Nonlinear programming without a penalty function or a filter

by N. I. M. Gould[1] and Ph. L. Toint[2]

Report 07/02 13th April 2007

[1] Oxford University Computing Laboratory,
Wolfson Building, Parks Road,
Oxford OX1 3QD, England.
Email: nick.gould@comlab.ox.ac.uk

[2] Department of Mathematics,
FUNDP-University of Namur,
61, rue de Bruxelles, B-5000 Namur, Belgium.
Email: philippe.toint@fundp.ac.be

Nonlinear programming without a penalty function or a filter

N. I. M. Gould and Ph. L. Toint

13th April 2007

Abstract

A new method is introduced for solving equality constrained nonlinear optimization problems. This method does not use a penalty function, nor a barrier or a filter, and yet can be proved to be globally convergent to first-order stationary points. It uses different trust-regions to cope with the nonlinearities of the objective function and the constraints, and allows inexact SQP steps that do not lie exactly in the nullspace of the local Jacobian. Preliminary numerical experiments on CUTEr problems indicate that the method performs well.

Keywords: Nonlinear optimization, equality constraints, numerical algorithms, global convergence.

1 Introduction

We consider the numerical solution of the equality constrained nonlinear optimization problem

$$\begin{cases} \min\limits_{x} & f(x) \\ & c(x) = 0, \end{cases} \tag{1.1}$$

where we assume that $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable and that f is bounded below on the feasible domain.

The present paper introduces a new method for the solution of (1.1), which belongs to the class of trust-region methods for constrained optimization, in the spirit of Omojokun (1989) in a Ph.D. thesis supervised by R. Byrd, and later developed by several authors, including Biegler, Nocedal and Schmid (1995), El-Alem (1995, 1999), Byrd, Gilbert and Nocedal (2000*a*), Byrd, Hribar and Nocedal (2000*b*), Liu and Yuan (2000) and Lalee, Nocedal and Plantenga (1998) (also see Chapter 15 of Conn, Gould and Toint, 2000).

The algorithm presented here has four main features. The first is that it attempts to consider the objective function and the constraints as independently as possible by using different models and trust regions for f and c. As is common to the methods cited, the steps are computed as a combination of normal and tangential components, the first aiming to reduce the constraint violation, and the second at reducing the objective function while retaining the improvement in violation by remaining in the plane tangent to the constraints, but only approximately so. This framework can thus be viewed as a sequential quadratic programming technique that allows for inexact tangential steps, which is the second main characteristic of our proposal (shared with Heinkenschloss and Vicente, 2001, and the recent paper by Byrd, Curtis and Nocedal, 2006). The third distinctive feature is that the algorithm is not compelled to compute both normal and tangential steps at every iteration, rather only to compute whichever is/are likely to improve feasibility and optimality significantly. Thus if an iterate is almost feasible, there is little point in trying to further improve feasibility while the objective value is far from optimal. The final central feature is that the algorithm does not use any merit function (penalty, barrier, or otherwise), thereby avoiding the practical problems associated with the setting of the

merit function parameters, but nor does it use the filter idea first proposed by Fletcher and Leyffer (2002). Instead, the convergence is driven by the *trust funnel*, a progressively decreasing limit on the permitted infeasibility of the successive iterates.

It is, in that sense and albeit very indirectly, reminiscent of the "flexible tolerance method" by Himmelblau (1972), but also of the "tolerance tube method" by Zoppke-Donaldson (1995) and the SQP method by Bielschowsky and Gomes (2006). All these methods use the idea of progressively reducing constraint violation to avoid using a penalty parameter. Both of the more modern algorithms are of the trust-region type, but differ significantly from our proposal. The first major difference is that they both require the tangential component of the step to lie exactly in the Jacobian's nullspace: they are thus "exact" rather than "inexact" SQP methods. The second is that they both use a single trust region to account simultaneously for constraint violation and objective function improvement. The third is that both limit constraint violation *a posteriori*, once the true nonlinear constraints have been evaluated, rather than attempting to limit its predicted value *a priori*. The "tolerance tube" method resorts to standard second-order correction steps when the iterates become too infeasible. No convergence seems to be available for the method, although the numerical results appear satisfactory. At variance, the method by Bielschowsky and Gomes (2006) is provably globally convergent to first-order critical points. It however involves a "restoration" phase (whose convergence is assumed) to achieve acceptable constraint violation in which the size of normal component of the step is restricted to be a fraction of the current infeasibility limit. This limit is updated using the gradient of the Lagrangian function, and the allowable fraction is itself computed from the norm of exact projection of the objective function gradient onto the nullspace of the constraints' Jacobian.

The paper is organized as follows. Section 2 introduces the new algorithm, whose convergence theory is presented in Section 3. Section 4 presents preliminary numerical results on CUTEr test problems; conclusions and perspectives are finally outlined in Section 5.

2 A trust-funnel algorithm

Let us measure, for any x, the constraint violation at x by

$$\theta(x) \stackrel{\rm def}{=} \tfrac{1}{2}\|c(x)\|^2 \tag{2.1}$$

where $\|\cdot\|$ denotes the Euclidean norm. Now consider iteration k, starting from the iterate x_k, for which we assume we know a bound $\theta_k^{\max}$ such that $\tfrac{1}{2}\|c(x_k)\|^2 < \theta_k^{\max}$.

Firstly, a *normal step* n_k is computed if the constraint violation is significant (in a sense to be defined shortly). This is achieved by reducing the Gauss-Newton approximation

$$\tfrac{1}{2}\|c_k + J_k n\|^2 \tag{2.2}$$

to $\theta(x_k+n_k)$—here we write $c_k \stackrel{\rm def}{=} c(x_k)$ and $J_k \stackrel{\rm def}{=} J(x_k)$ is the Jacobian of c at x_k—while requiring that n_k remains in the "normal trust region", i.e.,

$$n_k \in \mathcal{N}_k \stackrel{\rm def}{=} \{v \in \mathbb{R}^n \mid \|v\| \leq \Delta_k^c\}. \tag{2.3}$$

More formally, this Gauss-Newton-type step is computed by choosing n_k so that (2.2) is reduced sufficiently within $\mathcal{N}_k$ in the sense that

$$\delta_k^{c,n} \stackrel{\rm def}{=} \tfrac{1}{2}\|c_k\|^2 - \tfrac{1}{2}\|c_k + J_k n_k\|^2 \geq \kappa_{\rm nC}\|J_k^T c_k\| \min\left[\frac{\|J_k^T c_k\|}{1+\|W_k\|}, \Delta_k^c\right] \geq 0, \tag{2.4}$$

where $W_k = J_k^T J_k$ is the symmetric Gauss-Newton approximation of the Hessian of θ at x_k and $\kappa_{\rm nC} > 0$. Condition (2.4) is nothing but the familiar Cauchy condition for problem approximately minimizing (2.2) within the region $\mathcal{N}_k$. In addition, we also require the

normal step to be "normal", in that it mostly lies in the space spanned by the columns of the matrix J_k^T by imposing that

$$\|n_k\| \leq \kappa_{\rm n}\|c_k\| \tag{2.5}$$

for some $\kappa_{\rm n} > 0$. These conditions on the normal step are very reasonable in practice, as it is known that they hold if, for instance, n_k is computed by applying one or more steps of a truncated conjugate-gradient method (see Toint, 1981, and Steihaug, 1983) to the minimization of the square of the linearized infeasibility. Note that the conditions (2.3), (2.4) and (2.5) allow us to choose a null normal step ($n_k = 0$) if x_k is feasible.

Having computed the normal step, we next consider if some improvement is possible on the objective function, while not jeopardizing the infeasibility reduction we have just obtained. Because of this latter constraint, it makes sense to remain in $\mathcal{N}_k$, the region where we believe that our model of constraint violation can be trusted, but we also need to trust the model of the objective function given, as is traditional in sequential quadratic programming (see Section 15.2 of Conn et al., 2000), by

$$m_k(x_k + n_k + t) = f_k + \langle g_k^{\rm N}, t\rangle + \tfrac{1}{2}\langle t, G_k t\rangle \tag{2.6}$$

where

$$g_n^{\rm N} \stackrel{\rm def}{=} g_k + G_k n_k, \tag{2.7}$$

where $f_k = f(x_k)$, $g_k = \nabla f(x_k)$ and where G_k is a symmetric approximation of the Hessian of the Lagrangian $\ell(x, y) = f(x) + \langle y, c(x)\rangle$ given by

$$G_k \stackrel{\rm def}{=} H_k + \sum_{i=1}^{m} [\hat{y}_k]_i C_{ik}. \tag{2.8}$$

In this last definition, H_k is a bounded symmetric approximation of $\nabla^2 f(x_k)$, the matrices C_{ik} are bounded symmetric approximations of the constraints' Hessians $\nabla_{xx} c_i(x_k)$ and the vector $\hat{y}_k$ may be viewed as an approximation of the local Lagrange multipliers, in the sense that we require that

$$\|\hat{y}_k\|\|c_k\| \leq \kappa_{\rm y} \tag{2.9}$$

for some $\kappa_{\rm y} > 0$. Note that this condition does not impose any practical size restriction on $\hat{y}_k$ close to the feasible set, and therefore typically allows the choice $\hat{y}_k = y_{k-1}$, for suitable multiplier estimates y_{k-1} computed during the previous iteration, when x_k is close to feasibility. We assume that (2.6) can be trusted as a representation of $f(x_k + n_k + t)$ provided the complete step $s = n_k + t$ belongs to

$$\mathcal{T}_k \stackrel{\rm def}{=} \{s \in \mathbb{R}^n \mid \|s\| \leq \Delta_k^f\}, \tag{2.10}$$

for some radius Δ_k^f. Thus our attempts to reduce (2.6) should be restricted to to the intersection of $\mathcal{N}_k$ and $\mathcal{T}_k$, which imposes that the *tangential step* t_k results in a complete step $s_k = n_k + t_k$ that satisfies the inclusion

$$s_k \in \mathcal{B}_k \stackrel{\rm def}{=} \mathcal{N}_k \cap \mathcal{T}_k \stackrel{\rm def}{=} \{s \in \mathbb{R}^n \mid \|s\| \leq \Delta_k\}, \tag{2.11}$$

where the radius Δ_k of $\mathcal{B}_k$ is thus given by

$$\Delta_k = \min[\Delta_k^c, \Delta_k^f]. \tag{2.12}$$

As a consequence, it makes sense to ask n_k to belong to $\mathcal{B}_k$ before attempting the computation of t_k, which we formalize by requiring that

$$\|n_k\| \leq \kappa_{\rm B}\Delta_k, \tag{2.13}$$

for some $\kappa_{\rm B} \in (0, 1)$. We note here that using two different trust-region radii can be considered as unusual, but is not unique. For instance, the SLIQUE algorithm described

by Byrd, Gould, Nocedal and Waltz (2004) also uses different radii, but for different models of the same function, rather than for two different functions.

We still have to specify what we mean by "reducing (2.6)", as we are essentially interested in the reduction in the hyperplane tangent to the constraints. In order to compute an approximate projected gradient at $x_k + n_k$, we first compute a new local estimate of the Lagrange multipliers y_k such that

$$\|y_k + [J_k^T]^I g_k^{\rm N}\| \leq \omega_1(\|c_k\|) \tag{2.14}$$

for some monotonic bounding function[(1)] ω_1, the superscript ^{I} denoting the Moore-Penrose generalized inverse, and such that

$$\|r_k\| \leq \kappa_{\rm nr}\|g_k^{\rm N}\| \tag{2.15}$$

for some $\kappa_{\rm nr} > 0$, and

$$\langle g_k^{\rm N}, r_k\rangle \geq 0, \tag{2.16}$$

where

$$r_k \stackrel{\rm def}{=} g_k^{\rm N} + J_k^T y_k \tag{2.17}$$

is an approximate projected gradient of the model m_k at $x_k + n_k$. Conditions (2.14)–(2.16) are reasonable since they are obviously satisfied by choosing y_k to be a solution of the least-squares problem

$$\min_y \tfrac{1}{2}\|g_k^{\rm N} + J_k^T y\|^2, \tag{2.18}$$

and thus, by continuity, by sufficiently good approximations of this solution. In practice, one can compute such an approximation by applying a Krylov space iterative method starting from $y = 0$. If the solution of (2.18) is accurate, r_k is the orthogonal projection of $g_k^{\rm N}$ onto the nullspace of J_k, which then motivates that we then require the tangent step to produce a reduction in the model m_k which is at least a fraction of that achieved by solving the modified Cauchy point subproblem

$$\min_{\substack{\tau>0 \\ x_k+n_k-\tau r_k\in\mathcal{B}_k}} m_k(x_k + n_k - \tau r_k), \tag{2.19}$$

where we have assumed that $\|r_k\| > 0$. We know from Section 8.1.5 of Conn et al. (2000) that this procedure ensures, for some $\kappa_{\rm tC1} \in (0, 1]$, the modified Cauchy condition

$$\delta_k^{f,t} \stackrel{\rm def}{=} m_k(x_k + n_k) - m_k(x_k + n_k + t_k) \geq \kappa_{\rm tC1}\pi_k \min\left[\frac{\pi_k}{1 + \|G_k\|}, \tau_k\|r_k\|\right] > 0 \tag{2.20}$$

on the decrease of the objective function model within $\mathcal{B}_k$, where we have set

$$\pi_k \stackrel{\rm def}{=} \frac{\langle g_k^{\rm N}, r_k\rangle}{\|r_k\|} \geq 0 \tag{2.21}$$

(by convention, we define $\pi_k = 0$ whenever $r_k = 0$), and where

$$\tau_k = \frac{-\beta_k + \sqrt{\beta_k^2 + \Delta_k^2 - \|n_k\|^2}}{\|r_k\|} \tag{2.22}$$

is the maximal steplength along $-r_k$ from $x_k + n_k$ which remains in the trust-region $\mathcal{B}_k$, where we have used the definition $\beta_k \stackrel{\rm def}{=} \langle n_k, r_k\rangle/\|r_k\|$. We then require that the length of that step is comparable to the radius of $\mathcal{B}_k$, in the sense that, for some $\kappa_{\rm r} \in (0, \sqrt{1 - \kappa_{\rm B}^2})$,

$$\tau_k\|r_k\| \geq \kappa_{\rm r}\Delta_k \tag{2.23}$$

[(1)] Here and later in this paper, a *bounding function* ω is defined to be a continuous function from $\mathbb{R}_+$ into $\mathbb{R}$ with the property that $\omega(t)$ converges to zero as t tends to zero.

When n_k lies purely in the range of J_k^T and the least-squares problem (2.18) is solved accurately, then $\beta_k = 0$ and (2.23) holds with $\kappa_{\rm r} = \sqrt{1-\kappa_{\rm B}^2}$ because of (2.13). Hence (2.23) must hold with a smaller value of $\kappa_{\rm r}$ if (2.18) is solved accurately enough. As a result, the modified Cauchy condition (2.20) may now be rewritten as

$$\delta_k^{f,t} \stackrel{\rm def}{=} m_k(x_k+n_k) - m_k(x_k+n_k+t_k) \geq \kappa_{\rm tC}\pi_k \min\left[\frac{\pi_k}{1+\|G_k\|}, \Delta_k\right] \tag{2.24}$$

with $\kappa_{\rm tC} \stackrel{\rm def}{=} \kappa_{\rm tC1}\kappa_{\rm r} \in (0,1)$. We see from (2.24) that π_k may be considered as an optimality measure in the sense that it measures how much decrease could be obtained locally along the negative of the approximate projected gradient r_k. This role as an optimality measure is confirmed in Lemma 3.2 below.

Our last requirement on the tangential step t_k is to ensure that it does not completely "undo" the improvement in linearized feasibility obtained from the normal step without good reason. We consider two possible situations. The first is when the predicted decrease in the objective function is substantial compared to its possible deterioration along the normal step and the step is not to large compared to the maximal allowable infeasibility, i.e. when both

$$\delta_k^{f,t} \geq -\bar\kappa_\delta \delta_k^{f,n} \stackrel{\rm def}{=} -\bar\kappa_\delta[m_k(x_k) - m_k(x_k+n_k)] \tag{2.25}$$

and

$$\|s_k\| \leq \kappa_\Delta \sqrt{\theta_k^{\max}}, \tag{2.26}$$

for some $\bar\kappa_\delta \in (0,1)$ and some $\kappa_\Delta > 0$. In this case, we allow more freedom in the linearized feasibility and merely require that

$$\tfrac{1}{2}\|c_k + J_k(n_k+t_k)\|^2 \leq \kappa_{\rm tt}\theta_k^{\max} \tag{2.27}$$

for some $\kappa_{\rm tt} \in (0,1)$. If, on the other hand, (2.25) or (2.26) fails, meaning that we cannot hope to trade some decrease in linearized feasibility for a large improvement in objective function value over a reasonable step, then we require that the tangential step satisfies

$$\|c_k + J_k(n_k+t_k)\|^2 \leq \kappa_{\rm nt}\|c_k\|^2 + (1-\kappa_{\rm nt})\|c_k + J_k n_k\|^2 \stackrel{\rm def}{=} \vartheta_k, \tag{2.28}$$

for some $\kappa_{\rm nt} \in (0,1)$. Note that this inequality is already satisfied at the end of the normal step since $\|c_k + J_k n_k\| \leq \|c_k\|$ and thus already provides a relaxation of the (linearized) feasibility requirement at $x_k + n_k$. Figure 2.1 on the following page illustrate the geometry of the various quantities involved in the construction of a step s_k satisfying (2.28)

Finally, we observe that a tangential step does not make too much sense if $r_k = 0$, and we do not compute any. By convention we choose to define $\pi_k = 0$ and $t_k = 0$ in this case. The situation is similar if π_k is small compared to the current infeasibility. Given a monotonic bounding function ω_2, we thus decide that if

$$\pi_k > \omega_2(\|c_k\|), \tag{2.29}$$

fails, then the current iterate is still too far from feasibility to worry about optimality, and we also skip the tangential step computation by setting $t_k = 0$.

