

# Componentwise fast convergence in the solution of full-rank systems of nonlinear equations<sup>1</sup>

Nicholas I. M. Gould<sup>2,3</sup>, Dominique Orban<sup>4</sup>, Annick Sartenaer<sup>5,6</sup> and Philippe L. Toint<sup>5,6</sup>

## ABSTRACT

The asymptotic convergence of parameterized variants of Newton's method for the solution of nonlinear systems of equations is considered. The original system is perturbed by a term involving the variables and a scalar parameter which is driven to zero as the iteration proceeds. The exact local solutions to the perturbed systems then form a differentiable path leading to a solution of the original system, the scalar parameter determining the progress along the path. A homotopy-type algorithm, which involves an inner iteration in which the perturbed systems are approximately solved, is outlined. It is shown that asymptotically, a single linear system is solved per update of the scalar parameter. It turns out that a *componentwise* Q-superlinear rate may be attained under standard assumptions, and that this rate may be made arbitrarily close to quadratic. Numerical experiments illustrate the results and we discuss the relationships that this method shares with interior methods in constrained optimization.

---

<sup>1</sup> This work was supported by the Belgian National Fund for Scientific Research, EPSRC and an MNRT Grant for joint Ph.D. Support.

<sup>2</sup> Computational Science and Engineering Department, Rutherford Appleton Laboratory, Chilton, Oxfordshire, OX11 0QX, England, EU.  
Email : n.gould@rl.ac.uk

<sup>3</sup> Current reports available from "<http://www.numerical.rl.ac.uk/reports/reports.html>".

<sup>4</sup> CERFACS, av. G. Coriolis, Toulouse, 31057 France  
Email: Dominique.Orban@cerfacs.fr

<sup>5</sup> Department of Mathematics, University of Namur,  
61, rue de Bruxelles, B-5000 Namur, Belgium, EU.  
Email : philippe.toint@fundp.ac.be and annick.sartenaer@fundp.ac.be

<sup>6</sup> Current reports available from "<ftp://thales.math.fundp.ac.be/pub/reports>".

Computational Science and Engineering Department  
Atlas Centre  
Rutherford Appleton Laboratory  
Oxon OX11 0QX  
September 8, 2000.

# 1 Introduction

We consider the solution of a nonlinear system of algebraic equations

$$F(x) = 0, \tag{1.1}$$

where  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a continuously differentiable function of the vector of unknowns  $x$ . Let  $f_i(x)$  be the  $i$ -th component function of  $F(x)$ . We assume that the Jacobian  $J(x)$  of  $F(x)$ , *i.e.* the matrix whose  $i$ -th row is  $\nabla_x f_i(x)^T$ , is full rank at a root of interest,  $x^*$ , and Lipschitz continuous close to this root, that is

**A1.**  $J(x^*)$  is non-singular, and

**A2.**  $J(x)$  is Lipschitz continuous in some open neighbourhood of  $x^*$ .

It is well-known that in this case, if  $x^*$  is a limit point of the Newton iteration

$$x^{k+1} = x^k - J^{-1}(x^k)F(x^k) \tag{1.2}$$

started from some arbitrary  $x^0$ , the complete sequence  $\{x^k\}$  converges *in norm* at a Q-quadratic rate, that is to say that there is a constant  $\kappa \geq 0$  and integer  $k_q$  for which

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^2} \leq \kappa \tag{1.3}$$

for all  $k \geq k_q$  (see, for example, [3, 8]), where  $\|\cdot\|$  denotes any norm on  $\mathbb{R}^n$ .

While this result is entirely satisfactory if all we are concerned about is a fast normwise convergence rate, it is less useful if our interest is in methods that attain a fast *componentwise* rate of convergence. Let us denote by  $z_i$  the  $i$ -th component of any vector  $z \in \mathbb{R}^n$ . It follows immediately from (1.3) that there exists  $\kappa_\infty > 0$  such that

$$|x_i^{k+1} - x_i^*| \leq \|x^{k+1} - x^*\|_\infty \leq \kappa_\infty \|x^k - x^*\|_\infty^2$$

for every component  $x_i^k$  of  $x^k$ , and thus that  $\{x_i^k\}$  converges R-quadratically. However this does not imply that each (or any) component converges Q-quadratically, and indeed this may not be the case, as we show in the following example.

**Example 1.1** For  $1 \leq l \leq n$ , let  $e_l$  be the  $l$ -th column of the identity matrix of order  $n$ . Consider the nonlinear system of equations

$$x_i^2 + x_{i+1} = 0, \quad \text{for } 1 \leq i \leq n-1, \quad \text{and } x_n^2 + x_1 = 0, \tag{1.4}$$

and suppose that the iterate  $x^k = x_l e_l$ , ( $|x_l| < 1$ ) for some index  $1 \leq l \leq n$ . It is straightforward to show (see the Appendix) that a Newton step from  $x^k$  results in the improved estimate

$$x^{k+1} = \begin{cases} x_l^2 e_{l+1} & \text{for } 1 \leq l \leq n-1 \\ x_n^2 e_1 & \text{for } l = n, \end{cases} \tag{1.5}$$

of the root  $x^* = 0$ . Hence, if the Newton iteration is started at  $x^0 = x_l e_l$  for some  $1 \leq l \leq n$ , the resulting sequence of iterates has the property that each component is non-zero precisely once in every cycle of  $n$  iterations. Thus, although the norm of the iterates converges Q-quadratically, this is not true of individual components—they do, however, converge at an  $n$ -step Q- $2^n$  rate.  $\square$

In this paper, we are interested in Newton-like methods that produce iterates whose components converge at the same Q-rate since then we can be sure that, asymptotically, each iteration results in an improvement in any particular component(s). This is particularly important in large-scale computations, where every iteration is expensive, and when only a subset of the variables are of interest. Our aim then is not to spend time on unnecessary computation, which might happen if the components of interest do not improve over the course of a number of iterations. For example, if  $x = (x_p, y)$  where  $x_p$  is the vector of primal variables in a basic SQP Newton method for constrained optimization, and where  $y$  is the vector of associated Lagrange multipliers, it is known that the pair  $(x_p, y)$  asymptotically converges at a Q-quadratic rate under reasonable assumptions, while often it is only the primal variables  $x_p$  that are of interest, and these may not inherit the overall rate of the combined pair. Similarly, in some control problems, the control  $u$  may be of higher interest than the state  $x$ , or the convergence properties of a particular combination of some components of  $u$  with some components of  $x$  is important, and this combination might not possess the same convergence properties as the complete solution pair  $(x, u)$ . Another example where convergence does not take place componentwise is in the solution to the Navier-Stokes equations when studying turbulent fluid flow around a complex geometry, say, an aircraft. In the far field, the flow settles after a few iterations and fast convergence occurs, whereas in the vicinity of the aircraft, convergence may be much slower and variables may be of widely different magnitude, because of the effects of the complexity of the flowfield and of the local mesh refinement. It may thus be desirable to encourage faster convergence in certain, predefined, regions of the mesh—such as boundary layers, wakes, recirculation regions and shock waves—and thus to have control over the convergence of certain, predefined, variables.

Recently, it has been shown that the iterates generated by certain interior-point methods for constrained optimization converge at a componentwise fast (almost quadratic) Q-rate [6]. The methods we consider here are a variation, and generalization, on this theme.

In what follows, we use the order notation for conciseness and clarity. If  $\{\alpha^k\}$  and  $\{\beta^k\}$  are two sequences of positive numbers converging to zero, we say that  $\alpha^k = o(\beta^k)$  if  $\lim_{k \rightarrow \infty} \alpha^k / \beta^k = 0$ , we say that  $\alpha^k = O(\beta^k)$  if there exists a constant  $\kappa > 0$  such that  $\alpha^k \leq \kappa \beta^k$  for all sufficiently large  $k$ , and we write  $\alpha^k = \Theta(\beta^k)$  if  $\alpha^k = O(\beta^k)$  and  $\beta^k = O(\alpha^k)$ . We also extend this notation and say that  $\alpha^k = \Theta(1)$  if there exist constants  $\kappa^L, \kappa^U > 0$  such that  $\kappa^L \leq \alpha^k \leq \kappa^U$  for all sufficiently large  $k$ .

