Math. Program., Ser. A (2009) 119:1–32
DOI 10.1007/s10107-007-0199-0

FULL LENGTH PAPER

An affine-scaling interior-point CBB method for box-constrained optimization

William W. Hager · Bernard A. Mair · Hongchao Zhang

Received: 22 March 2007 / Accepted: 12 October 2007 / Published online: 20 November 2007

Abstract We develop an affine-scaling algorithm for box-constrained optimization which has the property that each iterate is a scaled cyclic Barzilai–Borwein (CBB) gradient iterate that lies in the interior of the feasible set. Global convergence is established for a nonmonotone line search, while there is local R-linear convergence at a nondegenerate local minimizer where the second-order sufficient optimality conditions are satisfied. Numerical experiments show that the convergence speed is insensitive to problem conditioning. The algorithm is particularly well suited for image restoration problems which arise in positron emission tomography where the cost function can be infinite on the boundary of the feasible set.

Keywords Interior-point · Affine-scaling · Cyclic Barzilai–Borwein methods · CBB · PET · Image reconstruction · Global convergence · Local convergence

Mathematics Subject Classification (2000) 90C06 · 90C26 · 65Y20

This material is based upon work supported by the National Science Foundation under Grants 0203270, 0619080, and 0620286.

W. W. Hager (✉) · B. A. Mair
Department of Mathematics, University of Florida, PO Box 118105, Gainesville, FL 32611-8105, USA
e-mail: hager@math.ufl.edu
URL: http://www.math.ufl.edu/∼hager

B. A. Mair
e-mail: bam@math.ufl.edu
URL: http://www.math.ufl.edu/∼bam

H. Zhang
Institute for Mathematics and its Applications (IMA), University of Minnesota,
400 Lind Hall, 207 Church Street S.E., Minneapolis, MN 55455-0436, USA
e-mail: hozhang@ima.umn.edu

1 Introduction

In this paper we develop an interior point algorithm for the box-constrained optimization problem

$$\min \{f(\mathbf{x}) : \mathbf{x} \in \mathcal{B}\}, \tag{1.1}$$

where f is a real-valued, continuously differentiable function defined on the set

$$\mathcal{B} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{l} \le \mathbf{x} \le \mathbf{u}\}. \tag{1.2}$$

Here $\mathbf{l} < \mathbf{u}$ and possibly, $l_i = -\infty$ or $u_i = \infty$. Initially, to simplify the exposition, we will focus on the special case

$$\min \{f(\mathbf{x}) : \mathbf{x} \ge \mathbf{0}\}. \tag{1.3}$$

The algorithm starts at a point $\mathbf{x}_1$ in the interior of the feasible set, and generates a sequence $\mathbf{x}_k$, $k \ge 2$, by the following rule:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + s_k \mathbf{d}_k \tag{1.4}$$

where $s_k \in (0, 1]$ is a positive stepsize and the ith component of $\mathbf{d}_k$ is given by

$$d_{ki} = -\left(\frac{1}{\lambda_k + g_i^+(\mathbf{x}_k)/x_{ki}}\right) g_i(\mathbf{x}_k). \tag{1.5}$$

Here λ_k is a positive scalar, $g_i(\mathbf{x})$ is the ith component of the gradient $\nabla f(\mathbf{x})$, and $t^+ = \max\{0, t\}$ for any scalar t. We compute λ_k using a cyclic version of the Barzilai–Borwein (BB) stepsize rule [2] in which the same BB step is reused in several iterations. We call the algorithm (1.4)–(1.5) with this CBB choice for λ_k the affine scaling cyclic Barzilai–Borwein method (AS_CBB).

We now motivate the search direction $\mathbf{d}_k$ in (1.5). The first-order optimality conditions (KKT conditions) for (1.3) can be expressed in the following way:

$$\mathbf{X}^1(\mathbf{x}) \circ \mathbf{g}(\mathbf{x}) = \mathbf{0} \quad \text{and} \quad \mathbf{x} \ge \mathbf{0}, \tag{1.6}$$

where

$$X_i^1(\mathbf{x}) = \begin{cases} 1 & \text{if} \quad g_i(\mathbf{x}) \le 0, \\ x_i & \text{otherwise.} \end{cases} \tag{1.7}$$

Here "$\circ$" denotes the Hadamard (or component-wise) product of two vectors in $\mathbb{R}^n$. That is, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\mathbf{x} \circ \mathbf{y} \in \mathbb{R}^n$ and $(\mathbf{x} \circ \mathbf{y})_i = x_i y_i$, where x_i is the ith component of $\mathbf{x}$. For a convex optimization problem, the KKT conditions (1.6) are necessary and sufficient for optimality. Our algorithm for solving (1.3) is an iterative method to find a nonnegative solution to the nonlinear equation $\mathbf{X}^1(\mathbf{x}) \circ \mathbf{g}(\mathbf{x}) = \mathbf{0}$.

In Newton's method for solving $\mathbf{X}^1(\mathbf{x}) \circ \mathbf{g}(\mathbf{x}) = \mathbf{0}$, the Newton direction $\mathbf{d}_k$ is the solution of the linearized equation

$$\mathbf{H}(\mathbf{x}_k)\,\mathbf{d}_k = -\mathbf{X}^1(\mathbf{x}_k) \circ \mathbf{g}(\mathbf{x}_k), \tag{1.8}$$

where

$$\mathbf{H}(\mathbf{x}_k) = \mathrm{diag}(\mathbf{X}^1(\mathbf{x}_k))\,\nabla^2 f(\mathbf{x}_k) + \mathrm{diag}(\mathbf{g}^+(\mathbf{x}_k)). \tag{1.9}$$

Here $\mathrm{diag}(\mathbf{x})$ is an n by n diagonal matrix with ith diagonal element x_i and $\mathbf{g}^+$ is the vector whose i-component is g_i^+. In situations where $\nabla^2 f(\mathbf{x})$ is a huge, dense matrix, it can be time consuming to solve the linear system (1.8). For the approximation $\nabla^2 f(\mathbf{x}_k) \approx \lambda_k \mathbf{I}$, obtained by a quasi-Newton method for example, the corresponding approximation to the Newton search direction reduces to (1.5).

This algorithm emerged in the context of image reconstruction for positron emission tomography (PET) [7,6,22,24,25] where there is a large data set, and $\nabla^2 f(\mathbf{x})$ is a huge, relatively dense matrix. The penalized maximum likelihood reconstruction problem in PET imaging is equivalent to minimizing the following objective function:

$$\min_{\mathbf{x}\geq\mathbf{0}} \; f(\mathbf{x}) := \sum_{j=1}^{m}([\mathbf{A}\mathbf{x}]_j - b_j \log\,[\mathbf{A}\mathbf{x}]_j) + \; P(\mathbf{x}), \tag{1.10}$$

where $\mathbf{A}$ is an m by n probability matrix (columns are nonnegative and sum to 1, row sums are strictly positive), $\mathbf{x} \in \mathbb{R}^n$ represents the unknown image, $\mathbf{b} \in \mathbb{R}^m$ ($\mathbf{b} \geq \mathbf{0}$) is the emission data, and $P(\mathbf{x})$ is a convex (smoothing) penalty term. The penalty term is used to smooth the reconstructed image. Our convention is that $\log(0) = -\infty$. Due to the log term in the cost function for (1.10), the cost function could be infinite when an algorithm generates an iterate on the boundary of the feasible set. An advantage of AS_CBB is that the iterates always lie in the interior of the feasible set; consequently, no modification of the algorithm is needed to prevent an undefined value for the cost function.

The BB method is a quasi-Newton method in which the Hessian $\nabla^2 f(\mathbf{x}_k)$ in Newton's method $\mathbf{x}_{k+1} = \mathbf{x}_k - \nabla^2 f(\mathbf{x}_k)^{-1}\mathbf{g}(\mathbf{x}_k)$ is replaced by $\lambda_k \mathbf{I}$ where λ_k, for $k \geq 2$, is the solution to

$$\min_{\lambda\in\mathbb{R}} \;\|\lambda \mathbf{s}_{k-1} - \mathbf{y}_{k-1}\|_2.$$

Here $\mathbf{s}_{k-1} = \mathbf{x}_k - \mathbf{x}_{k-1}$, $\mathbf{y}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1}$, and $\mathbf{g}_k = \mathbf{g}(\mathbf{x}_k)$. To achieve global convergence, we need to bound the denominator in (1.5) away from zero. This leads to a modified formula

$$\lambda_k^{\mathrm{BB}} := \arg\min_{\lambda\geq\lambda_0} \;\|\lambda \mathbf{s}_{k-1} - \mathbf{y}_{k-1}\|_2 = \max\left\{\lambda_0, \frac{\mathbf{s}_{k-1}^{\mathsf{T}}\mathbf{y}_{k-1}}{\mathbf{s}_{k-1}^{\mathsf{T}}\mathbf{s}_{k-1}}\right\}, \tag{1.11}$$

where $k \geq 2$ and $\lambda_0 > 0$ is a fixed parameter. The starting parameter value λ_1^{BB} can be chosen freely, subject to the constraint $\lambda_1 \geq \lambda_0$; for example,

$$\lambda_1^{BB} = \max\{\lambda_0, \ \|\mathbf{g}_1\|_\infty\}.$$

The BB method is superior to the classic steepest descent method in both theory and practice [3,4,14,17,19]. For two-dimensional strongly convex quadratics, R-superlinear convergence was established [2]. The convergence speed increases when the Hessian becomes ill conditioned, unlike steepest descent [1]. In this paper, we will employ the cyclic BB method (CBB) in which the same BB step is reused several iterations. CBB is a special case of a general class of gradient-based algorithms first presented in [18]. The analysis of [18] yields global convergence for CBB when f is a strongly convex quadratic. In [11] R-linear convergence of CBB is established for general convex quadratics, while [13] establishes local R-linear convergence for general nonlinear objective functions. BB methods are locally convergent, not globally convergent, for general nonlinear objective functions. Global convergence is achieved by using a nonmonotone line search [5,13,14,17,26].

Recently, BB-type methods have been applied to constrained optimization. In [12, 20,28] the authors use gradient projection and active set techniques to extend BB-type methods to box constrained optimization. Our scheme (1.4)–(1.5) is an extension of the CBB method to constrained optimization which is closer in spirit to the affine scaling methods where the iterates are always in the interior of the feasible region. Affine-scaling was first proposed in [15] for linear and quadratic programming. It has been extensively developed by Coleman and Li [8–10] and others (see [16,21,23], for example). Most of these methods are Newton or trust-region type methods which require either the evaluation of Hessian or the solution of a linear system of equations in each iteration. The terminology "affine-scaling" which is used to describe these methods in [21] and [23], is also employed in our paper. In [21,23] the authors write the first-order optimality condition as a nonlinear equation which is a product between a nonlinear scaling matrix and the gradient of the objective function. Different choices for the nonlinear scaling matrix lead to different algorithms. Our AS_CBB method corresponds to a diagonal nonlinear scaling matrix with $\mathbf{X}^1(\mathbf{x})$, defined in (1.7), on the diagonal.

Recently, Zhang [29] proposed a general framework for monotone affine-scaling interior-point gradient methods for box constrained optimization, however, an efficient implementation of his methods are an open problem. When the dimension is very large, which often occurs in medical imaging research, evaluating the Hessian and solving the large system of equations which arise in affine scaling methods is time consuming, unless the problem has special structure. The CBB affine scaling method which we introduce does not require the solution of a linear system. Hence, the iterations can be performed relatively quickly.

The paper is organized as follows. In Sect. 2 we introduce the cyclic BB method and our nonmonotone line search. Since the BB method does not monotonically reduce the value of the cost function, a nonmonotone line search is needed to ensure that the line search is not truncated when the algorithm is converging [13,14,26]. In Sect. 3 various continuity properties for the AS_CBB method are established. In Sect. 4 we show that when λ_k is uniformly bounded away from 0 and ∞, then either the affine scaling algorithm (1.4)–(1.5) terminates at a KKT point in a finite number of iterations, or the KKT conditions (1.6) are satisfied in an asymptotic sense. As a special case, this yields global convergence for the BB choice of λ_k given in (1.11).

In Sects. 5–7, we establish local R-linear convergence of AS_CBB, at a nondegenerate local minimizer $\mathbf{x}^*$ which satisfies the second-order sufficient optimality condition. Nondegeneracy means that $g_i(\mathbf{x}^*) > 0$ whenever $x_i^* = 0$. The second-order sufficient optimality condition is that there exist $\alpha > 0$ such that

$$\mathbf{d}^\mathsf{T} \nabla^2 f(\mathbf{x}^*)\mathbf{d} \geq \alpha \|\mathbf{d}\|^2 \tag{1.12}$$

for all $\mathbf{d} \in \mathbb{R}^n$ with $d_i = 0$ when $x_i^* = 0$. A suitable choice for λ_0 in (1.11) is any positive scalar strictly smaller than α. In Sect. 5 we review our R-linear convergence results for the CBB method, and we show that the iterate components corresponding to active constraints converge to zero at a Q-quadratic rate. In Sect. 6 we develop comparison results between the convergence of the CBB iterates and the convergence of the AS_CBB iterates. R-linear convergence for AS_CBB is established in Sect. 7.

Two types of numerical experiments are presented in Sect. 9. In the first experiments, we observe that the convergence speed is relatively insensitive to problem conditioning. In the second experiment, we compare the performance of AS_CBB to that of a conjugate gradient-based active set algorithm. We observe that the AS_CBB scheme is initially faster than the conjugate gradient algorithm, however, as the iterations converge, the conjugate gradient algorithm is asymptotically faster.

