Science & Technology Facilities Council

Preprint

RAL-P-2014-001

An interior-point trust-funnel algorithm for nonlinear optimization

FE Curtis, NIM Gould, DP Robinson, PhL Toint

January 2014

Submitted for publication in Mathematical Programming

RAL Library
STFC Rutherford Appleton Laboratory
R61
Harwell Oxford
Didcot
OX11 0QX

Tel: +44(0)1235 445384
Fax: +44(0)1235 446403
email: libraryral@stfc.ac.uk

ISSN 1361- 4762

An interior-point trust-funnel algorithm for nonlinear optimization

Frank E. Curtis,[1,2] Nicholas I. M. Gould,[3,3] and Daniel P. Robinson[5,6] and Philippe L. Toint[7]

ABSTRACT

We present an interior-point trust-funnel algorithm for solving large-scale nonlinear optimization problems. The method is based on an approach proposed by Gould and Toint (*Math. Prog.*, 122(1):155-196, 2010) that focused on solving equality constrained problems. Our method, which is designed to solve problems with both equality and inequality constraints, achieves global convergence guarantees by combining a trust-region methodology with a funnel mechanism. The prominent features of our algorithm are that (i) the subproblems that define each search direction may be solved approximately, (ii) criticality measures for feasibility and optimality aid in determining which subset of computations will be performed during each iteration, (iii) no merit function or filter is used, (iv) inexact sequential quadratic optimization steps may be computed when advantageous, and (v) it may be implemented matrix-free so that derivative matrices need not be formed or factorized so long as matrix-vector products with them can be performed.

---

[1] Department of Industrial and Systems Engineering, Lehigh University,
Harold S. Mohler Laboratory, 200 West Packer Avenue Bethlehem, PA 18015-1582, USA.
Email : frank.e.curtis@gmail.com .

[2] This work was supported by U.S. Department of Energy grant DE–SC0010615 and U.S. National Science Foundation grant DMS–1016291.

[3] Scientific Computing Department, Rutherford Appleton Laboratory,
Chilton, Oxfordshire, OX11 0QX, England, EU. Email: nick.gould@stfc.ac.uk .
Current reports available from "http://www.numerical.rl.ac.uk/people/nimg/pubs.html".

[4] This work was supported by the EPSRC grant EP/I013067/1.

[5] Department of Applied Mathematics and Statistics, Johns Hopkins University,
100 Whitehead Hall, 3400 N. Charles Street, Baltimore, MD 21218-2682, USA.
Email : daniel.p.robinson@gmail.com .

[6] This work was supported by the U.S. National Science Foundation grant DMS–1217153.

[7] Namur Center for Complex Systems (naXys) and Department of Mathematics,
University of Namur, 61, rue de Bruxelles, B-5000 Namur, Belgium, EU.
Email : philippe.toint@unamur.be .
Current reports available from "http://www.fundp.ac.be/∼phtoint/pht/publications.html".

Scientific Computing Department
Rutherford Appleton Laboratory
Oxfordshire OX11 0QX

January 2nd, 2014

1 Introduction

In this paper, we introduce a method for solving optimization problems of the form

$$\underset{x\in\mathbb{R}^N}{\text{minimize}} \;\; f(x) \quad \text{subject to} \;\; c(x) \leq 0, \tag{NP}$$

where $f : \mathbb{R}^N \to \mathbb{R}$ and $c : \mathbb{R}^N \to \mathbb{R}^M$ are twice continuously differentiable. (Our method can also be applied when equality constraints are present, but, for simplicity in our discussion, they are suppressed in our algorithm development and analysis; see §5 for further discussion.) Our algorithm is designed to solve large-scale instances of (NP). In particular, it is designed to be matrix-free in the sense that an implementation of it only requires matrix-vector products with the constraint Jacobian, its transpose, symmetric approximations of the Hessian of the Lagrangian, and corresponding preconditioners. That is, iterative methods may be used to approximately solve each subproblem arising in the algorithm.