If  $\{z^k\}$  is a sequence in  $\mathbb{R}^n$ , and if  $\mathcal{K}$  is an index set, the subsequence of  $\{z^k\}$  indexed by  $\mathcal{K}$  will be denoted by  $\{z^k\}_{k \in \mathcal{K}}$ .

The paper is organized as follows. In §2, we describe how we intend to solve (1.1) and outline a generic algorithm. We then examine the fast asymptotic properties of this algorithm in §3 and give an explicit componentwise Q-rate of convergence. We discuss the method and its relations with interior-point methods for constrained optimization in §4. Finally, we illustrate the results numerically in §5 and conclude in §6.

## 2 A parameterized variant of Newton's method

Suppose  $h : \mathbb{R}_+ \rightarrow \mathbb{R}^n$  is a continuously differentiable vector-valued function of the positive real parameter  $\mu$  such that  $h(0) = 0$ . A common way of solving problem (1.1) is instead to consider solving a sequence of systems

$$F(x) = h(\mu), \quad (2.1)$$

as  $\mu$  is progressively driven to zero (see for instance [7, 11, 14]). At iteration  $k$ , a typical method sets the value of  $\mu$  to  $\mu^k$ , generates a starting point  $x_s^k$ , possibly from the solution  $x^k$  corresponding to  $\mu^{k-1}$ , and uses a globally convergent method—which we refer to as the *inner iteration*—to find  $x^{k+1}$  satisfying

$$\|F(x^{k+1}) - h(\mu^k)\| \leq \epsilon^k \quad (2.2)$$

for some small  $\epsilon^k > 0$ . The sequences  $\{\mu^k\}$  and  $\{\epsilon^k\}$  are chosen to converge to zero. Our major concern in what follows will be the choice of the starting point  $x_s^k$  rather than the details of the globally convergent procedure—we simply mention that a trust-region variant of Newton's method (see [2]) is suitable. The most obvious starting point is simply the result of the previous inner iteration,  $x^k$ . However, we intend to show that a superior point may be obtained as an *approximation* to the Newton point for the next parameterized subproblem

$$F(x) = h(\mu^k)$$

from  $x^k$ . Applying an approximate Newton correction, we set

$$x_s^k = x^k + s^k, \quad (2.3)$$

where  $s^k$  satisfies

$$J(x^k)s^k = h(\mu^k) - F(x^k) + r^k \quad \text{and} \quad \|r^k\| \leq \eta^k, \quad (2.4)$$

for some residual vector  $r^k$  and a suitable sequence  $\{\eta^k\}$  of positive values whose limit is zero. We might then examine the possible fast componentwise convergence of the sequence  $\{x^k\}$  to a solution of (1.1), under appropriate assumptions on  $h$ , when the starting point (2.3) is used at each iteration.

In the remainder of this paper, our intent is to analyze methods similar to (2.1), but to allow more generality on the perturbation function  $h$ , and to permit the use of approximate Jacobian matrices  $G(x^k) \approx J(x^k)$ . More precisely, consider the function  $h : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}^n$  of  $(x, \mu)$  and suppose it satisfies the following properties

**A3** The derivatives of  $h(x, \mu)$  with respect to  $x$  and  $\mu$  exist and are Lipschitz continuous over a neighbourhood of  $(x^*, 0)$ , and

**A4.**  $h(x^*, 0) = 0$ ,  $h(x, \mu)$  is nonzero for all  $x \neq x^*$  sufficiently close to  $x^*$  and all sufficiently small  $\mu > 0$ , and satisfies  $\nabla_x h(x^*, 0) = 0$ .

We briefly comment Assumption **A4** in §4. We shall henceforth consider the parameterized system

$$F(x) = h(x, \mu) \quad (2.5)$$

as a function of  $\mu$  as  $\mu$  decreases to zero. Since (2.1) is a special case of (2.5), we only consider and refer to (2.5) in the remainder of this paper. Let  $x(\mu)$  be the root of (2.5) for which  $\lim_{\mu \rightarrow 0} x(\mu) =$

$x^*$ . From the implicit-function theorem, such a root exists for sufficiently small  $\mu$  by virtue of **A1**–**A4** and may be expressed as an analytic function of  $\mu$ , as is shown in Lemma 3.1. The method outlined above is sometimes referred to as a *homotopy* method and the simple idea behind it is to track the path of exact solutions  $x(\mu)$  of (2.5) to  $x^*$ , using a convergent iterative method to solve (2.5) with fixed  $\mu$ . An overview of homotopy methods may be found in [7, 8, 14]. Deeper motivation appears in [10] together with numerous references to practical situations, basic work and sparse and dense linear algebra concerns in the course of the solution to (2.5).

In our more general context, we still use (2.3) where  $s^k$  is now required to satisfy

$$G(x^k)s^k = h(x^k, \mu^k) - F(x^k) + r^k \quad \text{and} \quad \|r^k\| \leq \eta^k, \quad (2.6)$$

for some residual vector  $r^k$  and a suitable sequence  $\{\eta^k\}$  of positive values whose limit is zero. In this equation, the matrix  $G(x^k)$  is either  $J(x^k)$  or is a suitable approximation thereof, and is required to satisfy the following condition

**A5.** There exists a constant  $\kappa_G > 0$  such that for all  $x^k$  sufficiently close to  $x^*$ ,

$$\|[G(x^k) - J(x^k)]s^k\| \leq \kappa_G \mu^{k-1} \|s^k\|, \quad (2.7)$$

where  $s^k$  is defined by (2.6).

Condition (2.7) is similar to the Dennis-Moré condition [3], often used in quasi-Newton methods, but is stronger in that it makes the accuracy of the Jacobian approximation along the normalized step explicitly dependent on  $\mu^{k-1}$ , that is, as we show in Lemma 3.3, on the distance to the solution. We further comment Assumption **A5** in §4.

Since no derivative of  $h$  with respect to  $x$  appears in (2.6), a possible interpretation of (2.3) is as an approximation to the Newton point for the next parameterized system  $F(x) = h(x, \mu^k)$  starting from  $x^k$ , where the derivatives of  $h$  with respect to  $x$ , or any approximation thereof, have been discarded from  $G(x)$ . This is justified if, for instance,  $\nabla_x h(x, \mu)$  is expensive to compute or is perhaps unavailable. An alternative interpretation is as an approximation to the Newton point for  $F(x) = h(x^k, \mu^k)$ . However, the absence of  $\nabla_x h(x, \mu)$  in (2.6) is a key point to the results to come and is strongly related to the asymptotic behaviour of interior methods in optimization, as we now comment. Our motivation is that using (2.3) and (2.6) may also be interpreted as a method that extrapolates along the path of exact solutions  $x(\mu)$  to (2.5), in a manner similar to that used in primal interior-point methods for constrained optimization. In that framework, it has been shown [12] that Newton's method applied to (2.5)—thus including the gradient  $\nabla_x h(x, \mu)$  or an approximation thereof—is likely to produce infeasible steps, no matter how close  $x$  is to  $x^*$ . One thus has to cut back quite severely, which considerably slows down the convergence. To circumvent this problem, Dussault [4] computes an extrapolation step to the trajectory  $x(\mu)$  at the current iterate  $x^k$ , say, followed by the Newton step that would have been taken from  $x^k$ —the sum of these two steps being similar to the step  $s^k$  given by (2.6)—a process which turns out to be superlinearly convergent. Note that in a primal-dual interior-point framework, the aforementioned problem does not arise, and the technique applied by Dussault is identical to Newton's method, which may result in a componentwise nearly quadratic Q-rate of convergence [6].