Notation Let $\mathbb{R}^n_+$ denote the positive orthant defined by

$$\mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}.$$

For any scalar t, $t^+ = \max\{0, t\}$, while for any vector $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}^+$ is the vector whose ith component is v_i^+. The gradient of $f(\mathbf{x})$, arranged as a column vector, is $\mathbf{g}(\mathbf{x})$. The Hadamard (or component-wise) product $\mathbf{x} \circ \mathbf{y}$ of two vectors $\mathbf{x}, \mathbf{y} \in \Re^n$ is the vector in $\mathbb{R}^n$ defined by $(\mathbf{x} \circ \mathbf{y})_i = x_i y_i$. diag$(\mathbf{x})$ is an n by n diagonal matrix with ith diagonal element x_i, and $\|\cdot\|$ is the Euclidean norm. The subscript k often represents the iteration number in an algorithm, and $\mathbf{g}_k$ stands for $\mathbf{g}(\mathbf{x}_k)$. We let x_{ki} denote the ith component of the iterate $\mathbf{x}_k$.

In the local convergence analysis, $\mathbf{x}^*$ denotes a given nondegenerate local minimizer of f for which the second-order sufficient optimality conditions (1.12) hold. The ball with center $\mathbf{x}$ and radius ρ is denoted $\mathcal{B}_\rho(\mathbf{x})$. The active set $\mathcal{A}$ is defined by

$$\mathcal{A} = \{i \in [1, n] : x_i^* = 0\}.$$

The nondegeneracy assumption implies that $g_i(\mathbf{x}^*) > 0$ when $i \in \mathcal{A}$. The number of elements in $\mathcal{A}$ is denoted $|\mathcal{A}|$ and the complement of $\mathcal{A}$ is $\mathcal{A}^c$. If $i \in \mathcal{A}^c$, then $x_i^* > 0$.

2 Cyclic BB method and the line search

Recently in [13], we have shown that better performance in the BB method is achieved when the same BB stepsize is reused for several iterations. We call this strategy the cyclic BB method (CBB). If $m \geq 1$ is the cycle length and $\ell \geq 0$ is the cycle number, then the cyclic choice for λ_k is

$$\lambda_{m\ell+i} = \lambda_{m\ell+1}^{\mathrm{BB}} \quad \text{for } i = 1, \ldots, m. \tag{2.1}$$

Of course, when the cycle length is 1, then $\lambda_k = \lambda_k^{\mathrm{BB}}$ for each k. If $g_i(\mathbf{x}_k) \leq 0$, then by (1.5), we have

$$d_{ki} = -\frac{1}{\lambda_k} g_i(\mathbf{x}_k).$$

If the iterates converge to a local minimizer $\mathbf{x}^*$ and $x_i^* > 0$, then by the first-order optimality conditions, $g_i(\mathbf{x}^*) = 0$; hence, the denominator in (1.5) approaches λ_k as k increases, and the iterates are closely approximated by the CBB iterates for which we recently establish [13] local R-linear convergence. Due to a scaling operation [second term in the denominator of (1.5)], the iterates always remain in the interior of the feasible set as will be proved in Lemma 3.4.

We now present a line search which ensures the global convergence of AS_CBB. Typically, the BB method does not monotonically reduce the value of the cost function, even in a neighborhood of a local minimizer where the iterates converge. Hence, in order to retain the original features of BB-type methods, a nonmonotone line search should be employed [13,14,26]. We will use the same type of nonmonotone line search developed in [20] for our CBB-gradient project method. The line search, shown in Fig. 1, makes use of the following local maximum function:

$$f_k^{\max} = \max\{f(\mathbf{x}_{k-i}) : 0 \leq i \leq \min(k-1, M-1)\}, \tag{2.2}$$

Initialize $k = 1$, $\mathbf{x}_1 > \mathbf{0}$ starting guess, and $f_0^r = f(\mathbf{x}_1)$.
While $\mathbf{x}_k$ is not a stationary point satisfying (1.6)

1. Let $\mathbf{d}_k$ be given by (1.5).
2. Choose f_k^r so that $f(\mathbf{x}_k) \leq f_k^r \leq \max\{f_{k-1}^r, f_k^{\max}\}$ and $f_k^r \leq f_k^{\max}$ infinitely often.
3. Let f_R be either f_k^r or $\min\{f_k^{\max}, f_k^r\}$. If $f(\mathbf{x}_k + \mathbf{d}_k) \leq f_R + \delta \mathbf{g}_k^{\mathsf{T}} \mathbf{d}_k$, then $s_k = 1$.
4. If $f(\mathbf{x}_k + \mathbf{d}_k) > f_R + \delta \mathbf{g}_k^{\mathsf{T}} \mathbf{d}_k$, then $s_k = \eta^j$ where $j > 0$ is the smallest integer such that

$$f(\mathbf{x}_k + \eta^j \mathbf{d}_k) \leq f_R + \eta^j \delta \mathbf{g}_k^{\mathsf{T}} \mathbf{d}_k. \tag{2.3}$$

5. Set $\mathbf{x}_{k+1} = \mathbf{x}_k + s_k \mathbf{d}_k$ and $k = k + 1$.

End

Fig. 1 A nonmonotone line search

where $M > 0$ is a fixed integer, the memory. In the line search of Fig. 1, $\delta \in (0, 1)$ and $\eta \in (0, 1)$ are the Armijo line search parameters. The condition $f_k^r \geq f(\mathbf{x}_k)$ ensures that the Armijo line search of Step 4 can be satisfied, and the requirement that "$f_k^r \leq f_k^{\max}$ infinitely often" in Step 2 is needed in our global convergence proof. This requirement is easily fulfilled; for example, $f_k^r = f_k^{\max}$ every L iterations. Another strategy, closer in spirit to the one used in the numerical experiments, is to choose a decrease parameter $\Delta > 0$ and an integer $L > 0$ and set $f_k^r = f_k^{\max}$ if

$$f(\mathbf{x}_{k-L}) - f(\mathbf{x}_k) \leq \Delta.$$

As we will show in Lemma 3.4, the fact that $s_k \leq 1$ implies that $\mathbf{x}_{k+1} > \mathbf{0}$ when $\mathbf{x}_k$ does not satisfy the KKT conditions.

3 Continuity properties

We begin with the following observation:

Proposition 3.1 *If f is continuously differentiable and $\mathbf{X}^1(\cdot)$ is defined by (1.7), then the map*

$$\mathbf{X}^1(\cdot) \circ \mathbf{g}(\cdot) : \mathbb{R}^n_+ \to \mathbb{R}^n$$

is continuous.

Proof If either $g_i(\mathbf{x}) < 0$ or $g_i(\mathbf{x}) > 0$, then both g_i and X_i^1 are continuous at $\mathbf{x}$, and hence, the product $g_i(\cdot)X_i^1(\cdot)$ is continuous at $\mathbf{x}$. Suppose that $g_i(\mathbf{x}) = 0$. By the definition of $\mathbf{X}^1$, we have

$$|g_i(\mathbf{y})X_i^1(\mathbf{y})| \leq \max\{1, y_i\}|g_i(\mathbf{y})|$$

for any $\mathbf{y} \geq \mathbf{0}$ and $i \in [1, n]$. Since $g_i(\cdot)$ is continuous and $g_i(\mathbf{x}) = 0$, it follows that $g_i(\mathbf{y})X_i^1(\mathbf{y})$ approaches zeros as $\mathbf{y}$ approaches $\mathbf{x}$. Hence, $\mathbf{g} \circ \mathbf{X}^1$ is continuous everywhere. □

The AS_CBB search directions have the following property:

Proposition 3.2 *Suppose f is twice continuously differentiable on the domain $\mathbf{x} \geq \mathbf{0}$ and the parameter λ_k in (1.5) satisfies*

$$0 < \lambda_0 := \inf_{k \geq 1} \lambda_k \leq \sup_{k \geq 1} \lambda_k := \lambda_{\max} < \infty. \tag{3.1}$$

If $\{\mathbf{x}_k\}_{k=1}^{\infty}$ is a bounded sequence with $\mathbf{x}_k > \mathbf{0}$ and $\mathbf{g}(\mathbf{x}_k) \neq \mathbf{0}$ for each k, then

$$\lim_{k \to \infty} \mathbf{d}_k = \mathbf{0} \quad \text{if and only if} \quad \lim_{k \to \infty} \mathbf{X}^1(\mathbf{x}_k) \circ \mathbf{g}(\mathbf{x}_k) = \mathbf{0}, \tag{3.2}$$

where $\mathbf{d}_k$ and $\mathbf{X}^1(\cdot)$ are defined in (1.5) and (1.7), respectively.

Remark When $\mathbf{x}_k > 0$, the condition $\mathbf{g}(\mathbf{x}_k) \neq \mathbf{0}$ implies that $\mathbf{d}_k \neq \mathbf{0}$ and $\mathbf{x}_k$ is not a KKT point. In other words, the algorithm (1.4)–(1.5) does not stop at iteration k. Suppose λ_k is given by the BB formula (1.11). Applying the fundamental theorem of calculus to the difference $\mathbf{y}_k = \mathbf{g}_k - \mathbf{g}_{k-1}$, we obtain

$$\frac{\mathbf{s}_{k-1}^{\mathsf{T}}\mathbf{y}_{k-1}}{\mathbf{s}_{k-1}^{\mathsf{T}}\mathbf{s}_{k-1}} = \frac{\mathbf{s}_{k-1}^{\mathsf{T}}\left(\int_0^1 \nabla^2 f(\mathbf{x}_{k-1} + t\,\mathbf{s}_{k-1})dt\right)\mathbf{s}_{k-1}}{\mathbf{s}_{k-1}^{\mathsf{T}}\mathbf{s}_{k-1}} \leq \overline{\lambda}, \tag{3.3}$$

where $\overline{\lambda}$ is the largest eigenvalue of $\nabla^2 f$ over any bounded, convex set containing the sequence $\mathbf{x}_k$, $k \geq 1$. The denominator $\mathbf{s}_{k-1}^{\mathsf{T}}\mathbf{s}_{k-1}$ in (3.3) cannot vanish since $\mathbf{g}(\mathbf{x}_k) \neq \mathbf{0}$ for each k. Hence, when λ_k is given by (1.11), we have

$$\lambda_0 \leq \lambda_k^{\mathrm{BB}} \leq \max\,\{\lambda_0, \overline{\lambda}\}$$

for each $k > 1$. It follows that the cyclic choice (2.1) for λ_k satisfies the same inequality:

$$\lambda_0 \leq \lambda_k \leq \max\,\{\lambda_0, \overline{\lambda}\}. \tag{3.4}$$

This shows that when λ_k is given by either (1.11) or the cyclic choice (2.1), the bounds (3.1) are satisfied automatically.

Proof To prove Proposition 3.2, we show that for any $i \in [1, n]$,

$$\lim_{k\to\infty} d_{ki} = 0 \quad \text{if and only if} \quad \lim_{k\to\infty} X_i^1(\mathbf{x}_k) g_i(\mathbf{x}_k) = 0, \tag{3.5}$$

in which case the proposition follows immediately. By (1.5), we have

$$x_{ki}\, g_i(\mathbf{x}_k) = -d_{ki}(\lambda_k x_{ki} + g_i^+(\mathbf{x}_k)). \tag{3.6}$$

Since the $\mathbf{x}_k$ are bounded, f is continuously differentiable, and (3.1) holds, the factor $\lambda_k X_i^1(\mathbf{x}_k) + g_i^+(\mathbf{x}_k)$ is bounded. Hence, if d_{ki} tends to zero, then $X_i^1(\mathbf{x}_k) g_i(\mathbf{x}_k)$ tends to zero.

Conversely, suppose that $X_i^1(\mathbf{x}_k)\, g_i(\mathbf{x}_k)$ tends to zero. In this case, we can write

$$\{1, 2, \ldots\} = \mathcal{K}_1 \cup \mathcal{K}_2,$$

where either $\mathcal{K}_1$ or $\mathcal{K}_2$ may be empty,

$$\text{(a)} \lim_{k\in\mathcal{K}_1} g_i(\mathbf{x}_k) = 0 \quad \text{and} \quad \text{(b)} \lim_{k\in\mathcal{K}_2} X_i^1(\mathbf{x}_k) = 0.$$

If $\mathcal{K}_1$ has an infinite number of elements, then (1.5) and (3.4) imply that d_{ki} tends to 0 for $k \in \mathcal{K}_1$ approaching ∞. If $\mathcal{K}_2$ has an infinite number of elements, then for

$k \in \mathcal{K}_2$ with k sufficiently large, we have $X_i^1(\mathbf{x}_k) = x_{ki}$ and $g_i^+(\mathbf{x}_k) = g_i(\mathbf{x}_k) > 0$. Consequently, (1.5) can be rewritten

$$d_{ki} = -\frac{x_{ki}}{1 + \lambda_k x_{ki}/g_i(\mathbf{x}_k)}, \quad k \in \mathcal{K}_2.$$

By (b) both $X_i^1(\mathbf{x}_k) = x_{ki}$ and d_{ki} tend to as $k \in \mathcal{K}_2$ tends to ∞. Hence, the entire sequence $\{d_{ki} : k \geq 1\}$ approaches $\mathbf{0}$, which completes the proof of (3.5). □

We now show that at a KKT point $\mathbf{x}^*$, $\mathbf{d}_k$ approaches $\mathbf{0}$ as $\mathbf{x}_k$ approaches $\mathbf{x}^*$.

Lemma 3.3 *If f is continuously differentiable and $\mathbf{x}^*$ is a KKT point for* (1.3)*, then for any $\epsilon > 0$, there exists a $\delta > 0$ such that for all $\mathbf{x}_k > \mathbf{0}$ and $\lambda_k \geq \lambda_0 > 0$, we have $\|\mathbf{d}_k\| \leq \epsilon$ whenever $\|\mathbf{x}^* - \mathbf{x}_k\| \leq \delta$.*

Proof Since $\mathbf{x}^*$ is a KKT point for (1.3), we have $\mathbf{g}(\mathbf{x}^*) \circ \mathbf{X}^1(\mathbf{x}^*) = \mathbf{0}$. Hence, either

$$\text{(a) } g_i(\mathbf{x}^*) = 0 \quad \text{or} \quad \text{(b) } g_i(\mathbf{x}^*) > 0 \text{ and } X_i^1(\mathbf{x}^*) = x_i^* = 0$$

for each i. From the definition of $\mathbf{d}_k$, it follows that for any $\mathbf{x}_k > \mathbf{0}$, we have $|d_{ki}| \leq |g_i(\mathbf{x}_k)|/\lambda_0$. If (a) holds, then $|d_{ki}|$ tends to 0 as $\mathbf{x}_k$ approaches $\mathbf{x}^*$. If (b) holds, then for $\mathbf{x}_k$ in a neighborhood of $\mathbf{x}^*$, we have $g_i^+(\mathbf{x}_k) = g_i(\mathbf{x}_k) > 0$ and $X_i^1(\mathbf{x}_k) = x_{ki}$. Hence, $|d_{ki}| \leq X_i^1(\mathbf{x}_k) = x_{ki}$. Again, $|d_{ki}|$ tends to 0 as $\mathbf{x}_k$ approaches $\mathbf{x}^*$. □

We now show that the search directions $\mathbf{d}_k$ satisfy a sufficient descent property and if $\mathbf{x}_k > \mathbf{0}$, then $\mathbf{x}_k + \mathbf{d}_k > \mathbf{0}$.