The method we propose utilizes components of both interior-point (IP) and sequential quadratic optimization (commonly known as SQP) methods. Algorithms of this type are often referred to as barrier-SQP methods. The interior-point aspects of our algorithm allow us to avoid the combinatorial explosion that may occur within, say, an active-set approach. The efficiency of interior-point methods for solving linear and convex quadratic optimization problems has been well-established [1, 7, 12–14, 18, 19, 26, 30, 32]. Extending these methods for solving nonlinear problems has been the subject of research for decades [3, 4, 6, 15, 33–37] and numerical evidence illustrates strong performance. We follow an approach similar to Byrd et. al. [3, 4] and solve a sequence of so-called barrier subproblems for decreasing values of the barrier parameter. This means that we must solve a sequence of equality constrained subproblems, and these may be solved efficiently with an SQP-based method. It is well known that traditional SQP methods are very efficient for solving small- to medium-sized optimization problems [8, 9, 16, 17], while more recently proposed SQP methods utilize exact second derivatives and are, in theory, capable of solving large problems [21–23, 31]. Preliminary results when solving small- to medium-sized problems is promising, but their effectiveness on large problems has not yet been confirmed. There have, however, been several proposed SQP strategies that have proved capable of solving large equality constrained problems [2, 25, 29].

In this paper, we use the trust-funnel approach originally described in [25], and then corrected in [24], as the basis for solving a sequence of equality constrained barrier subproblems that arise in an interior-point framework. We note, however, that a naïve implementation of the SQP method described in [24, 25] within an interior-point paradigm may result in a method for which the establishment of convergence guarantees is elusive. This is a consequence of the fact that interior-point methods—as their name suggests—require the algorithm iterates to remain in the strict interior of the feasible region associated with the inequality constraints, while the method in [24, 25] does not innately possess the mechanisms necessary to avoid the boundary of the feasible region in this context. In this paper, we describe modifications of this trust-funnel method that are appropriate for our interior-point setting. These modifications include imposing explicit constraints in the trust-region subproblems to ensure that the iterates remain in the strict interior of the feasible region, and the incorporation of scaled trust-region constraints and optimality measures. Scalings of these types have been used previously in interior-point methods; e.g., see [3, 6].

The paper is organized as follows. In Section 2, we introduce our trust-funnel algorithm for solving the barrier subproblem in an interior-point approach. In Section 3, we prove that our trust-funnel algorithm will terminate finitely with arbitrarily small positive tolerances on the criticality measures. In Section 4, we consider convergence of the barrier subproblem solutions for a decreasing sequence of the barrier parameter. Finally, conclusions are provided in Section 5.

Notation

The gradient and Hessian of f at x are written as $g(x)$ and $\nabla_{xx}f(x)$ respectively. The $M \times N$ matrix $J(x)$ represents the Jacobian of the constraint function c evaluated at x, with its jth row being $\nabla c_j(x)^T$. The matrix $\nabla_{xx}c_j(x)$ is the Hessian of c_j evaluated at x. We let e denote the vector of all ones and I denote the identity matrix, both of whose dimensions are determined by the context in which they are used. Given a vector $s \in \mathbb{R}^M$, $[s]_j$ is the jth element of s and $S := \text{diag}([s]_1, [s]_2, \ldots, [s]_M)$. A forcing function $\omega : [0, \infty) \to [0, \infty)$ is defined as any continuous and strictly increasing function that satisfies $\omega(0) = 0$.

Preliminaries

We make the following assumption throughout the paper.

Assumption 1.1. *The functions f and c are twice continuously differentiable.*

Problem (NP) is not solved directly by our algorithm. Rather, we introduce a vector of slack variables $s \in \mathbb{R}^M$ and solve the equivalent optimization problem

$$\underset{x\in\mathbb{R}^N, s\in\mathbb{R}^M}{\text{minimize}} \quad f(x) \quad \text{subject to} \quad c(x,s) := c(x) + s = 0, \quad s \geq 0. \tag{NPs}$$

The following definition gives first-order stationarity conditions for (NPs) [27, 28].