Let  $\{\mu^k\}$ ,  $\{\varepsilon^k\}$  and  $\{\eta^k\}$  be strictly decreasing sequences of parameters whose limit are zero, and consider the iteration given in Algorithm 2.1 for solving (1.1) by way of (2.5).

**Algorithm 2.1: Parameterized root-finding procedure**

Given  $x^0 \in \mathbb{R}^n$ ,  $\mu^0 > 0$ ,  $\epsilon^0 > 0$  and  $\eta^0 > 0$ , set  $k = 0$  and perform the following steps:

**Step 1** [*Starting point*]. Generate the starting point  $x_s^k$  from  $x^k$  using (2.3) and (2.6).

**Step 2** [*Inner iteration*]. Starting from  $x_s^k$ , find  $x^{k+1}$  satisfying

$$\|F(x^{k+1}) - h(x^{k+1}, \mu^k)\| \leq \epsilon^k, \quad (2.8)$$

stopping at the first inner iterate for which this condition is satisfied.

**Step 3** [*Update*]. Update the parameters  $\mu^k$ ,  $\epsilon^k$  and  $\eta^k$  in such a way that they form strictly decreasing sequences converging to zero. Set  $k \leftarrow k + 1$ . Return to Step 1.

For future reference, we write equation (2.8) at iteration  $k - 1$ , that is

$$\|F(x^k) - h(x^k, \mu^{k-1})\| \leq \epsilon^{k-1}. \quad (2.9)$$

It follows immediately from (2.8) or (2.9) that if a subsequence  $\{x^k\}_{k \in \mathcal{K}}$ , defined by some index set  $\mathcal{K}$ , has a limit  $x^\infty$ , then  $x^\infty$  is a root of (1.1).

In order to find a point satisfying (2.8), we may for instance apply Newton's method to (2.5). However, in what follows, we will not be concerned with the exact mechanism of the inner iteration, merely that it starts from  $x_s^k$  and returns as soon as a point  $x^{k+1}$  satisfying (2.8) is discovered.

Notice that the step  $s^k$  computed in Step 1 of Algorithm 2.1 might have been rejected by the mechanism of the inner iteration, perhaps by a linesearch or trust-region acceptance rule, had we used (2.6) as inner iteration procedure, and started from  $x^k$ . By computing such a step outside the inner iteration, we may choose to ignore such (often overly cautious) restrictions. Notice also that the flexibility implied by (2.6) may be exploited by allowing early termination of iterative methods (such as Krylov subspace methods) applied to the system

$$G(x^k)s = h(x^k, \mu^k) - F(x^k)$$

in the unknown  $s$ .

Our aim is thus to propose suitable values for the parameters  $\{\mu^k\}$  and the accuracy tolerances  $\{\epsilon^k\}$  and  $\{\eta^k\}$ , and to examine the effect such choices have on the rate of convergence of the overall iteration.

### 3 Fast local convergence

In this section, we propose a set of updating rules for Step 3 of Algorithm 2.1. While it is possible to derive more abstract conditions on general parameters  $\{\mu^k\}$  and accuracy tolerances  $\{\epsilon^k\}$  and  $\{\eta^k\}$  that permit fast convergence, we restrict ourselves here to the update

$$\mu^{k+1} = \tau_\mu^k (\mu^k)^{1+\theta_\mu^k} \quad (3.1)$$

and accuracy tolerances

$$\epsilon^k = \tau_\epsilon^k (\mu^k)^{1+\theta_\epsilon^k} \text{ and} \quad (3.2)$$

$$\eta^k = \tau_\eta^k (\mu^k)^{1+\theta_\eta^k}, \quad (3.3)$$

for appropriate scale factors

$$0 < \tau_{\min} \leq \tau_\mu^k, \tau_\epsilon^k, \tau_\eta^k \leq \tau_{\max}, \quad (3.4)$$

and exponents

$$\theta_\mu^k, \theta_\epsilon^k, \theta_\eta^k \geq 0, \quad (3.5)$$

where  $\tau_{\min}$  and  $\tau_{\max}$  are given constants. Our aim is to show that there are suitable values of these exponents (3.5) that allow fast componentwise convergence. As asymptotic properties are sought, boundedness of the scale factors (3.4) is sufficient, and we will not be concerned by their exact values.

Our first result considers the behaviour of  $x(\mu)$  near  $x^*$ .

**Lemma 3.1** Suppose that **A1–A4** hold. Then

$$x(\mu) = x^* + \mu x' + O(\mu^2), \quad (3.6)$$

for all sufficiently small  $\mu$ , where

$$x' = J(x^*)^{-1} \nabla_\mu h(x^*, 0). \quad (3.7)$$

**Proof.** The definition of  $x(\mu)$  and (2.5) give that

$$F(x(\mu)) = h(x(\mu), \mu).$$

By the implicit-function theorem and **A1–A4**, there exists a vector  $x'$  such that  $x(\mu)$  may be expressed in the form

$$x(\mu) = x^* + \mu x' + \varepsilon(\mu) \text{ where } \|\varepsilon(\mu)\| = o(\mu). \quad (3.8)$$

Note that (3.8) immediately implies that

$$x(\mu) - x^* = O(\mu). \quad (3.9)$$

A Taylor expansion of  $F(x(\mu))$  involving (3.8) and (3.9) may be combined with **A2** and the identity  $F(x^*) = 0$  to give

$$\begin{aligned} F(x(\mu)) &= F(x^*) + J(x^*)(\mu x' + \varepsilon(\mu)) + O(\mu^2) \\ &= \mu J(x^*)x' + J(x^*)\varepsilon(\mu) + O(\mu^2). \end{aligned} \quad (3.10)$$

On the other hand, a second-order expansion of  $h(x(\mu), \mu)$  about  $(x^*, 0)$  gives

$$\begin{aligned} h(x(\mu), \mu) &= h(x^*, 0) + \nabla_x h(x^*, 0)(x(\mu) - x^*) + \nabla_\mu h(x^*, 0)\mu + O(\mu^2) \\ &= \nabla_\mu h(x^*, 0)\mu + O(\mu^2), \end{aligned} \quad (3.11)$$

where we used **A3**, **A4** and (3.9). A matched asymptotic expansion between (3.10) and (3.11) yields (3.7) and, using **A1**, the estimate  $\|\varepsilon(\mu)\| = O(\mu^2)$ . Such an estimate combines with (3.8) to give (3.6).  $\square$

In the remainder of this section, we consider a subsequence  $\{x^k\}_{k \in \mathcal{K}}$  converging to  $x^*$ . Unless otherwise specified, the following results involve only  $k \in \mathcal{K}$ . Our main theorem will then be concerned with the whole sequence  $\{x^k\}$ . Before proceeding, we need the following technical result, inspired by [6, Lemma 3.1] and [13, Lemma 3.1], that will provide a bound, given in Lemma 3.3, on the distance between  $x^k$  satisfying (2.9) and  $x^*$  for all sufficiently large  $k \in \mathcal{K}$ .

**Lemma 3.2** Suppose that **A1**–**A4** hold and that the vector  $x(\mu, u) \in \mathbb{R}^n$  is defined implicitly as the solution to the nonlinear system

$$\Psi(x, \mu, u) \stackrel{\text{def}}{=} F(x) - h(x, \mu) - u = 0, \quad (3.12)$$

for  $\mu \in \mathbb{R}_+$  and  $u \in \mathbb{R}^n$ . Then there exists a constant  $\delta > 0$  such that the following statements hold.