Lemma 3.4 *Suppose f is twice continuously differentiable on the domain $\mathbf{x} \geq \mathbf{0}$. If $\mathbf{x}_k > \mathbf{0}$, $\mathbf{g}(\mathbf{x}_k) \neq \mathbf{0}$, and $\lambda_k \geq \lambda_0 > 0$, then $\mathbf{d}_k$ defined in* (1.5) *satisfies*

$$\mathbf{d}_k^\mathsf{T} \mathbf{g}(\mathbf{x}_k) \leq -\lambda_k \|\mathbf{d}_k\|^2 < 0 \quad \textit{and} \quad \mathbf{x}_k + \mathbf{d}_k > \mathbf{0}. \tag{3.7}$$

Proof Since $\mathbf{x}_k > \mathbf{0}$, it follows that $\mathbf{X}^1(\mathbf{x}_k) > \mathbf{0}$. By (1.5), we have

$$\begin{aligned}
\mathbf{d}_k^\mathsf{T} \mathbf{g}(\mathbf{x}_k) &= -\sum_{i=1}^n \frac{g_i(\mathbf{x}_k)^2}{\lambda_k + g_i^+(\mathbf{x}_k)/x_{ki}} \\
&= -\sum_{i=1}^n (\lambda_k + g_i^+(\mathbf{x}_k)/x_{ki}) d_{ki}^2 \\
&\leq -\min_{1 \leq i \leq n} (\lambda_k + g_i^+(\mathbf{x}_k)/x_{ki}) \|\mathbf{d}_k\|^2 \\
&\leq -\lambda_k \|\mathbf{d}_k\|^2,
\end{aligned}$$

which gives the first inequality in (3.7). Since $\mathbf{g}(\mathbf{x}_k) \neq \mathbf{0}$, $\mathbf{d}_k \neq \mathbf{0}$. Since $\lambda_k \geq \lambda_0 > 0$, it follows that $-\lambda_k \|\mathbf{d}_k\|^2 < 0$, which gives the first strict inequality in (3.7). Finally, since $\lambda_k > 0$, we have

$$d_{ki} = -\frac{g_i(\mathbf{x}_k)}{\lambda_k + g_i^+(\mathbf{x}_k)/x_{ki}} = \begin{cases} -\dfrac{g_i(\mathbf{x}_k)}{\lambda_k} \geq 0 & \text{if } \; g_i(\mathbf{x}_k) \leq 0, \\ -\dfrac{g_i(\mathbf{x}_k)}{\lambda_k + g_i(\mathbf{x}_k)/x_{ki}} > -x_{ki} & \text{otherwise.} \end{cases}$$

Hence, $\mathbf{x}_k + \mathbf{d}_k > 0$, which gives the last inequality in (3.7). □

4 Global convergence

The continuity properties developed in the previous section are now used to prove the global convergence of AS_CBB.

Theorem 4.1 *Suppose f is twice continuously differentiable and the following level set $\mathcal{L}$ is bounded*:

$$\mathcal{L} = \{\mathbf{x} \geq \mathbf{0} : f(\mathbf{x}) \leq f(\mathbf{x}_1)\}. \tag{4.1}$$

The affine scaling algorithm (1.4)–(1.5) *with the nonmonotone line search of Fig.* 1 *and with λ_k satisfying* (3.1) *either terminates in a finite number of iterations at a KKT point, or*

$$\liminf_{k\to\infty} \|\mathbf{d}_k\| = 0 = \liminf_{k\to\infty} \|\mathbf{g}(\mathbf{x}_k) \circ \mathbf{X}^1(\mathbf{x}_k)\|. \tag{4.2}$$

Proof By Lemma 3.4, the search direction $\mathbf{d}_k$ in Step 1 of the line search is a descent direction. Since $f_k^r \geq f(\mathbf{x}_k)$ and $\delta < 1$, the Armijo line search condition (Fig. 1) is fulfilled for j sufficiently large. We now show that $\mathbf{x}_k \in \mathcal{L}$ for each k. Since $f_1^{\max} = f_0^r = f(\mathbf{x}_1)$, Step 2 of line search implies that $f_1^r \leq f(\mathbf{x}_1)$. Proceeding by induction, suppose that for some $k \geq 1$, we have

$$f_j^r \leq f(\mathbf{x}_1) \quad \text{and} \quad f_j^{\max} \leq f(\mathbf{x}_1) \tag{4.3}$$

for all $j \in [1, k]$. Since $\mathbf{d}_k$ is a direction of descent, it follows from Steps 3 and 4 of the line search and the induction hypothesis that

$$f(\mathbf{x}_{k+1}) \leq f_k^r \leq f(\mathbf{x}_1). \tag{4.4}$$

Hence, $f_{k+1}^{\max} \leq f(\mathbf{x}_1)$ and $f_{k+1}^r \leq \max\{f_{k+1}^{\max}, f_k^r\} \leq f(\mathbf{x}_1)$. This completes the induction. Thus (4.3) holds for all j. Consequently, $f_R \leq f(\mathbf{x}_1)$ in Steps 3 and 4 of the line search. Since $\mathbf{d}_k$ is a descent search direction, Steps 3 and 4 imply that $f(\mathbf{x}_k) \leq f(\mathbf{x}_1)$. Hence, $\mathbf{x}_k \in \mathcal{L}$ for each k.

By Lemma 3.4, each of the iterates $\mathbf{x}_k$ is strictly positive. By (1.5), $|d_{ki}| \le |g_i(\mathbf{x}_k)|/\lambda_0$. Since $\mathcal{L}$ is bounded and $\mathbf{x}_k \in \mathcal{L}$ for each k, we have

$$d_{\max} = \max_{k \ge 1} \|\mathbf{d}_k\| < \infty.$$

If $\bar{\mathcal{L}}$ is the collection of $\mathbf{x} \ge \mathbf{0}$ whose distance to $\mathcal{L}$ is at most $d_{\max}$, then ∇f is Lipschitz continuous on $\bar{\mathcal{L}}$. As shown in [27, Lemma 2.1], we have

$$s_k \ge \min\left\{1,\ \left(\frac{2\eta(1-\delta)}{L}\right)\frac{|\mathbf{g}(\mathbf{x}_k)^{\mathsf{T}}\mathbf{d}_k|}{\|\mathbf{d}_k\|^2}\right\} \tag{4.5}$$

for all k, where L is the Lipschitz constant for ∇f on $\bar{\mathcal{L}}$. Combining (3.4), (3.7), and (4.5) gives

$$s_k \ge \min\left\{1,\ \left(\frac{2\eta(1-\delta)\lambda_0)}{L}\right)\right\} := c. \tag{4.6}$$

Steps 3 and 4 of the line search and (3.4) yield

$$f(\mathbf{x}_{k+1}) \le f_k^r + \delta c \mathbf{g}_k^{\mathsf{T}}\mathbf{d}_k \le f_k^r - \delta c \lambda_k \|\mathbf{d}_k\|^2 \le f_k^r - \delta c \lambda_0 \|\mathbf{d}_k\|^2. \tag{4.7}$$

To prove that $\liminf\limits_{k\to\infty} \|\mathbf{d}_k\| = 0$, we suppose that to the contrary, there exists a constant $\gamma > 0$ such that $\|\mathbf{d}_k\| \ge \gamma$ for all k sufficiently large. By (4.7), we have

$$f(\mathbf{x}_{k+1}) \le f_k^r - \tau, \quad \text{where } \tau = \delta c \lambda_0 \gamma^2, \tag{4.8}$$

for all k. Let k_i, $i = 0, 1, \ldots$, denote an increasing sequence of integers with the property that $f_j^r \le f_j^{\max}$ for $j = k_i$ and $f_j^r \le f_{j-1}^r$ when $k_i < j < k_{i+1}$. Such a sequence exists by the requirement on f_k^r given in Step 2 of the line search. Hence, we have

$$f_j^r \le f_{k_i}^r \le f_{k_i}^{\max}, \quad \text{when } k_i \le j < k_{i+1}. \tag{4.9}$$

By (4.8) it follows that

$$f(\mathbf{x}_j) \le f_{j-1}^r - \tau \le f_{k_i}^{\max} - \tau \quad \text{when } k_i < j \le k_{i+1}.$$

Consequently, $f_{k_{i+1}}^{\max} \le f_{k_i}^{\max}$. Since $f_j^r \le f_j^{\max}$ for $j = k_{i+1}$, it follows that

$$f_{k_{i+1}}^r \le f_{k_{i+1}}^{\max} \le f_{k_i}^{\max}. \tag{4.10}$$

Hence, if $a = k_{i_1}$ and $b = k_{i_2}$ where $i_1 > i_2$ and $a - b > M$ [M is the memory in (2.2)], then by (4.9)–(4.10), we have

$$f_a^{\max} = \max_{0 \le j < M} f(\mathbf{x}_{a-j}) \le \max_{1 \le j \le M} f_{a-j}^r - \tau \le f_b^{\max} - \tau.$$

Since the sequence k_i, $i = 0, 1, \ldots$, is infinite by Step 2, a subsequence of $f_{k_i}^{\max}$ tends to $-\infty$. On the other hand, since $\mathcal{L}$ is bounded, f is bounded from below. Hence, there is a contradiction, and $\liminf_{k\to\infty} \|\mathbf{d}_k\| = 0$. Proposition 3.2 completes the proof. □

When f is strongly convex, there is a unique minimizer for (1.3), and we can strengthen the statement of Theorem 4.1.

Theorem 4.2 *Suppose f is twice continuously differentiable and strongly convex, and there is a positive integer L with the property that for each k, there exists $j \in [k, k+L)$ such that $f_j^r \le f_j^{\max}$. Then the affine scaling algorithm* (1.4)–(1.5) *with the nonmonotone line search of Fig.* 1 *and with λ_k satisfying* (3.1) *converges to the global minimizer $\mathbf{x}^*$ of* (1.3).

Proof Since f is strongly convex, the level set $\mathcal{L}$ in (4.1) is bounded. Hence, the assumptions of Theorem 4.1 are satisfied. At the start of the proof of Theorem 4.1, we showed that $f(\mathbf{x}_k) \le f(\mathbf{x}_1)$ for each k. Since $\mathcal{L}$ is bounded, the $\mathbf{x}_k$ lie in a bounded set. If the iterates do not reach $\mathbf{x}^*$ in a finite number of steps, then by (4.2) and by the compactness of $\mathcal{L}$, there exists an infinite sequence $l_1 < l_2 < \cdots$ such that $\mathbf{d}(\mathbf{x}_{l_j})$ approaches $\mathbf{0}$ and $\{\mathbf{x}_{l_j}\}$ approaches a limit $\bar{\mathbf{x}}$ as j tends to ∞. By Propositions 3.1 and 3.2, $\mathbf{g}(\bar{\mathbf{x}}) \circ \mathbf{X}^1(\bar{\mathbf{x}}) = \mathbf{0}$ and $\bar{\mathbf{x}} \ge \mathbf{0}$. Consequently, $\bar{\mathbf{x}}$ satisfies the first-order optimality conditions for (1.3). Since f is strongly convex, $\bar{\mathbf{x}} = \mathbf{x}^*$.

The goal in the remainder of the proof is to show that the sequence of iterates $\mathbf{x}_k$, $k \ge 1$, converges to $\mathbf{x}^*$. Given an integer $N \ge 0$, we first establish the following continuity property:

(P) For any $\epsilon > 0$, there exists a $\delta > 0$ such that $\|\mathbf{x}_j - \mathbf{x}^*\| \le \epsilon$ for all $j \in [k, k+N]$ whenever $\|\mathbf{x}_k - \mathbf{x}^*\| \le \delta$.

The proof is by induction on N. Clearly, the result is true when $N = 0$. Assume that this holds for some $N > 0$. By Lemma 3.3, we know that for δ_1 sufficiently small, $\|\mathbf{d}_j\| \le \epsilon/2$ when $\|\mathbf{x}_j - \mathbf{x}^*\| \le \delta_1$. Since the step size $s_j \in (0, 1]$, we have

$$\|\mathbf{x}_{j+1} - \mathbf{x}_j\| = s_j \|\mathbf{d}_j\| \le \|\mathbf{d}_j\| \le \epsilon/2 \tag{4.11}$$

when $\|\mathbf{x}_j - \mathbf{x}^*\| \le \delta_1$. Choose δ small enough that

$$\|\mathbf{x}_j - \mathbf{x}^*\| \le \min\{\delta_1, \epsilon/2\} \text{ for all } j \in [k, k+N] \tag{4.12}$$

whenever $\|\mathbf{x}_k - \mathbf{x}^*\| \le \delta$. The triangle inequality, (4.11), and (4.12) yield

$$\|\mathbf{x}_{j+1} - \mathbf{x}^*\| \le \|\mathbf{x}_{j+1} - \mathbf{x}_j\| + \|\mathbf{x}_j - \mathbf{x}^*\| \le \epsilon \tag{4.13}$$

for $j = k + N$ when $\|\mathbf{x}_k - \mathbf{x}^*\| \le \delta$. Combining (4.12) and (4.13), the induction step is complete.