In the same spirit, the attentive reader may have observed that we have imposed the current violation to be "significant" as a condition to compute the normal step n_k, but didn't specify what we formally meant, because our optimality measure π_k was not defined at that point. We now complete our description by requiring that, for some bounding function ω_3, we require the computation of the normal step only when

$$\|c_k\| \geq \omega_3(\pi_{k-1}) \tag{2.30}$$

when $k > 0$. If (2.30) fails, we remain free to compute a normal step, but we may also skip it. In this latter case, we simply set $n_k = 0$. For technical reasons which will become clear below, we impose the additional conditions that

$$\omega_3(t) = 0 \Longleftrightarrow t = 0 \quad \text{and} \quad \omega_2(\omega_3(t)) \leq \kappa_\omega t \tag{2.31}$$

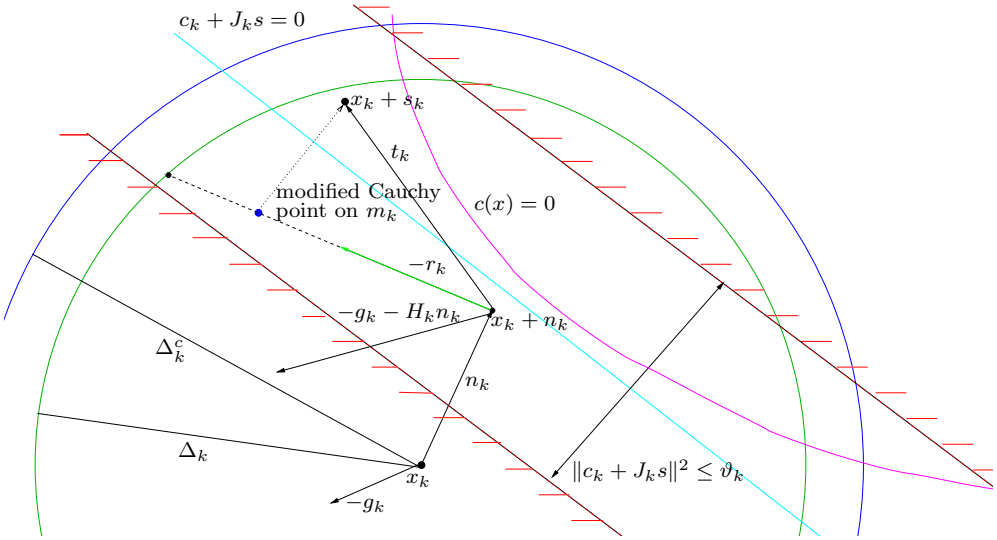

Figure 2.1: The components of a step s_k satisfying (2.28) in the case where $\Delta_k^f = \Delta_k^c$.

for all $t \geq 0$ and for some $\kappa_\omega \in (0, 1)$.

While (2.29) and (2.30) together provide considerable flexibility in our algorithm in that a normal or tangential step is only computed when relevant, our setting also produce the possibility that both these conditions fail. In this case, we have that $s_k = n_k + t_k$ is identically zero, and the sole computation in the iteration is that of the new Lagrange multiplier y_k; we will actually show that such behaviour cannot persist unless x_k is optimal.

Once we have computed the step s_k and the trial point

$$x_k^+ \stackrel{\text{def}}{=} x_k + s_k \tag{2.32}$$

completely, we are left with the task of accepting or rejecting it. Our proposal is based on the distinction between *f-iterations* and *c-iterations*, in the spirit of Fletcher and Leyffer (2002), Fletcher, Leyffer and Toint (2002*b*) or Fletcher, Gould, Leyffer, Toint and Wächter (2002*a*). Assuming that $s_k \neq 0$, we will say that iteration k is an f-iteration if a nonzero tangential step t_k has been computed and if

$$\delta_k^f \stackrel{\text{def}}{=} m_k(x_k) - m_k(x_k + s_k) \geq \kappa_\delta \delta_k^{f,t} \tag{2.33}$$

with $\kappa_\delta = 1 - 1/\bar{\kappa}_\delta$, and

$$\theta(x_k^+) \leq \theta_k^{\max}. \tag{2.34}$$

If $s_k \neq 0$ and one of (2.33) or (2.34) fails or if no tangential has been computed, because (2.13) or (2.29) fails, iteration k is said to be a c-iteration. Inequality (2.33) indicates that the improvement in the objective function obtained in the tangential step is not negligible compared to the change in f resulting from the normal step, while at the same time, keeping feasibility within reasonable bounds, as expressed by (2.34). Thus the iteration's expected major achievement is, in this case, a decrease in the value of the objective function f, hence its name. If (2.33) fails, then the expected major achievement (or failure) of iteration k is, *a contrario*, to improve feasibility, which is also the case when the step only contains its normal component. Finally, if $s_k = 0$, iteration k is said to be a y-iteration because the only computation potentially performed is that of a new vector of Lagrange multiplier estimates. The main idea behind the technique we propose for accepting the trial point is to measure whether the major expected achievement of the iteration has been realized.

- If iteration k is a f-iteration, we accept the trial point if the achieved objective function reduction is comparable to its predicted value. More formally, the trial point is accepted (i.e., $x_{k+1} = x_k^+$) if

$$\rho_k^f \stackrel{\rm def}{=} \frac{f(x_k) - f(x_k^+)}{\delta_k^f} \geq \eta_1 \tag{2.35}$$

  and rejected (i.e., $x_{k+1} = x_k$) otherwise. The radius of $\mathcal{T}_k$ is then updated by

$$\Delta_{k+1}^f \in \begin{cases} [\Delta_k^f, \infty) & \text{if } \rho_k^f \geq \eta_2, \\ [\gamma_2 \Delta_k^f, \Delta_k^f] & \text{if } \rho_k^f \in [\eta_1, \eta_2), \\ [\gamma_1 \Delta_k^f, \gamma_2 \Delta_k^f] & \text{if } \rho_k^f < \eta_1, \end{cases} \tag{2.36}$$

  where the constants η_1, η_2, γ_1, and γ_2 are given and satisfy the conditions $0 < \eta_1 \leq \eta_2 < 1$ and $0 < \gamma_1 \leq \gamma_2 < 1$, as is usual for trust-region methods. The radius of $\mathcal{N}_k$ is possibly increased if feasibility is maintained well within its prescribed bounds, in the sense that

$$\Delta_{k+1}^c \in [\Delta_k^c, +\infty) \quad \text{if} \quad \theta(x_k^+) \leq \eta_3 \theta_k^{\max} \quad \text{and} \quad \rho_k^f \geq \eta_1 \tag{2.37}$$

  for some constant $\eta_3 \in (0, 1)$, or

$$\Delta_{k+1}^c = \Delta_k^c \tag{2.38}$$

  otherwise. The value of the maximal infeasibility measure is also left unchanged, that is $\theta_{k+1}^{\max} = \theta_k^{\max}$. Note that (2.33) implies that $\delta_k^f > 0$ because $\delta_k^{f,t} > 0$ unless x_k is first-order critical, and hence that condition (2.35) is well-defined.

- If iteration k is a c-iteration, we accept the trial point if the achieved improvement in feasibility is comparable to its predicted value $\delta_k^c \stackrel{\rm def}{=} \frac{1}{2}\|c_k\|^2 - \frac{1}{2}\|c_k + J_k s_k\|^2$, and if the latter is itself comparable to its predicted decrease along the normal step, that is if

$$\delta_k^c \geq \kappa_{\rm cn} \delta_k^{c,n} \quad \text{and} \quad \rho_k^c \stackrel{\rm def}{=} \frac{\theta(x_k) - \theta(x_k^+)}{\delta_k^c} \geq \eta_1 \tag{2.39}$$

  for some $\kappa_{\rm cn} \in (0, 1)$. If (2.39) fails, the trial point is rejected. The radius of $\mathcal{N}_k$ is then updated by

$$\Delta_{k+1}^c \in \begin{cases} [\Delta_k^c, \infty) & \text{if } \rho_k^c \geq \eta_2 & \text{and} & \delta_k^c \geq \kappa_{\rm cn} \delta_k^{c,n}, \\ [\gamma_2 \Delta_k^c, \Delta_k^c] & \text{if } \rho_k^c \in [\eta_1, \eta_2) & \text{and} & \delta_k^c \geq \kappa_{\rm cn} \delta_k^{c,n}, \\ [\gamma_1 \Delta_k^c, \gamma_2 \Delta_k^c] & \text{if } \rho_k^c < \eta_1 & \text{or} & \delta_k^c < \kappa_{\rm cn} \delta_k^{c,n}. \end{cases} \tag{2.40}$$

  and that of $\mathcal{T}_k$ is unchanged: $\Delta_{k+1}^f = \Delta_k^f$. We also update the value of the maximal infeasibility by

$$\theta_{k+1}^{\max} = \begin{cases} \max\left[\kappa_{\rm tx1} \theta_k^{\max}, \theta(x_k^+) + \kappa_{\rm tx2}(\theta(x_k) - \theta(x_k^+))\right] & \text{if (2.39) holds} \\ \theta_k^{\max} & \text{otherwise,} \end{cases} \tag{2.41}$$

  for some $\kappa_{\rm tx1} \in (0, 1)$ and $\kappa_{\rm tx2} \in (0, 1)$.

- If iteration k is a y-iteration, we do not have any other choice than to restart with $x_{k+1} = x_k$ using the new multipliers. We then define

$$\Delta_{k+1}^f = \Delta_k^f \quad \text{and} \quad \Delta_{k+1}^c = \Delta_k^c \tag{2.42}$$

  and keep the current value of the maximal infeasibility $\theta_{k+1}^{\max} = \theta_k^{\max}$.

We are now ready to state our complete algorithm, Algorithm 2.1 on the next page.

Algorithm 2.1: Trust-funnel Algorithm

Step 0: Initialization. An initial point x_0, an initial vector of multipliers y_{-1} and positive initial trust-region radii Δ_0^f and Δ_0^c are given. Define $\theta_0^{\max} = \max[\kappa_{\rm ca}, \kappa_{\rm cr}\theta(x_0)]$ for some constants $\kappa_{\rm ca} > 0$ and $\kappa_{\rm cr} > 1$. Set $k = 0$.

Step 1: Normal step. Possibly compute a normal step n_k that sufficiently reduces the linearized infeasibility (in the sense that (2.4) holds), under the constraint that (2.3) and (2.5) also hold. This computation must be performed if $k = 0$ or (2.30) holds when $k > 0$.

If (2.30) fails and n_k has not been computed, set $n_k = 0$.

Step 2: Tangential step. If (2.13) holds, then

Step 2.1: select a vector $\hat{y}_k$ satisfying (2.9) and define G_k by (2.8);

Step 2.2: compute y_k and r_k satisfying (2.14)–(2.17) and (2.23);

Step 2.3: If (2.29) holds, compute a tangential step t_k that sufficiently reduces the model (2.6) (in the sense that (2.24) holds), preserves linearized feasibility enough to ensure either all of (2.25)–(2.27) or (2.28), and such that the complete step $s_k = n_k + t_k$ satisfies (2.11).

If (2.13) fails, set $y_k = 0$. In this case or if (2.29) fails, set $t_k = 0$ and $s_k = n_k$. In all cases, define $x_k^+ = x_k + s_k$.

Step 3: Conclude a y-iteration. If $s_k = 0$, then

Step 3.1: accept $x_k^+ = x_k$;

Step 3.2: define $\Delta_{k+1}^f = \Delta_k^f$ and $\Delta_{k+1}^c = \Delta_k^c$;

Step 3.3: set $\theta_{k+1}^{\max} = \theta_k^{\max}$.

Step 4: Conclude an f-iteration. If $t_k \neq 0$ and (2.33) and (2.34) hold,

Step 4.1: accept x_k^+ if (2.35) holds;

Step 4.2: update Δ_k^f according to (2.36) and Δ_k^c according to (2.37)–(2.38);

Step 4.3: set $\theta_{k+1}^{\max} = \theta_k^{\max}$.

Step 5: Conclude a c-iteration. If $s_k \neq 0$ and either $t_k = 0$ or (2.33) or (2.34) fail(s),

Step 5.1: accept x_k^+ if (2.39) holds;

Step 5.2: update Δ_k^c according to (2.40);

Step 5.3: update the maximal infeasibility $\theta_k^{\max}$ using (2.41).

Step 5: Prepare for the next iteration. If x_k^+ has been accepted, set $x_{k+1} = x_k^+$, else set $x_{k+1} = x_k$. Increment k by one and go to Step 1.

We now comment on Algorithm 2.1. If either (2.35) or (2.39) holds, iteration k is called *successful*. It is said to be *very successful* if either $\rho_k^f \geq \eta_2$ or $\rho_k^c \geq \eta_2$, in which case none of the trust-region radii is decreased. We also define the following useful index sets:

$$\mathcal{S} \stackrel{\rm def}{=} \{k \mid x_{k+1} = x_k^+\}, \tag{2.43}$$

the set of successful iterations,

$$\mathcal{Y} \stackrel{\rm def}{=} \{k \mid s_k = 0\}, \quad \mathcal{F} \stackrel{\rm def}{=} \{k \mid t_k \neq 0 \text{ and (2.33) and (2.34) hold}\} \quad \text{and} \quad \mathcal{C} \stackrel{\rm def}{=} \mathbb{N}\backslash(\mathcal{Y}\cup\mathcal{F}),$$

the sets of y-, f- and c-iterations. We further divide this last set into

$$\mathcal{C}_w = \mathcal{C} \cap \{k \mid t_k \neq 0 \text{ and (2.25)–(2.27) hold}\} \quad \text{and} \quad \mathcal{C}_t = \mathcal{C} \setminus \mathcal{C}_w. \tag{2.44}$$

Note that (2.28) must hold for $k \in \mathcal{C}_t$.

We first verify that our algorithm is well-defined by deducing a useful "Cauchy-like" condition on the predicted reduction in the infeasibility measure $\theta(x)$ (whose gradient is $J(x)^T c(x)$) over each complete iteration outside $\mathcal{Y} \cup \mathcal{C}_w$.

Lemma 2.1 *For all $k \notin \mathcal{Y} \cup \mathcal{C}_w$, we have that*

$$\delta_k^c \geq \kappa_{nC2} \|J_k^T c_k\| \min\left[\frac{\|J_k^T c_k\|}{1+\|W_k\|}, \Delta_k^c\right] \geq 0, \tag{2.45}$$

for some $\kappa_{nC2} > 0$.

Proof. We first note that our assumption on k implies that (2.28) holds for each k such that $t_k \neq 0$. In this case, we easily verify that

$$\begin{aligned} 2\delta_k^c &= \|c_k\|^2 - \|c_k + J_k s_k\|^2 \\ &\geq \|c_k\|^2 - \kappa_{\rm nt}\|c_k\|^2 - (1-\kappa_{\rm nt})\|c_k + J_k n_k\|^2 \\ &= (1-\kappa_{\rm nt})\left[\|c_k\|^2 - \|c_k + J_k n_k\|^2\right] \\ &\geq 2(1-\kappa_{\rm nt})\kappa_{\rm nC}\|J_k^T c_k\| \min\left[\frac{\|J_k^T c_k\|}{1+\|W_k\|}, \Delta_k^c\right], \end{aligned}$$

where we have used (2.28) and (2.4) successively. The inequality (2.45) then results from the definition $\kappa_{\rm nC2} = (1-\kappa_{\rm nt})\kappa_{\rm nC}$. If, on the other hand, $t_k = 0$, then (2.45) directly follows from (2.4) with $\kappa_{\rm nC2} = \kappa_{\rm nC}$. □

Note that, provided $s_k \neq 0$, this result ensures that the ratio in the second part of (2.39) is well defined provided $\|J_k^T c_k\| > 0$. Conversely, if $\|c_k\| = 0$, then iteration k must be an f-iteration, and (2.39) is irrelevant. If $\|J_k^T c_k\| = 0$, but $\|c_k\| = 0$, then x_k is an infeasible stationary point of θ, an undesirable situation on which we comment below. We next show a simple useful property of y-iterations.

Lemma 2.2 *For all $k \in \mathcal{Y}$,*

$$\pi_k \leq \kappa_\omega \pi_{k-1}.$$

Proof. This immediately results from the fact that both (2.30) and (2.29) must fail at y-iterations, yielding that $\pi_k \leq \omega_2(\|c_k\|) \leq \omega_2(\omega_3(\pi_{k-1}))$ where we used the monotonicity of ω_2. The desired conclusion follows from the second part of (2.31). □

We conclude this section by stating an important direct consequence of the definition of our algorithm.

Lemma 2.3 *The sequence $\{\theta_k^{\max}\}$ is monotonically decreasing and the inequality*

$$0 \leq \theta(x_j) < \theta_k^{\max} \tag{2.46}$$

holds for all $j \geq k$.

Proof. This results from the initial definition of $\theta_0^{\max}$ in Step 0, the inequality (2.34) (which holds at f-iterations), the fact that $\theta_k^{\max}$ is only updated by formula (2.41) at successful c-iterations, at which Lemma 2.1 ensures that $\delta_k^c > 0$. □

The monotonicity of sequence $\{\theta_k^{\max}\}$ is what drives the algorithm towards feasibility and, ultimately, to optimality: the iterates can be thought as flowing towards a critical point through a funnel centered on the feasible set. Hence the algorithm's name. Note finally that Lemma 2.3 implies that

$$x_k \in \mathcal{L} \stackrel{\rm def}{=} \{x \in \mathbb{R}^n \mid \theta(x) \leq \theta_0^{\max}\}$$

for all $k \geq 0$.