- (i) The derivatives of  $x(\mu, u)$  with respect to  $\mu$  and  $u$  exist and are Lipschitz-continuous functions over the neighbourhood

$$\mathcal{N}_\delta \stackrel{\text{def}}{=} \{(\mu, u) \in \mathbb{R}_+ \times \mathbb{R}^n \mid \mu + \|u\| \leq \delta\}.$$

- (ii) If  $(\mu_1, u_1), (\mu_2, u_2) \in \mathcal{N}_\delta$ , we have

$$\|x(\mu_1, u_1) - x(\mu_2, u_2)\| = \Theta(|\mu_1 - \mu_2| + \|u_1 - u_2\|). \quad (3.13)$$

**Proof.** Since **A1** and **A4** imply that the Jacobian matrix  $\nabla_x \Psi(x^*, 0, 0) = J(x^*)$  is non-singular, since  $\Psi(x^*, 0, 0) = 0$  and since **A2** and **A3** imply that  $\Psi(x, \mu, u)$  has Lipschitz-continuous partial derivatives in a neighbourhood of  $(x^*, 0, 0)$ , (i) follows from the implicit-function theorem. Moreover, still from the implicit-function theorem, the Jacobian of the function  $x(\mu, u)$  found in (i) is given by

$$\nabla_{\mu, u} x(\mu, u) = [J(x(\mu, u)) - \nabla_x h(x(\mu, u), \mu)]^{-1} \begin{pmatrix} \nabla_\mu h(x(\mu, u), \mu) & I_n \end{pmatrix},$$

where  $I_n$  is the identity matrix of size  $n$ . If  $(\mu_1, u_1), (\mu_2, u_2) \in \mathcal{N}_\delta$ , a Taylor expansion thus yields

$$x(\mu_1, u_1) - x(\mu_2, u_2) = \nabla_{\mu, u} x(\mu_2, u_2) \begin{pmatrix} \mu_1 - \mu_2 \\ u_1 - u_2 \end{pmatrix} + O\left(\left\| \begin{pmatrix} \mu_1 - \mu_2 \\ u_1 - u_2 \end{pmatrix} \right\|^2\right). \quad (3.14)$$

Notice that if  $(\mu, u)$  is close to  $(0, 0)$ , the Jacobian  $\nabla_{\mu, u} x(\mu, u)$  remains uniformly bounded and has full-rank in a neighbourhood of  $(0, 0)$  because of **A1**–**A4**. Its smallest singular value thus remains bounded away from zero, and the first term on the right-hand side of (3.14) is therefore dominant. These two last facts combine with (3.14) to give (3.13).  $\square$



Note that because of Lemma 3.1, the exact solution  $x(\mu)$  to (2.5) is unique for all sufficiently small  $\mu$ . Therefore, we have that  $x(\mu, 0) = x(\mu)$ , with  $x(\mu, 0)$  as defined in Lemma 3.2.

Lemma 3.2 now allows us to derive the following bound on the distance between  $x^k$  and  $x^*$ .

**Lemma 3.3** Suppose that **A1–A4** hold and that  $\{x^k\}_{k \in \mathcal{K}} \rightarrow x^*$  where  $\{x^k\}$  is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5). Then, for sufficiently large  $k \in \mathcal{K}$ ,

$$\|x^k - x^*\| = O(\mu^{k-1}). \quad (3.15)$$

**Proof.** Inequality (2.9) may be written as

$$F(x^k) - h(x^k, \mu^{k-1}) = u^k \quad \text{with} \quad \|u^k\| \leq \epsilon^{k-1}.$$

For sufficiently large  $k \in \mathcal{K}$ , we have from the fact that  $\mu^k \rightarrow 0$ , (2.9) and (3.2) that  $(\mu^{k-1}, u^k)$  lies in the neighbourhood  $\mathcal{N}_\delta$  defined in Lemma 3.2. We may therefore apply (3.13) with the parameters  $(\mu^{k-1}, u^k)$  and  $(0, 0)$  to obtain

$$\|x^k - x^*\| = \Theta(\mu^{k-1} + \|u^k\|) = O(\mu^{k-1} + \epsilon^{k-1}) = O(\mu^{k-1}),$$

where we used the identities  $x(\mu^{k-1}, u^k) = x^k$ ,  $x(0, 0) = x^*$  and (3.2).  $\square$

We now give an upper bound on the size of the step  $s^k$  computed from (2.6).

**Lemma 3.4** Suppose that **A1–A5** hold, that  $\{x^k\}_{k \in \mathcal{K}} \rightarrow x^*$  where  $\{x^k\}$  is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5), and that the step  $s^k$  is computed from (2.6). Then, for sufficiently large  $k \in \mathcal{K}$ ,

$$\|s^k\| = O(\mu^{k-1}). \quad (3.16)$$

**Proof.** Assumptions **A1**, **A2** and **A5** together imply that  $G(x^k)$  is non-singular and

$$\|G^{-1}(x^k)\| \leq \kappa_1 \quad (3.17)$$

for some  $\kappa_1 > 0$  and all sufficiently large  $k \in \mathcal{K}$ , while (2.6) shows that

$$\begin{aligned} s^k &= G^{-1}(x^k) \left( h(x^k, \mu^k) - F(x^k) + r^k \right) \\ &= G^{-1}(x^k) \left( h(x^k, \mu^k) - h(x^k, \mu^{k-1}) + h(x^k, \mu^{k-1}) - F(x^k) + r^k \right), \end{aligned} \quad (3.18)$$

where  $\|r^k\| \leq \eta^k$ . In addition, **A3**, **A4**, (3.15) and the fact that  $\mu^k < \mu^{k-1}$  imply that for sufficiently large  $k \in \mathcal{K}$ ,

$$h(x^k, \mu^{k-1}) = \nabla_\mu h(x^*, 0) \mu^{k-1} + O\left((\mu^{k-1})^2\right), \quad (3.19)$$

and

$$h(x^k, \mu^k) = \nabla_{\mu} h(x^*, 0) \mu^k + O\left((\mu^{k-1})^2\right), \quad (3.20)$$

so that

$$h(x^k, \mu^k) - h(x^k, \mu^{k-1}) = O(|\mu^k - \mu^{k-1}|) + O\left((\mu^{k-1})^2\right) = O(\mu^{k-1}). \quad (3.21)$$

Combining (3.17) and (3.18) with (2.9), (3.2), (3.3) and (3.21), we thus find that

$$\begin{aligned} \|s^k\| &\leq \kappa_{\text{I}} \left( \|h(x^k, \mu^k) - h(x^k, \mu^{k-1})\| + \epsilon^{k-1} + \eta^k \right) \\ &\leq \kappa_{\text{I}} \left( O(\mu^{k-1}) + \epsilon^{k-1} + \eta^k \right) \\ &= O(\mu^{k-1}), \end{aligned} \quad (3.22)$$

which proves (3.16).  $\square$

Our next result shows the benefit of using the advanced starting point (2.3) with  $s^k$  computed from (2.6).

**Lemma 3.5** Suppose that **A1–A5** hold and that  $\{x^k\}_{k \in \mathcal{K}} \rightarrow x^*$  where  $\{x^k\}$  is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5). Then, if

$$\liminf_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \frac{\min(2, 1 + \theta_{\eta}^k)}{(1 + \theta_{\epsilon}^k)(1 + \theta_{\mu}^{k-1})} > 1, \quad (3.23)$$

it follows that

$$\|F(x_{\text{S}}^k) - h(x_{\text{S}}^k, \mu^k)\| \leq \epsilon^k, \quad (3.24)$$

for all sufficiently large  $k \in \mathcal{K}$ , where the starting point  $x_{\text{S}}^k$  is computed as in Step 1 of Algorithm 2.1.