Since f is continuous, for any $\Delta > 0$, there exists an $\epsilon > 0$ with the property that

$$|f(\mathbf{x}) - f(\mathbf{x}^*)| \le \Delta \text{ whenever } \|\mathbf{x} - \mathbf{x}^*\| \le \epsilon.$$

Take $N = M + L$ where L appears in the statement of the theorem and M is the memory in (2.2), and choose δ in accordance with (P). Choose j large enough that $\|\mathbf{x}_{l_j} - \mathbf{x}^*\| \le \delta$. Hence, we have

$$f(\mathbf{x}_k) \le f(\mathbf{x}^*) + \Delta \quad \text{for all } k \in [l_j, l_j + M + L].$$

By the definition of $f_k^{\max}$,

$$f_k^{\max} \le f(\mathbf{x}^*) + \Delta \quad \text{for all } k \in [l_j + M, l_j + M + L]. \tag{4.14}$$

As at the end of the proof of Theorem 4.1, beneath (4.8), let k_i, $i = 0, 1, \ldots$, denote an increasing sequence of integers with the property that $f_j^r \le f_j^{\max}$ for $j = k_i$ and $f_j^r \le f_{j-1}^r$ when $k_i < j < k_{i+1}$. By (4.10) we have

$$f_{k_{i+1}}^{\max} \le f_{k_i}^{\max} \tag{4.15}$$

for each i. The assumption that for each k, there exists $j \in [k, k + L)$ such that $f_j^r \le f_j^{\max}$ implies that

$$k_{i+1} - k_i \le L. \tag{4.16}$$

By (4.16) there exists some $k_i \in [l_j + M, l_j + M + L]$ for each l_j. By (4.14),

$$f_{k_i}^{\max} \le f(\mathbf{x}^*) + \Delta. \tag{4.17}$$

Since Δ was arbitrary, it follows from (4.15) and (4.17) that

$$\lim_{i \to \infty} f_{k_i}^{\max} = f(\mathbf{x}^*); \tag{4.18}$$

the convergence is monotone by (4.15). By the choice of k_i and by the inequality $f(\mathbf{x}_k) \le f_k^r$ in Step 2, we have

$$f(\mathbf{x}_k) \le f_k^r \le f_{k_i}^{\max} \quad \text{for all } k \ge k_i. \tag{4.19}$$

Combining (4.18) and (4.19),

$$\lim_{k \to \infty} f(\mathbf{x}_k) = f(\mathbf{x}^*). \tag{4.20}$$

Since $\mathbf{x}^*$ is the stationary point,

$$\mathbf{g}(\mathbf{x}^*)^{\mathsf{T}}(\mathbf{x} - \mathbf{x}^*) \ge 0, \quad \text{for all } \mathbf{x} \ge \mathbf{0}. \tag{4.21}$$

Since f is strongly convex, there exists $\gamma > 0$ such that

$$f(\mathbf{x}_k) \ge f(\mathbf{x}^*) + \mathbf{g}^{\mathsf{T}}(\mathbf{x}^*)(\mathbf{x}_k - \mathbf{x}^*) + \gamma \|\mathbf{x}_k - \mathbf{x}^*\|^2.$$

Combining this with (4.21) gives

$$f(\mathbf{x}_k) \geq f(\mathbf{x}^*) + \gamma \|\mathbf{x}_k - \mathbf{x}^*\|^2,$$

Referring to (4.20), we conclude that $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}^*$. □

5 Overview of the linear convergence analysis

In [13] we prove that CBB is locally R-linearly convergent for unconstrained optimization at a strict local minimizer. We will prove local R-linear convergence for AS_CBB by comparing the AS_CBB iterates to CBB iterates obtained by fixing $x_i = 0$ for $i \in \mathcal{A}$, the set of active indices at a local minimizer $\mathbf{x}^*$. Our analysis applies to the unit step version of (1.4); that is, we assume the iterates are given by

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k, \quad d_{ki} = -\left(\frac{1}{\lambda_k + g_i^+(\mathbf{x}_k)/x_{ki}}\right) g_i(\mathbf{x}_k). \tag{5.1}$$

Throughout the analysis, the letter c is used to denote a generic constant which is independent of k.

To begin, we first observe that the components of the AS_CBB iterates corresponding to active constraints at $\mathbf{x}^*$ decay to zero Q-quadratically. Given any $\mathbf{x} \in \mathbb{R}^n$, let $\hat{\mathbf{x}}$ denote the vector obtained by replacing with 0 the components associated with active indicates. That is,

$$\hat{x}_i = \begin{cases} x_i & \text{if } i \in \mathcal{A}^c, \\ 0 & \text{if } i \in \mathcal{A}. \end{cases}$$

Thus $\mathbf{x} - \hat{\mathbf{x}}$ is the vector with components

$$x_i - \hat{x}_i = \begin{cases} 0 & \text{if } i \in \mathcal{A}^c, \\ x_i & \text{if } i \in \mathcal{A}. \end{cases}$$

Proposition 5.1 *Suppose that for some $\rho > 0$, f is twice continuously differentiable on the domain*

$$\mathcal{B}_\rho(\mathbf{x}^*) \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\}. \tag{5.2}$$

If $\mathbf{x}_k$, $k \geq 1$, is a sequence generated by the AS_CBB algorithm (5.1), and $\Lambda \geq \lambda_0$ is a fixed scalar, then there exist positive constants ρ_1 and c_1 with the following property: If $\mathbf{x}_k \in \mathcal{B}_{\rho_1}(\mathbf{x}^*)$ and $\mathbf{x}_k > \mathbf{0}$, then

$$\|\mathbf{d}_k\| \leq c_1 \|\mathbf{x}_k - \mathbf{x}^*\|;$$

moreover, if $g_i(\mathbf{x}^*) > 0$ *when* $\mathbf{x}_i^* = 0$ (*that is,* $\mathbf{x}^*$ *is a nongenerate local minimizer*) *and* $\lambda_k \le \Lambda$*, then*

$$\|\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}\| \le c_1 \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|^2.$$

Proof Choose $\rho_1 > 0$ small enough that $\rho_1 \le \rho$ and $g_i(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathcal{B}_{\rho_1}(\mathbf{x}^*)$ and $i \in \mathcal{A}$. For $i \in \mathcal{A}$ and $\mathbf{x}_k \in \mathcal{B}_{\rho_1}(\mathbf{x}^*)$, it follows from (1.5) that $0 \le d_{ki} \le x_{ki}$, and

$$\sum_{i \in \mathcal{A}} d_{ki}^2 \le \sum_{i \in \mathcal{A}} x_{ki}^2 \le \|\mathbf{x}_k - \mathbf{x}^*\|^2. \tag{5.3}$$

If $i \in \mathcal{A}^c$, then

$$|d_{ki}| \le |g_i(\mathbf{x}_k)|/\lambda_0 = |g_i(\mathbf{x}_k) - g_i(\mathbf{x}^*)|/\lambda_0 \le \mu \|\mathbf{x}_k - \mathbf{x}^*\|/\lambda_0,$$

where μ is a Lipschitz constant for $\mathbf{g}$ on the set (5.2). Hence,

$$\|\mathbf{d}_k\|^2 \le (1 + \mu^2 |\mathcal{A}^c|/\lambda_0^2) \|\mathbf{x}_k - \mathbf{x}^*\|^2,$$

which establishes the first half of the proposition.

Choose $\epsilon > 0$ such that $g_i(\mathbf{x}) > \epsilon$ for all $\mathbf{x} \in \mathcal{B}_{\rho_1}(\mathbf{x}^*)$ and $i \in \mathcal{A}$. For the active components, it follows from (1.5) that

$$\begin{aligned}
\|\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}\|^2 = \sum_{i \in \mathcal{A}} x_{(k+1)i}^2 &= \sum_{i \in \mathcal{A}} \left(x_{ki} - \frac{x_{ki}}{1 + \lambda_k x_{ki}/g_{ki}} \right)^2 \\
&\le \sum_{i \in \mathcal{A}} \left(\frac{\lambda_k x_{ki}^2}{g_{ki}} \right)^2 \\
&\le (\Lambda/\epsilon)^2 \sum_{i \in \mathcal{A}} x_{ki}^4 \\
&\le (\Lambda/\epsilon)^2 \left(\sum_{i \in \mathcal{A}} x_{ki}^2 \right)^2 = (\Lambda/\epsilon)^2 \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|^4.
\end{aligned}$$

This establishes the second half of the proposition. □

Proposition 5.1 establishes Q-quadratic convergence of active components of the AS_CBB iterates. To analyze the convergence of the components corresponding to the inactive indices, we compare CBB iterates to AS_CBB iterates. The CBB algorithm is the same as the AS_CBB algorithm except that the second term in the denominator of (1.5) is deleted. In other words, the CBB step is given by

$$\mathbf{d}_k = -\left(\frac{1}{\lambda_k} \right) \mathbf{g}_k.$$

In [13] we establish the following local convergence result for the CBB iterates:

Lemma 5.2 *If $F : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable in a neighborhood of a local minimizer $\mathbf{x}^*$, and the Hessian $\nabla^2 F(\mathbf{x}^*)$ is positive definite, then there exist positive constants δ and β, and a positive constant $\gamma < 1$ with the property that for all starting points $\mathbf{x}_0, \mathbf{x}_1 \in \mathcal{B}_\delta(\mathbf{x}^*)$, $\mathbf{x}_0 \neq \mathbf{x}_1$ and λ_1 given by (1.11), the CBB iterates satisfy*

$$\|\mathbf{x}_k - \mathbf{x}^*\| \le \beta\gamma^k \|\mathbf{x}_1 - \mathbf{x}^*\|. \tag{5.4}$$

For any fixed $k \ge 1$, we will compare an AS_CBB iterate $\mathbf{x}_{k+j}$ to a CBB iterate $\mathbf{z}_{k,j}$ starting from $\hat{\mathbf{x}}_k$. We introduce the following notation: $\nu(k) = 1 + m\lfloor (k-1)/m \rfloor$, where $\lfloor r \rfloor$ denotes the largest integer j such that $j \le r$. With this notation, the CBB parameter defined in (2.1) can expressed $\lambda_k = \lambda^{\mathrm{BB}}_{\nu(k)}$.

The comparison iterate is defined by

$$\begin{aligned} \mathbf{z}_{k,0} &= \hat{\mathbf{x}}_k \\ \mathbf{z}_{k,j+1} &= \mathbf{z}_{k,j} - \alpha_{k,j}\hat{\mathbf{g}}(\mathbf{z}_{k,j}) \end{aligned} \tag{5.5}$$

$$\alpha_{k,j} = \begin{cases} 1/\lambda_k & \text{if} \quad \nu(k+j) = \nu(k), \\ \dfrac{1}{\widehat{\lambda}_{k+j}} & \text{otherwise.} \end{cases} \tag{5.6}$$

The stepsize parameter $\widehat{\lambda}_{k+j}$ is defined like λ_k in the AS_CBB iteration except that $\mathbf{x}$ is replaced by $\mathbf{z}$. More precisely, we define

$$\widehat{\lambda}^{\mathrm{BB}}_{k+j} = \frac{\mathbf{v}_{j-1}^{\mathsf{T}}\mathbf{w}_{j-1}}{\mathbf{v}_{j-1}^{\mathsf{T}}\mathbf{v}_{j-1}}, \quad \mathbf{v}_{j-1} = \mathbf{z}_{k,j} - \mathbf{z}_{k,j-1}, \quad \mathbf{w}_{j-1} = \mathbf{g}(\mathbf{z}_{k,j}) - \mathbf{g}(\mathbf{z}_{k,j-1}),$$

and $\widehat{\lambda}_{k+j} = \widehat{\lambda}^{\mathrm{BB}}_{\nu(k+j)}$.

Notice that the comparison iterate is based on a single starting point $\hat{\mathbf{x}}_k$ since $\alpha_{k,0}$ is obtained from the AS_CBB iterate. The comparison CBB iteration starts out as a modified CBB iteration where the stepsize takes the top value in (5.6) until $\nu(k+j) > \nu(k)$, at which point the bottom value in (5.6) is used. By the definition of ν, we have $j \le m$ when the switch to the bottom expression takes place.

As an application of our local convergence result for the CBB iteration, we now prove the following:

Proposition 5.3 *Suppose that for some $\rho > 0$, f is twice continuously differentiable on the domain*

$$\mathcal{B}_\rho(\mathbf{x}^*) \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \ge \mathbf{0}\}, \tag{5.7}$$

and the second-order sufficient optimality condition (1.12) is satisfied. Then there exist $\delta > 0$ and an integer $N > 0$ such that for all starting points $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$, the CBB iterates generated by (5.5)–(5.6) satisfy

$$\mathbf{z}_{k,j} \in \mathcal{B}_\rho(\mathbf{x}^*) \quad \textit{and} \quad \mathbf{z}_{k,j} \ge \mathbf{0} \textit{ for } j \ge 0 \textit{ and} \tag{5.8}$$

$$\|\mathbf{z}_{k,j} - \mathbf{x}^*\| \le \frac{1}{2}\|\mathbf{z}_{k,0} - \mathbf{x}^*\| \; for \; j \ge N. \tag{5.9}$$

Proof Let μ be a Lipschitz constant for $\mathbf{g}$ on the set (5.7). Let $\delta_1 > 0$ be the minimum of ρ and the parameter δ of Lemma 5.2 associated with the function of $n - |\mathcal{A}|$ variables obtained by setting the active components of $\mathbf{x}$ to zero in $f(\mathbf{x})$; in other words, F corresponds to the function $f(\hat{\mathbf{x}})$. Choose δ_1 smaller if necessary to ensure that

$$\hat{\mathbf{x}} \ge \mathbf{0} \text{ whenever } \mathbf{x} \in \mathcal{B}_{\delta_1}(\mathbf{x}^*). \tag{5.10}$$

Let $\delta > 0$ be any scalar small enough that

$$\beta(1 + \mu\lambda_0^{-1})^m \delta < \delta_1 \le \rho, \tag{5.11}$$

where m is the cycle length of the CBB algorithm and $\beta \ge 1$ is the constant in (5.4). Since $\hat{\mathbf{g}}(\mathbf{x}^*) = \mathbf{0}$, if follows from (5.5) and (5.6) that

$$\|\mathbf{z}_{k,j+1} - \mathbf{x}^*\| \le \left(1 + \frac{\mu}{\lambda_0}\right)\|\mathbf{z}_{k,j} - \mathbf{x}^*\|$$

if $\nu(k+j) = \nu(k)$ and $\mathbf{z}_{k,j} \in \mathcal{B}_\rho(\mathbf{x}^*)$. Hence, if $\mathbf{z}_{k,0} \in \mathcal{B}_\delta(\mathbf{x}^*)$, then by (5.11), both $\mathbf{z}_{k,j+1}$ and $\mathbf{z}_{k,j} \in \mathcal{B}_{\delta_1}(\mathbf{x}^*)$ when $\nu(k+j) = \nu(k)$ since $j < m$ in this case. Consequently, if J is the smallest integer for which $\nu(k+J) > \nu(k)$, then $J \le m$; and if $\mathbf{z}_{k,0} \in \mathcal{B}_\delta(\mathbf{x}^*)$, then $\mathbf{z}_{k,J-1}$ and $\mathbf{z}_{k,J} \in \mathcal{B}_{\delta_2}(\mathbf{x}^*)$, where $\delta_2 = (1 + \mu\lambda_0^{-1})^m\delta$. By (5.11) and (1.12), we can apply Lemma 5.2 to the CBB iterates that start from $\mathbf{z}_{k,J}$ and $\mathbf{z}_{k,J-1}$ (note that the components of $\mathbf{z}_{k,j}$ associated with the active indices $\mathcal{A}$ always vanish since the components of both $\mathbf{z}_{k,0}$ and $\hat{\mathbf{g}}$ associated with active indices vanish). By (5.10), (5.11), and Lemma 5.2, it follows that for N sufficiently large and for $j \ge N$, (5.8)–(5.9) hold. □

As a corollary of Proposition 5.3, the comparison iterates are well-defined in the sense that they remain inside a sphere where f is differentiable whenever they start from a point $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$.