Definition 1.1 (First-order KKT point for problem (NPs))**.** *The vector triple (x, s, y) is a first-order KKT point for problem* (NPs) *if it satisfies*

$$g(x) + J(x)^T y = 0, \quad c(x,s) = 0, \quad Sy = 0, \quad \textit{and} \quad (s,y) \geq 0.$$

To solve (NPs), we compute a sequence of (approximate) minimizers of the barrier subproblem

$$\underset{x\in\mathbb{R}^N, s\in\mathbb{R}^M}{\text{minimize}} \quad f(x,s) \quad \text{subject to} \quad c(x,s) = 0, \quad s > 0, \tag{1.1}$$

where for each fixed value of $\mu > 0$ we define the barrier function

$$f(x,s) := f(x) - \mu \sum_{i=1}^{M} \ln([s]_i). \tag{1.2}$$

Given a Lagrange multiplier estimate vector y for the constraint $c(x, s) = 0$, the Lagrangian function associated with (1.1) and its gradient with respect to (x, s) are given by

$$\mathcal{L}(x,s,y) := f(x,s) + c(x,s)^T y \quad \text{and} \quad \nabla_{(x,s)}\mathcal{L}(x,s,y) := \nabla f(x,s) + J(x,s)^T y,$$

where we define $J(x,s) := \nabla c(x,s) = \begin{pmatrix} J(x) & I \end{pmatrix}$ to represent the Jacobian of $c(x, s)$ with respect to (x, s). A primal-dual point (x, s, y) is called a first-order KKT point of the barrier problem (1.1) if it satisfies

$$\nabla_{(x,s)}\mathcal{L}(x,s,y) = 0, \quad c(x,s) = 0, \quad \text{and} \quad (s,y) > 0.$$

Multiplying the second block of the first equation by S leads to the following equivalent definition.

Definition 1.2 (First-order KKT point for the barrier subproblem (1.1))**.** *The vector triple (x, s, y) is a first-order KKT-point for the barrier subproblem* (1.1) *if it satisfies*

$$g(x) + J(x)^T y = 0, \quad Sy = \mu e, \quad c(x,s) = 0, \quad \textit{and} \quad (s,y) > 0.$$

A comparison of Definitions 1.1 and 1.2 suggests that KKT points of the barrier subproblem become increasingly accurate solutions to problem (NPs) for decreasing values of the barrier parameter μ.

2 A Trust-Funnel Algorithm for Solving the Barrier Subproblem

In this section, we present our trust-funnel algorithm for (approximately) solving the barrier subproblem (1.1) for a fixed value of the barrier parameter $\mu > 0$. As μ is fixed for a particular instance of (1.1), the dependence on μ of certain quantities in this section is ignored. However, these dependences—in particular, with respect to criticality tolerances that are employed in the algorithm—will be a central focus in §4 that addresses the "outer" algorithm for solving problem (NPs).

2.1 Algorithm overview

Our method generates a sequence $\{(x_k, s_k, y_k)\}$ of primal, slack, and dual variables. In addition, defining the measure of constraint violation

$$v(x, s) := \|c(x, s)\|_2, \tag{2.3}$$

our method maintains a monotonically decreasing sequence of positive scalars $\{v_k^{\max}\}$ such that

$$s_k > 0, \;\; c(x_k, s_k) \geq 0, \;\; v_k := v(x_k, s_k) \leq v_k^{\max}, \;\; \text{and} \;\; v_{k+1}^{\max} \leq v_k^{\max} \;\; \text{for all } k \geq 0. \tag{2.4}$$

We require $s_0 > 0$, and the restriction that $s_k > 0$ is maintained via explicit constraints imposed on all search direction calculations. Additionally, we ensure that $c(x_k, s_k) \geq 0$ holds at the beginning of iteration k by incorporating a *slack reset* procedure that sets

$$[s_k]_i \leftarrow \begin{cases} [s_k]_i & \text{if } [c(x_k, s_k)]_i \geq 0, \\ -[c(x_k)]_i & \text{otherwise.} \end{cases} \tag{2.5}$$

If we let s_k^{prior} denote the value of s_k *prior* to the slack reset in iteration k, then it follows that

$$v_k \leq v(x_k, s_k^{\text{prior}}), \quad s_k^{\text{prior}} \leq s_k, \quad \text{and} \quad f(x_k, s_k) \leq f(x_k, s_k^{\text{prior}}). \tag{2.6}$$