3 Global convergence to first-order critical points

Before starting our convergence analysis, we recall our assumption that both f and c are twice continuously differentiable. Moreover, we also assume that there exists a constant $\kappa_{\rm H}$ such that, for all for all ξ in $\bigcup_{k\geq 0}[x_k, x_k^+] \cup \mathcal{L}$, all k and all $i \in \{1, \ldots, m\}$,

$$1 + \max[\, \|g_k\|, \|\nabla_{xx} f(\xi)\|, \|\nabla_{xx} c_i(\xi)\|, \|J(\xi)\|, \|H_k\|, \|C_{ik}\|\,] \leq \kappa_{\rm H}. \tag{3.1}$$

When H_k and C_{ik} are chosen as $\nabla_{xx} f(x_k)$ and $\nabla_{xx} c_i(x_k)$, respectively, this last assumption is for instance satisfied if the first and second derivatives of f and c are uniformly bounded, or, because of continuity, if the sequences $\{x_k\}$ and $\{x_k^+\}$ remain in a bounded domain of $\mathbb{R}^n$.

We finally complete our set of assumptions by supposing that

$$f(x) \geq f_{\rm low} \quad \text{for all} \quad x \in \mathcal{L}. \tag{3.2}$$

This assumption is often realistic and is, for instance, satisfied if the smallest singular value of the constraint Jacobian $J(x)$ is uniformly bounded away from zero. Observe that (3.2) obviously holds by continuity if we assume that all iterates remain in a bounded domain.

We first state some useful consequences of (3.1).

Lemma 3.1 *For all k,*

$$1 + \|W_k\| \leq \kappa_{_H}^2, \tag{3.3}$$

$$\|g_k^{_N}\| \ \leq (1 + \kappa_{_n}\sqrt{2\theta_0^{\max}} + m\kappa_{_n}\kappa_{_y})\kappa_{_H} \stackrel{\rm def}{=} \kappa_{_g} \tag{3.4}$$

Proof. The first inequality immediately follows from

$$1 + \|W_k\| = 1 + \|J_k\|^2 \leq (1 + \|J_k\|)^2 \leq \kappa_{\rm H}^2,$$

where the last inequality is deduced from (3.1). The bound (3.4) is obtained from (2.7), the inequality

$$\|g_k^{\rm N}\| \leq \|g_k\| + \|G_k\|\,\|n_k\| \leq \|g_k\| + \kappa_{\rm n}\big[\|H_k\|\,\|c_k\| + m\|\hat{y}_k\|\,\|c_k\| \max_{i=1,\ldots,m} \|C_{i,k}\|\big],$$

Lemma 2.3, (2.9) and (3.1). □

We also establish a useful sufficient condition for first-order criticality.

Lemma 3.2 *Assume that, for some infinite subsequence indexed by $\mathcal{K}$,*

$$\lim_{k\to\infty, k\in\mathcal{K}} \|c_k\| = 0. \tag{3.5}$$

Then

$$\lim_{k\to\infty,k\in\mathcal{K}} g_k^N = \lim_{k\to\infty,k\in\mathcal{K}} g_k. \tag{3.6}$$

If, in addition,

$$\lim_{k\to\infty,k\in\mathcal{K}} \pi_k = 0, \tag{3.7}$$

then

$$\lim_{k\to\infty,k\in\mathcal{K}} g_k + J_k^T y_k = 0 \quad \textit{and} \quad \lim_{k\to\infty,k\in\mathcal{K}} \|P_k g_k\| = 0, \tag{3.8}$$

where P_k is the orthogonal projection onto the nullspace of J_k, and all limit points of the sequence $\{x_k\}_{k\in\mathcal{K}}$ (if any) are first-order critical.

Proof. Combining the uniform bound (3.4) with (2.15), we obtain that the sequence $\{\|r_k\|\}_{\mathcal{K}}$ is uniformly bounded and therefore can be considered as the union of convergent subsequences. Moreover, because of (2.5), the limit (3.5) first implies that

$$\lim_{k\to\infty,k\in\mathcal{K}} n_k = 0, \tag{3.9}$$

which then implies with (2.9) and (3.1) that (3.6) holds. This limit, together with (2.14) and (2.17), ensures that

$$\lim_{k\to\infty,k\in\mathcal{P}} r_k = \lim_{k\to\infty,k\in\mathcal{P}} [g_k + J_k^T y_k] = \lim_{k\to\infty,k\in\mathcal{P}} [g_k - J_k^T [J_k^T]^I g_k] = \lim_{k\to\infty,k\in\mathcal{P}} P_k g_k, \tag{3.10}$$

where we have restricted our attention on a particular subsequence indexed by $\mathcal{P} \subseteq \mathcal{K}$ such that the limit in the left-hand side is well-defined. Assume now that this limit is a nonzero vector. Then, using now (2.21), (3.9), (3.6) and the hermitian and idempotent nature of P_k, we have that

$$\begin{aligned} \lim_{k\to\infty,k\in\mathcal{P}} \pi_k &= \lim_{k\to\infty,k\in\mathcal{P}} \frac{\langle g_k, r_k\rangle}{\|r_k\|} = \lim_{k\to\infty,k\in\mathcal{P}} \frac{\langle g_k, P_k g_k\rangle}{\|P_k g_k\|} \\ &= \lim_{k\to\infty,k\in\mathcal{P}} \frac{\langle P_k g_k, P_k g_k\rangle}{\|P_k g_k\|} = \lim_{k\to\infty,k\in\mathcal{P}} \|P_k g_k\|. \end{aligned} \tag{3.11}$$

But (3.7) implies that this latter limit is zero, and (3.10) also gives that r_k must converge to zero along $\mathcal{P}$, which is impossible. Hence $\lim_{k\to\infty,k\in\mathcal{P}} r_k = 0$ and the desired conclusion then follows from (3.10). □

This lemma indicates that all we need to show for first-order global convergence are the two limits (3.5) and (3.7) for an index set $\mathcal{K}$ as large as possible. Unfortunately, and as is unavoidable with local methods for constrained optimization, our algorithm may fail to produce (3.5)–(3.7) and, instead, end up being trapped by a local infeasible stationary of the infeasibility measure $\theta(x)$. If $x_\diamond$ is such a point, then

$$J(x_\diamond)^T c(x_\diamond) = 0 \quad \text{with} \quad c(x_\diamond) \neq 0.$$

If started from $x_\diamond$, Algorithm 2.1 will fail to progress towards feasibility, as no suitable normal step can be found in Step 1. A less unlikely scenario, where there exists a subsequence indexed by $\mathcal{Z}$ such that

$$\lim_{k\to\infty,k\in\mathcal{Z}} \|J_k^T c_k\| = 0 \quad \text{with} \quad \liminf_{k\to\infty,k\in\mathcal{Z}} \|c_k\| > 0, \tag{3.12}$$

indicates the approach of such an infeasible stationary point. In both cases, restarting the whole algorithm from a different starting point might be the best strategy. Barring this undesirable situation, we would however like to show that our algorithm converges to first-order critical points for (1.1), whenever uniform asymptotic convexity of $\theta(x)$ in

the orthogonal of the nullspace of J_k is obtained when feasibility is approached. More specifically, we assume from now on that, for some small constant $\kappa_{\rm c} \in (0,1)$,

$$\text{there exists } \kappa_{\rm J} \in (0,1) \text{ such that } \sigma_{\min}(J_k) \geq \kappa_{\rm J} \text{ whenever } \|c(x_k)\| \leq \kappa_{\rm c}, \tag{3.13}$$

where $\sigma_{\min}(A)$ is the smallest positive singular value of the matrix A. It is important to note that this assumption holds by continuity if $J(x)$ is Lipschitz continuous and $\sigma_{\min}(J(x))$ uniformly bounded away from zero on the feasible set, in which case the Jacobian of the constraints has constant rank over this set. This assumption also ensures that, for any subsequence indexed by $\mathcal{K}$ such that (3.5) holds, $k_1 > 0$ exists such that for $k \geq k_1, k \in \mathcal{K}$,

$$\|J_k s_k\| \geq \kappa_{\rm J} \|s_k^R\| \tag{3.14}$$

where $s_k^R \stackrel{\rm def}{=} (I - P_k)s_k$ is the projection of s_k onto the range space of J_k^T. We also obtain the following useful bound.

Lemma 3.3 *There exists a constant $\kappa_G > \kappa_H$ such that, $1 + \|G_k\| \leq \kappa_G$ for every k.*

Proof. In view of (2.14), of the monotonicity of ω_1, (2.9) and (3.4), (3.13) yields, when $\|c_k\| \leq \kappa_{\rm c}$, that

$$\|\hat{y}_k\| \leq \omega_1(\|c_k\|) + \frac{\|g_k^{\rm N}\|}{\kappa_{\rm J}} \leq \omega_1(\kappa_{\rm c}) + \frac{\kappa_{\rm g}}{\kappa_{\rm J}}.$$

On the other hand, if when $\|c_k\| \geq \kappa_{\rm c}$, then (2.9) gives that

$$\|\hat{y}_k\| \leq \frac{\kappa_{\rm y}}{\|c_k\|} \leq \frac{\kappa_{\rm y}}{\kappa_{\rm c}}.$$

Hence the desired conclusion follows from (2.8) and (3.1), with

$$\kappa_{\rm G} \stackrel{\rm def}{=} \kappa_{\rm H} + m\kappa_{\rm H} \max\left[\omega_1(\kappa_{\rm c}) + \frac{\kappa_{\rm g}}{\kappa_{\rm J}}, \frac{\kappa_{\rm y}}{\kappa_{\rm c}}\right] > \kappa_{\rm H}.$$

□

As for most of the existing theory for convergence of trust-region methods, we also make use of the following direct consequence of Taylor's theorem.

Lemma 3.4 *For all k,*

$$|f(x_k^+) - m_k(x_k^+)| \leq \kappa_G \Delta_k^2, \tag{3.15}$$

and

$$|\,\|c(x_k^+)\|^2 - \|c_k + J_k s_k\|^2| \leq 2\kappa_C [\Delta_k^{\rm c}]^2, \tag{3.16}$$

with $\kappa_C = \kappa_H^2 + m\kappa_H\sqrt{2\theta_0^{\max}} > \kappa_H$.

Proof. The first inequality follows from Lemma 3.3, the fact that $f(x)$ is twice continuously differentiable and the fact that (2.11) and (2.12) give the bound

$$\|s_k\| \leq \Delta_k \leq \Delta_k^c \tag{3.17}$$

(see Theorem 6.4.1 in Conn et al., 2000). Similarly, the second inequality follows from the fact that $\theta(x)$ is twice continuously differentiable with its Hessian given by

$$\nabla_{xx}\theta(x) = J(x)^T J(x) + \sum_{i=1}^{m} c_i(x) \nabla_{xx} c_i(x), \tag{3.18}$$

(3.1), Lemma 2.3 and (3.17). □

The same type of reasoning also allows us to deduce that all c-iterations are in $\mathcal{C}_t$ for Δ_k^c sufficiently small.

Lemma 3.5 *Assume that $k \in \mathcal{C}$ and that*

$$\Delta_k^c \leq \frac{2(1-\kappa_{tt})}{\kappa_H \kappa_\Delta (\sqrt{2m} + (2m+1)\kappa_H \kappa_\Delta)} \stackrel{\rm def}{=} \kappa_{\mathcal{C}} \tag{3.19}$$

Then $k \in \mathcal{C}_t$.

Proof. Consider some $k \in \mathcal{C}$. Using the mean-value theorem, we obtain that

$$\theta(x_k^+) = \theta_k + \langle J_k^t c_k, s_k \rangle + \tfrac{1}{2}\langle s_k, \nabla_{xx}\theta(\xi_k) s_k \rangle$$

for some $\xi_k \in [x_k, x_k^+]$, which implies, in view of (3.18), that

$$\theta(x_k^+) = \theta_k + \langle c_k, J_k s_k \rangle + \tfrac{1}{2}\|J(\xi_k)s_k\|^2 + \tfrac{1}{2}\sum_{i=1}^m c_i(\xi_k)\langle s_k, \nabla_{xx} c_i(\xi_k) s_k \rangle. \tag{3.20}$$

A further application of the mean-value theorem then gives that

$$c_i(\xi_k) = c_i(x_k) + \langle e_i, J(\mu_k)(\xi_k - x_k)\rangle = c_i(x_k) + \langle J(\mu_k)^T e_i, \xi_k - x_k \rangle$$

for some $\mu_k \in [0, \xi_k]$. Summing on all constraints and using the triangle inequality, (3.1) (twice), the bound $\|\xi_k - x_k\| \leq \|s_k\|$ and Lemma 2.3, we thus obtain that

$$\begin{aligned} \left| \sum_{i=1}^m c_i(\xi_k)\langle s_k, \nabla_{xx} c_i(\xi_k) s_k \rangle \right| &\leq \left[\|c(x_k)\|_1 + \kappa_{\rm H}\|s_k\| \right] \kappa_{\rm H}\|s_k\|^2 \\ &\leq \kappa_{\rm H}\sqrt{m}\,\|c(x_k)\|\,\|s_k\|^2 + \kappa_{\rm H}^2\|s_k\|^3 \\ &\leq \kappa_{\rm H}\sqrt{2m\,\theta_k^{\max}}\,\|s_k\|^2 + \kappa_{\rm H}^2\|s_k\|^3 \end{aligned}$$

Substituting this inequality into (3.20), we deduce that

$$\begin{aligned} \theta(x_k^+) \leq\ & \tfrac{1}{2}\|c_k + J_k s_k\|^2 + \tfrac{1}{2}\left[\|J(\xi_k)s_k\|^2 - \|J_k s_k\|^2 \right] \\ & + \tfrac{1}{2}\kappa_{\rm H}\sqrt{2m\,\theta_k^{\max}}\,\|s_k\|^2 + \tfrac{1}{2}\kappa_{\rm H}^2\|s_k\|^3 \end{aligned} \tag{3.21}$$

Define now $\phi_k(x) \stackrel{\rm def}{=} \frac{1}{2}\|J(x)s_k\|^2$. Then a simple calculation shows that

$$\nabla_x \phi_k(x) = \sum_{i=1}^m [J(x)s_k]_i \nabla_{xx} c_i(x) s_k.$$

Using this relation, the mean-value theorem again and (3.1), we obtain that

$$\begin{aligned} |\phi_k(\xi_k) - \phi_k(x_k)| &= |\langle \xi_k - x_k, \nabla_x \phi_k(\zeta_k)\rangle| \\ &= |\langle \xi_k - x_k, \textstyle\sum_{i=1}^m [J(\zeta_k)s_k]_i \nabla_{xx} c_i(\zeta_k) s_k \rangle| \\ &\leq \sum_{i=1}^m \|\xi_k - x_k\|\,\|\nabla_{xx} c_i(\zeta_k)\|\,\|J(\zeta_k)\|\,\|s_k\|^2 \\ &\leq m\,\kappa_{\rm H}^2\|s_k\|^3 \end{aligned}$$

for some $\zeta_k \in [x_k, \xi_k] \subseteq [x_k, x_k + s_k]$. We therefore obtain that

$$\tfrac{1}{2}\left| \|J(\xi_k)s_k\|^2 - \|J_k s_k\|^2 \right| = |\phi_k(\xi_k) - \phi_k(x_k)| \leq m\,\kappa_{\rm H}^2\|s_k\|^3. \tag{3.22}$$

Assume now that $k \in C_w$. Then, using (3.21), (2.27), (3.22), (2.26), (2.11) and (3.19) successively, we obtain that

$$\begin{aligned}
\theta(x_k^+) &\leq \tfrac{1}{2}\|c_k + J_k s_k\|^2 + \tfrac{1}{2}\big[\,\|J(\xi_k)s_k\|^2 - \|J_k s_k\|^2\big] \\
&\quad + \tfrac{1}{2}\kappa_{\rm H}\sqrt{2m\,\theta_k^{\max}}\,\|s_k\|^2 + \tfrac{1}{2}\kappa_{\rm H}^2\|s_k\|^3 \\
&\leq \kappa_{\rm tt}\theta_k^{\max} + (m+\tfrac{1}{2})\,\kappa_{\rm H}^2\|s_k\|^3 + \tfrac{1}{2}\kappa_{\rm H}\sqrt{2m}\sqrt{\theta_k^{\max}}\,\|s_k\|^2 \qquad (3.23)\\
&\leq \kappa_{\rm tt}\theta_k^{\max} + (m+\tfrac{1}{2})\,\kappa_{\rm H}^2\kappa_\Delta^2\,\theta_k^{\max}\Delta_k^c + \tfrac{1}{2}\kappa_\Delta\kappa_{\rm H}\sqrt{2m}\,\theta_k^{\max}\Delta_k^c \\
&\leq \theta_k^{\max}.
\end{aligned}$$

On the other hand, the fact that $k \in \mathcal{C}_w$ ensures that (2.25) holds, and thus, using the definition of $\bar{\kappa}_\delta$, that

$$(1-\kappa_\delta)\delta_k^{f,t} \geq -\delta_k^{f,n},$$

which in turn yields that

$$\delta_k^f = \delta_k^{f,n} + \delta_k^{f,n} \geq \kappa_\delta \delta_k^{f,t}.$$

But this last inequality and (3.23) show that both (2.33) and (2.34) hold at iteration k. Since a tangential step was computed at this iteration, we obtain that $k \in \mathcal{F}$, which is a contradiction because $k \in \mathcal{C}$. Hence our assumption that $k \in \mathcal{C}_w$ is impossible and the desired conclusion follows. □

Lemmas 3.4 and 3.5 have the following useful consequences.

Lemma 3.6 *Assume that $k \in \mathcal{F}$ and that*

$$\Delta_k \leq \frac{\kappa_\delta \kappa_{tC}\pi_k(1-\eta_2)}{\kappa_G}. \qquad (3.24)$$

Then $\rho_k^f \geq \eta_2$, iteration k is very successful and $\Delta_{k+1}^f \geq \Delta_k^f$. Similarly, if $k \in \mathcal{C}$ and

$$\Delta_k^c \leq \min\left[\kappa_{\mathcal{C}}, \frac{\kappa_{nC2}\|J_k^T c_k\|(1-\eta_2)}{\kappa_C}\right]. \qquad (3.25)$$

Then $\rho_k^c \geq \eta_2$, iteration k is very successful and $\Delta_{k+1}^c \geq \Delta_k^c$.