6 Comparison between CBB and AS_CBB iterates

Our proof of local R-linear convergence for the AS_CBB algorithm is based on the local R-linear convergence of the CBB algorithm, as indicated in Lemma 5.2 and Proposition 5.3, and the Q-quadratic convergence of the components of the iterates associated with active constraints as given in Proposition 5.1. In our next lemma, we estimate the distance between the CBB iterate $\mathbf{z}_{k,j}$ and the AS_CBB iterate $\mathbf{x}_{k+j}$.

Lemma 6.1 *Suppose that for some $\rho > 0$, f is twice continuously differentiable on the domain*

$$\mathcal{B}_\rho(\mathbf{x}^*) \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \ge \mathbf{0}\}, \tag{6.1}$$

where $\mathbf{x}^*$ *is a nondegenerate local minimizer satisfying the second-order sufficient optimality condition* (1.12). *Let* r *be small enough that* $r \leq \min\{\rho_1, \rho/2\}$ *where* ρ_1 *is the parameter of Proposition* 5.1, *and*

$$\mathbf{d}^{\mathsf{T}} \nabla^2 f(\mathbf{x}) \mathbf{d} \geq (\alpha/2) \|\mathbf{d}\|^2 \tag{6.2}$$

for all $\mathbf{x} \in \mathcal{B}_{2r}(\mathbf{x}^*)$ *and* $\mathbf{d} \in \mathbb{R}^n$ *with* $d_i = 0$ *when* $i \in \mathcal{A}$. *Assume that* $0 < \lambda_0 \leq \alpha/4$ *and let* $\Lambda \geq \lambda_0$ *be a fixed scalar. If* $\mathbf{x}_{k+j}$, $j \geq 0$, *is a sequence generated by the AS_CBB algorithm* (5.1), *then for any fixed positive integer* N, *there exist positive constants* δ *and* c_2 *with the following property: For any* $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$ *satisfying*

$$\max\{|x_{ki}| : i \in \mathcal{A}\} \leq \|\hat{\mathbf{x}}_k - \mathbf{x}^*\|^{3/2} \text{ and } \lambda_k \leq \Lambda, \tag{6.3}$$

and for any $\ell \in [0, N]$, *if*

$$\|\mathbf{z}_{k,j} - \mathbf{x}^*\| \geq \frac{1}{2} \|\mathbf{z}_{k,0} - \mathbf{x}^*\| \quad \text{for all } j \in [0, \max\{0, \ell - 1\}], \tag{6.4}$$

then we have

$$\mathbf{x}_{k+j} \in \mathcal{B}_r(\mathbf{x}^*) \text{ and } \|\mathbf{x}_{k+j} - \mathbf{z}_{k,j}\| \leq c_2 \|\mathbf{x}_k - \mathbf{x}^*\|^{3/2} \tag{6.5}$$

for all $j \in [0, \ell]$.

Proof To facilitate the proof, we show, in addition to (6.5), that

$$\|\mathbf{s}_{k+j}\| \leq c_2 \|\mathbf{x}_k - \mathbf{x}^*\| \text{ and} \tag{6.6}$$

$$\max \ \{|\alpha_{(k+j)i} - \alpha_{k,j}| : i \in \mathcal{A}^c\} \leq c_2 \|\mathbf{x}_k - \mathbf{x}^*\|^{1/2} \tag{6.7}$$

for all $j \in [0, \ell]$, where $\alpha_{k,j}$ is defined in (5.6) and $\boldsymbol{\alpha}_k$ is the coefficient of the gradient in (1.5):

$$\alpha_{ki} = -\frac{1}{\lambda_k + g_i^+(\mathbf{x}_k)/x_{ki}} \tag{6.8}$$

The proof of (6.5)–(6.7) is by induction on ℓ. We begin with $\ell = 0$, in which case $j = 0$. Choose $\delta < r$. Hence, the left side of (6.5) holds trivially. By the initial condition $\mathbf{z}_{k,0} = \hat{\mathbf{x}}_k$ and (6.3), we have

$$\begin{aligned} \|\mathbf{x}_k - \mathbf{z}_{k,0}\|^2 = \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|^2 = \sum_{i \in \mathcal{A}} x_{ki}^2 \\ \leq (|\mathcal{A}|) \|\hat{\mathbf{x}}_k - \mathbf{x}^*\|^3 \leq (|\mathcal{A}|) \|\mathbf{x}_k - \mathbf{x}^*\|^3, \end{aligned} \tag{6.9}$$

which gives the right side of (6.5), $c_2 = |\mathcal{A}|^{1/2}$. Since $\delta < r \leq \rho_1$, it follows from Proposition 5.1 that for $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$,

$$\|\mathbf{s}_k\| = \|\mathbf{d}_k\| \leq c_1 \|\mathbf{x}_k - \mathbf{x}^*\|, \tag{6.10}$$

which implies that (6.6) holds, when c_2 is increased if necessary. Take δ smaller if necessary and choose ϵ so that

$$\min\{x_i : i \in \mathcal{A}^c, \mathbf{x} \in \mathcal{B}_\delta(\mathbf{x}^*)\} \geq \epsilon > 0.$$

Equations (5.6) and (6.8) imply that for any $i \in \mathcal{A}^c$ and $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$,

$$|\alpha_{ki} - \hat{\alpha}_{k,0}| = \begin{cases} 0 & \text{if} \quad g_{ki} \leq 0, \\ \left|\dfrac{1}{\lambda_k + g_{ki}/x_{ki}} - \dfrac{1}{\lambda_k}\right| & \text{otherwise.} \end{cases} \tag{6.11}$$

If $g_{ki} > 0$, then we have

$$|\alpha_{ki} - \hat{\alpha}_{k,0}| = \left|\frac{1}{\lambda_k + g_{ki}/x_{ki}} - \frac{1}{\lambda_k}\right| \leq \frac{|g_{ki}|}{\epsilon(\lambda_k)^2} \leq \frac{|g_{ki}|}{\epsilon\lambda_0^2}. \tag{6.12}$$

Since

$$|g_{ki}| \leq \|\hat{\mathbf{g}}(\mathbf{x}_k) - \hat{\mathbf{g}}(\mathbf{x}^*)\| \leq \|\mathbf{g}(\mathbf{x}_k) - \mathbf{g}(\mathbf{x}^*)\| \leq \mu\|\mathbf{x}_k - \mathbf{x}^*\|, \tag{6.13}$$

where μ is a Lipschitz constant for $\mathbf{g}$ on the set (6.1), (6.7) follows from (6.11)–(6.13) when δ is sufficiently small .

Now, proceeding by induction, suppose that for some $L \in [1, N)$, (6.5)–(6.7) hold for all $\ell \in [0, L]$ when the conditions (6.3)–(6.4) are satisfied. We wish to show that for a suitable choice of δ and c_2, we can replace L by $L + 1$. Hence, let us suppose that (6.4) holds for all $j \in [0, L]$. By the induction hypothesis and (6.6),

$$\begin{aligned} \|\mathbf{x}_{k+L+1} - \mathbf{x}^*\| &\leq \|\mathbf{x}_k - \mathbf{x}^*\| + \sum_{i=0}^{L} \|\mathbf{s}_{k+i}\| \\ &\leq (1 + (L+1)c_2)\|\mathbf{x}_k - \mathbf{x}^*\| \leq (1 + (L+1)c_2)\delta. \end{aligned} \tag{6.14}$$

Consequently, by choosing δ smaller if necessary, we have $\mathbf{x}_{k+L+1} \in \mathcal{B}_r(\mathbf{x}^*)$ when $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$.

By the triangle inequality, (6.14) with L replaced by $L - 1$, and Proposition 5.1, we have

$$\begin{aligned} \|\mathbf{x}_{k+L+1} - \mathbf{z}_{k,L+1}\| &\leq \|\mathbf{x}_{k+L+1} - \hat{\mathbf{x}}_{k+L+1}\| + \|\hat{\mathbf{x}}_{k+L+1} - \mathbf{z}_{k,L+1}\| \\ &\leq c_1\|\mathbf{x}_{k+L} - \hat{\mathbf{x}}_{k+L}\|^2 + \|\hat{\mathbf{x}}_{k+L+1} - \mathbf{z}_{k,L+1}\| \\ &\leq c_1\|\mathbf{x}_{k+L} - \mathbf{x}^*\|^2 + \|\hat{\mathbf{x}}_{k+L+1} - \mathbf{z}_{k,L+1}\| \\ &\leq c_1(1 + Lc_2)^2\|\mathbf{x}_k - \mathbf{x}^*\|^2 + \|\hat{\mathbf{x}}_{k+L+1} - \mathbf{z}_{k,L+1}\|. \end{aligned} \tag{6.15}$$

By the definition of the AS_CBB and CBB iterates, and by (6.5) for $j = L$,

$$\begin{aligned}\|\hat{\mathbf{x}}_{k+L+1} - \mathbf{z}_{k,L+1}\| &\le \|\hat{\mathbf{x}}_{k+L} - \mathbf{z}_{k,L}\| + \|\hat{\mathbf{s}}_{k+L} - \alpha_{k,L}\hat{\mathbf{g}}(\mathbf{z}_{k,L})\| \\ &\le \|\mathbf{x}_{k+L} - \mathbf{z}_{k,L}\| + \|\hat{\mathbf{s}}_{k+L} - \alpha_{k,L}\hat{\mathbf{g}}(\mathbf{z}_{k,L})\| \\ &\le c_2\|\mathbf{x}_k - \mathbf{x}^*\|^{3/2} + \|\hat{\mathbf{s}}_{k+L} - \alpha_{k,L}\hat{\mathbf{g}}(\mathbf{z}_{k,L})\|. \end{aligned} \tag{6.16}$$

For $i \in \mathcal{A}^c$, we have

$$\begin{aligned}|(\hat{\mathbf{s}}_{k+L} - \alpha_{k,L}\hat{\mathbf{g}}(\mathbf{z}_{k,L}))_i| &= |\alpha_{(k+L)i}g_i(\mathbf{x}_{k+L}) - \alpha_{k,L}g_i(\mathbf{z}_{k,L})| \\ &\le (|\alpha_{(k+L)i}|)|g_i(\mathbf{x}_{k+L}) - \mathbf{g}_i(\mathbf{z}_{k,L})| \\ &\quad + (|\alpha_{k,L} - \alpha_{(k+L)i}|)|g_i(\mathbf{z}_{k,L})|. \end{aligned} \tag{6.17}$$

Suppose that δ is small enough that $c_2\delta^{3/2} \le r$. It follows from (6.5) that $\mathbf{z}_{k,j} \in \mathcal{B}_{2r}(\mathbf{x}^*)$ for all $j \in [0, L]$. We apply the Lipschitz continuity of $\mathbf{g}$ on $\mathcal{B}_{2r}(\mathbf{x}^*)$ to (6.17) to obtain

$$\begin{aligned}&|(\hat{\mathbf{s}}_{k+L} - \alpha_{k,L}\hat{\mathbf{g}}(\mathbf{z}_{k,L}))_i| \\ &\quad \le \mu\lambda_0^{-1}\|\mathbf{x}_{k+L} - \mathbf{z}_{k,L}\| + c_2\|\mathbf{x}_k - \mathbf{x}^*\|^{1/2}\|g_i(\mathbf{z}_{k,L}) - g_i(\mathbf{x}^*)\| \\ &\quad \le \mu\lambda_0^{-1}c_2\|\mathbf{x}_k - \mathbf{x}^*\|^{3/2} + c_2\mu\|\mathbf{x}_k - \mathbf{x}^*\|^{1/2}\|\mathbf{z}_{k,L} - \mathbf{x}^*\| \end{aligned} \tag{6.18}$$

since $|\alpha_{ki}| \le 1/\lambda_0$ for any k and both (6.5) and (6.7) hold for $j = L$. By the triangle inequality, (6.5), and (6.14), we have

$$\begin{aligned}\|\mathbf{z}_{k,L} - \mathbf{x}^*\| &\le \|\mathbf{z}_{k,L} - \mathbf{x}_{k+L}\| + \|\mathbf{x}_{k+L} - \mathbf{x}^*\| \\ &\le c_2\|\mathbf{x}_k - \mathbf{x}^*\|^{3/2} + (1 + Lc_2)\|\mathbf{x}_k - \mathbf{x}^*\|. \end{aligned} \tag{6.19}$$

Combine (6.15)–(6.19) to obtain (6.5) for $j = L + 1$, assuming δ is sufficiently small. As a consequence, $\mathbf{z}_{k,L+1} \in \mathcal{B}_{2r}(\mathbf{x}^*)$.