**Proof.** It follows from **A2**, (2.3), (2.6) and **A5** that for all sufficiently large  $k \in \mathcal{K}$ ,

$$F(x_{\text{S}}^k) = F(x^k + s^k) \quad (3.25)$$

$$= F(x^k) + J(x^k)s^k + O\left(\|s^k\|^2\right) \quad (3.26)$$

$$= F(x^k) + G(x^k)s^k + [J(x^k) - G(x^k)]s^k + O\left(\|s^k\|^2\right) \quad (3.27)$$

$$= h(x^k, \mu^k) + r^k + O(\mu^{k-1}\|s^k\|) + O\left(\|s^k\|^2\right) \quad (3.28)$$

$$= h(x^k, \mu^k) + O(\eta^k) + O(\mu^{k-1}\|s^k\|) + O\left(\|s^k\|^2\right). \quad (3.29)$$

Hence (3.3), Lemma 3.4, (3.29) and the fact that  $\mu^k < \mu^{k-1}$  reveal that

$$\|F(x_{\text{S}}^k) - h(x^k, \mu^k)\| = O(\eta^k) + O\left((\mu^{k-1})^2\right) = O\left((\mu^{k-1})^{\min(2, 1 + \theta_{\eta}^k)}\right). \quad (3.30)$$

On the other hand, using **A3**, **A4** and the expansions

$$h(x^k, \mu^k) = \nabla_{\mu} h(x^*, 0) \mu^k + O\left((\mu^{k-1})^2\right) \quad \text{and} \quad h(x_{\text{S}}^k, \mu^k) = \nabla_{\mu} h(x^*, 0) \mu^k + O\left((\mu^{k-1})^2\right),$$

where we have used (3.15) and (3.16), we have

$$h(x^k, \mu^k) - h(x_s^k, \mu^k) = O\left((\mu^{k-1})^2\right). \quad (3.31)$$

Thus, combining (3.30) and (3.31), we have

$$\begin{aligned} \|F(x_s^k) - h(x_s^k, \mu^k)\| &\leq \|F(x_s^k) - h(x^k, \mu^k)\| + \|h(x^k, \mu^k) - h(x_s^k, \mu^k)\| \\ &= O\left((\mu^{k-1})^{\min(2, 1+\theta_\eta^k)}\right) + O\left((\mu^{k-1})^2\right) \\ &= O\left((\mu^{k-1})^{\min(2, 1+\theta_\eta^k)}\right). \end{aligned} \quad (3.32)$$

Equalities (3.1) and (3.2) give

$$\epsilon^k = \tau_\epsilon^k (\mu^k)^{1+\theta_\epsilon^k} = \tau_\epsilon^k (\tau_\mu^{k-1})^{1+\theta_\epsilon^k} (\mu^{k-1})^{(1+\theta_\mu^{k-1})(1+\theta_\epsilon^k)}, \quad (3.33)$$

and observe that because of (3.5), (3.23) is equivalent to

$$\min(2, 1 + \theta_\eta^k) \geq (1 + \theta_\epsilon^k)(1 + \theta_\mu^{k-1}) + \nu, \quad (3.34)$$

for all sufficiently large  $k \in \mathcal{K}$  and for some constant  $\nu > 0$ . The required bound (3.24) now follows, from (3.32), (3.33) and (3.34).  $\square$

It follows immediately from Lemma 3.5 that for all sufficiently large  $k \in \mathcal{K}$ , the starting point  $x_s^k$  satisfies the inner-iteration stopping rule (2.8) provided that the exponents  $\theta_\mu^{k-1}$ ,  $\theta_\epsilon^k$ , and  $\theta_\eta^k$  satisfy (3.23), and hence that we will choose, for all sufficiently large  $k \in \mathcal{K}$ ,

$$x^{k+1} = x_s^k, \quad (3.35)$$

since Step 2 of Algorithm 2.1 requires that the first point satisfying (2.8) is taken.

We now show that, up to at least first order, the asymptotic form (3.6) of  $x(\mu)$  is inherited by any  $x^k$  close to  $x^*$ , satisfying (2.9).

**Lemma 3.6** Suppose that **A1–A4** hold, that  $\{x^k\}_{k \in \mathcal{K}} \rightarrow x^*$  where  $\{x^k\}$  is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5) and that

$$\liminf_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \theta_\epsilon^k > 0. \quad (3.36)$$

Then for all sufficiently large  $k \in \mathcal{K}$ ,

$$x^k = x^* + \mu^{k-1}x' + o(\mu^{k-1}), \quad (3.37)$$

where  $x'$  is given by (3.7).

**Proof.** Inequality (2.9) may be written as

$$F(x^k) = h(x^k, \mu^{k-1}) + u^k, \quad \text{where } \|u^k\| \leq \epsilon^{k-1}.$$

It follows from **A2**, the fact that  $x^*$  is a root of (1.1) and (3.15), that

$$F(x^k) = J(x^*)(x^k - x^*) + w^k, \quad \text{where } \|w^k\| = O(\|x^k - x^*\|^2) = O((\mu^{k-1})^2),$$

for all sufficiently large  $k \in \mathcal{K}$ . Hence, **A1** implies that

$$x^k - x^* = J^{-1}(x^*)(h(x^k, \mu^{k-1}) + u^k - w^k). \quad (3.38)$$

On the other hand, we have from **A3**, **A4** and (3.15) that

$$h(x^k, \mu^{k-1}) = \nabla_{\mu} h(x^*, 0) \mu^{k-1} + O((\mu^{k-1})^2), \quad (3.39)$$

which combines with (3.38) to give

$$x^k - x^* = J(x^*)^{-1} \left( \nabla_{\mu} h(x^*, 0) \mu^{k-1} + u^k - w^k + O((\mu^{k-1})^2) \right). \quad (3.40)$$

Observe now that (3.36) is equivalent to  $\theta_{\epsilon}^k \geq \theta_{\epsilon}^*$  for some  $\theta_{\epsilon}^* > 0$  and for all sufficiently large  $k \in \mathcal{K}$ , which, together with (3.2) implies that  $\epsilon^{k-1} = o(\mu^{k-1})$ . Combining this observation with the identities (3.7) and (3.40), we obtain (3.37).  $\square$

Finally, we show that an estimate of the form (3.37) continues to hold at  $x^{k+1}$ .

**Lemma 3.7** Suppose that **A1–A5** hold, that  $\{x^k\}_{k \in \mathcal{K}} \rightarrow x^*$  where  $\{x^k\}$  is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5), that (3.23) and (3.36) hold, that we have

$$\liminf_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \theta_{\eta}^k > 0, \quad (3.41)$$

and

$$\limsup_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \theta_{\mu}^{k-1} < 1. \quad (3.42)$$

Then for all sufficiently large  $k \in \mathcal{K}$ ,

$$x^{k+1} = x^* + \mu^k x' + o(\mu^k), \quad (3.43)$$

where  $x'$  is given by (3.7).

**Proof.** Applying the arguments used to derive (3.20) and (3.29), it follows from (2.3), (2.6), **A2–A5**, (3.35), Lemmas 3.3 and 3.4, and the inequality  $\mu^k < \mu^{k-1}$  that

$$\begin{aligned} F(x^{k+1}) &= F(x^k + s^k) \\ &= h(x^k, \mu^k) + O(\eta^k) + O(\mu^{k-1} \|s^k\|) + O(\|s^k\|^2) \\ &= \nabla_{\mu} h(x^*, 0) \mu^k + O(\eta^k) + O((\mu^{k-1})^2), \end{aligned} \quad (3.44)$$

for sufficiently large  $k \in \mathcal{K}$ . On the other hand, using **A2**, the identity  $F(x^*) = 0$ , (3.15), (3.16) and the inequality  $\|x^{k+1} - x^*\| \leq \|x^k - x^*\| + \|s^k\| = O(\mu^{k-1})$ , a second-order Taylor

expansion gives

$$\begin{aligned} F(x^{k+1}) &= F(x^*) + J(x^*)(x^{k+1} - x^*) + O(\|x^{k+1} - x^*\|^2) \\ &= J(x^*)(x^{k+1} - x^*) + O((\mu^{k-1})^2). \end{aligned} \quad (3.45)$$