Proof. The proof of both statements is identical to that of Theorem 6.4.2 of Conn et al. (2000) for the objective functions $f(x)$ and $\theta(x)$, respectively. In the first case, one uses (2.24), (2.33) and (3.15). In the second, one first notices that (3.25) implies, in view of Lemma 3.5, that $k \in \mathcal{C}_t$ and thus that (2.45) holds. This last inequality in then used together with (3.1), (3.16) and the bound (3.3) to deduce the second conclusion. □

The mechanism for updating the trust-region radii then implies the next crucial lemma, where we show that the radius of either trust region cannot become arbitrarily small compared to the considered criticality measure for dual and primal feasibility.

Lemma 3.7 *Assume that, for some $\epsilon_f > 0$,*

$$\pi_k \geq \epsilon_f \text{ for all } k \in \mathcal{F}. \qquad (3.26)$$

Then, for all k,

$$\Delta_k^f \geq \gamma_1 \min\left[\frac{\kappa_\delta\kappa_{tC}\epsilon_f(1-\eta_2)}{\kappa_G}, \Delta_0^f\right] \stackrel{\rm def}{=} \epsilon_{\mathcal{F}}. \qquad (3.27)$$

Similarly, assume that, for some $\epsilon_\theta > 0$,

$$\|J_k^T c_k\| \geq \epsilon_\theta \text{ for all } k \in \mathcal{C}. \qquad (3.28)$$

Then, for all k,

$$\Delta_k^c \geq \gamma_1 \min\left[\kappa_{\mathcal{C}}, \frac{\kappa_{nC2}\epsilon_\theta(1-\eta_2)}{\kappa_C}, \Delta_0^c\right] \stackrel{\rm def}{=} \epsilon_{\mathcal{C}}. \qquad (3.29)$$

Proof. Again the two statements are proved in the same manner, and immediately result from the mechanism of the algorithm, Lemma 3.6 and the inequality $\Delta_k \leq \Delta_k^f$, given that Δ_k^f is only updated at f-iterations and Δ_k^c is only updated at c-iterations. □

We now start our analysis proper by considering the case where the number of successful iterations is finite.

Lemma 3.8 *Assume that $|\mathcal{S}| < +\infty$. Then there exists an x_* and a y_* such that $x_k = x_*$ and $y_k = y_*$ for all sufficiently large k, and either*

$$J(x_*)^T c(x_*) = 0 \quad \textit{and} \quad c(x_*) \neq 0,$$

or

$$P_* g(x_*) = 0 \quad \textit{and} \quad c(x_*) = 0,$$

where P_* is the orthogonal projection onto the nullspace of $J(x_*)$.

Proof. The existence of a suitable x_* immediately results from the mechanism of the algorithm and the finiteness of $\mathcal{S}$, which implies that $x_* = x_{k_s+j}$ for all $j \geq 1$, where k_s is the index of the last successful iteration.

Assume first that there are infinitely many c-iterations. This yields that Δ_k^c is decreased in (2.40) at every such iteration for $k \geq k_s$ and therefore that $\{\Delta_k^c\}$ converges to zero, because it is never increased at y-iterations or unsuccessful f-iterations. Lemma 3.5 then implies that all c-iterations are in $\mathcal{C}_t$ for k large enough. Since, for such a k, $\|J_k^T c_k\| = \|J(x_*)^T c(x_*)\|$ for all $k > k_s$, this in turn implies, in view of the second statement of Lemma 3.7, that $\|J(x_*)^T c(x_*)\| = 0$. If x_* is not feasible, then we obtain the first of the two possibilities listed in the lemma's statement. If, on the other hand, $c(x_*) = 0$, we have, from (2.5), that $n_k = 0$ and thus that $\delta_k^f = \delta_k^{f,t} \geq 0$ for all k sufficiently large. Hence (2.33) holds for k large. Moreover, we also obtain from (2.28) (which must hold for k large because $\mathcal{C}$ is asymptotically equal to $\mathcal{C}_t$) that $\|c_k + J_k s_k\| = 0$ and also, since $\theta_k^{\max}$ is only reduced at successful c-iterations, that $\theta_k^{\max} = \theta_*^{\max} > 0$ for all k sufficiently large. Combining these observations, we then obtain from Lemma 3.4 that

$$\theta(x_k^+) = \theta(x_k^+) - \tfrac{1}{2}\|c_k + J_k s_k\|^2 \leq \kappa_{\rm H}^2 [\Delta_k^c]^2 \leq \theta_k^{\max}$$

(and (2.34) holds) for all sufficiently large k. Thus we have that t_k must be zero for all $k \in \mathcal{C}$ sufficiently large. Since we already know that $n_k = 0$ for all k large enough, we thus obtain that $s_k = 0$ for these k and all iterations must eventually be y-iterations. Hence our assumption that there are infinitely many c-iterations is impossible.

Assume now that $\mathcal{C}$ is finite but $\mathcal{F}$ infinite. Since there must be an infinite number of unsuccessful f-iterations k_s, and since the radii are not updated at y-iterations, we obtain that $\{\Delta_k^f\}$, and hence $\{\Delta_k\}$, converge to zero. Using now the first statement of Lemma 3.7, we conclude that, for all k sufficiently large, $\pi_k = 0$ and, because (2.29) holds at f-iterations, $\|c_k\| = 0$. Thus $c(x_*) = 0$. As above, the second of the lemma's statements then holds because of this equality, the fact that $\pi_k = 0$ for all large k and Lemma 3.2.

Assume finally that $\mathcal{C} \cup \mathcal{F}$ is finite. Thus all iterations must be y-iterations for k large enough. In view of Lemma 2.2, we must then obtain that $\pi_* = 0$. But the fact that $n_k = 0$ for all large k, the first part of (2.31) and (2.30) then imply that $c(x_*) = 0$. The second of the lemma's statements then again holds because of Lemma 3.2.

□

This bound is central in the next result, directly inspired of Lemma 6.5.1 of Conn et al. (2000).

Lemma 3.9 *Assume that (3.13) holds and that $\mathcal{K}$ is the index of a subsequence such that (3.5) holds and $\mathcal{K} \cap \mathcal{C}_w \cap \mathcal{Y} = \emptyset$. Then there exists a $k_2 > 0$ such that, for $k \geq k_1, k \in \mathcal{K}$,*

$$\|s_k^R\| \leq \frac{2}{\kappa_j^2} \|J_k^T c_k\| \tag{3.30}$$

and

$$\delta_k^c \geq \kappa_R \|s_k^R\|^2, \tag{3.31}$$

where κ_R is a positive constant.

Proof. The proof of (3.30) is identical to that of Lemma 6.5.1 in Conn et al. (2000) (applied on the minimization of $\theta(x)$ in the range space of J_k^T), taking into account that the smallest eigenvalue of W_k is bounded below by $\kappa_{\rm J}^2$ for $k \geq k_1$ because of (3.14). Substituting now (3.30) in (2.45) (which must hold since $k \notin \mathcal{Y} \cup \mathcal{C}_w$) and using (3.3) then yields that

$$\delta_k^c \geq \tfrac{1}{2}\kappa_{\rm J}^2 \kappa_{\rm nC2} \|s_k^R\| \min\left[\frac{\kappa_{\rm J}^2 \|s_k^R\|}{2\kappa_{\rm H}^2}, \Delta_k^c\right],$$

which in turn gives (3.31) by using the bound $\|s_k^R\| \leq \|s_k\| \leq \Delta_k^c$ with

$$\kappa_{\rm R} \stackrel{\rm def}{=} \tfrac{1}{2}\kappa_{\rm J}^2 \kappa_{\rm nC2} \min\left[\frac{\kappa_{\rm J}^2}{2\kappa_{\rm H}^2}, 1\right].$$

□

We then prove that iterations in $\mathcal{C}_t$ must be very successful when the feasible set is approached.

Lemma 3.10 *Assume and (3.13) holds and that $\mathcal{K}$ is the index of a subsequence such that (3.5) holds and $\mathcal{K} \cap \mathcal{Y} = \emptyset$. Then, for all $k \in \mathcal{K} \cap \mathcal{C}_t$ sufficiently large, $\rho_k^c \geq \eta_2$, iteration k is very successful and $\Delta_{k+1}^c \geq \Delta_k^c$.*

Proof. The limit (3.5) and (3.1) imply that $\|J_k^T c_k\|$ converges to zero in $\mathcal{K}$. Since $k \notin \mathcal{C}_w$, (3.30) holds and we may use it to obtain that

$$\lim_{k \to \infty, k \in \mathcal{K} \cap \mathcal{C}_t} \|s_k^R\| = 0.$$

Combining this limit with (3.31) and using Lemma 6.5.3 of Conn et al. (2000), we deduce that $\rho_k^c \geq \eta_2$ for $k \in \mathcal{K} \cap \mathcal{C}_t$ sufficiently large. This implies that Δ_k^c is never decreased for $k \in \mathcal{K} \cap \mathcal{C}_t$ large enough. □

We now return to the convergence properties of our algorithm, and, having covered in Lemma 3.8 the case of finitely many successful iterations, we consider the case where there are infinitely many of those. We start by assuming that they are all f-iterations for k large.

Lemma 3.11 *Assume that (3.13) holds, that $|\mathcal{S}| = +\infty$ and that $|\mathcal{C} \cap \mathcal{S}| < +\infty$. Then there exists an infinite subsequence indexed by $\mathcal{K}$ such that*

$$\lim_{k \to \infty, k \in \mathcal{K}} \|c_k\| = 0. \tag{3.32}$$

and

$$\lim_{k \to \infty, k \in \mathcal{K}} \pi_k = 0. \tag{3.33}$$

Proof. As a consequence of our assumptions, we immediately obtain that all successful iterations must belong to $\mathcal{F}$ for k sufficiently large, and that there are infinitely many of them. We also deduce that the sequence $\{f(x_k)\}$ is monotonically decreasing for large enough k. Assume now, for the purpose of deriving a contradiction, that (3.26) holds. Then (2.24), (2.33), (3.1) and (3.27) together give that, for all $k \in \mathcal{S}$ sufficiently large,

$$\delta_k^f \geq \kappa_\delta \kappa_{\rm tC} \epsilon_f \min\left[\frac{\epsilon_f}{\kappa_{\rm G}}, \min[\Delta_k^c, \epsilon_f]\right]. \tag{3.34}$$

Assume now that there exists an infinite subsequence indexed by $\mathcal{K}_f \subseteq \mathcal{S}$ such that $\{\Delta_k^c\}$ converges to zero in $\mathcal{K}_f$. Since Δ_k^c is only decreased at unsuccessful c-iterations, this in turn implies that there is a subsequence of such iterations indexed $\mathcal{K}_c \subseteq \mathcal{C} \setminus \mathcal{S}$ with Δ_k^c converging to zero. Because of Lemma 3.5, we may also assume, without loss of generality, that $\mathcal{K}_c \subseteq \mathcal{C}_t \setminus \mathcal{S}$. Lemma 3.10 then gives that $\|c_k\|$, and thus, because of (3.13), $\|J_k^T c_k\|$, must be bounded away from zero along $\mathcal{K}_c$. The second statement of Lemma 3.6 and the fact that Δ_k^c is arbitrarily small for k sufficiently large in $\mathcal{K}_c$ then ensure that iteration k must be very successful for $k \in \mathcal{K}_c$ large enough, which is impossible. We therefore conclude that the sequence $\mathcal{K}_f$ described above cannot exist, and hence that there is an $\epsilon_* > 0$ such that $\Delta_k^c \geq \epsilon_*$ for $k \in \mathcal{S}$. Substituting this bound in (3.34) yields that

$$\delta_k^f \geq \kappa_\delta \kappa_{\rm tC} \epsilon_f \min\left[\frac{\epsilon_f}{\kappa_{\rm G}}, \min[\epsilon_*, \epsilon_f]\right] > 0. \tag{3.35}$$

But we also have that

$$f(x_{k_0}) - f(x_k) = \sum_{j=k_0, j\in\mathcal{S}}^{k-1} [f(x_j) - f(x_{j+1})] \geq \eta_1 \sum_{j=k_0, j\in\mathcal{S}}^{k-1} \delta_j^f. \tag{3.36}$$

This bound combined with (3.35) and the identity $|\mathcal{F} \cap \mathcal{S}| = +\infty$ then implies that f is unbounded below, which, in view of (2.46), contradicts (3.2). Hence (3.26) is impossible and we deduce that

$$\liminf_{k\to\infty} \pi_k = 0, \tag{3.37}$$

Let $\mathcal{K}$ be the index of a subsequence such that (3.37) holds as a true limit, immediately giving (3.33). The fact that all successful iterations must eventually be f-iterations implies (2.29) and we may thus deduce from (3.37), that (3.32) must hold. □

After considering the case where the number of successful c-iterations is finite, we now turn to the situation where it is infinite. We first deduce, in the next two lemmas, global convergence for the problem of minimizing θ.

Lemma 3.12 *Assume that* $|\mathcal{C} \cap \mathcal{S}| = +\infty$. *Then,*

$$\liminf_{k\to\infty, k\in\mathcal{C}} \|J_k^T c_k\| = 0. \tag{3.38}$$

Proof. Assume, for the purpose of deriving a contradiction, that (3.28) holds. Observe that the value of $\theta_k^{\max}$ is updated (and reduced) in (2.41) at each of the infinitely many iterations indexed by $\mathcal{C} \cap \mathcal{S}$.

Let us first assume that the maximum in (2.41) is attained infinitely often by the first term. Since $\kappa_{\rm tx1} < 1$, we deduce that

$$\lim_{k\to\infty} \theta_k^{\max} = 0.$$

Using the uniform boundedness of the constraint Jacobian (3.1) and (2.46), we then immediately deduce from this limit that

$$\lim_{k\to\infty} \|J_k^T c_k\| \leq \kappa_{\rm H} \lim_{k\to\infty} \|c_k\| \leq \kappa_{\rm H} \lim_{k\to\infty} \theta_k^{\max} = 0,$$

which is impossible in view of (3.28). Hence the maximum in (2.41) can only be attained a finite number of times by the first term. Now let $k \in \mathcal{C} \cap \mathcal{S}$ be the index of an iteration where the maximum is attained by the second term. Combining (2.45), (3.3), (3.28) and (3.29), we obtain that

$$\begin{aligned} \theta_k^{\max} - \theta_{k+1}^{\max} &\geq \theta(x_k) - \theta_{k+1}^{\max} \\ &\geq (1-\kappa_{\text{tx2}})\big[\theta(x_k) - \theta(x_{k+1})\big] \\ &\geq (1-\kappa_{\text{tx2}})\eta_1\delta_k^c \\ &\geq (1-\kappa_{\text{tx2}})\eta_1\kappa_{\text{nC2}}\epsilon_\theta \min\left[\frac{\epsilon_\theta}{\kappa_{\text{H}}^2}, \epsilon_{\mathcal{C}}\right] \\ &> 0. \end{aligned} \tag{3.39}$$

Since the value of $\theta_k^{\max}$ is monotonic, this last inequality and the infinite nature of $|\mathcal{C}\cap\mathcal{S}|$ implies that the sequence $\{\theta_k^{\max}\}$ is unbounded below, which obviously contradicts (2.46). Hence, the maximum in (2.41) cannot either be attained infinitely often by the second term. We must therefore conclude that our initial assumption (3.28) is impossible, which gives (3.38). □

Lemma 3.13 *Assume that $|\mathcal{C} \cap \mathcal{S}| = +\infty$. Then either there exists a subsequence of iterates approaching infeasible stationary point(s) of $\theta(x)$ in the sense that there is a subsequence indexed by $\mathcal{Z}$ such that (3.12) holds, or we have that*

$$\lim_{k\to\infty} \|c_k\| = 0. \tag{3.40}$$

and there exists an $\epsilon_* > 0$ such that

$$\Delta_k^c \geq \epsilon_*, \tag{3.41}$$

for all $k \in \mathcal{C}$ sufficiently large.

Proof. Assume that no $\mathcal{Z}$ exists such that (3.12) holds. Then Lemma 3.12 implies that there must exist an infinite subsequence indexed by $\mathcal{G} \subseteq \mathcal{C} \cap \mathcal{S}$ such that

$$\lim_{k\to\infty,k\in\mathcal{G}} \|J_k^T c_k\| = \lim_{k\to\infty,k\in\mathcal{G}} \|c_k\| = \lim_{k\to\infty,k\in\mathcal{G}} \theta(x_k) = 0. \tag{3.42}$$

As above, we immediately conclude from the inequality $\kappa_{\text{tx1}} < 1$ and (2.41) that

$$\lim_{k\to\infty} \theta_k^{\max} = 0 \tag{3.43}$$

and thus, in view of (2.46) that (3.40) holds if the maximum in (2.41) is attained infinitely often in $\mathcal{G}$ by the first term. If this is not the case, we deduce from (2.41) that

$$\lim_{k\to\infty,k\in\mathcal{G}} \theta_{k+1}^{\max} \leq \lim_{k\to\infty,k\in\mathcal{G}} \theta(x_k) = 0.$$

and thus, because of the monotonicity of the sequence $\{\theta_k^{\max}\}$, that (3.43) and (3.40) again hold.

Lemma 3.10 (with $\mathcal{K} = \mathbb{N}$) and (3.40) then imply that $\Delta_{k+1}^c \geq \Delta_k^c$ for all $k \in \mathcal{C}_t$. In addition, Lemma 3.5 ensures that Δ_k^c is bounded below by a constant for all $k \in \mathcal{C}_w = \mathcal{C} \setminus \mathcal{C}_t$. These two observations and the fact that Δ_k^c is only decreased for $k \in \mathcal{C}$ finally imply (3.41). □

Observe that it is not crucial that $\theta_k^{\max}$ is updated at every iteration in $\mathcal{C} \cap \mathcal{S}$, but rather that such updates occur infinitely often in a subset of this set along which $\|J_k^T c_k\|$ converges to zero. Other mechanisms to guarantee this property are possible, such as updating $\theta_k^{\max}$

every p iteration in $\mathcal{C}\cap\mathcal{S}$ at which $\|J_k^T c_k\|$ decreases. Relaxed scheme of this type may have the advantage of not pushing $\theta_k^{\max}$ too quickly to zero, therefore allowing more freedom for f-iterations.