To complete the induction step, we must verify (6.6) and (6.7) for $j = L + 1$. Since $\mathbf{x}_{k+L+1} \in \mathcal{B}_r(\mathbf{x}^*)$, it follows from Proposition 5.1 and (6.14) that

$$\|\mathbf{s}_{k+L+1}\| \le c_1\|\mathbf{x}_{k+L+1} - \mathbf{x}^*\| \le c_1(1 + (L + 1)c_2)\|\mathbf{x}_k - \mathbf{x}^*\|,$$

which gives (6.6) for $j = L + 1$.

Finally, let us focus on (6.7). If $\nu(k + L + 1) = \nu(k)$, then by the same analysis used in (6.11)–(6.13), (6.7) holds for $j = L + 1$. If $\nu(k + L + 1) > \nu(k)$, then there exists an index $j \in (0, L]$ such that

$$\widehat{\lambda}_{k+L+1} = \frac{\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}}{\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{v}_{k+j}} \quad \text{and} \quad \lambda_{k+L+1} = \frac{\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j}}{\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{s}_{k+j}}.$$

By the triangle inequality,

$$\|\mathbf{s}_{k+j} - \mathbf{v}_{k+j}\| \le \|\mathbf{x}_{k+j+1} - \mathbf{z}_{k,j+1}\| + \|\mathbf{x}_{k+j} - \mathbf{z}_{k,j}\|.$$

Since (6.5) has already been established for $j \in [0, L+1]$, it follows that

$$\|\mathbf{s}_{k+j} - \mathbf{v}_{k+j}\| \le 2c_2\|\mathbf{x}_k - \mathbf{x}^*\|^{3/2}. \tag{6.20}$$

Combining this with (6.6) gives

$$\begin{aligned} |\mathbf{s}_{k+i}^{\mathsf{T}}\mathbf{s}_{k+i} - \mathbf{v}_{k+i}^{\mathsf{T}}\mathbf{v}_{k+j}| &= \left|2\mathbf{s}_{k+j}^{\mathsf{T}}(\mathbf{s}_{k+j} - \mathbf{v}_{k+j}) - \|\mathbf{v}_{k+j} - \mathbf{s}_{k+j}\|^2\right| \\ &\le 4c_2^2(\|\mathbf{x}_k - \mathbf{x}^*\|^{5/2} + \|\mathbf{x}_k - \mathbf{x}^*\|^3) \\ &\le 4c_2^2(1+\sqrt{r})\|\mathbf{x}_k - \mathbf{x}^*\|^{5/2}. \end{aligned} \tag{6.21}$$

Let Λ_1 denote an upper bound for the largest eigenvalue for $\nabla^2 f(\mathbf{x})$ over $\mathbf{x} \in \mathcal{B}_{2r}(\mathbf{x}^*)$. Since $\mathbf{z}_{k,j} \in \mathcal{B}_{2r}(\mathbf{x}^*)$ for all $j \in [0, L]$, a Taylor expansion yields

$$\alpha/2 \le |\widehat{\lambda}_{k+j}| \le \Lambda_1, \tag{6.22}$$

which implies that

$$|\alpha_{k,j}| = 1/|\widehat{\lambda}_{k+j}| \ge 1/\Lambda_1.$$

By (6.2), it follows that

$$\|\hat{\mathbf{g}}(\mathbf{z}_{k,j})\| = \|\hat{\mathbf{g}}(\mathbf{z}_{k,j}) - \hat{\mathbf{g}}(\mathbf{x}^*)\| \ge (\alpha/2)\|\mathbf{z}_{k,j} - \mathbf{x}^*\|.$$

As a consequence of these relations, we have

$$\begin{aligned} \|\mathbf{v}_{k+j}\| &= \|\alpha_{k,j}\hat{\mathbf{g}}_{k,j}\| = \|\alpha_{k,j}(\hat{\mathbf{g}}(\mathbf{z}_{k,j}) - \hat{\mathbf{g}}(\mathbf{x}^*))\| \\ &\ge \frac{\alpha}{2\Lambda_1}\|\mathbf{z}_{k,j} - \mathbf{x}^*\|. \end{aligned} \tag{6.23}$$

Applying (6.4) with $j = L$ and (6.9) gives

$$\begin{aligned} \|\mathbf{z}_{k,j} - \mathbf{x}^*\| &\ge \frac{1}{2}\|\mathbf{z}_{k,0} - \mathbf{x}^*\| = \frac{1}{2}\|\hat{\mathbf{x}}_k - \mathbf{x}^*\| \\ &\ge \frac{1}{2}(\|\mathbf{x}_k - \mathbf{x}^*\| - \|\hat{\mathbf{x}}_k - \mathbf{x}_k\|) \ge \frac{1}{2}\|\mathbf{x}_k - \mathbf{x}^*\|(1 - |\mathcal{A}|^{\frac{1}{2}}\sqrt{\delta}). \end{aligned} \tag{6.24}$$

By (6.23)–(6.24) and for δ sufficiently small, there exists a scalar $\sigma > 0$ such that

$$\|\mathbf{v}_{k+j}\| \ge \sigma\|\mathbf{x}_k - \mathbf{x}^*\|. \tag{6.25}$$

Combining (6.21) and (6.25), we obtain

$$\left|1 - \frac{\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{s}_{k+j}}{\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{v}_{k+j}}\right| = \frac{|\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{s}_{k+j} - \mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{v}_{k+j}|}{\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{v}_{k+j}} \le c\|\mathbf{x}_k - \mathbf{x}^*\|^{1/2}, \tag{6.26}$$

where c is a generic constant.

Next, we estimate the difference in the numerators of λ_{k+j} and $\widehat{\lambda}_{k+j}$. Since the components of $\mathbf{z}_{k,j}$ associated with $i \in \mathcal{A}$ vanish, it follows that

$$\begin{aligned} \mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j} - \mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j} &= \mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j} - \mathbf{v}_{k+j}^{\mathsf{T}}(\mathbf{g}(\mathbf{z}_{k,j+1}) - \mathbf{g}(\mathbf{z}_{k,j})) \\ &= \mathbf{s}_{k+j}^{\mathsf{T}}(\mathbf{y}_{k+j} - (\mathbf{g}(\mathbf{z}_{k,j+1}) - \mathbf{g}(\mathbf{z}_{k,j}))) \\ &\quad + (\mathbf{s}_{k+j} - \mathbf{v}_{k+j})^{\mathsf{T}}(\mathbf{g}(\mathbf{z}_{k,j+1}) - \mathbf{g}(\mathbf{z}_{k,j})). \end{aligned} \tag{6.27}$$

By (6.6) and (6.20) for $j \in [0, L+1]$, the Lipschitz continuity of $\mathbf{g}$, and the fact that $\mathbf{z}_{k,j} \in \mathcal{B}_{2r}(\mathbf{x}^*)$, we have

$$\begin{aligned} |(\mathbf{s}_{k+j} - \mathbf{v}_{k+j})^{\mathsf{T}}(\mathbf{g}(\mathbf{z}_{k,j+1}) - \mathbf{g}(\mathbf{z}_{k,j}))| &\le \mu\|\mathbf{s}_{k+j} - \mathbf{v}_{k+j}\|\|\mathbf{v}_{k+j}\| \\ &\le \mu\|\mathbf{s}_{k+j} - \mathbf{v}_{k+j}\|(\|\mathbf{s}_{k+j} - \mathbf{v}_{k+j}\| + \|\mathbf{s}_{k+j}\|) \\ &\le 2\mu c_2^2(2\sqrt{\delta} + 1)\|\mathbf{x}_k - \mathbf{x}^*\|^{5/2}. \end{aligned} \tag{6.28}$$

Also, by (6.5) and (6.6) for all $j \in [0, L+1]$, we have

$$\begin{aligned} &|\mathbf{s}_{k+j}^{\mathsf{T}}(\mathbf{y}_{k+j} - (\mathbf{g}(\mathbf{z}_{k,j+1}) - \mathbf{g}(\mathbf{z}_{k,j})))| \\ &\quad \le \|\mathbf{s}_{k+j}\|\Big(\|\mathbf{g}(\mathbf{x}_{k+j+1}) - \mathbf{g}(\mathbf{z}_{k,j+1})\| + \|\mathbf{g}(\mathbf{x}_{k+j}) - \mathbf{g}(\mathbf{z}_{k,j})\|\Big) \\ &\quad \le 2\mu c_2^2\|\mathbf{x}_k - \mathbf{x}^*\|^{5/2}. \end{aligned} \tag{6.29}$$

Combining (6.27), (6.28) and (6.29) yields

$$\left|\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j} - \mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}\right| \le c\|\mathbf{x}_k - \mathbf{x}^*\|^{5/2}, \tag{6.30}$$

where c is a generic constant.

On the other hand, since $\mathbf{z}_{k,j} \in \mathcal{B}_{2r}(\mathbf{z}^*)$, it follows from (6.25) that

$$\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j} \ge \alpha\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{v}_{k+j} \ge \alpha\sigma^2\|\mathbf{x}_k - \mathbf{x}^*\|^2.$$

Combining this with (6.30) gives

$$\left|1 - \frac{\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j}}{\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}}\right| = \frac{|\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j} - \mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}|}{\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}} \le c\|\mathbf{x}_k - \mathbf{x}^*\|^{1/2},$$

where c is a generic constant. Hence, $\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j}/\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}$ is close to 1 for δ sufficiently small. Consequently, the reciprocal $\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}/\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j}$ is also close to 1. More precisely, we have

$$\left|1 - \frac{\mathbf{v}_{k+j}^{\mathsf{T}}\mathbf{w}_{k+j}}{\mathbf{s}_{k+j}^{\mathsf{T}}\mathbf{y}_{k+j}}\right| \le c\|\mathbf{x}_k - \mathbf{x}^*\|^{1/2}, \tag{6.31}$$

where c is a generic constant. We utilize (6.22) to obtain

$$\begin{aligned}
|\lambda_{k+L+1}^{-1} - \widehat{\lambda}_{k+L+1}^{-1}| &= \left| \frac{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{s}_{k+j}}{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{y}_{k+j}} - \frac{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{v}_{k+j}}{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{w}_{k+j}} \right| \\
&= \widehat{\lambda}_{k+L+1}^{-1} \left| 1 - \left(\frac{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{s}_{k+j}}{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{v}_{k+j}} \right) \left(\frac{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{w}_{k+j}}{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{y}_{k+j}} \right) \right| \\
&\le \frac{2}{\alpha} \left| 1 - \left(\frac{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{s}_{k+j}}{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{v}_{k+j}} \right) \left(\frac{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{w}_{k+j}}{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{y}_{k+j}} \right) \right| \\
&= \frac{2}{\alpha} |a(1-b) + b| \le \frac{2}{\alpha} (|a| + |b| + |ab|), \qquad (6.32)
\end{aligned}$$

where

$$a = 1 - \frac{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{s}_{k+j}}{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{v}_{k+j}} \quad \text{and} \quad b = 1 - \frac{\mathbf{v}_{k+j}^{\mathsf{T}} \mathbf{w}_{k+j}}{\mathbf{s}_{k+j}^{\mathsf{T}} \mathbf{y}_{k+j}}.$$

Together, (6.26), (6.31), and (6.32) yield

$$|\lambda_{k+L+1}^{-1} - \widehat{\lambda}_{k+L+1}^{-1}| \le c \|\mathbf{x}_k - \mathbf{x}^*\|^{1/2} \qquad (6.33)$$

for δ sufficiently small. By (6.22), $\widehat{\lambda}_{k+L+1}^{-1} \le 2/\alpha$. Hence, for δ sufficiently small, $\lambda_{k+L+1}^{-1} \le 3/\alpha$, or equivalently, $\lambda_{k+L+1} \ge \alpha/3$. Since $\lambda_0 \le \alpha/4$, we conclude that

$$\max\{\lambda_0, \lambda_{k+L+1}\} = \lambda_{k+L+1} \qquad (6.34)$$

for δ sufficiently small. It follows that

$$|\alpha_{(k+L+1)i} - \alpha_{k,L+1}| = \left| \frac{1}{\lambda_{k+L+1} + g_i^+(\mathbf{x}_k)/x_{ki}} - \widehat{\lambda}_{k,L+1}^{-1} \right|.$$

Therefore, by (6.14), (6.33) and (6.34) and for any $i \in \mathcal{A}(x^*)^c$, we have

$$\begin{aligned}
|\alpha_{(k+L+1)i} - \alpha_{k,L+1}| &= \begin{cases} |\lambda_{k+L+1}^{-1} - \widehat{\lambda}_{k,L+1}^{-1}| & \text{if } g_{(k+L+1)i} \le 0, \\ \left| \dfrac{1}{\lambda_{k+L+1} + g_i(\mathbf{x}_k)/x_{ki}} - \widehat{\lambda}_{k,L+1}^{-1} \right| & \text{otherwise}, \end{cases} \\
&\le |\lambda_{k+L+1}^{-1} - \widehat{\lambda}_{k,L+1}^{-1}| + c \|\mathbf{x}_{k+L+1} - \mathbf{x}^*\| \\
&\le c \|\mathbf{x}_k - \mathbf{x}^*\|^{1/2}
\end{aligned}$$

for δ sufficiently small. This establishes (6.7) for $\ell = L + 1$. Hence, the induction step is complete. □

7 Linear convergence for AS_CBB

In this section, we combine the local R-linear convergence of the CBB algorithm, the comparison between CBB and AS_CBB iterates, as given in Lemma 6.1, and the Q-quadratic decay of AS_CBB active components, as given in Proposition 5.1, to obtain local R-linear convergence for AS_CBB.