Observe that because of (3.5), (3.42) is equivalent to  $0 \leq \theta_\mu^{k-1} \leq \theta_\mu^*$  for some  $\theta_\mu^* < 1$  and for all sufficiently large  $k \in \mathcal{K}$ , which combines with (3.1) to give  $(\mu^{k-1})^2 = o(\mu^k)$ . Similarly, (3.41) is equivalent to  $\theta_\eta^k \geq \theta_\eta^*$  for some  $\theta_\eta^* > 0$  and for all sufficiently large  $k \in \mathcal{K}$ , which, together with (3.3) ensures that  $\eta^k = \tau_\eta^k (\mu^k)^{1+\theta_\eta^k} = o(\mu^k)$ . Using these observations and combining (3.44) and (3.45), **A1** implies that

$$\begin{aligned} x^{k+1} - x^* &= \mu^k J(x^*)^{-1} \nabla_\mu h(x^*, 0) + O(\eta^k) + O((\mu^{k-1})^2) \\ &= \mu^k J(x^*)^{-1} \nabla_\mu h(x^*, 0) + o(\mu^k), \end{aligned}$$

which completes the proof.  $\square$

Note that there is a fundamental difference between Lemma 3.6 and Lemma 3.7, which is that (3.37) is concerned with an iterate  $x^k$  belonging to a converging subsequence indexed by  $\mathcal{K}$ , while in (3.43), there is no assumption that  $k+1 \in \mathcal{K}$ . This fact is the key to the proof of our main result.

**Theorem 3.8** Suppose that **A1–A5** hold, that  $\{x^k\}_{k \in \mathcal{K}} \rightarrow x^*$  where  $\{x^k\}$  is a sequence generated by Algorithm 2.1 using the updating rules (3.1)–(3.5) and that (3.23), (3.36), (3.41) and (3.42) are satisfied. Then the *complete* sequence  $\{x^k\}$  converges to  $x^*$  and for all sufficiently large  $k$ , we have

$$(i) \quad \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^{1+\theta_\mu^{k-1}}} = \Theta(1) \text{ provided that } x' \neq 0; \text{ and}$$

$$(ii) \quad \frac{|x_i^{k+1} - x_i^*|}{|x_i^k - x_i^*|^{1+\theta_\mu^{k-1}}} = \Theta(1) \text{ for every component } i = 1, \dots, n \text{ for which } x_i' \neq 0.$$

As a consequence, the entire sequence  $\{x^k\}$  converges to  $x^*$  at a (componentwise) Q-superlinear rate determined by  $\limsup_{k \rightarrow \infty} \theta_\mu^k$ .

**Proof.** It follows from Lemma 3.7 and the fact that  $\mu^k \rightarrow 0$  that the sequence  $\{x^{k+1}\}_{k \in \mathcal{K}}$  also converges to  $x^*$ . Applying again Lemma 3.6 and Lemma 3.7 to the subsequence of  $\{x^k\}$  defined by the index set  $\mathcal{K}^+ = \mathcal{K} \cup \{k+1 \mid k \in \mathcal{K}\}$ , we conclude that the sequence  $\{x^{k+1}\}_{k \in \mathcal{K}^+}$  also converges to  $x^*$ . An inductive argument thus shows that the whole sequence  $\{x^k\}$  converges to  $x^*$ . To prove (i), let us assume that  $x' \neq 0$ . We then have from Lemma 3.6 and Lemma 3.7 that  $\|x^k - x^*\| = \Theta(\mu^{k-1})$  and  $\|x^{k+1} - x^*\| = \Theta(\mu^k)$ , which we combine with (3.1) to obtain the desired result. We prove (ii) similarly, by considering equations (3.37) and (3.43) componentwise.  $\square$

From Theorem 3.8, the asymptotic componentwise Q-rate of convergence of  $\{x^k\}$  to  $x^*$  is determined by the upper limit  $\limsup_{k \rightarrow \infty} \theta_\mu^k$ . Under the assumption that  $x_i' \neq 0$  for all  $i = 1, \dots, n$ , and according to the restrictions (3.23), (3.36), (3.41) and (3.42), a componentwise Q-rate of convergence arbitrarily close to 2 is achieved by choosing  $\{\theta_\mu^k\}$  and  $\{\theta_\epsilon^k\}$  so that  $\limsup_{k \rightarrow \infty} \theta_\mu^k = 1 - \alpha$ ,  $\liminf_{k \rightarrow \infty} \theta_\epsilon^k = \beta$  and  $\theta_\eta^k > 1$  for all sufficiently large  $k$ , where  $\alpha, \beta \in (0, 1)$  can be chosen as small as desired. A Q-rate of (just under) 1.5 may be obtained by setting  $\theta_\mu^k = 0.5$ ,  $\theta_\epsilon^k = 0.01$  for all  $k$  and  $\theta_\eta^k$  such that  $\liminf_{k \rightarrow \infty} \theta_\eta^k > 0.515$ .

## 4 Discussion

So far, we have not commented on appropriate choices for  $h(x, \mu)$ . Perhaps the most obvious and simplest choice is the linear function

$$h(x, \mu) = \mu e, \quad (4.1)$$

where  $e$  is the vector of all ones, for which  $\nabla_\mu h(x^*, 0) = e$ . A similar choice is made in primal-dual interior-point methods where the complementarity conditions are perturbed and where we choose

$$h(x, \mu) = \mu \tilde{e}, \quad (4.2)$$

where  $\tilde{e}$  is made of zeros and ones. We refer the interested reader to [6] for more details on a primal-dual interior-point method in relation with the componentwise convergence of the iterates. In view of Theorem 3.8, ideally we would like to use an  $h(x, \mu)$  for which every component of  $x'$  is nonzero. We have no such guarantee if we use (4.1) or (4.2). For instance,

$$\nabla_\mu h(x^*, 0) = J(x^*)e, \quad (4.3)$$

such as would arise if

$$h(x, \mu) = \mu J(x^*)e, \quad (4.4)$$

would be ideal for then  $x' = e$ , but of course depends on the (unknown)  $J(x^*)$ . If  $F$  is twice continuously differentiable, the choice

$$h(x, \mu) = \mu J(x)e, \quad (4.5)$$

which also aims for  $x' = e$ , is more similar to a primal interior-point method, and differs from (4.4) in a number of respects. In the framework of interior methods, the most obvious of these is probably that if  $\nabla_x[F(x) - h(x, \mu)]$  is independent of  $\mu$ , then so is the region of convergence of the method (2.6) [1, Lemma 2.1], [8, §3.2.12]; for (4.2), the size of this region is  $\Theta(1)$  while for (4.5), it is  $\Theta(\mu)$  and shrinks as convergence occurs [9]. This behaviour is observed in the numerical tests of Section 5. A particular form of  $h(x, \mu)$  may emerge naturally from the problem itself, or may need to have properties dictated by the problem, for instance, a higher level of nonlinearity in  $\mu$ . See for instance [10] for references to such work. If we are only interested in the convergence properties of certain components then, in view of Theorem 3.8 (ii), the perturbation function  $h(x, \mu)$  may be chosen such that the asymptotic behaviour of these components is controlled, *i.e.* such that the corresponding components of the vector  $x'$  are nonzero. In any case, we certainly want to choose  $h(x, \mu)$  and  $\mu^0$  such that a solution to  $F(x) = h(x, \mu^0)$  is either known or easy to find.

Interestingly, our approach can easily be applied to the framework of nonlinear least-squares problems, where one solves

$$\min_x \frac{1}{2} \|R(x)\|_2^2 \quad (4.6)$$

for some smooth function  $R$  from  $\mathbb{R}^n$  into  $\mathbb{R}^m$  with  $m \geq n$ . If we assume that the Jacobian of  $R$ ,  $\nabla_x R(x)$  is of full rank at the solution  $x^*$  of (4.6), the normal equations then provide a link to our approach. Indeed, we may replace (2.8) by

$$\|\nabla_x R(x^{k+1})^T R(x^{k+1}) - h(x^{k+1}, \mu^k)\| \leq \epsilon^k$$

and compute the step to  $x_s^k$  using

$$G(x^k)s^k = h(x^k, \mu^k) - \nabla_x R(x^k)^T R(x^k) + r^k \quad \text{and} \quad \|r^k\| \leq \eta^k,$$

where now  $G(x^k) \approx \nabla_x R(x^k)^T \nabla_x R(x^k)$ .