Our next result analyzes some technical consequences of the fact that there might be an infinite number of c-iterations. In particular, it indicates that feasibility improves linearly at c-iterations for sufficiently large k, and hence that these iterations must play a diminishing role as k increases.

Lemma 3.14 *Assume that (3.13) holds, that $|\mathcal{C}\cap\mathcal{S}| = +\infty$ and that no subsequence exists such that (3.12) holds. Then (3.40) holds and*

$$\lim_{k\to\infty} n_k = 0, \tag{3.44}$$

and

$$\lim_{k\to\infty} \delta_k^{f,n} = 0, \tag{3.45}$$

where $\delta_k^{f,n} \stackrel{\rm def}{=} m_k(x_k) - m_k(x_k+n_k)$. Moreover (3.41) holds for $k\in\mathcal{C}$ sufficiently large. In addition, we have that for $k\in\mathcal{C}\cap\mathcal{S}$ sufficiently large,

$$\theta_{k+1} < \kappa_\theta \theta_k \tag{3.46}$$

and

$$\theta_{k+1}^{\max} \leq \kappa_{\theta m}\theta_k^{\max} \tag{3.47}$$

for some $\kappa_\theta\in(0,1)$ and some $\kappa_{\theta m}\in(0,1)$.

Proof. We first note that (3.40) holds because of Lemma 3.13. The limit (3.40) and (2.5) then give that (3.44) holds, while (3.45) then follows from the identity

$$\delta_k^{f,n} = \langle g_k, n_k\rangle + \tfrac{1}{2}\langle n_k, G_k n_k\rangle, \tag{3.48}$$

the Cauchy-Schwarz inequality, (3.40), Lemma 3.3 and (3.4). Finally, Lemma 3.13 implies that (3.41) holds for all $k\in\mathcal{C}$ sufficiently large.

If we now restrict our attention to $k\in\mathcal{C}\cap\mathcal{S}$, we also obtain, using (2.39), (3.40), (2.4), (3.13) and (3.41), that

$$\begin{aligned}
\theta_k - \theta_{k+1} &\geq \eta_1\kappa_{\rm cn}\kappa_{\rm nC}\|J_k^T c_k\| \min\left[\frac{\|J_k^T c_k\|}{1+\|W_k\|}, \Delta_k^c\right] \\
&\geq \frac{\eta_1\kappa_{\rm cn}\kappa_{\rm nC}\kappa_{\rm J}^2}{\kappa_{\rm H}^2}\|c_k\|^2 \\
&= \frac{2\eta_1\kappa_{\rm cn}\kappa_{\rm nC}\kappa_{\rm J}^2}{\kappa_{\rm H}^2}\theta_k,
\end{aligned} \tag{3.49}$$

which gives (3.46) with $\kappa_\theta \stackrel{\rm def}{=} 1 - 2\eta_1\kappa_{\rm cn}\kappa_{\rm nC}\kappa_{\rm J}^2/\kappa_{\rm H}^2 \in (0,1)$, where this last inclusion follows from the fact that $\theta_k \geq \theta_k - \theta_{k+1}$ and (3.49). We now observe that $\theta_k^{\max}$ is decreased in (2.41) at every successful c-iteration, yielding that, for $k\in\mathcal{C}\cap\mathcal{S}$ large enough,

$$\begin{aligned}
\theta_{k+1}^{\max} &= \max\left[\kappa_{\rm tx1}\theta_j^{\max}, \theta(x_k) - (1-\kappa_{\rm tx2})(\theta(x_k) - \theta(x_k^+))\right] \\
&\leq \max\left[\kappa_{\rm tx1}\theta_k^{\max}, \theta(x_k) - (1-\kappa_{\rm tx2})(1-\kappa_\theta)\theta(x_k)\right] \\
&\leq \max[\kappa_{\rm tx1}, 1-(1-\kappa_\theta)(1-\kappa_{\rm tx2})]\theta_k^{\max} \\
&= \kappa_{\theta m}\theta_k^{\max},
\end{aligned}$$

where we have used (3.46) and Lemma 2.3 to deduce the last inequalities, and where we have defined $\kappa_{\theta m} \stackrel{\rm def}{=} \max[\kappa_{\rm tx1}, 1-(1-\kappa_\theta)(1-\kappa_{\rm tx2})] \in (0,1)$. This yields (3.47) and concludes the proof. □

Convergence of the criticality measure π_k to zero then follows for a subsequence of iterations, as we now prove.

Lemma 3.15 *Assume that (3.13) holds and that* $|\mathcal{C} \cap \mathcal{S}| = +\infty$. *Then either there is a subsequence indexed by* $\mathcal{Z}$ *such that (3.12) holds, or (3.40) holds and*

$$\liminf_{k\to\infty} \pi_k = 0. \tag{3.50}$$

Proof. Assume that no subsequence exists such that (3.12) holds. We may then apply Lemma 3.14 and deduce that (3.40), (3.44), (3.45) hold and that (3.41) also hold for all $k \in \mathcal{C}$ sufficiently large.

Assume now, again for the purpose of deriving a contradiction, that the inequality (3.26) is satisfied for all k sufficiently large. This last inequality and Lemma 3.7 then guarantee that (3.27) holds for all k sufficiently large, which, with (3.41), also yields that, for $k \in \mathcal{C}$ large enough,

$$\Delta_k \geq \min[\epsilon_*, \epsilon_{\mathcal{F}}] > 0. \tag{3.51}$$

The next step in our proof is to observe that, if iteration k is a successful c-iteration, then (2.34) must hold because of (2.46). The successful c-iterations thus asymptotically come in two types:

1. iterations for which the tangential step has not been computed,
2. iterations for which (2.33) fails.

Assume first that there is an infinite number of successful c-iterations of type 1. Iterations of this type happen because either (2.13) or (2.29) fails, the latter being impossible since both (3.26) and (3.40) hold. But (2.13) cannot fail either for k sufficiently large because of (3.44) and (3.51). Hence this situation is impossible.

Assume otherwise that there is an infinite number of successful c-iterations of type 2. Since (2.33) does not hold, we deduce that, for the relevant indices k,

$$\delta_k^f = \delta_k^{f,t} + \delta_k^{f,n} < \kappa_\delta \delta_k^{f,t}$$

and thus, using the fact that (2.24) ensures the non-negativity of $\delta_k^{f,t}$, that

$$0 \leq \delta_k^{f,t} \leq \frac{|\delta_k^{f,n}|}{1-\kappa_\delta} \stackrel{\rm def}{=} \hat{\kappa}_\delta |\delta_k^{f,n}|. \tag{3.52}$$

We may then invoke (3.45) to deduce that $\delta_k^{f,t}$ converges to zero. However this is impossible since $\delta_k^{f,t}$ satisfies (2.24) and thus must be bounded away from zero because of (3.1), (3.26) and (3.51).

We may therefore conclude that an impossible situation occurs for infinite subsequences of each of the two types of successful c-iterations. This in turn implies that $|\mathcal{C} \cap \mathcal{S}|$ is finite, which is also a contradiction. Our assumption (3.26) is therefore impossible, and (3.50) follows. □

We now combine our results so far and state a first important convergence property of our algorithm.

Theorem 3.16 *As long as infeasible stationary points are avoided, there exists a subsequence indexed by* $\mathcal{K}$ *such that (3.5), (3.7) and (3.8) hold, and thus at least one limit point of the sequence* $\{x_k\}$ *(if any) is first-order critical. Moreover, we also have that (3.40) holds when* $|\mathcal{C} \cap \mathcal{S}| = +\infty$.

Proof. The desired conclusions immediately follow from Lemmas 3.2, 3.8, 3.11, 3.13, 3.15. □

Our intention is now to prove that the complete sequences $\{\pi_k\}$ and $\{\|P_k g_k\|\}$ both converge to zero, rather than merely subsequences. The first step to achieve this objective is to prove that the projection $P(x)$ onto the nullspace of the Jacobian $J(x)$ is Lipschitz continuous when x is sufficiently close to the feasible domain.

Lemma 3.17 *There exists a constant $\kappa_P > 0$ such that, for all x_1 and x_2 satisfying $\max\left[\|c(x_1)\|, \|c(x_2)\|\right] \leq \kappa_c$, we have that*

$$\|P(x_1) - P(x_2)\| \leq \kappa_P \|x_1 - x_2\|. \tag{3.53}$$

Proof. Because of (3.13) and our assumption on $c(x_1)$ and $c(x_2)$, we know that

$$P(x_i) = I - J(x_i)^T [J(x_i)J(x_i)^T]^{-1} J(x_i) \quad (i = 1, 2) \tag{3.54}$$

Denoting $J_1 \stackrel{\rm def}{=} J(x_1)$ and $J_2 \stackrel{\rm def}{=} J(x_2)$, we first observe that

$$[J_1 J_1^T]^{-1} - [J_2 J_2^T]^{-1} = [J_1 J_1^T]^{-1}\big((J_1 - J_2)J_1^T - J_2(J_1 - J_2)^T\big)[J_2 J_2^T]^{-1}. \tag{3.55}$$

But the mean-value theorem and (3.1) imply that, for $i = 1, \ldots, m$,

$$\begin{aligned} \|\nabla_x c_i(x_{k_1}) - \nabla_x c_i(g_{k_2})\| &\leq \left\| \int_0^1 \nabla_{xx} c_i(x_{k_1} + t(x_{k_2} - x_{k_1})(x_{k_1} - x_{k_2})\, dt \right\| \\ &\leq \max_{t\in[0,1]} \|\nabla_{xx} c_i(x_{k_1} + t(x_{k_2} - x_{k_1})\|\, \|x_{k_1} - x_{k_2}\| \\ &\leq \kappa_{\rm H} \|x_{k_1} - x_{k_2}\|, \end{aligned}$$

which in turn yields that

$$\|(J_1 - J_2)^T\| = \|J_1 - J_2\| \leq m\kappa_{\rm H} \|x_1 - x_2\|. \tag{3.56}$$

Hence, using (3.55), (3.1) and (3.13), we obtain that

$$\|[J_1 J_1^T]^{-1} - [J_2 J_2^T]^{-1}\| \leq \frac{2m\kappa_{\rm H}^2}{\kappa_{\rm J}^4} \|x_1 - x_2\|. \tag{3.57}$$

Computing now the difference between $P(x_1)$ and $P(x_2)$ and using (3.54), we deduce that

$$\begin{aligned} P(x_1) - P(x_2) &= J_1^T [J_1 J_1^T]^{-1}(J_1 - J_2) + (J_2 - J_1)^T [J_2 J_2^T]^{-1} J_2 \\ &\quad + J_1^T \left([J_1 J_1^T]^{-1} - [J_2 J_2^T]^{-1}\right) J_2 \end{aligned}$$

and thus, using (3.1) and (3.13) again with (3.56) and (3.57),

$$\|P(x_1) - P(x_2)\| \leq \frac{m\kappa_{\rm H}^2}{\kappa_{\rm J}^2} \|x_1 - x_2\| + \frac{m\kappa_{\rm H}^2}{\kappa_{\rm J}^2} \|x_1 - x_2\| + \frac{2m\kappa_{\rm H}^4}{\kappa_{\rm J}^4} \|x_1 - x_2\|.$$

This then yields (3.53) with $\kappa_{\rm L} = \frac{2m\kappa_{\rm H}^2}{\kappa_{\rm J}^2}\left(1 + \frac{\kappa_{\rm H}^2}{\kappa_{\rm J}^2}\right)$. □

We now refine our interpretation of the criticality measure π_k, and verify that it approximates the norm of the projected gradient when the constraint violation is small enough.

Lemma 3.18 *Assume that*

$$\min\left[\tfrac{1}{2}\|P_k g_k\|, \tfrac{1}{12}\|P_k g_k\|^2\right] > \kappa_H \kappa_G \kappa_n \|c_k\| + \omega_1(\|c_k\|). \tag{3.58}$$

Then we have that

$$\pi_k = \psi_k \|P_k g_k\| \tag{3.59}$$

for some $\psi_k \in [\frac{1}{9}, \frac{11}{3}]$.

Proof. From (2.17) and (2.14), we know that

$$r_k = P_k(g_k + G_k n_k) + \omega_1(\|c_k\|)u$$

for some normalized u, and thus, using (2.21),

$$\pi_k \left(\|P_k g_k + P_k G_k n_k + \omega_1(\|c_k\|)u\|\right) = \pi_k \|r_k\| = \langle g_k, r_k\rangle + \langle G_k n_k, r_k\rangle. \tag{3.60}$$

Now, using the triangle inequality, (3.1), (2.5), (3.58) and the bound $\kappa_{\rm H} \geq 1$, we verify that

$$\|G_k n_k + \omega_1(\|c_k\|)u\| \leq \kappa_{\rm G}\kappa_{\rm n}\|c_k\| + \omega_1(\|c_k\|) < \tfrac{1}{2}\|P_k g_k\|$$

and hence

$$\|r_k\| = \|P_k g_k + P_k G_k n_k + \omega_1(\|c_k\|)u\| = \|P_k g_k\|(1+\alpha_k)$$

with $|\alpha_k| < \frac{1}{2}$. Substituting this relation in (3.60) and using the symmetric and idempotent nature of the orthogonal projection P_k, we obtain that

$$\pi_k \;=\; \frac{1}{1+\alpha_k}\frac{\langle g_k, P_k g_k\rangle}{\|P_k g_k\|} + \frac{\langle g_k, P_k G_k n_k + \omega_1(\|c_k\|)u\rangle}{(1+\alpha_k)\|P_k g_k\|} + \frac{\langle G_k n_k, r_k\rangle}{\|r_k\|}$$

But the Cauchy-Schwarz inequality, (2.5), (3.1), the bounds $\|P_k\| \leq 1$ and $\kappa_{\rm H} \geq 1$ and (3.58) then ensure that

$$\left|\frac{\langle G_k n_k, r_k\rangle}{\|r_k\|}\right| \leq \kappa_{\rm G}\kappa_{\rm n}\|c_k\| < \tfrac{1}{2}\|P_k g_k\|$$

and that

$$\left|\frac{\langle g_k, P_k G_k n_k + \omega_1(\|c_k\|)u\rangle}{(1+\alpha_k)\|P_k g_k\|}\right| \leq \frac{\kappa_{\rm H}\kappa_{\rm G}\kappa_{\rm n}\|c_k\| + \omega_1(\|c_k\|)}{(1+\alpha_k)\|P_k g_k\|} < \frac{1}{12(1+\alpha_k)}\|P_k g_k\|.$$

Hence we deduce that, for some $\beta_k \in [-\frac{1}{2}, \frac{1}{2}]$ and some $\zeta_k \in [-\frac{1}{12}, \frac{1}{12}]$,

$$\pi_k = \frac{1+\zeta_k}{1+\alpha_k}\|P_k g_k\| + \beta_k\|P_k g_k\| = \frac{1+\zeta_k+\beta_k+\alpha_k\beta_k}{1+\alpha_k}\|P_k g_k\|.$$

This in turn yields (3.59) because

$$\psi_k \stackrel{\rm def}{=} \frac{1+\zeta_k+\beta_k+\alpha_k\beta_k}{1+\alpha_k} \in \left[\tfrac{1}{9}, \tfrac{11}{3}\right]$$

for all $(\alpha_k, \beta_k) \in [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{12}, \frac{1}{12}]$. □

The preceding result ensures the following simple but useful technical consequence.

Lemma 3.19 *Assume that $\epsilon > 0$ is given and that*

$$\kappa_H\kappa_G\kappa_n\|c_k\| + \omega_1(\|c_k\|) \leq \epsilon. \tag{3.61}$$

Then, for any $\alpha > \frac{1}{5}$,

$$\min\left[\tfrac{1}{2}\|P_k g_k\|, \tfrac{1}{12}\|P_k g_k\|^2\right] \geq 5\alpha\epsilon \quad \textit{implies that} \quad \pi_k \geq \alpha\epsilon.$$

Proof. Assume first that (3.58) fails. We then obtain, using (3.61), that

$$5\alpha\epsilon \leq \min\left[\tfrac{1}{2}\|P_k g_k\|, \tfrac{1}{12}\|P_k g_k\|^2\right] \leq \kappa_{\rm H}\kappa_{\rm G}\kappa_{\rm n}\|c_k\| + \omega_1(\|c_k\|) \leq \epsilon,$$

which is impossible because $\alpha > \frac{1}{5}$. Hence (3.58) must hold. In this case, we see, using Lemma 3.18, that

$$\tfrac{1}{2}\pi_k = \tfrac{1}{2}\psi_k\|P_k g_k\| \geq \psi_k \min\left[\tfrac{1}{2}\|P_k g_k\|, \tfrac{1}{12}\|P_k g_k\|^2\right] \geq \tfrac{5}{9}\alpha\epsilon > \tfrac{1}{2}\alpha\epsilon,$$

as desired. □

We now examine the consequences of the existence of a subsequence of consecutive f-iterations where π_k is bounded away from zero.