Theorem 7.1 *Suppose that for some $\rho > 0$, f is twice continuously differentiable on the domain*

$$\mathcal{B}_\rho(\mathbf{x}^*) \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\},$$

where $\mathbf{x}^*$ is a nondegenerate local minimizer satisfying the second-order sufficient optimality condition (1.12). Let λ_0 be chosen in accordance with Lemma 6.1. Then there exist positive scalars δ and η, and a positive scalar $\gamma < 1$ with the property that for all starting points $\mathbf{x}_0, \mathbf{x}_1 \in \mathcal{B}_\delta(\mathbf{x}^*)$, $\mathbf{x}_0 \neq \mathbf{x}_1$, and with λ_1 given by (1.11), the AS_CBB iterates generated by (5.1) satisfy

$$\|\mathbf{x}_k - \mathbf{x}^*\| \leq \eta\gamma^k \|\mathbf{x}_1 - \mathbf{x}^*\|. \tag{7.1}$$

Proof Let $N > 0$ be the integer given in Proposition 5.3 and let r be the constant in Lemma 6.1. Choose r smaller if necessary to ensure that

$$r \leq \min\left\{\frac{1}{2c_1}, \frac{1}{182c_1^2(1+c_1)^3}\right\}, \tag{7.2}$$

where c_1 is the constant of Proposition 5.1. Let δ_1 be the minimum of the δ's in Proposition 5.3 and Lemma 6.1. Let Λ be the absolute largest eigenvalue for $\nabla^2 f(\mathbf{x})$ over $\mathbf{x} \in \mathcal{B}_{2r}$ and let c_2 be the constant in (6.5) and (6.6). The theorem is established for the following choice of δ:

$$\delta = \min\left\{1, \delta_1, \frac{r}{1+c_1}, \frac{1}{16c_2^2}\right\}. \tag{7.3}$$

If $\mathbf{x}_j \in \mathcal{B}_r(\mathbf{x}^*)$ for all $j \in [0, k]$, then $\lambda_j \leq \Lambda$ for all $j \in [1, k]$. By Proposition 5.1 and the fact that $r \leq 1/(2c_1)$, we have

$$\|\mathbf{x}_{j+1} - \hat{\mathbf{x}}_{j+1}\| \leq \frac{1}{2}\|\mathbf{x}_j - \hat{\mathbf{x}}_j\| \leq \cdots \leq \left\{\begin{array}{l}\left(\frac{1}{2}\right)^j \\ \left(\frac{1}{2}\right)^j \|\mathbf{x}_1 - \mathbf{x}^*\|\end{array}\right. \tag{7.4}$$

for all $j \in [0, k]$. The top inequality is due the fact that $\mathbf{x}_1 \in \mathcal{B}_\delta(\mathbf{x}^*)$ with $\delta \leq 1$, and the bottom inequality is due to the relation $\|\mathbf{x}_1 - \hat{\mathbf{x}}_1\| \leq \|\mathbf{x}_1 - \mathbf{x}^*\|$.

Suppose that $\mathbf{x}_j \in \mathcal{B}_r(\mathbf{x}^*)$ for all $j \in [0, k-1]$, and

$$\max\{|x_{ji}| : i \in \mathcal{A}\} > \|\hat{\mathbf{x}}_j - \mathbf{x}^*\|^{3/2} \tag{7.5}$$

for $j = k$. We now show that

$$\|\mathbf{x}_j - \mathbf{x}^*\| \le \left(\frac{1}{2}\right)^{j-1} \|\mathbf{x}_1 - \mathbf{x}^*\|, \text{ for } j = k \text{ or } j = k+1. \tag{7.6}$$

Observe that

$$\begin{aligned}
\|\mathbf{x}_k - \mathbf{x}^*\|^2 &\le \max\{|x_{ki}|^{4/3} : i \in \mathcal{A}\} + \sum_{i \in \mathcal{A}} x_{ki}^2 \\
&\le \left(\sum_{i \in \mathcal{A}} x_{ki}^2\right)^{2/3} + \sum_{i \in \mathcal{A}} x_{ki}^2 \le 2\left(\sum_{i \in \mathcal{A}} x_{ki}^2\right)^{2/3} \\
&\le 2\|\mathbf{x}_k - \hat{\mathbf{x}}_k\|^{4/3} \le 2(c_1\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\|^2)^{4/3} = 2c_1^{4/3}\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\|^{8/3}.
\end{aligned}$$

The first inequality is due to (7.5), the second inequality is from (7.4), and the fourth inequality is from Proposition 5.1. Taking square roots gives

$$\begin{aligned}
\|\mathbf{x}_k - \mathbf{x}^*\| \le \sqrt{2}c_1^{2/3}\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\|^{4/3} &= (\sqrt{2}c_2^{2/3}\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\|^{1/3})\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\| \\
&\le (\sqrt{2}c_2^{2/3}\|\mathbf{x}_{k-1} - \mathbf{x}^*\|^{1/3})\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\| \\
&\le (\sqrt{2}c_2^{2/3} r^{1/3})\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\|.
\end{aligned}$$

The constraint $r \le 1/[182c_1^2(1+c_1)^3]$ in (7.2) ensures that

$$\sqrt{2}c_2^{2/3} r^{1/3} \le 1/[4(1+c_1)].$$

This together with (7.4) imply

$$\begin{aligned}
\|\mathbf{x}_k - \mathbf{x}^*\| &\le \left(\frac{1}{4(1+c_1)}\right)\|\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}\| \\
&\le \frac{1}{4(1+c_1)}\left(\frac{1}{2}\right)^{k-2}\|\mathbf{x}_1 - \hat{\mathbf{x}}_1\| \le \frac{1}{2(1+c_1)}\left(\frac{1}{2}\right)^{k-1}\|\mathbf{x}_1 - \mathbf{x}^*\|.
\end{aligned} \tag{7.7}$$

This establishes (7.6) for $j = k$. Proposition 5.1 and (7.7) give

$$\begin{aligned}
\|\mathbf{x}_{k+1} - \mathbf{x}^*\| &\le \|\mathbf{x}_k - \mathbf{x}^*\| + \|\mathbf{s}_k\| \le (1+c_1)\|\mathbf{x}_k - \mathbf{x}^*\| \\
&\le \left(\frac{1}{2}\right)^k \|\mathbf{x}_1 - \mathbf{x}^*\|,
\end{aligned} \tag{7.8}$$

which establishes (7.6) for $j = k+1$.

Now, suppose that $\mathbf{x}_j \in \mathcal{B}_r(\mathbf{x}^*)$ for all $j \in [0, k-1]$, and for some integer $\ell > k$, (7.5) holds for all $j \in [k, \ell)$. Since $\delta \le r$, it follows from repeated application of (7.6) that

$$\mathbf{x}_j \in \mathcal{B}_\delta(\mathbf{x}^*) \quad \text{and} \quad \|\mathbf{x}_j - \mathbf{x}^*\| \le \left(\frac{1}{2}\right)^{j-1} \|\mathbf{x}_1 - \mathbf{x}^*\| \text{ for all } j \in [k, \ell]. \tag{7.9}$$

On the other hand, suppose that $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$ and (7.5) is violated for $j = k$. Let $\ell > 0$ be the smallest integer with the property that

$$\|\mathbf{z}_{k,\ell} - \mathbf{x}^*\| \le \frac{1}{2}\|\mathbf{z}_{k,0} - \mathbf{x}^*\|.$$

By Proposition 5.3, $\ell \le N$. By Lemma 6.1, (6.5) and (6.6) hold and we have

$$\begin{aligned} \|\mathbf{x}_{k+\ell} - \mathbf{x}^*\| &\le \|\mathbf{x}_{k+\ell} - \mathbf{z}_{k,\ell}\| + \|\mathbf{z}_{k,\ell} - \mathbf{x}^*\| \\ &\le c_2\|\mathbf{x}_k - \mathbf{x}^*\|^{3/2} + \frac{1}{2}\|\mathbf{z}_{k,0} - \mathbf{x}^*\| \\ &= c_2\|\mathbf{x}_k - \mathbf{x}^*\|^{3/2} + \frac{1}{2}\|\mathbf{x}_k - \mathbf{x}^*\| \\ &\le (c_2\sqrt{\delta} + 1/2)\|\mathbf{x}_k - \mathbf{x}^*\| \\ &\le \frac{3}{4}\|\mathbf{x}_k - \mathbf{x}^*\|. \end{aligned} \tag{7.10}$$

The last inequality is due to the fact that $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$ where $\delta \le 1/(16c_2^2)$. The last inequality also implies that $\mathbf{x}_{k+\ell} \in \mathcal{B}_\delta(\mathbf{x}^*)$. Moreover, by (6.5) we have $\mathbf{x}_j \in \mathcal{B}_r(\mathbf{x}^*)$ and $\lambda_j \le \Lambda$ for all $j \in [k, k+\ell]$.

Starting with $k_1 = 1$, we apply either (7.9) or (7.10) to generate a sequence of iterates $\mathbf{x}_{k_i}$, $i = 1, 2, \ldots$, for which the error tends to zero at a geometric rate. If the sequence has reached $k = k_i$, then k_{i+1} is obtained in the following way:

A. If (7.5) holds for $j = k_i$, then $k_{i+1} = \ell$ where ℓ (possibly infinite) is chosen so that (7.5) holds for all $j \in [k_i, \ell)$ and

$$\max\{|x_{\ell i}| : i \in \mathcal{A}\} \le \|\hat{\mathbf{x}}_\ell - \mathbf{x}^*\|^{3/2}.$$

B. If (7.5) is violated for $j = k_i$, then $k_{i+1} = k_i + \ell$ where ℓ is chosen so that (7.10) holds for $k = k_i$.

If k_{i+1} is chosen in accordance with Rule B, then since $k_{i+1} - k_i \le N$, we have

$$\begin{aligned} \|\mathbf{x}_{k_{i+1}} - \mathbf{x}^*\| &\le \left(\frac{3}{4}\right)\|\mathbf{x}_{k_i} - \mathbf{x}^*\| \le \left(\frac{3}{4}\right)^{(k_{i+1}-k_i)/N} \|\mathbf{x}_{k_i} - \mathbf{x}^*\| \\ &= \gamma^{k_{i+1}-k_i}\|\mathbf{x}_{k_i} - \mathbf{x}^*\|, \quad \text{where } \gamma = \left(\frac{3}{4}\right)^{1/N}. \end{aligned} \tag{7.11}$$

If k_{i+1} is chosen in accordance with Rule A, then by (7.9) and the fact that $\gamma \ge 1/2$, we have

$$\|\mathbf{x}_k - \mathbf{x}^*\| \le \gamma^{k-1}\|\mathbf{x}_1 - \mathbf{x}^*\| = \gamma^{k-k_1}\|\mathbf{x}_1 - \mathbf{x}^*\| \quad \text{for all } k \in [k_i, k_{i+1}]. \tag{7.12}$$

Together, (7.11) and (7.12) imply that for each i,

$$\|\mathbf{x}_{k_i} - \mathbf{x}^*\| \le \gamma^{k_i - k_1} \|\mathbf{x}_1 - \mathbf{x}^*\| = \gamma^{-1}\gamma^{k_i}\|\mathbf{x}_1 - \mathbf{x}^*\|. \tag{7.13}$$

By Lemma 6.1, we know that for any $\mathbf{x}_k \in \mathcal{B}_\delta(\mathbf{x}^*)$ satisfying (6.3)–(6.4), the relations (6.5)–(6.6) hold for all $j \in [0, \ell]$. Moreover, since $j \le \ell \le N$, it follows from (6.14) that

$$\|\mathbf{x}_{k+j} - \mathbf{x}^*\| \le (1 + Nc_2)\|\mathbf{x}_k - \mathbf{x}^*\|.$$

If $k \in [k_i, k_{i+1})$ where k_{i+1} is chosen in accordance with Rule B, then we have

$$\begin{aligned} \|\mathbf{x}_k - \mathbf{x}^*\| &\le (1 + Nc_2)\|\mathbf{x}_{k_i} - \mathbf{x}^*\| \\ &\le \gamma^{-1}(1 + Nc_2)\gamma^{k_i}\|\mathbf{x}_1 - \mathbf{x}^*\| \le \eta\gamma^k\|\mathbf{x}_1 - \mathbf{x}^*\|, \end{aligned} \tag{7.14}$$

where $\eta = (1 + Nc_2)/\gamma^{N+1}$. If $k \in [k_i, k_{i+1})$ where k_{i+1} is chosen in accordance with Rule A, then by (7.9), we have

$$\|\mathbf{x}_k - \mathbf{x}^*\| \le \left(\frac{1}{2}\right)^{k-1} \|\mathbf{x}_1 - \mathbf{x}^*\| \le \gamma^{k-1}\|\mathbf{x}_1 - \mathbf{x}^*\| \le \eta\gamma^k\|\mathbf{x}_1 - \mathbf{x}^*\|. \tag{7.15}$$

Together, (7.13)–(7.15) complete the proof. □

8 Box constraints

We now generalize AS_CBB to handle the box constrained problem (1.1). In this case, the definition of $\mathbf{X}^1$ in (1.7) is replaced by

$$X_i(\mathbf{x}) = \begin{cases} u_i - x_i & \text{if } \; g_i(\mathbf{x}) \le 0, \\ x_i - l_i & \text{otherwise.} \end{cases}$$

With the convention that $\infty \times 0 = 0$, the KKT conditions can be expressed

$$\mathbf{X}(\mathbf{x}) \circ \mathbf{g}(\mathbf{x}) = \mathbf{0} \quad \text{and} \quad \mathbf{l} \le \mathbf{x} \le \mathbf{u}.$$

With the convention that $1/\infty = 0$, the new approximation to the Newton search direction is

$$d_{ki} = -\frac{1}{\lambda_k + |g_i(\mathbf{x}_k)|/X_i(\mathbf{x}_k)}\, g_i(\mathbf{x}_k).$$

In the special case considered earlier where $l_i = 0$ and $u_i = \infty$, we have $X_i(x) = \infty$ when $g_i(\mathbf{x}) \le 0$ and $|g_i(\mathbf{x})|/X_i(\mathbf{x}) = 0$, exactly as in (1.5).