That is, both the barrier function and constraint violation decrease as a result of the (trivial) slack reset computation (2.5). We explicitly enforce $v_k \leq v_k^{\max}$ with the updating strategy discussed in Section 2.4. Finally, the sequence $\{v_k^{\max}\}$ is positive and monotonically decreasing by construction and guides the iterates toward feasibility; the set of points permitted by the gradually narrowing region defined by $v(x, s) \leq v_k^{\max}$ is called the *funnel* [24, 25]. Overall, the claims in (2.4) are formally established in Section 2.6.

Given the current estimate (x_k, s_k) of a solution of (1.1), a trial step $d_k := (d_k^x, d_k^s)$ is computed as the sum of a "normal" step $n_k := (n_k^x, n_k^s)$ and a "tangential" step $t_k := (t_k^x, t_k^s)$, i.e.,

$$d_k = \begin{pmatrix} d_k^x \\ d_k^s \end{pmatrix} = \begin{pmatrix} n_k^x \\ n_k^s \end{pmatrix} + \begin{pmatrix} t_k^x \\ t_k^s \end{pmatrix} = n_k + t_k.$$

The normal step is computed to (approximately) minimize a Gauss-Newton model of v at (x_k, s_k); thus, it has the purpose of reducing linearized infeasibility. The tangential step t_k is intended to reduce the barrier function (1.2) and is calculated as an (approximate) minimizer of a quadratic model of the barrier function within an appropriate subspace that does not undo the improvement in reducing linearized infeasibility achieved by n_k. Once $d_k = n_k + t_k$ is computed, an attempt to decrease the constraint violation and/or barrier function is made, where the decision of which to consider is based

on quantities that reflect the overall merit of the constituent steps. We discuss these details in turn in the following subsections.

2.2 The normal step

The normal step is designed to predict a reduction in constraint violation as measured by v defined in (2.3). To achieve this goal, we compute the normal step $n_k := (n_k^x, n_k^s)$ as an approximate solution of

$$\underset{n=(n^x,n^s)}{\text{minimize}} \; m_k^v(n) \;\; \text{subject to} \;\; \|P_k^{-1}n\|_2 \leq \min\{\delta_k^v, \kappa_{\text{n}}\pi_k^v\}, \;\; s_k + n^s \geq \kappa_{\text{fbn}} s_k, \tag{2.7}$$

where $\kappa_{\text{n}} \geq 1$ and $\kappa_{\text{fbn}} \in (0,1)$ are constants, $\delta_k^v > 0$ is a dynamic algorithm parameter, and we define

$$m_k^v(n) := \|c(x_k,s_k) + J(x_k,s_k)n\|_2 \quad \text{and} \quad P_k := \begin{pmatrix} I & 0 \\ 0 & S_k \end{pmatrix} \tag{2.8}$$

along with the "v-criticality" measures

$$\pi_k^v := \pi^v(x_k,s_k) := \|P_k J(x_k,s_k)^T c(x_k,s_k)\|_2 \quad \text{and} \quad \chi_k^v := \chi^v(x_k,s_k) := \begin{cases} \pi_k^v/v_k & \text{if } v_k \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

The quantities π_k^v and χ_k^v serve as criticality measures at (x_k, s_k) for minimizing v subject to the slacks being nonnegative. The scaling matrix P_k is important in the trust region constraint since it assists in keeping the iterates within the nonnegative orthant; it restricts $[n_k^s]_j$ to be relatively small when $[s_k]_j$ is close to zero. Overall, problem (2.7) involves the local minimization of the norm of a Gauss-Newton approximation of v at (x_k, s_k) subject to a trust-region constraint and a fraction-to-the-boundary rule.