Lemma 3.20 *Assume that there exist $k_1 \in \mathcal{S}$ and $k_2 \in \mathcal{S}$ with $k_2 > k_1$ such that all successful iterations between k_1 and $k_2 - 1$ are f-iterations, i.e.*

$$\{k_1, \ldots, k_2 - 1\} \cap \mathcal{S} \subseteq \mathcal{F}, \tag{3.62}$$

with the property that

$$\pi_j \geq \epsilon \quad \textit{for all} \quad j \in \{k_1, \ldots, k_2 - 1\} \cap \mathcal{S} \tag{3.63}$$

for some $\epsilon > 0$. Assume furthermore that

$$f(x_{k_1}) - f(x_{k_2}) \leq \frac{\eta_1 \kappa_\delta \kappa_{tC} \epsilon^2}{2\kappa_G}. \tag{3.64}$$

Then

$$\|x_{k_1} - x_{k_2}\| \leq \frac{1}{\eta_1 \kappa_\delta \kappa_{tC} \epsilon} \big[f(x_{k_1}) - f(x_{k_2})\big]. \tag{3.65}$$

Proof. Consider a successful iteration j in the range $k_1, \ldots, k_2 - 1$ and note that the sequence $\{f(x_j)\}_{j=k_1}^{k_2}$ is monotonically decreasing. We then deduce from (2.11), (2.24), (2.33) and (3.63) that

$$\delta_j^f \geq \kappa_\delta \kappa_{\text{tC}} \pi_j \min\left[\frac{\pi_j}{1 + \|G_j\|}, \Delta_j\right] \geq \kappa_\delta \kappa_{\text{tC}} \epsilon \min\left[\frac{\epsilon}{\kappa_{\text{G}}}, \Delta_j\right].$$

Hence, since $j \in \mathcal{S}$, (2.35) implies that

$$f(x_j) - f(x_{j+1}) \geq \eta_1 \delta_k^f \geq \eta_1 \kappa_\delta \kappa_{\text{tC}} \epsilon \min\left[\frac{\epsilon}{\kappa_{\text{G}}}, \Delta_j\right]. \tag{3.66}$$

But the bound (3.64) and the inequality $f(x_j) - f(x_{j+1}) \leq f(x_{k_1}) - f(x_{k_2})$ yield together that the minimum in the right-hand side of (3.66) must be achieved by the second term. This in turn implies that

$$\|x_j - x_{j+1}\| \leq \Delta_j \leq \frac{1}{\eta_1 \kappa_\delta \kappa_{\text{tC}} \epsilon} \big[f(x_j) - f(x_{j+1})\big].$$

Summing now over all successful iterations from k_1 to $k_2 - 1$ and using the triangle inequality, we therefore obtain that

$$\|x_{k_1} - x_{k_2}\| \leq \sum_{j=k_1, j\in\mathcal{S}}^{k_2-1} \|x_j - x_{j+1}\| \leq \frac{1}{\eta_1 \kappa_\delta \kappa_{\text{tC}} \epsilon} \sum_{j=k_1, j\in\mathcal{S}}^{k_2-1} \big[f(x_j) - f(x_{j+1})\big]$$

and (3.65) follows. □

Our next step is to extend Lemma 3.11 by showing that the constraint violation goes to zero not only along the subsequence for which the criticality π_k goes to zero, but actually along the complete sequence of iterates.

Lemma 3.21 *Assume that $|\mathcal{C} \cap \mathcal{S}| < +\infty$ and that $|\mathcal{S}| = +\infty$, and that ω_2 is strictly increasing on $[0, t_\omega]$ for some $t_\omega > 0$. Then*

$$\lim_{k\to\infty} \|c_k\| = 0.$$

Proof. Let k_0 be the index of the last successful iteration in $\mathcal{C}$ (or -1 if there is none). Thus all successful iterations beyond k_0 must be f-iterations. In this case, we know that the sequence $\{f(x_k)\}$ is monotonically decreasing (by the mechanism of the algorithm) and bounded below by $f_{\rm low}$ because of (3.2); it is thus convergent to some limit $f_* \geq f_{\rm low}$. Assume first that there exists a subsequence indexed by $\mathcal{K}_c \subseteq \mathcal{F} \cap \mathcal{S}$ such that

$$\|c_k\| \geq \epsilon_0$$

for some $\epsilon_0 > 0$ and all $k \in \mathcal{K}_c$ with $k > k_0$. Because of (2.29) and the monotonicity of ω_2, we then deduce that

$$\pi_k \geq \omega_2(\epsilon_0)$$

for all $k \in \mathcal{K}_c$ with $k > k_0$. On the other hand, Lemma 3.11 implies the existence of an infinite subsequence $\mathcal{K}$ such that (3.5) and (3.7) both hold. We now choose an $\epsilon > 0$ small enough to ensure that

$$\epsilon \leq \min\left[\tfrac{1}{2}\omega_2(\epsilon_0), t_\omega\right] \quad \text{and} \quad \omega_2^{-1}(\epsilon) + \tfrac{1}{4}\epsilon \leq \tfrac{1}{2}\epsilon_0. \tag{3.67}$$

(Note that the first part of the condition and our assumption on ω_2 ensures that this bounding function is invertible for all $t \leq \epsilon$.) We next choose an index $k_1 \in \mathcal{K}_c$ large enough to ensure that $k_1 > k_0$ and also that

$$f_{k_1} - f_* \leq \min\left[\frac{\eta_1 \kappa_\delta \kappa_{\rm tC} \epsilon^2}{2\kappa_{\rm G}}, \frac{\eta_1 \kappa_\delta \kappa_{\rm tC} \epsilon^2}{4\kappa_{\rm H}}\right], \tag{3.68}$$

which is possible since $\{f(x_k)\}$ converges in a monotonically decreasing manner to f_*. We finally select k_2 to be the first index in $\mathcal{K}$ after k_1 such that

$$\pi_j \geq \epsilon \quad \text{for all} \quad k_1 \leq j < k_2, j \in \mathcal{S}, \quad \text{and} \quad \pi_{k_2} < \epsilon. \tag{3.69}$$

Because $f(x_1) - f(x_{k_2}) \leq f(x_{k_1}) - f_*$ and (3.68), we may then apply Lemma 3.20 to the iterations k_1 and k_2, and deduce that (3.65) holds, and therefore, using (3.64), that

$$\|x_{k_1} - x_{k_2}\| \leq \frac{\epsilon}{4\kappa_{\rm H}}.$$

Thus, using the vector-valued mean-value theorem, we the obtain that

$$\begin{aligned}
\|c_{k_1} - c_{k_2}\| &\leq \left\| \int_0^1 J(x_{k1} + t(x_{k_2} - x_{k_1}))(x_{k_1} - x_{k_2})\, dt \right\| \\
&\leq \max_{t\in[0,1]} \|J(x_{k1} + t(x_{k_2} - x_{k_1}))\|\, \|x_{k_1} - x_{k_2}\| \\
&\leq \kappa_{\rm H} \|x_{k_1} - x_{k_2}\| \\
&\leq \tfrac{1}{4}\epsilon
\end{aligned}$$

As a consequence, using the triangle inequality, the fact that $\omega_2(\|c_{k_2}\|) \leq \pi_{k_2}$ (since $k_2 \in \mathcal{F}$) and the second part of (3.67), we deduce that

$$\epsilon_0 \leq \|c_{k_1}\| \leq \|c_{k_2}\| + \tfrac{1}{4}\epsilon \leq \omega_2^{-1}(\pi_{k_2}) + \tfrac{1}{4}\epsilon \leq \omega_2^{-1}(\epsilon) + \tfrac{1}{4}\epsilon \leq \tfrac{1}{2}\epsilon_0$$

which is a contradiction. Hence our initial assumption on the existence of the subsequence $\mathcal{K}_c$ is impossible and $\|c_k\|$ must converge to zero, as required. □

We finally strengthen the convergence results obtained in Theorem 3.16 by avoiding taking limits along subsequences.

Theorem 3.22 *Assume that (3.13) holds and that ω_2 is strictly increasing in $[0, t_\omega]$ for some $t_\omega > 0$. Then, we have that, either there exists a subsequence indexed by $\mathcal{Z}$ such that (3.12) holds, or*

$$\lim_{k\to\infty} \|c_k\| = 0 \quad \textit{and} \quad \lim_{k\to\infty} \|P_k g_k\| = 0, \tag{3.70}$$

and all limit points of the sequence $\{x_k\}$ (if any) are first-order critical.

Proof. Assume that no subsequence exists such that (3.12) holds. If there are only finitely many successful iterations, the desired conclusion directly follows from Theorem 3.8. Assume therefore that $|\mathcal{S}| = +\infty$ and immediately note that the first limit in (3.70) follows from Theorem 3.16. Thus we only need to prove the second limit in (3.70) when they are infinitely many successful iterations.

For this purpose, assume, with the objective of deriving a contradiction, that there exists an infinite subsequence indexed by $\mathcal{K}$ such that, for some $\epsilon \in (0, \frac{1}{5})$,

$$\min\left[\tfrac{1}{2}\|P_k g_k\|, \tfrac{1}{12}\|P_k g_k\|^2\right] \geq 10\epsilon \qquad \text{for all} \quad k \in \mathcal{K}. \tag{3.71}$$

Now choose $k_1 \in \mathcal{K}$ large enough to ensure that, for all $k \geq k_1$, (3.61) holds,

$$\|c_k\| \leq \min\left[\frac{2\kappa_{\rm H}}{\kappa_{\rm n}\kappa_{\rm G}}, \kappa_{\rm c}\right], \tag{3.72}$$

and

$$\omega_2(\|c_k\|) \leq \tfrac{1}{2}\epsilon. \tag{3.73}$$

If $|\,\mathcal{C}\cap\mathcal{S}\,| = +\infty$, we also require that the conclusions of Lemma 3.14 apply, that

$$|\delta_k^{f,n}| \leq \frac{\kappa_{\rm tC}\epsilon^2}{2\hat{\kappa}_\delta\kappa_{\rm G}} \tag{3.74}$$

for all $k \geq k_1$, and that

$$\sqrt{2\theta_{k_1}^{\max}} \leq \frac{\eta_1\kappa_\delta\kappa_{\rm tC}(1-\sqrt{\kappa_{\theta m}})\epsilon^2}{4\kappa_{\rm H}^3(\kappa_{\rm P}+1)\kappa_{\rm n}\kappa_0} \tag{3.75}$$

(where $\kappa_0 \stackrel{\rm def}{=} \max\left[1, \frac{2\hat{\kappa}_\delta\kappa_{\rm H}}{\kappa_{\rm tC}\epsilon}\right]$), which is possible because of Lemma 3.14. Conversely, if $|\,\mathcal{C}\cap\mathcal{S}\,| < +\infty$, we require that k_1 is larger than the index of the last successful c-iteration. Observe that, because of (3.61) and Lemma 3.19 (with $\alpha = 2$), (3.71) implies that

$$\pi_{k_1} \geq 2\epsilon > 0. \tag{3.76}$$

We now choose k_2 to be the (first) successful iteration after k_1 such that

$$\pi_{k_2} < \epsilon, \tag{3.77}$$

which we know must exist because of Theorem 3.16. Note that this last inequality, (3.61) and Lemma 3.19 (with $\alpha = 1$) then give that

$$\min\left[\tfrac{1}{2}\|P_{k_2} g_{k_2}\|, \tfrac{1}{12}\|P_{k_2} g_{k_2}\|^2\right] \leq 5\epsilon. \tag{3.78}$$

Our choice of k_1 and k_2 also yields that

$$\pi_j \geq \epsilon \quad \text{for} \quad k_1 \leq j < k_2. \tag{3.79}$$

Assume now that $|\,\mathcal{C}\cap\mathcal{S}\,| = +\infty$ and consider an iteration $j \in \mathcal{C}\cap\mathcal{S}$ with $k_1 \leq j < k_2$, and note that (2.29) must hold at such an iteration because of (3.73) and (3.79). Assume first that (2.13) also holds and thus that the tangential step t_j is computed.

We know from (3.46) and Lemma 2.3 that $\theta_j^+ \leq \kappa_\theta \theta_j \leq \kappa_\theta \theta_j^{\max}$. Hence (2.34) holds. As a consequence (2.33) must fail and we obtain that

$$\hat{\kappa}_\delta |\delta_j^{f,n}| > \delta_j^{f,t} \geq \kappa_{\rm tC} \epsilon \min\left[\frac{\epsilon}{\kappa_{\rm G}}, \Delta_j\right]$$

where we used (2.24), (3.52), (3.79) and Lemma 3.3. But (3.74) then implies that the minimum in the last right-hand side must be achieved by the second term, and hence, using (2.11), that

$$\|s_j\| \leq \Delta_k \leq \frac{\hat{\kappa}_\delta}{\kappa_{\rm tC}\epsilon} |\delta_j^{f,n}|. \tag{3.80}$$

Using now successively the definition of $\delta_j^{f,n}$ (as in (3.48)), the Cauchy-Schwarz inequality, (3.1), (2.5) and (3.72), we deduce that

$$\begin{aligned} |\delta_j^{f,n}| &= |\langle g_j, n_j\rangle + \tfrac{1}{2}\langle n_j, G_j n_j\rangle| \\ &\leq \|g_j\|\,\|n_j\| + \tfrac{1}{2}\|G_j\|\,\|n_j\|^2 \\ &\leq (\kappa_{\rm H} + \tfrac{1}{2}\kappa_{\rm G}\kappa_{\rm n}\|c_j\|)\|n_j\| \\ &\leq 2\kappa_{\rm H}\|n_j\|. \end{aligned}$$

Combining the last bound with (3.80), we find that

$$\|s_j\| \leq \frac{2\hat{\kappa}_\delta \kappa_{\rm H}}{\kappa_{\rm tC}\epsilon}\|n_j\|.$$

Conversely, if (2.13) does not hold, we have that $t_j = 0$ and hence $s_j = n_j$. As a consequence, we obtain that, for every $j \in \mathcal{C} \cap \mathcal{S}$ such that $k_1 \leq j < k_2$,

$$\|s_j\| \leq \max\left[1, \frac{2\hat{\kappa}_\delta \kappa_{\rm H}}{\kappa_{\rm tC}\epsilon}\right]\|n_j\| \leq \kappa_{\rm n}\kappa_0\|c_j\| \leq \kappa_{\rm n}\kappa_0\sqrt{2\theta_j^{\max}} \tag{3.81}$$

where (2.5) and Lemma 2.3 are used to obtain the last two inequalities. Remembering now Lemma 3.14 and the fact that $\theta_j^{\max}$ is unchanged at iterations outside $\mathcal{C} \cap \mathcal{S}$, we thus deduce that, for any $k_3 \geq k_1$,

$$\begin{aligned} \sum_{j=k_1, j\in\mathcal{C}\cap\mathcal{S}}^{k_3} \|s_j\| &\leq \kappa_{\rm n}\kappa_0 \sum_{j=k_1, j\in\mathcal{C}\cap\mathcal{S}}^{k_3} \sqrt{2\theta_j^{\max}} \\ &\leq \kappa_{\rm n}\kappa_0\sqrt{2\theta_{k_1}^{\max}} \sum_{j=k_1, j\in\mathcal{C}\cap\mathcal{S}}^{k_3} \kappa_{\theta m}^{\frac{1}{2}|\mathcal{C}\cap\{k_1^c,\ldots,k_3\}|} \\ &\leq \kappa_{\rm n}\kappa_0\sqrt{2\theta_{k_1}^{\max}} \sum_{j=0}^{\infty} \kappa_{\theta m}^{j/2} \\ &\leq \frac{\kappa_{\rm n}\kappa_0}{1-\sqrt{\kappa_{\theta m}}}\sqrt{2\theta_{k_1}^{\max}} \end{aligned} \tag{3.82}$$

But this last bound, (3.75) and the inequality $\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon \leq 4\kappa_{\rm H}$ then yield that, for any $k_3 \geq k_1$,

$$\sum_{j=k_1, j\in\mathcal{C}\cap\mathcal{S}}^{k_3} \|s_j\| \leq \frac{\epsilon}{\kappa_{\rm H}^2(\kappa_{\rm P}+1)}. \tag{3.83}$$

Note that this bound is valid irrespective of k_3. Using the mean value theorem, we now obtain that

$$|\, f(x_j) - f(x_{j+1})\,| = |\,\langle g_j, s_j\rangle + \tfrac{1}{2}\langle s_j, \nabla_{xx} f(\xi_j) s_j\rangle\,| \leq \kappa_{\rm H}\|s_j\| + \tfrac{1}{2}\kappa_{\rm H}\|s_k\|^2$$

for some $\xi_j \in [x_j, x_{j+1})$, and where we have used the Cauchy-Schwarz inequality and (3.1) to deduce the last inequality. But (2.5), (3.81) and condition (3.72) then imply that

$$\|s_j\| + \tfrac{1}{2}\|s_k\|^2 \leq \|s_j\|(1 + \tfrac{1}{2}\kappa_0\|n_j\|) \leq \|s_j\|(1 + \tfrac{1}{2}\kappa_0\kappa_{\rm n}\|c_j\|) \leq 2\|s_j\|$$

and hence, using (3.82), that

$$\sum_{j=1,j\in\mathcal{C}\cap\mathcal{S}}^{k_3} |\, f(x_j) - f(x_{j+1}) \,| \leq 2\kappa_{\rm H} \sum_{j=k_1,j\in\mathcal{C}\cap\mathcal{S}}^{k_3} \|s_j\| \leq \frac{2\kappa_{\rm n}\kappa_{\rm H}\kappa_0}{1-\sqrt{\kappa_{\theta m}}}\sqrt{2\theta_{k_1}^{\max}}.$$

Taking the limit for k_3 going to infinity, we see, using (3.75), that

$$\sum_{j=k_1,j\in\mathcal{C}\cap\mathcal{S}}^{\infty} |\, f(x_j) - f(x_{j+1}) \,| \leq \frac{2\kappa_{\rm n}\kappa_{\rm H}\kappa_0}{1-\sqrt{\kappa_{\theta m}}}\sqrt{2\theta_{k_1}^{\max}} \leq \frac{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon^2}{2\kappa_{\rm H}^2(\kappa_{\rm P}+1)}. \tag{3.84}$$

Note that this bound remains valid if $|\,\mathcal{C}\cap\mathcal{S}\,| < +\infty$ since the sum on the left-hand side is empty in that case.