9 Experimental studies

In our first set of numerical experiments, we investigate how the convergence speed of AS_CBB depends on the problem condition number. We consider the linear least square problem:

$$\min\{\|\mathbf{Ax} - \mathbf{b}\|^2 : \mathbf{x} \geq \mathbf{0}\},$$

where the elements of $\mathbf{b}$ are chosen randomly on the interval $[-1, 1]$, and $\mathbf{A}$ is a 20 by 10 matrix with 10% of its elements nonzero, and with condition number varying between 10 and 10^8. Hence, the condition number of the matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ associated with the quadratic objective function varies between 100 and 10^{16}. Our matrix $\mathbf{A}$ was generated using MATLAB's `sprand` command. The syntax is as follows:

```
A = sprand (m, n, density, 1/condA)
```

where `m` = 20, `n` = 10, `density` = 0.1, and `condA` had values between 10 and 10^8. The vector $\mathbf{b}$ was generated using MATLAB's `rand` command with the following syntax:

```
b = 2*rand (m, 1) - 1.
```

Hence, $\mathbf{b}$ typically lies outside the column space of $\mathbf{A}$. In our implementation of AS_CBB, the cycle length in the cyclic BB iteration (2.1) was $m = 4$, while the memory in the nonmonotone line search was $M = 8$. In Fig. 2 we plot the quantity

$$\log_{10}(\|(\mathbf{x}_k - \mathbf{g}_k)^+ - \mathbf{x}_k\|_\infty)$$

versus the iteration number, where $\|\cdot\|_\infty$ is the maximum absolute component of a vector.

These results indicate that the convergence speed is relatively insensitive to the problem conditioning. The matrices $\mathbf{A}$ and the vector $\mathbf{b}$ associated with the plots in Fig. 2 can be found at the following web site for the paper:

http://www.math.ufl.edu/~hager/papers/PET

In the second set of numerical experiments, we compare the convergence speed of AS_CBB to that of the conjugate gradient-based active set algorithm ASA developed in [20] using an image reconstruction problem that arises in Positron Emission Tomography. The data, obtained from a PET scan of the thorax, and a Fortran code to evaluate the cost function and the gradient are also available at the web site for the paper.

Figure 3 plots $\log_{10}$(error) versus the iteration number. Although the conjugate gradient-based algorithm reaches the 10^{-6} error level in less than 500 iterations, AS_CBB is able to reduce the error more quickly during the initial iterations; for example, the error is less than 10^{-1} within 16 iterations, while ASA uses 50 iterations to reach the same error level. Since the line search in the conjugate gradient routine

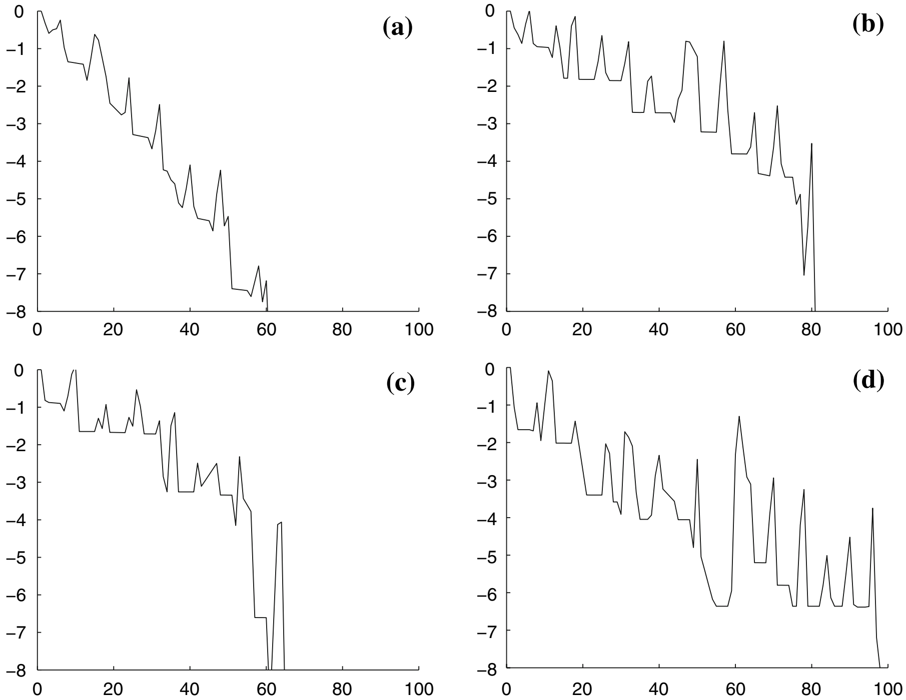

Fig. 2 Iteration number versus $\log_{10}$(error) for matrices with various condition numbers: **a** 10^2, **b** 10^4, **c** 10^8, **d** 10^{16}

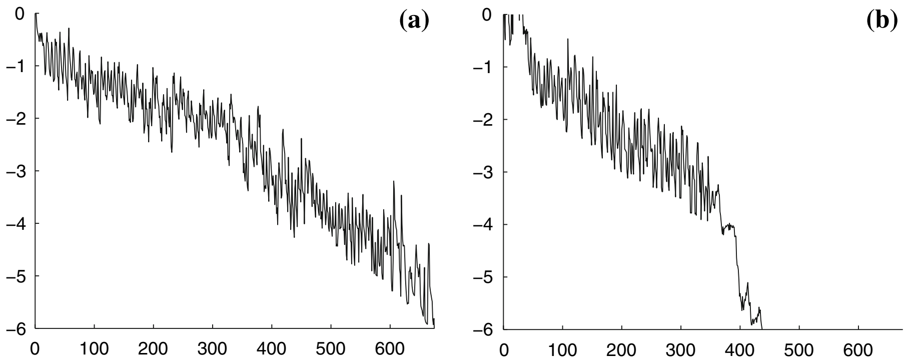

Fig. 3 Iteration number versus $\log_{10}$(error) for **a** AS_CBB and **b** ASA (**a** conjugate gradient-based active set algorithm)

requires both the function and the gradient (to satisfy the Wolfe conditions), while the Armijo line search in AS_CBB only requires the function value, AS_CBB is initially much faster than the conjugate gradient code; there are fewer iterations to achieve a given error tolerance and each iteration requires fewer gradient evaluations. Due to the asymptotic superiority of the conjugate gradient code, it eventually surpasses

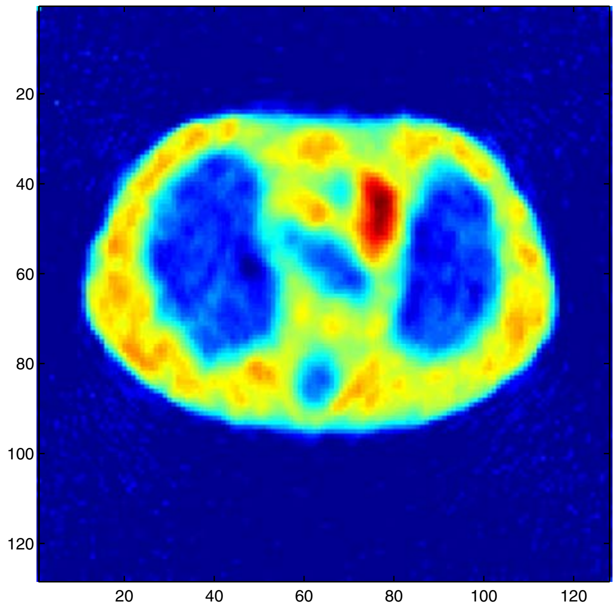

Fig. 4 Reconstructed image for error tolerance 10^{-4}

AS_CBB in speed as the error tolerance becomes tiny. However, in practice a low accuracy solution of (1.10) is often sufficient. The reconstructed image corresponding to the error tolerance 10^{-4} appears in Fig. 4

10 Conclusions

We develop a new affine scaling algorithm for box constrained optimization. This algorithm was obtained by applying Newton's method to the first-order optimality conditions and approximating the Hessian matrix at iteration k by $\lambda_k \mathbf{I}$ where λ_k is obtained from a quasi-Newton condition. This approximation to the Hessian leads to a scaled iterate which lies in the interior of the feasible set. This feature is especially useful for problems where the cost function is infinite on the boundary of the feasible set. We obtain λ_k using a cyclic Barzilai–Borwein stepsize given by (1.11) and (2.1). By Theorem 4.1, the resulting affine-scaling cyclic Barzilai–Borwein algorithm AS_CBB is globally convergent when implemented using a nonmonotone line search shown in Fig. 1. By Theorem 7.1, the algorithm is locally R-linearly convergent at a nondegenerate local minimizer where the second-order sufficient optimality condition holds. As seen in Fig. 2, the convergence of AS_CBB is relatively insensitive to the problem condition number. We compared performance to that of the asymptotically faster active set algorithm ASA using a test problem which arises in Positron Emission Tomography (PET). AS_CBB is initially much faster than ASA, while ASA is faster in the limit, as the error tolerance becomes tiny.

References

1. Akaike, H.: On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. Ann. Inst. Stat. Math. Tokyo **11**, 1–17 (1959)
2. Barzilai, J., Borwein, J.M.: Two point step size gradient methods. IMA J. Numer. Anal. **8**, 141–148 (1988)
3. Birgin, E.G., Biloti, R., Tygel, M., Santos, L.T.: Restricted optimization: a clue to a fast and accurate implementation of the common reflection surface stack method. J. Appl. Geophys. **42**, 143–155 (1999)
4. Birgin, E.G., Chambouleyron, I., Martínez, J.M.: Estimation of the optical constants and the thickness of thin films using unconstrained optimization. J. Comput. Phys. **151**, 862–880 (1999)
5. Birgin, E.G., Martínez, J.M., Raydan, M.: Nonmonotone spectral projected gradient methods for convex sets. SIAM J. Optim. **10**, 1196–1211 (2000)
6. Chang, J.-H., Anderson, J.M.M., Mair, B.A.: An accelerated penalized maximum likelihood algorithm for positron emission tomography. IEEE Trans. Nucl. Sci. **54**(5), 1648–1659
7. Chang, J.-H., Anderson, J.M.M., Votaw, J.R.: Regularized image reconstruction algorithms for positron emission tomography. IEEE Trans. Med. Imag. **23**, 1165–1195 (2004)
8. Coleman, T.F., Li, Y.: On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. Math. Prog. **67**, 189–224 (1994)
9. Coleman, T.F., Li, Y.: An interior trust region approach for nonlinear minimization subject to bounds. SIAM J. Optim. **6**, 418–445 (1996)
10. Coleman, T.F., Li, Y.: A trust region and affine scaling interior point method for nonconvex minimization with linear inequality constraints. Math. Program. **88**, 1–31 (2000)
11. Dai, Y.H.: Alternate stepsize gradient method. Optimization **52**, 395–415 (2003)
12. Dai, Y.H., Fletcher, R.: Projected Barzilai–Borwein methods for large-scale box-constrained quadratic programming. Numer. Math. **100**, 21–47 (2005)
13. Dai, Y.H., Hager, W.W., Schittkowski, K., Zhang, H.: The cyclic Barzilai–Borwein method for unconstrained optimization. IMA J. Numer. Anal. **26**, 604–627 (2006)
14. Dai, Y.H., Zhang, H.: An adaptive two-point stepsize gradient algorithm. Numer. Algorithms **27**, 377–385 (2001)
15. Dikin, I.I.: Iterative solution of problems of linear and quadratic programming. Sov. Math. Dokl. **8**, 674–675 (1967)
16. Dostál, Z.: Box constrained quadratic programming with proportioning and projections. SIAM J. Optim. **7**, 871–887 (1997)
17. Fletcher, R.: On the Barzilai–Borwein method. Technical Report, Department of Mathematics, University of Dundee, Dundee (2001)
18. Friedlander, A., Martínez, J.M., Molina, B., Raydan, M.: Gradient method with retards and generalizations. SIAM J. Numer. Anal. **36**, 275–289 (1999)
19. Glunt, W., Hayden, T.L., Raydan, M.: Molecular conformations from distance matrices. J. Comput. Chem. **14**, 114–120 (1993)
20. Hager, W.W., Zhang, H.: A new active set algorithm for box constrained optimization. SIAM J. Optim. **17**, 526–557 (2006)
21. Heinkenschloss, M., Ulbrich, M., Ulbrich, S.: Superlinear and quadratic convergence of affine-scaling interior-point Newton methods for problems with simple bounds without strict complementarity assumption. Math. Program. **86**, 615–635 (1999)
22. Johnson, C.A., Seidel, J., Sofer, A.: Interior-point methodology for 3-D PET reconstruction. IEEE Trans. Med. Imag. **19**, 271–285 (2000)
23. Kanzow, C., Klug, A.: On affine-scaling interior-point Newton methods for nonlinear minimization with bound constraints. Comput. Optim. Appl. **35**, 177–197 (2006)
24. Kaufman, L.: Implementing and accelerating the EM algorithm for positron emission tomography. IEEE Trans. Med. Imag. **6**, 37–51 (1987)
25. Kaufman, L.: Maximum likelihood, least squares, and penalized least squares for PET. IEEE Trans. Med. Imag. **12**, 200–214 (1993)
26. Raydan, M.: The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. SIAM J. Optim. **7**, 26–33 (1997)
27. Zhang, H., Hager, W.W.: A nonmonotone line search technique and its application to unconstrained optimization. SIAM J. Optim. **14**, 1043–1056 (2004)

28. Zhang, H., Hager, W.W.: PACBB: a projected adaptive cyclic Barzilai–Borwein method for box constrained optimization. In: Hager, W.W., Huang, S.-J., Pardalos, P.M., Prokopyev, O.A. (eds.) Multiscale Optimization Methods and Applications, pp. 387–392. Springer, New York (2005)
29. Zhang, Y.: Interior-point gradient methods with diagonal-scalings for simple-bound constrained optimization. Technical Report, TR04-06, Department of Computational and Applied Mathematics. Rice University, Houston (2004)