It is not necessarily prudent to compute a normal step in every iteration. Indeed, computing a normal step may be wasteful if the current iterate is nearly feasible and computational efforts may be better spent on computing a new Lagrange multiplier estimate or tangential step. In our algorithm, we only require a normal step to be computed when either our v-criticality measure π_k^v is sufficiently large relative to our previous "f-criticality" measure π_{k-1}^f (defined in (2.25) in the next subsection), or when v_k is sufficiently large relative to $v_k^{\max}$ (see (2.4)). Specifically, for some $\kappa_{\text{vv}} \in (0,1)$ and forcing function ω_n (and with $\pi_{-1}^f := 0$), we require the computation of a normal step if either

$$\pi_k^v > \omega_n(\pi_{k-1}^f) \quad \text{or} \quad v_k \geq \kappa_{\text{vv}} v_k^{\max}. \tag{2.10}$$

(If (2.10) does not hold, but $\pi_k^v > 0$, then one may still consider computing a normal step since the fact that $\pi_k^v > 0$ implies that the computation would be well-defined. However, in such cases, a normal step is not necessary for our convergence analysis.) When a normal step is not computed, we set $n_k \leftarrow 0$.

By an approximate solution to (2.7), we mean that $n_k := (n_k^x, n_k^s)$—when it is computed—should be feasible for (2.7) and yield a decrease in m_k^v no less than that achieved along a scaled steepest descent direction for $(m_k^v)^2$. The scaled steepest descent direction that we employ for this comparison is derived in the following manner. Performing the change of variables $n^{\text{P}} := P_k^{-1}n$ so that the trust-region constraint becomes $\|n^{\text{P}}\|_2 \leq \min\{\delta_k^v, \kappa_{\text{n}}\pi_k^v\}$, the transformed problem for minimizing $(m_k^v)^2$ has the conventional ℓ_2-norm steepest descent direction $-P_k J(x_k,s_k)^T c(x_k,s_k)$. Returning to the original space gives the scaled steepest descent direction $-P_k^2 J(x_k,s_k)^T c(x_k,s_k)$. For (2.7), we define the Cauchy step $n_k^{\text{C}} = (n_k^{\text{C}x}, n_k^{\text{C}s})$ as the minimizer of the objective of (2.7) in this scaled steepest descent direction, i.e.,

$$n_k^{\text{C}} := n_k^{\text{C}}(\alpha_{\text{N}}^{\text{C}}), \;\; \text{where} \;\; n_k^{\text{C}}(\alpha) := \begin{pmatrix} n_k^{\text{C}x}(\alpha) \\ n_k^{\text{C}s}(\alpha) \end{pmatrix} := -\alpha P_k^2 J(x_k,s_k)^T c(x_k,s_k) \tag{2.11}$$

and $\alpha_{\mathrm{N}}^{\mathrm{C}}$ is the solution to

$$\underset{\alpha\geq 0}{\text{minimize}}\ \ m_k^v(n_k^{\mathrm{C}}(\alpha))\ \ \text{subject to}\ \ \|P_k^{-1}n_k^{\mathrm{C}}(\alpha)\|_2 \leq \min\{\delta_k^v, \kappa_{\mathrm{n}}\pi_k^v\},\ \ s_k + n_k^{\mathrm{C}s}(\alpha) \geq \kappa_{\mathrm{fbn}}s_k. \tag{2.12}$$

We show in Lemma 2.8 that the decrease in m_k^v obtained by n_k^{C} is positive.

Overall, when (2.10) holds, we require a normal step n_k satisfying the constraints of (2.7), i.e.,

$$\|P_k^{-1}n_k\|_2 \leq \min\{\delta_k^v, \kappa_{\mathrm{n}}\pi_k^v\},\quad s_k + n_k^s \geq \kappa_{\mathrm{fbn}}s_k, \tag{2.13}$$

along with

$$\Delta m_k^{v,n} := m_k^v(0) - m_k^v(n_k) \geq m_k^v(0) - m_k^v(n_k^{\mathrm{C}}). \tag{2.14}$$

Many steps satisfy (2.13) and (2.14) with the simplest choice being $n_k = n_k^{\mathrm{C}}$.

