+\Delta\delta(x))}{((1+\Delta\delta(x))^2-1)^{3/2}}. \tag{155}$$

Upon manipulating the expressions on the right hand side of equations equations (152)-(155) we obtain the following bounds

$$|\Phi(x)^{'}| \le \left|\frac{\delta(x)^{'}}{\delta(x)}\right|, \tag{156}$$

$$|\Phi(x)^{''}| \le \left|\frac{\delta(x)^{''}}{\delta(x)}\right|+\left|\frac{\delta(x)^{''}}{\delta(x)}\right|^2(1+\Delta)^2, \tag{157}$$

$$\left|K^{'}\right| \le \left|\frac{\delta(x)^{'}}{\delta(x)}\right|\left(\frac{1+\Delta}{4}\right)\frac{1}{(2\Delta\eta)^{\frac{1}{4}}}, \tag{158}$$

$$\left|K^{''}\right| \le \frac{1}{(2\Delta\eta)^{\frac{1}{4}}}\left(\left(\frac{5}{16}(1+\Delta)^2+\frac{1}{2}\right)\left|\frac{\delta(x)^{'}}{\delta(x)}\right|^2+\frac{1+\Delta}{4}\left|\frac{\delta(x)^{''}}{\delta(x)}\right|\right). \tag{159}$$

In these expressions, η is the lower bound on $\delta(x)$ defined in equation (46). Now, we recall from equation (46), that $\left|\frac{\delta'(x)}{\delta(x)}\right|$ and $\left|\frac{\delta(x)''}{\delta(x)}\right|$ are bounded for all x. Upon introducing these bounds in the right hand side of equation (150) and performing the change of variable $x = \beta k$ we obtain

$$|r(k, \delta)| \leq \beta^2 M e^{\int_0^{k+1} \Phi(\delta(\beta t))dt}, \quad k \geq 0. \tag{160}$$

References

[1] G.H.Golub and R.S.Varga. Chebyshev semi-iterative methods, successive over-relaxation iterative methods, and second-order richardson iterative methods, parts i and ii. *Numer. Math.*, 3:147–168, 1961.

[2] E. Giladi. *Hybrid Numerical Asymptotic Methods.* PhD thesis, Stanford University, 1995.

[3] E.Giladi. On the interplay between inner and outer iterations for a class of iterative methods. In *Proceedings of the Colorado conference on iterative methods*, April 1994.

[4] G.H. Golub and M.L. Overton. The convergence of inexact chebyshev and richardson iterative methods for solving linear systems. *Numer. Math.*, 53:571–593, 1988.

[5] R.A.Nicolaides. On the local convergence of certain two step iterative procedures. *Numer. Math.*, 24:95–101, 1975.

[6] V.Pereyra. Accelerating the convergence of discretization algorithms. *SIAM J. Numer. Anal.*, 4:508–533, 1967.

[7] N.K.Nichols. On the convergence of two-stage iterative process for solving linear equations. *SIAM J. Numer. Anal.*, 10:460–469, 1973.

[8] S.C.Eisenstat R.S.Dembo and T.Steihaug. Inexact newton methods. *SIAM J. Numer. Anal.*, 19:400–408, 1982.

[9] A.Frommer and D.B.Szyld. H-splittings and two-stage iterative methods. *Numerische Mathematik*, 63:345–356, 1992.

[10] H. C. Elman and G. H. Golub. Inexact and preconditioned uzawa algorithms for saddle point problems. *SIAM Numer. Anal.*, 31:1645–1661, 1994.

[11] T. A. Manteuffel. The tchebyshev iteration for nonsymmetric linear systems. *Numer. Math.*, 28:307–327, 1977.

[12] E. Giladi and J. B. Keller. Eulerian number asymptotics. *Proceedings of the London Royal Society, Series A*, 445:291–303, 1994.

[13] C. M. Bender and S.Orszag. *Advanced Mathematical Methods for Scientists and Engineers.* Mc. Graw Hill, 1978.

[14] J.B.Keller. Rays, waves and asymptotics. *Bull. Am. Math. Soc.*, 84:727–750, 1978.

[15] C. Knessl. The wkb approximation to the g/m/m queue. *SIAM J. Appl. Math.*, 51:1119–1133, 1991.

[16] G.F.Carrier and C.E.Pearson. *Ordinary Differential Equations.* SIAM, 1991.

[17] T. A. Manteuffel. Adaptive procedure for estimation of parameters for the nonsymmetric tchebychev iteration. *Numer. Math.*, 31:187–208, 1978.

[18] R.S.Varga. *Matrix iterative analysis*. Prentice-Hall, 1962.