We finally briefly comment on the assumptions we made in §2. While the first part of Assumption **A4** is clearly implied by our intent to use a homotopy method, the second part limits the level of dependence of  $h$  on  $x$ , at least close to  $x^*$ , which is in line with our initial motivation where  $h$  is independent of  $x$ . It is easily verified that the examples of this section and of §5 satisfy **A4**. Regarding **A5**, one could argue that having the right-hand side of (2.7) depend on  $\mu$  is not intuitive since its left-hand side does not. In view of Lemma 3.4, it is possible to replace the right-hand side of (2.7) by  $\kappa_G \|s^k\|^2$ —which would be stronger since the bound  $\mu^{k-1} = O(\|s^k\|)$  does not necessarily hold—or even by  $\kappa_G \|s^k\|^{1+\theta_s^k}$  for some  $\theta_s^k \geq \theta_s^* > 0$ . This last choice would however slightly modify condition (3.23) which we would need to replace by

$$\liminf_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \frac{\min(2, 1 + \theta_\eta^k, 1 + \theta_s^k)}{(1 + \theta_\epsilon^k)(1 + \theta_\mu^{k-1})} > 1. \quad (4.7)$$

A weaker assumption would be to replace the right-hand side of (2.7) by  $\kappa_G (\mu^{k-1})^2$ , or, weaker still,  $\kappa_G (\mu^{k-1})^{1+\theta_G^k}$  for some  $\theta_G^k \geq \theta_G^* > 0$ , in which case (3.23) becomes

$$\liminf_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \frac{\min(2, 1 + \theta_\eta^k, 1 + \theta_G^k)}{(1 + \theta_\epsilon^k)(1 + \theta_\mu^{k-1})} > 1.$$

In this last case again, the left and right-hand sides of (2.7) would not involve quantities that are directly comparable.

## 5 Numerical experiments

In this section we illustrate the above results on Example 1.1 when  $n = 5$  and when  $G(x^k) = J(x^k)$  for all  $k$ . The starting point is chosen as  $x^0 = 0.8 e_3$ . The following figures indicate the values taken by the components of the vector  $x$  and by  $\|x\|_2$  along the iterations when solving the problem given in Example 1.1 for the indicated functions  $h(x, \mu)$ . The first row of each table indicates the starting values and the iteration counter increases when reading the table from top to bottom. In Figure 5.1, a pure Newton scheme is applied, which corresponds to  $h(x, \mu) = 0$ , and we observe the behaviour described in Example 1.1. In this and subsequent experiments, all computation is performed in extended-precision arithmetic under Mathematica, so as to see the

predicted rates of convergence. Figure 5.2 corresponds to the case where we set  $h(x, \mu) = \mu J(x^*)e$ , where  $\mu^0 = 0.9$  and is updated according to  $\mu^{k+1} = (\mu^k)^{1.9}$ , which corresponds to using  $\theta_\mu^k = 0.9$  for all  $k$ . We also use the values  $\epsilon^0 = \|F(x^0) - h(x^0, \mu^0)\|$ ,  $\theta_\epsilon^k = 0.05$ ,  $r^k = 0$ —i.e. the system (2.6) is solved exactly—and  $\tau_{\min} = \tau_\mu^k = \tau_\epsilon^k = \tau_\eta^k = \tau_{\max} = 1$  for all  $k$ . In this case, a single step (2.6) suffices for each value of  $\mu^k$ , because  $x^k$  happens to lie inside the region of convergence of Newton’s method for all  $k$ . In Figure 5.3, we set  $h(x, \mu) = \mu J(x)e$  and use the same rules for  $\mu$ . In the inner iteration, we choose to take steps given by (2.6) as well. This time, two steps (2.6) are required for each of the first 7 iterations—a possible reason for this is that the region of convergence of Newton’s method shrinks faster as  $x^*$  is approached than when using  $h(x, \mu) = \mu J(x^*)e$ . For the remaining iterations, a single step is sufficient. Notice that a single step suffices precisely where the (nearly) quadratic convergence shows up. It is also worth noticing that  $\|x\|_2$  converges almost quadratically to zero in all cases, but that this behaviour appears later in the course of the iterations whenever  $h(x, \mu) \neq 0$ .

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$\ x\ _2$
0	0	0.8	0	0	0.8
0	0	0	0.64	0	0.64
0	0	0	0	0.4096	0.4096
0.16777216	0	0	0	0	0.16777216
0	0.02814749767	0	0	0	0.02814749767
0	0	0.0007922816251	0	0	0.0007922816251
0	0	0	6.277101735e-7	0	6.277101735e-7
0	0	0	0	3.94020062e-13	3.94020062e-13
1.552518092e-25	0	0	0	0	1.552518092e-25
0	2.410312427e-50	0	0	0	2.410312427e-50
0	0	5.809605995e-100	0	0	5.809605995e-100
0	0	0	3.375152182e-199	0	3.375152182e-199

Figure 5.1:  $h(x, \mu) = 0$ 

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$\ x\ _2$
0	0	0.8	0	0	0.8
0.8185793185	0.8185793185	0.8185793185	0.1488524089	0.8185793185	1.643911628
0.4926351154	0.5471713671	0.4578868715	0.6040597546	0.5259466409	1.180361938
0.3392381302	0.3939035662	0.3537864395	0.3711274953	0.4019772598	0.8335038742
0.2255449862	0.2153839477	0.2388073667	0.2095195862	0.2355470167	0.5036572229
0.09159450042	0.08317375999	0.08418223925	0.09044286051	0.07961993546	0.1921283982
0.01130186677	0.01335473277	0.01173190562	0.01214697118	0.01301825532	0.02758156213
0.0002449105253	0.000203454507	0.000254172933	0.0002129319505	0.000223634159	0.0005112054717
6.691793911e-8	7.688192705e-8	5.829599195e-8	8.150778483e-8	6.223884381e-8	1.558988667e-7
5.590081365e-15	6.194418554e-15	7.627238484e-15	5.114830515e-15	8.359926894e-15	1.496143264e-14
1.57772181e-28	1.191179967e-28	1.277954666e-28	1.459392675e-28	1.135959704e-28	2.993579025e-28
0.	0.	0.	0.	0.	0.

Figure 5.2:  $h(x, \mu) = \mu J(x^*)e$



$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$\ x\ _2$
0	0	0.8	0	0	0.8
0.988252	1.21087	0.846408	1.44309	1.12423	2.55062
0.806001	1.25925	0.807841	1.09401	1.20735	2.35426
0.448428	0.833458	0.822022	0.514291	1.03327	1.704
0.783414	-0.000733957	0.557764	0.387957	0.240596	1.06454
0.0796238	0.10213	0.322878	0.0421053	0.24526	0.427722
0.00702223	0.0119832	0.0070138	0.00736868	0.0177941	0.0247594
0.0000812764	0.000081415	0.0000812741	0.0000813083	0.0000812747	0.000181814
$2.3538e-8$	$2.35383e-8$	$2.35609e-8$	$2.35379e-8$	$2.35435e-8$	$5.26454e-8$
$2.27071e-15$	$2.27045e-15$	$2.27046e-15$	$2.27152e-15$	$2.27044e-15$	$5.07748e-15$
$9.348e-29$	$9.348e-29$	$9.348e-29$	$9.348e-29$	$9.348e-29$	$2.09028e-28$
0.	0.	0.	0.	0.	0.