2.3 Lagrange multipliers and the tangential step

Having dealt with the normal step, we now consider computing estimates of an optimal Lagrange multiplier vector and/or a tangential step. The multiplier estimates, if computed, are intended to (approximately) minimize a measure of criticality for the barrier subproblem (1.1) that takes into account changes in the problem function values that are predicted by the normal step. The tangential step, if computed, is designed to reduce the barrier function without having too adverse an effect on the reduction in linearized infeasibility predicted by the normal step. Since the conditions imposed on the multiplier estimates and tangential step are intertwined—e.g., the computed multiplier estimates are required to have a well-defined Cauchy point for the tangential step subproblem—we consider their computations together in this subsection. Our motivation in this section is to compute quantities related to those in a traditional SQP approach applied to subproblem (1.1).

Given the kth estimate y_k of an optimal Lagrange multiplier vector, a traditional SQP trial step associated with the barrier subproblem (1.1) is defined as the solution (when it exists) of

$$\underset{d=(d^x,d^s)}{\text{minimize}}\ \ f(x_k,s_k)+\nabla f(x_k,s_k)^T d+\tfrac{1}{2}d^T\nabla_{(x,s)(x,s)}\mathcal{L}(x_k,s_k,y_k)d\quad\text{subject to}\quad c(x_k,s_k)+J(x_k,s_k)d=0.$$

It may be verified that a solution $d=(d^x,d^s)$ of this subproblem (when it exists) satisfies

$$\begin{pmatrix}\nabla_{xx}\mathcal{L}(x_k,y_k) & J(x_k)^T & 0\\ J(x_k) & 0 & I\\ 0 & S_k & \mu S_k^{-1}\end{pmatrix}\begin{pmatrix}d^x\\ y\\ d^s\end{pmatrix} = -\begin{pmatrix}g(x_k)\\ c(x_k,s_k)\\ -\mu e\end{pmatrix}, \tag{2.15}$$

where y is an estimate of an optimal Lagrange multiplier vector for the constraint $c(x_k,s_k)+J(x_k,s_k)d=0$. The SQP step generated in this fashion is often called a *primal* step since the dual vector y_k does not appear in (2.15) other than in the Hessian $\nabla_{xx}\mathcal{L}$. We can instead compute a *primal-dual* step by applying Newton's Method to the conditions in Definition 1.2, which leads to

$$\begin{pmatrix}\nabla_{xx}\mathcal{L}(x_k,y_k) & J(x_k)^T & 0\\ J(x_k) & 0 & I\\ 0 & S_k & Y_k\end{pmatrix}\begin{pmatrix}d^x\\ y\\ d^s\end{pmatrix} = -\begin{pmatrix}g(x_k)\\ c(x_k,s_k)\\ -\mu e\end{pmatrix}. \tag{2.16}$$

This system is identical to (2.15), except that the $(3,3)$-block now contains dual information. It is also easily verified that a solution of (2.16) is a KKT point for

$$\underset{d=(d^x,d^s)}{\text{minimize}}\quad f(x_k,s_k)+\nabla f(x_k,s_k)^T d+\tfrac{1}{2}d^T H(x_k,s_k,y_k)d\quad\text{subject to}\quad c(x_k,s_k)+J(x_k,s_k)d=0,$$

where

$$H(x_k, s_k, y_k) := \begin{pmatrix} \nabla_{xx}\mathcal{L}(x_k, y_k) & 0 \\ 0 & Y_k S_k^{-1} \end{pmatrix}.$$

The previous paragraph, along with the widely accepted view that the *primal-dual* approach is generally superior to the *primal* approach in practice, motivates us to approximate the barrier function (1.2) with

$$m_k^f(d) := f(x_k, s_k) + \nabla f(x_k, s_k)^T d + \tfrac{1}{2} d^T G_k d, \tag{2.17}$$

where, for all k, we define

$$G_k := \begin{pmatrix} \nabla_{xx}\mathcal{L}(x_k, y_k^{\text{B}}) & 0 \\ 0 & D_k \end{pmatrix} \tag{2.18}$$

with y_k^{B} a (bounded) Lagrange multiplier vector satisfying

$$[y_k^{\text{B}}]_i > 0 \ \text{ for all } i \in \{1, 2, \ldots, M\} \quad \text{and} \quad \|y_k^{\text{B}}\|_2 \leq \kappa_{\text{y}} \ \text{ for some scalar } \kappa_{\text{y}} > 0, \tag{2.19}$$

and choose $D_k \approx Y_k S_k^{-1}$ as a positive-definite diagonal matrix satisfying

$$\|D_k\|_2 \leq \kappa_{\text{D}} \ \text{ for some scalar } \kappa_{\text{D}} > 0. \tag{2.20}$$