We now observe that the objective function is decreased at every successful f-iteration and the total decrease, from iteration k_1 on, cannot exceed the maximum value of $f(x_k)$ for $k \geq k_1$ minus the lower bound $f_{\rm low}$ specified by (3.2). Moreover the maximum of $f(x_k)$ beyond iteration k_1 cannot itself exceed $f(x_{k1})$ augmented by the total increase occuring at all c-iterations beyond k_1, which is given by (3.84). As a consequence, we may conclude that

$$\begin{aligned}\sum_{j=k_1,j\in\mathcal{S}}^{\infty} |\, f(x_j) - f(x_{j+1}) \,| &= \sum_{j=k_1,j\in\mathcal{F}\cap\mathcal{S}}^{\infty} \big[\, f(x_j) - f(x_{j+1}) \,\big] + \sum_{j=k_1,j\in\mathcal{C}\cap\mathcal{S}}^{\infty} |\, f(x_j) - f(x_{j+1}) \,| \\ &\leq \left[f(x_{k_1}) + \frac{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon}{2\kappa_{\rm H}^2(\kappa_{\rm P}+1)} - f_{\rm low}\right] + \frac{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon}{2\kappa_{\rm H}^2(\kappa_{\rm P}+1)},\end{aligned}$$

which in turn implies that

$$\sum_{j=0,j\in\mathcal{S}}^{\infty} |\, f(x_j) - f(x_{j+1}) \,| < +\infty \quad \text{and} \quad \lim_{\ell\to\infty} \sum_{j=\ell,j\in\mathcal{S}}^{\infty} |\, f(x_j) - f(x_{j+1}) \,| = 0.$$

Because of this last limit, we may therefore possibly increase $k_1 \in \mathcal{K}$ (and k_2 accordingly) to ensure that

$$\sum_{j=k_1,j\in\mathcal{S}}^{\infty} |\, f(x_j) - f(x_{j+1}) \,| \leq \min\left[\frac{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon^2}{2\kappa_{\rm G}}, \frac{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon^2}{2\kappa_{\rm H}^2(\kappa_{\rm P}+1)}\right] \tag{3.85}$$

in addition to (3.61), (3.72), (3.73), as well as the conclusions of Lemma 3.14, (3.74) and (3.75) if $|\,\mathcal{C}\cap\mathcal{S}\,| = +\infty$.

Consider now a range of consecutive successful f-iterations (i.e. a range containing at least one successful f-iteration and no successful c-iteration), indexed by $\{k_a, \ldots, k_b - 1\}$. Observe that (3.85) gives that

$$f(x_{k_a}) - f(x_{k_b}) \leq \frac{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon^2}{2\kappa_{\rm G}}.$$

Then, using Lemma 3.20 (which is applicable because of (3.79) and this last bound), we deduce that

$$\|x_{k_a} - x_{k_b}\| \leq \tfrac{1}{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon}\big[f(x_{k_a}) - f(x_{k_b})\big].$$

We now sum on all disjoint sequences $\{k_{a,\ell}, \ldots, k_{b,\ell}\}_{\ell=1}^{p}$ of this type between k_1 and $k_2 - 1$ (if any), and find that

$$\sum_{j=k_1,j\in\mathcal{F}\cap\mathcal{S}}^{k_2-1} \|x_j - x_{j+1}\| = \sum_{\ell=1}^{p} \|x_{k_{a,\ell}} - x_{k_{b,\ell}}\| \leq \frac{1}{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon} \sum_{\ell=1}^{p} \big[f(x_{k_{a,\ell}}) - f(x_{k_{b,\ell}})\big]. \tag{3.86}$$

We now decompose this last sum and obtain, using (3.84) and (3.85), that

$$\begin{aligned}\sum_{\ell=1}^{p}\big[f(x_{k_{a,\ell}})-f(x_{k_{b,\ell}})\big] &\leq \sum_{\ell=1}^{\infty}\big[f(x_{k_{a,\ell}})-f(x_{k_{b,\ell}})\big]\\ &= \sum_{j=k_1,j\in\mathcal{F}\cap\mathcal{S}}^{\infty}\big[f(x_j)-f(x_{j+1})\big]\\ &= \sum_{j=k_1,j\in\mathcal{S}}^{\infty}\big[f(x_j)-f(x_{j+1})\big]-\sum_{j=k_1,j\in\mathcal{C}\cap\mathcal{S}}^{\infty}\big[f(x_j)-f(x_{j+1})\big]\\ &\leq \sum_{j=k_1,j\in\mathcal{S}}^{\infty}|f(x_j)-f(x_{j+1})|+\sum_{j=k_1,j\in\mathcal{C}\cap\mathcal{S}}^{\infty}\big|f(x_j)-f(x_{j+1})\big|\\ &\leq \frac{\eta_1\kappa_\delta\kappa_{\rm tC}\epsilon^2}{\kappa_{\rm H}^2(\kappa_{\rm P}+1)}\end{aligned}$$

Substituting this inequality in (3.86), we obtain that

$$\sum_{j=k_1,j\in\mathcal{F}\cap\mathcal{S}}^{k_2-1}\|x_j-x_{j+1}\|\leq\frac{\epsilon}{\kappa_{\rm H}^2(\kappa_{\rm P}+1)}$$

and thus, using the triangle inequality and (3.83) with $k_3=k_2-1$, that

$$\|x_{k_1}-x_{k_2}\|\leq\sum_{j=k_1,j\in\mathcal{C}\cap\mathcal{S}}^{k_2-1}\|x_j-x_{j+1}\|+\sum_{j=k_1,j\in\mathcal{F}\cap\mathcal{S}}^{k_2-1}\|x_j-x_{j+1}\|\leq\frac{2\epsilon}{\kappa_{\rm H}^2(\kappa_{\rm P}+1)}. \quad (3.87)$$

We now return to considering the sizes of the projected gradients at iterations k_1 and k_2. We know from (3.71), (3.78) and the triangle inequality that

$$\begin{aligned}\|P_{k_1}g_{k_1}\|-\|P_{k_2}g_{k_2}\| &\leq \|P_{k_1}g_{k_1}-P_{k_2}g_{k_2}\|\\ &\leq \|(P_{k_1}-P_{k_2})g_{k_1}\|+\|P_{k_2}(g_{k_1}-g_{k_1})\|\\ &\leq \|P_{k_1}-P_{k_2}\|\,\|g_{k_1}\|+\|P_{k_2}\|\,\|g_{k_1}-g_{k_1}\|.\end{aligned}$$

In view of (3.72), we may now apply Lemma 3.17 and, recalling that the norm of an orthogonal projection is bounded above by one, deduce that

$$\|P_{k_1}g_{k_1}\|-\|P_{k_2}g_{k_2}\|\leq\kappa_{\rm P}\kappa_{\rm H}\|x_{k_1}-x_{k_2}\|+\|g_{k_1}-g_{k_1}\|, \quad (3.88)$$

where we have used (3.1) to bound $\|g_{k_1}\|$. But the vector-valued mean-value theorem ensures that

$$\begin{aligned}\|g_{k_1}-g_{k_2}\| &\leq \left\|\int_0^1\nabla_{xx}f(x_{k_1}+t(x_{k_2}-x_{k_1}))(x_{k_1}-x_{k_2})\,dt\right\|\\ &\leq \max_{t\in[0,1]}\|\nabla_{xx}f(x_{k_1}+t(x_{k_2}-x_{k_1}))\|\,\|x_{k_1}-x_{k_2}\|\\ &\leq \kappa_{\rm H}\|x_{k_1}-x_{k_2}\|,\end{aligned}$$

where we also used (3.1). Substituting this last inequality in (3.88) and using (3.87), we finally obtain that

$$\|P_{k_1}g_{k_1}\|-\|P_{k_2}g_{k_2}\|\leq\kappa_{\rm H}(\kappa_{\rm P}+1)\|x_{k_1}-x_{k_2}\|\leq\frac{2\epsilon}{\kappa_{\rm H}}. \quad (3.89)$$

Observe now that the inequality $\epsilon\leq\frac{1}{5}$ and (3.78) imply together that

$$\|P_{k_2}g_{k_2}\|\leq 10\epsilon\leq 2 \quad\text{or}\quad \|P_{k_2}g_{k_2}\|^2\leq 60\epsilon\leq 12<16,$$

which in turn implies that

$$\|P_{k_2} g_{k_2}\| < 4 \tag{3.90}$$

and thus that

$$\min\left[\tfrac{1}{2}\|P_{k_2} g_{k_2}\|, \tfrac{1}{12}\|P_{k_2} g_{k_2}\|^2 \right] = \tfrac{1}{12}\|P_{k_2} g_{k_2}\|^2. \tag{3.91}$$

Suppose now that

$$\|P_{k_1} g_{k_1}\| \leq 6, \tag{3.92}$$

in which case

$$\min\left[\tfrac{1}{2}\|P_{k_1} g_{k_1}\|, \tfrac{1}{12}\|P_{k_1} g_{k_1}\|^2 \right] = \tfrac{1}{12}\|P_{k_1} g_{k_1}\|^2. \tag{3.93}$$

Then, successively using (3.71), (3.78), (3.93), (3.91), the bound of one on the norm of orthogonal projections, (3.1) and (3.89), we conclude that

$$\begin{aligned}
5\epsilon &\leq \min\left[\tfrac{1}{2}\|P_{k_1} g_{k_1}\|, \tfrac{1}{12}\|P_{k_1} g_{k_1}\|^2 \right] - \min\left[\tfrac{1}{2}\|P_{k_2} g_{k_2}\|, \tfrac{1}{12}\|P_{k_2} g_{k_2}\|^2 \right] \\
&= \tfrac{1}{12}\left[\|P_{k_1} g_{k_1}\|^2 - \|P_{k_2} g_{k_2}\|^2 \right] \\
&= \tfrac{1}{12}\left[\|P_{k_1} g_{k_1}\| + \|P_{k_2} g_{k_2}\| \right]\left[\|P_{k_1} g_{k_1}\| - \|P_{k_2} g_{k_2}\| \right] \\
&\leq \tfrac{1}{6}\kappa_{\rm H}\left[\|P_{k_1} g_{k_1}\| - \|P_{k_2} g_{k_2}\| \right] \\
&\leq \tfrac{1}{3}\epsilon
\end{aligned}$$

which is impossible. Hence (3.92) must be false. Combining now this observation with (3.90), we obtain that

$$2 < \|P_{k_1} g_{k_1}\| - \|P_{k_2} g_{k_2}\| \leq \frac{2\epsilon}{\kappa_{\rm H}},$$

which is again impossible. Hence our assumption (3.71) is itself impossible and the second limit of (3.70) must hold. □

We end our theoretical developments at this point, but the theory and results presented so far suggest some comments.

1. Assumption (3.2) is not really crucial in the sense that one may apply *c*-iterations (by temporarily setting $f \equiv 0$ and keeping $\hat{y}_k = 0$) *a priori* (hence reducing infeasibility) to reduce the domain. If a global lower bound on the objective function value on the feasible domain is known, a comparison of the infeasibility and objective function value at the starting point may be useful to decide whether pure *c*-iterations should be applied first, or if the complete algorithm can be applied directly from the starting point.

2. When the Jacobian J_k is full-rank, we can rewrite the test (2.14) in the form

$$\|J_k(g_k^{\rm N} + J_k^T y_k)\| \leq \omega_1(\|c_k\|),$$

which provides an implementable version of (2.14).

3. Remarkably, convergence of trust-region methods for unconstrained optimization may be obtained as a by-product of the results presented here. Indeed, if there are no constraints, the algorithm reduces to the basic trust-region method by setting $\theta_0^{\max} = \kappa_{\rm ca}$, and, for every k, $n_k = 0$, $y_k = 0$, $\hat{y}_k = 0$, $r_k = g_k$. Since $\pi_k = \|g_k\|$, we have that $\pi_k > \omega_2(0) = 0$ and a non-zero t_k is always computed. Moreover, every iteration is then an *f*-iteration with $\delta_k^f = \delta_k^{f,t}$ at which we choose, as allowed by (2.37), not to update the (irrelevant) $\Delta_k^{\rm c}$.

4. Obviously, one could use $G_k = H_k$ and still obtain global convergence. The vector $\hat{y}_k$ then becomes irrelevant. This is particularly apt when the constraints are linear.

5. The tangential step is only required to satisfy the modified Cauchy condition (2.24), but there is no theoretical need to compute the associated modified Cauchy point (the solution of (2.19)). If one considers that t_k results from an iterative process starting (and possibly ending) at this modified Cauchy point, it is then necessary to ensure that this point satisfies either (2.28) or (2.25)-(2.26)-(2.27). A possible technique is to first solve (2.18) accurately enough to ensure that

$$\|c_k + J_k(n_k - \tau_k r_k)\|^2 \leq \kappa_{\rm tt}\theta_k^{\max}, \tag{3.94}$$

which is possible since it holds trivially if (2.18) is solved exactly, because then $J_k r_k = 0$ by construction and $\vartheta_k < \|c_k + J_k n_k\|^2$. As soon as (3.94) holds, then the modified Cauchy point can be computed and (2.26) and (2.25) tested. If any of these fail, then the solution of (2.18) must be continued to ensure that

$$\|c_k + J_k(n_k - \tau_k r_k)\|^2 \leq \vartheta_k$$

and a new, improved, modified Cauchy point can then be found along $-r_k$ at which (2.28) holds.

6. It is interesting to observe that the conditions (2.27) or (2.28) happen to be irrelevant for successful f-iterations in the theory discussed above. For such iterations, the role of limiting the acceptable infeasibility is played by (2.34). Indeed (2.27) is only used to show that (2.34) also holds for small enough Δ_k^c, which then implies that the considered iteration is an f-iteration. Condition (2.28) is crucial in establishing (2.45) in Lemma 2.1, but this global Cauchy condition on the feasibility improvement in only used for c-iterations (in Lemmas 3.6, 3.9 and 3.12). Finally, (2.28) is also used in Lemma 3.8, but again only for c-iterations.

   In a situation where evaluating the value of the infeasibility measure θ is cheap and the tangential step is computed by an iterative process, it may be possible to detect that (2.33) holds before the end of this process, and then simply replace conditions (2.27)/(2.28) by the verification that (2.34) holds. Of course, if (2.35) then fails or if (2.34) cannot be enforced, then the iteration has to be handled as an unsucessful c-iteration, since we can no longer turn it into a successful c-iteration for which (2.27)/(2.28) is meaningful.

7. The preliminary numerical experience discussed in the next section has shown that our algorithm, like many SQP methods, might suffer from the Maratos effect. A well documented cure for this problem (see Mayne and Polak, 1982, Coleman and Conn, 1982, or Section 15.3.2 of Conn et al., 2000) is to use second-order correction steps. In our context, we define a such a step $s_k^{\rm C}$ as a step performed from $x_k + s_k$ to correct for an unsuccessful f-iteration, and such that

$$\|s_k + s_k^{\rm C}\| \leq \Delta_k \tag{3.95}$$

and

$$\theta(x_k + s_k + s_k^{\rm C}) \leq \theta_k^{\max}. \tag{3.96}$$

   Of course, for the f-iteration using the augmented step $s_k + s_k^{\rm C}$ to be successful , we still require, extending (2.35), that

$$\rho_k^{\rm C} \stackrel{\rm def}{=} \frac{f(x_k) - f(x_k + s_k + s_k^{\rm C})}{m_k(x_k) - m_k(x_k + s_k)} \geq \eta_1. \tag{3.97}$$

   Using the comment just made on the irrelevant nature of (2.27) or (2.28) for successful f-iterations, we may now verify that the convergence theory presented above is not modified by the presence of these correction steps. Indeed, a successful iteration using the augmented step satisfies all the conditions required for a successful f-iteration where $m_k(x + s_k)$ is then interpreted, in the spirit of Section 10.4.2 in

Conn et al. (2000), as a prediction of $f(x_k + s_k + s_k^{\rm C})$ and where the infeasibility-limiting condition (2.34) is replaced by (3.96).

In practice, a second-order correction is often computed by producing a step $s_k^{\rm C}$ that reduces infeasibility, typically by "projecting" the trial point lying in or close to the nullspace of $J(x_k)$ onto the actual feasible set. In this case, $s_k^{\rm C}$ not only improves feasibility (ensuring (3.96)), but often makes $m_k(x_k + s_k)$ to be a better prediction of the value of $f(x_k + s_k + s_k^{\rm M})$ than of $f(x_k + s_k)$ (which tends to make the iteration acceptable in (3.97)). Because $\|s_k^{\rm C}\|$ is then of the order of $\|s_k\|^2$, condition (3.95) usually follows from (2.11).

The authors are well aware that many theoretical questions remain open at this stage of analysis, such as convergence to second-order critical points, rate of convergence and worst-case complexity analysis. Furthermore, the many degrees of freedom in the algorithm provide considerable room for implementation.

4 Preliminary numerical experience

An initial (experimental) implementation of the method presented above has been produced, in which the following choices are made:

- The normal step n_k is computed by applying the truncated conjugate-gradient algorithm to attempt to minimize the Gauss-Newton model (2.2), stopping as soon as the gradient of (2.2) has been reduced in norm by a factor 10^{-12}, or an iterate crosses the trust-region boundary—in which case the path of iterates is retraced to find the point at which it crosses—or more than $m + 1$ iterations have been performed.

- The multipliers y_k are computed by solving the normal equations

$$J_k J_k^T y_k = -J_k g_k^{\rm N}$$

  if J_k is of full rank, or approximated by applying the conjugate-gradient method to (2.18) otherwise. In the latter case, as before the iteration is terminated as soon as the gradient of (2.18) has been reduced in norm by a factor 10^{-12}, or more than $m + 1$ iterations have been performed.

- An exact tangential step is computed using the generalized Lanczos trust-region (GLTR) method of Gould, Lucidi, Roma and Toint (1999). In particular, the projected, preconditioned variant proposed by Gould, Hribar and Nocedal (2001), in which iterates are kept in the null space of J_k by operations involving factors of the coefficient matrix

$$\begin{pmatrix} I & \bar{J}_k^T \\ \bar{J}_k & 0 \end{pmatrix}$$

  obtained from the sparse symmetric, indefinite factorization package `MA57` (Duff, 2004), is used. Here $\bar{J}_k$ comprises rows of J_k after redundancies have been identified and removed, exactly as described in Cartis and Gould (2006).