Figure 5.3:  $h(x, \mu) = \mu J(x)e$ 

## 6 Conclusion

In this paper, we have studied a homotopy-type method for solving square systems of non-singular, nonlinear equations. The method exhibits an asymptotic convergence rate that is essentially as fast as that of Newton's method but has the property that all the components of the iterates converge to their limit at comparable rates. This property may be desirable in a number of practical applications. The results are illustrated on an example which exhibits a normwise Q-quadratic convergence which does not occur componentwise when a pure Newton scheme is applied. The method relies on the computation of a suitable starting point for the modified Newton iteration, which turns out to readily satisfy the stopping conditions. Asymptotically, a single step is then sufficient and leads to a componentwise Q-superlinear convergence rate, which may be as close to quadratic as desired. It is noticeable that the results do not depend on how the exponents  $\theta_\epsilon^k$ ,  $\theta_\mu^k$  and  $\theta_\eta^k$  are updated, merely that they should satisfy the inequalities (3.23), (3.36), (3.41) and (3.42). The method leaves freedom to the user in two major respects. The first, and most important, is the choice of the perturbation function  $h(x, \mu)$  and its possible dependence on  $x$ , subject to **A4**. The second is to permit inexact Jacobians. The method presented in this paper was inspired by similar behaviour in primal-dual interior-point methods [6] and exterior penalty methods [5] for constrained optimization. It is still an issue to determine whether a particular  $h(x, \mu)$  is advantageous, or should be avoided; in particular, there might or might not exist a perturbation function  $h(x, \mu)$  that depends on  $x$  and which is better than the usual perturbation in primal interior-point methods for linear and nonlinear programming. The application of the technique presented in this paper to nonlinear least-squares problems is also the object of continuing research.

## Acknowledgements

The authors wish to thank John Dennis for his useful comments on earlier drafts of this work.

## Appendix.

We justify the claims made in Example 1.1.

**Theorem A.0.1** A single Newton iteration for the nonlinear system (1.4) from the point  $x^k = x_l e_l$ , for some index  $1 \leq l \leq n$ , results in the point  $x^{k+1}$  satisfying (1.5).

**Proof.** On writing (1.4) as

$$F(x) = \begin{pmatrix} x_1^2 + x_2 \\ x_2^2 + x_3 \\ \dots \\ x_{n-1}^2 + x_n \\ x_n^2 + x_1 \end{pmatrix} = 0,$$

it follows that

$$J(x) = \begin{pmatrix} 2x_1 & 1 & & & \\ & 2x_2 & 1 & & \\ & & \cdot & \cdot & \\ & & & 2x_{n-1} & 1 \\ 1 & & & & 2x_n \end{pmatrix} = P + 2 \sum_{i=1}^n x_i e_i e_i^T,$$

where  $P$  is the permutation matrix

$$P = \begin{pmatrix} & & & & 1 \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ 1 & & & & \end{pmatrix}.$$

At the point of interest,

$$J(x_l e_l) = P + 2x_l e_l e_l^T$$

and hence the Sherman-Morrison-Woodbury formula along with the fact that  $P$  is a permutation, and thus that  $P^{-1} = P^T$ , reveals that

$$J^{-1}(x_l e_l) = P^T - \frac{2x_l}{1 + 2x_l e_l^T P^T e_l} P^T e_l (P e_l)^T.$$

Note that

$$P e_l = \begin{cases} e_n & \text{for } l = 1 \\ e_{l-1} & \text{for } 2 \leq l \leq n \end{cases} \quad \text{and} \quad P^T e_l = \begin{cases} e_{l+1} & \text{for } 1 \leq l \leq n-1 \\ e_1 & \text{for } l = n. \end{cases}$$

There are now three cases to consider.

(i). If  $l = 1$ ,

$$J^{-1}(x_1 e_1) = P^T - \frac{2x_1}{1 + 2x_1 e_1^T e_2} e_2 e_n^T = P^T - 2x_1 e_2 e_n^T.$$

But, since  $F(x_1 e_1) = x_1^2 e_1 + x_1 e_n$ , the Newton correction is

$$-J^{-1}(x_1 e_1)F(x_1 e_1) = -(P^T - 2x_1 e_2 e_n^T)(x_1^2 e_1 + x_1 e_n) = x_1^2 e_2 - x_1 e_1,$$

and hence  $x^{k+1} = x_1^2 e_2$ .

(ii). Secondly, if  $1 < l < n$ ,

$$J^{-1}(x_l e_l) = P^T - \frac{2x_l}{1 + 2x_l e_l^T e_{l+1}} e_{l+1} e_{l-1}^T = P^T - 2x_l e_{l+1} e_{l-1}^T.$$

Since  $F(x_l e_l) = x_l^2 e_l + x_l e_{l-1}$ , the Newton correction in this case is

$$-J^{-1}(x_l e_l)F(x_l e_l) = -(P^T - 2x_l e_{l+1} e_{l-1}^T)(x_l^2 e_l + x_l e_{l-1}) = x_l^2 e_{l+1} - x_l e_l,$$

and hence  $x^{k+1} = x_l^2 e_{l+1}$ .

(iii). Finally, if  $l = n$ ,

$$J^{-1}(x_n e_n) = P^T - \frac{2x_n}{1 + 2x_n e_n^T e_1} e_1 e_{n-1}^T = P^T - 2x_n e_1 e_{n-1}^T.$$

Since  $F(x_n e_n) = x_n^2 e_n + x_n e_{n-1}$ , the Newton correction is now

$$-J^{-1}(x_n e_n)F(x_n e_n) = -(P^T - 2x_n e_1 e_{n-1}^T)(x_n^2 e_n + x_n e_{n-1}) = x_n^2 e_1 - x_n e_n,$$

and hence  $x^{k+1} = x_n^2 e_1$ . □

## References

- [1] R. H. Byrd, G. Liu, and J. Nocedal. On the local behavior of an interior point method for nonlinear programming. In D. J. Higham D. F. Griffiths and G. A. Watson, editors, *Numerical Analysis 1997*, number 380 in Pitman Research Notes in Mathematics Series, pages 37–56, Harlow, Essex, England, 1998. Longman Scientific & Technical.
- [2] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust Region Methods*. SIAM, Philadelphia, USA, 2000.
- [3] J. E. Dennis and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. Prentice-Hall, Englewood Cliffs, USA, 1983. Reprinted as *Classics in Applied Mathematics 16*, SIAM, Philadelphia, USA.
- [4] J.-P. Dussault. Numerical stability and efficiency of penalty algorithms. *SIAM Journal on Numerical Analysis*, 32(1):296–317, 1995.
- [5] N. I. M. Gould. On the convergence of a sequential penalty function method for constrained minimization. *SIAM Journal on Numerical Analysis*, 26(1):107–126, 1989.
- [6] N. I. M. Gould, D. Orban, A. Sartenaer, and Ph. L. Toint. Superlinear Convergence of Primal-Dual Interior Point Algorithms for Nonlinear Programming. Technical Report TR/PA/00/20, CERFACS, Toulouse, France, 2000.

- [7] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer series in Operations Research. Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA, 1999.
- [8] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*. Academic Press, New York, 1970.
- [9] M. C. Villalobos, R. A. Tapia, and Y. Zhang. The Sphere of Convergence of Newton's Method on Two Equivalent Systems from Nonlinear Programming. Technical Report TR9913, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA, 1999.
- [10] L. T. Watson. Numerical linear algebra aspects of globally convergent homotopy methods. *SIAM Review*, 28(4):529–545, 1986.
- [11] L. T. Watson. Globally convergent homotopy methods: A tutorial. *Appl. Math. Comput.*, 31BK:369–396, 1989.
- [12] M. H. Wright. Why a pure primal Newton barrier step may be infeasible. *SIAM Journal on Optimization*, 5(1):1–12, 1995.
- [13] S. J. Wright and F. Jarre. The role of linear objective functions in barrier methods. *Mathematical Programming*, 84(2):357–373, 1999.
- [14] W. I. Zangwill and C. B. Garcia. *Pathways to solutions, fixed points and equilibria*. Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1981.