- Of the many constants involved in the algorithm, we chose

$$\kappa_{\rm ca} = 1000,\ \ \kappa_{\rm cr} = 2,\ \ \kappa_{\rm B} = 0.9,\ \ \kappa_\delta = 0.1,\ \ \kappa_{\rm tx1} = 0.9,\ \ \kappa_{\rm tx2} = 0.5,$$
$$\kappa_{\rm n} = 10^2,\ \ \kappa_{\rm nC} = 0.5,\ \ \kappa_{\rm tC1} = 0.5,\ \ \eta_1 = 0.1,\ \ \eta_2 = 0.9,\ \ \text{and}\ \ \eta_3 = 0.5.$$

  The other constants $(\kappa_{\rm nr}, \kappa_{\rm r}, \kappa_{\rm tt}, \kappa_{\rm nt}, \kappa_{\rm cn})$ are not needed for our simplified algorithm. Each bounding function used is $\omega(\alpha) = 0.01 \min(1, \alpha^2)$. The radius update (2.36) is implemented as

$$\Delta_{k+1}^f \in \begin{cases} 2\Delta_k^f & \text{if } \rho_k^f \geq \eta_2, \\ \Delta_k^f & \text{if } \rho_k^f \in [\eta_1, \eta_2), \\ 0.5\Delta_k^f & \text{if } \rho_k^f < \eta_1, \end{cases}$$

while that for (2.40) is

$$\Delta_{k+1}^c \in \begin{cases} 2\Delta_k^c & \text{if } \rho_k^c \geq \eta_2 \quad \text{and} \quad \delta_k^c \geq \kappa_{\rm cn}\delta_k^{c,n}, \\ \Delta_k^c & \text{if } \rho_k^c \in [\eta_1, \eta_2) \quad \text{and} \quad \delta_k^c \geq \kappa_{\rm cn}\delta_k^{c,n}, \\ 0.5\Delta_k^c & \text{if } \rho_k^c < \eta_1 \quad \text{or} \quad \delta_k^c < \kappa_{\rm cn}\delta_k^{c,n}. \end{cases}$$

Thus, at this stage, the possibility of computing inexact SQP steps has not been implemented, primarily because we have yet to finalise the details. Nevertheless, we believe that it is of interest to see whether the basic trust-funnel convergence mechanism we have developed shows promise.

Thus we give the results obtained by applying our algorithm to all of the equality constrained problems from the CUTEr collection (see Gould, Orban and Toint, 2003); for those whose dimensions may be adjusted, we chose small variants simply so as not to overload our computing environment. For each problem, in Table 4.1 we report its number of variables (n), its number of constraints (m), the number if iterations required for convergence (iter), the number of gradient evaluations (ngeval), the final objective function and constraint values (f and c) and the cpu-time spent (time). The algorithm is stopped as soon as the norms of primal and dual infeasibility $\|c(x_k)\|$ and $\|g(x_k) + J^T(x_k)y_k\|$ are both smaller than 10^{-5}. An upper limit on the number of iterations was set to 1000.

All of our experiments were performed on a single processor of a 3.06 GHz Dell Precision 650 Workstation. Our algorithm, Algorithm 2.1 on page 8, was implemented as a Matlab M-file, and the tests performed using Matlab 7.2.

Name	n	m	iter	ngeval	f	c	time
ALLINITC	2	1	9	8	-1.00e+00	9.27e-06	0.10
BT1	2	1	211	182	-1.00e-00	3.55e-13	1.82
BT2	3	1	12	12	3.26e-02	4.47e-06	0.11
BT3	5	3	6	6	4.09e+00	4.22e-15	0.05
BT4	3	2	9	7	-4.55e+01	1.33e-08	0.07
BT5	3	2	11	8	9.62e+02	7.45e-09	0.11
BT6	5	2	21	11	2.77e-01	1.39e-06	0.17
BT7	5	3	31	20	3.06e+02	3.34e-07	0.23
BT8	5	2	10	10	1.00e+00	3.81e-06	0.07
BT9	4	2	19	16	-1.00e+00	5.04e-06	0.24
BT10	2	2	7	7	-1.00e+00	4.18e-09	0.06
BT11	5	3	10	9	8.25e-01	2.99e-08	0.09
BT12	5	3	7	7	6.19e+00	8.30e-08	0.07
BYRDSPHR	3	2	8	7	-4.68e+00	1.74e-10	0.09
DATA1	90	73	1000	735	-1.50e+01	2.49e-08	54.28
DIXCHLNG	10	5	11	10	2.47e+03	2.97e-07	0.14
EIGENA2	110	55	4	4	0.00e+00	0.00e+00	0.07
EIGENACO	110	55	4	4	0.00e+00	0.00e+00	0.11
EIGENB2	110	55	3	3	1.80e+01	0.00e+00	0.07
EIGENBCO	110	55	3	3	9.00e+00	0.00e+00	0.08
EIGENC2	462	231	14	9	6.51e-11	2.77e-06	25.90
EIGENCCO	462	231	30	19	6.81e-28	2.88e-06	79.93
ELEC	150	50	110	54	1.06e+03	2.32e-09	2.97
FCCU	19	8	5	5	1.11e+01	3.55e-15	0.06
GRIDNETE	60	36	7	7	3.96e+01	1.78e-15	0.33
GRIDNETH	264	144	7	7	5.71e+01	2.07e-07	1.49
HS6	2	1	12	10	3.41e-21	1.66e-07	0.08
HS7	2	1	10	8	-1.73e+00	9.67e-10	0.08
HS8	2	2	8	6	-1.00e+00	3.62e-11	0.06
HS9	2	1	4	3	-5.00e-01	0.00e+00	0.10
HS26	3	1	17	14	7.37e-11	4.86e-06	0.15
HS27	3	1	10	10	4.00e-02	1.61e-06	0.13
HS28	3	1	3	3	0.00e+00	1.33e-15	0.04

Table 4.1: Results for the Trust-funnel Algorithm

Name	n	m	iter	ngeval	f	c	time
HS39	4	2	19	16	-1.00e+00	5.04e-06	0.26
HS40	4	3	5	5	-2.50e-01	1.69e-06	0.06
HS42	4	2	6	6	1.39e+01	6.21e-06	0.06
HS46	5	2	16	16	1.11e-09	7.92e-06	0.16
HS47	5	3	22	16	4.06e-10	5.41e-07	0.21
HS47	5	3	22	16	4.06e-10	5.41e-07	0.21
HS48	5	2	3	3	0.00e+00	1.55e-15	0.02
HS49	5	2	16	16	6.96e-09	1.33e-15	0.12
HS50	5	3	9	9	4.93e-32	2.66e-15	0.08
HS51	2	1	6	6	0.00e+00	0.00e+00	0.04
HS51	5	3	3	3	0.00e+00	0.00e+00	0.02
HS52	5	3	3	3	5.33e+00	4.01e-15	0.03
HS55SIM	1	1	2	2	0.00e+00	0.00e+00	0.03
HS56	7	4	125	63	-3.46e+00	2.42e-09	1.51
HS61	3	2	8	7	-1.44e+02	7.57e-06	0.09
HS77	5	2	21	11	2.42e-01	1.03e-06	0.16
HS78	5	3	6	5	-2.92e+00	2.24e-08	0.05
HS79	5	3	6	6	7.88e-02	1.34e-08	0.05
HS100LNP	7	2	15	9	6.81e+02	2.00e-09	0.16
HS111LNP	10	3	17	13	-4.78e+01	1.63e-06	0.17
KOPPEL	12	6	4	3	4.50e+00	2.98e-09	0.03
LCH	150	1	101	43	-4.23e+00	3.01e-07	4.16
LUKVLE1	100	98	8	8	6.23e+00	1.00e-08	0.21
LUKVLE2	100	49	1000	998	-4.04e+41	3.85e-06	15.30
LUKVLE3	100	2	11	11	2.76e+01	3.63e-08	0.25
LUKVLE4	100	49	1000	741	-1.63e+16	5.43e-07	16.13
LUKVLE6	99	49	16	16	6.04e+03	3.62e-06	0.47
LUKVLE7	100	4	13	11	-2.59e+01	1.11e-07	0.24
LUKVLE8	100	98	285	284	1.06e+04	1.63e-06	10.73
LUKVLE9	100	6	248	106	1.12e+01	9.56e-06	4.25
LUKVLE10	100	98	24	14	3.49e+01	7.40e-07	0.60
LUKVLE11	98	64	44	31	5.97e+02	7.88e-07	1.34
LUKVLE12	97	72	1000	942	2.73e+03	6.92e-01	7.45
LUKVLE13	98	64	15	15	7.90e+02	8.55e-07	0.72
LUKVLE14	98	64	40	27	1.04e+03	2.55e-09	1.17
LUKVLE15	97	72	69	42	2.51e-09	2.19e-06	3.13
LUKVLE16	97	72	24	21	1.47e+02	4.44e-06	0.40
LUKVLE17	97	72	31	29	3.05e+02	2.44e-06	0.86
LUKVLE18	97	72	28	18	1.05e+02	1.59e-06	0.57
LUKVLESC	98	64	17	17	1.05e-03	6.25e-08	0.43
MARATOS2	2	1	5	4	-1.00e+00	3.19e-08	0.05
MARATOS	2	1	5	4	-1.00e+00	3.19e-08	0.04
MWRIGHT	5	3	14	10	1.29e+00	8.68e-06	0.10
ORTHRDM2	203	100	12	8	7.78e+00	5.96e-08	0.42
ORTHRDS2	103	50	43	36	1.22e+02	6.29e-06	1.43
ORTHREGA	133	64	55	33	3.50e+02	5.62e-08	1.94
ORTHREGB	27	6	3	3	0.00e+00	1.06e-09	0.05
ORTHREGC	105	50	36	11	1.98e+00	9.20e-07	0.89
ORTHREGD	103	50	17	10	1.56e+01	1.14e-10	0.30
ORTHRGDM	23	10	27	24	3.90e+01	5.32e-06	0.34
ORTHRGDS	155	76	13	11	2.34e+01	1.25e-12	0.56
ROBOT	14	2	11	6	1.03e-29	1.15e-06	0.09
S316-322	2	1	3	3	3.34e+02	2.22e-16	0.02
WOODSNE	4	4	1000	996	-8.93e+00	1.00e+00	8.15

Table 4.1: Results for the Trust-funnel Algorithm (continued)

We are well aware that the numerical experience reported here is preliminary, and does not yet exploit all the interesting features of the algorithm. The results are however more than acceptable in most cases at this stage, which is encouraging and motivates further numerical (and theoretical) developments. Nevertheless, the algorithm did not solve five problems: of these `LUKVLE2` and `LUKVLE2` are reported to be unbounded from below, while `LUKVLE12` and `WOODSNE` are (at least locally) infeasible. Only `DATA1` genuinely failed, the algorithm not being able to reduce the dual feasibility below (roughly) 1.0e-4, This problem (along with `LUKVLE17`, `LUKVLE18`, `ORTHRDS2`, `ORTHREGD` and `ORTHRGDM`) has a severely rank-deficient Jacobian at the critical point found. One further problem `BT1` proved difficult, but further investigation revealed that this is was to the Maratos effect—adding a second-order correction along the lines discussed at the end of the previous section cured this immediately (the problem was then solved in 12 iterations).

5 Conclusion and perspectives

We have presented a new SQP algorithm for the solution of the equality constrained nonlinear programming problem, that avoids the use of penalty or barrier parameters and that allows for inexact tangential steps. Convergence to first-order critical point has been proved and preliminary numerical experience reported which motivates further research.

A first line of work is the inclusion of a multi-dimensional filter mechanism (see Gould, Leyffer and Toint, 2005) in the algorithm, with the objective to make the constraint on decreasing infeasibility more flexible. A second interesting development is the inclusion of bound or more general inequalities in the present framework. On a more practical level, further work is necessary to fully exploit the potentialities of the new method in allowing for inexact tangential steps and also to refine the current implementation by better tuning of the algorithmic parameters and the incorporation of preconditioning.

Acknowledgments

The authors are grateful to Professor Y. Yuan and his students at the Chinese Academy of Sciences for the organization of the ICNAO2006 conference in Beijing, which provided an excellent environment for the derivation of some of the results presented here. They are also pleased to acknowledge the hospitality of CERFACS (Toulouse) where discussions with S. Gratton, D. Orban and A. Sartenaer stimulated the work that led to this paper. The Matlab interface to `MA57` developed by Mario Arioli (RAL) is also gratefully acknowledged.

References

L. T. Biegler, J. Nocedal, and C. Schmid. A reduced Hessian method for large-scale constrained optimization. *SIAM Journal on Optimization*, **5**(2), 314–347, 1995.

R. H. Bielschowsky and F. A. M. Gomes. Dynamical control of infeasibility in nonlinearly constrained optimization. Technical Report 23/06, Department of Applied Mathematics, IMECC-UNICAMP, Campinas, Brasil, 2006.

R. H. Byrd, F. E. Curtis, and J. Nocedal. Inexact SQP methods for equality constrained optimization. Technical report, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, Illinois, USA, November 2006.

R. H. Byrd, J. Ch. Gilbert, and J. Nocedal. A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming, Series A*, **89**(1), 149–186, 2000*a*.

R. H. Byrd, N. I. M. Gould, J. Nocedal, and R. A. Waltz. An algorithm for nonlinear optimization using linear programming and equality constrained subproblems. *Mathematical Programming, Series B*, **100**(1), 27–48, 2004.

R. H. Byrd, M. E. Hribar, and J. Nocedal. An interior point algorithm for large scale nonlinear programming. *SIAM Journal on Optimization*, **9**(4), 877–900, 2000*b*.

C. Cartis and N. I. M. Gould. Finding a point in the relative interior of a polyhedron. Technical Report RAL-TR-2006-016, Rutherford Appleton Laboratory, Chilton, Oxfordshire, England, 2006.

T. F. Coleman and A. R. Conn. Nonlinear programming via an exact penalty function method : Asymptotic analysis. *Mathematical Programming*, **24**(3), 123–136, 1982.

A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust-Region Methods*. Number 01 *in* 'MPS-SIAM Series on Optimization'. SIAM, Philadelphia, USA, 2000.

I. S. Duff. MA57 - a code for the solution of sparse symmetric definite and indefinite systems. *ACM Transactions on Mathematical Software*, **30**(2), 118–144, 2004.

M. El-Alem. Global convergence without the assumption of linear independence for a trust-region algorithm for constrained optimization. *Journal of Optimization Theory and Applications*, **87**(3), 563–577, 1995.

M. El-Alem. A global convergence theory for a general class of trust-region-based algorithms for constrained optimization without assuming regularity. *SIAM Journal on Optimization*, **9**(4), 965–990, 1999.

R. Fletcher and S. Leyffer. Nonlinear programming without a penalty function. *Mathematical Programming*, **91**(2), 239–269, 2002.

R. Fletcher, N. I. M. Gould, S. Leyffer, Ph. L. Toint, and A. Wächter. Global convergence of trust-region SQP-filter algorithms for nonlinear programming. *SIAM Journal on Optimization*, **13**(3), 635–659, 2002*a*.

R. Fletcher, S. Leyffer, and Ph. L. Toint. On the global convergence of a filter-SQP algorithm. *SIAM Journal on Optimization*, **13**(1), 44–59, 2002*b*.

N. I. M. Gould, M. E. Hribar, and J. Nocedal. On the solution of equality constrained quadratic problems arising in optimization. *SIAM Journal on Scientific Computing*, **23**(4), 1375–1394, 2001.

N. I. M. Gould, S. Leyffer, and Ph. L. Toint. A multidimensional filter algorithm for nonlinear equations and nonlinear least-squares. *SIAM Journal on Optimization*, **15**(1), 17–38, 2005.

N. I. M. Gould, S. Lucidi, M. Roma, and Ph. L. Toint. Solving the trust-region subproblem using the Lanczos method. *SIAM Journal on Optimization*, **9**(2), 504–525, 1999.

N. I. M. Gould, D. Orban, and Ph. L. Toint. CUTEr, a constrained and unconstrained testing environment, revisited. *ACM Transactions on Mathematical Software*, **29**(4), 373–394, 2003.

M. Heinkenschloss and L. N. Vicente. Analysis of inexact trust region SQP algorithms. *SIAM Journal on Optimization*, **12**(2), 283–302, 2001.

D. M. Himmelblau. *Applied Nonlinear Programming*. McGraw-Hill, New-York, 1972.

M. Lalee, J. Nocedal, and T. D. Plantenga. On the implementation of an algorithm for large-scale equality constrained optimization. *SIAM Journal on Optimization*, **8**(3), 682–706, 1998.

X. Liu and Y. Yuan. A robust trust-region algorithm for solving general nonlinear programming problems. *SIAM Journal on Scientific Computing*, **22**, 517–534, 2000.

D. Q. Mayne and E. Polak. A superlinearly convergent algorithm for constrained optimization problems. *Mathematical Programming Studies*, **16**, 45–61, 1982.

E. O. Omojokun. *Trust region algorithms for optimization with nonlinear equality and inequality constraints*. PhD thesis, University of Colorado, Boulder, Colorado, USA, 1989.

T. Steihaug. The conjugate gradient method and trust regions in large scale optimization. *SIAM Journal on Numerical Analysis*, **20**(3), 626–637, 1983.

Ph. L. Toint. Towards an efficient sparsity exploiting Newton method for minimization. *in* I. S. Duff, ed., 'Sparse Matrices and Their Uses', pp. 57–88, London, 1981. Academic Press.

C. Zoppke-Donaldson. *A Tolerance-Tube Approach to Sequential Quadratic Programming with Applications*. PhD thesis, Department of Mathematics and Computer Science, University of Dundee, Dundee, Scotland, UK, 1995.