Overall, our goal is to compute a tangential step t_k that satisfies $m_k^f(n_k + t_k) \leq m_k^f(n_k)$ and lies approximately in the null space of the constraint Jacobian $J(x_k, s_k)$ so as not to undo the predicted gain in linearized feasibility provided by the normal step. This latter requirement implies that improvement in the barrier function should be sought within the trust-region $\{d : \|P_k^{-1} d\|_2 \leq \delta_k^v\}$, since it is only within this region that the linearized constraint model is believed to be trustworthy. In addition, we assume that the barrier function model m_k^f may be trusted as a faithful representation within the trust-region $\{d : \|P_k^{-1} d\|_2 \leq \delta_k^f\}$ for a given tangential trust-region radius $\delta_k^f > 0$. Consequently, we use

$$\|P_k^{-1} n_k\|_2 \leq \kappa_{\text{B}} \min\{\delta_k^v, \delta_k^f\} \ \text{ with } \ \kappa_{\text{B}} \in (0, 1) \tag{2.21}$$

as a necessary condition for computing a tangential step. If (2.21) is satisfied, then we require the computation of a new Lagrange multiplier estimate and, potentially, a tangential step. Otherwise, we set $y_k \leftarrow y_{k-1}$ and $t_k \leftarrow 0$ since the cost of computing new multipliers and a tangential step may be wasteful.

When (2.21) is satisfied, we seek an approximate solution of the tangential step subproblem

$$\begin{aligned} &\underset{t=(t^x, t^s)}{\text{minimize}} \ \ m_k^f(n_k + t) \\ &\text{subject to } \ J(x_k, s_k) t = 0, \ \ \|P_k^{-1}(n_k + t)\|_2 \leq \min\{\delta_k^v, \delta_k^f\}, \ \ s_k + n_k^s + t^s \geq \kappa_{\text{fbt}}(s_k + n_k^s) \end{aligned} \tag{2.22}$$

for some $\kappa_{\text{fbt}} \in (0, 1)$. Observing the change of variables $t^{\text{P}} = P_k^{-1} t$, this subproblem is equivalent to

$$\begin{aligned} &\underset{t^{\text{P}}=(t^{\text{Px}}, t^{\text{Ps}})}{\text{minimize}} \ \ m_k^f(n_k + P_k t^{\text{P}}) \\ &\text{subject to } \ J(x_k, s_k) P_k t^{\text{P}} = 0, \ \ \|P_k^{-1} n_k + t^{\text{P}}\|_2 \leq \min\{\delta_k^v, \delta_k^f\}, \ \ t^{\text{Ps}} \geq (\kappa_{\text{fbt}} - 1)(e + S_k^{-1} n_k^s). \end{aligned}$$

To define an appropriate Cauchy point, we first compute approximate least-squares multipliers corresponding to the scaled subproblem at $t^{\text{P}} = 0$, i.e., we compute y_k as an approximate solution of

$$\underset{y \in \mathbb{R}^M}{\text{minimize}} \ \ m_k^{\mathcal{L}}(y), \ \ \text{where } \ m_k^{\mathcal{L}}(y) := \tfrac{1}{2} \|P_k(\nabla m_k^f(n_k) + J(x_k, s_k)^T y)\|_2^2. \tag{2.23}$$

Scaling the resulting (approximate) projected gradient back into the original space, we obtain the

(approximate) oblique projected gradient

$$r_k := r_k(y_k) := P_k^2(\nabla m_k^f(n_k) + J(x_k, s_k)^T y_k) \tag{2.24}$$

and the related f-criticality measures

$$\pi_k^f := \pi_k^f(y_k) := \|P_k(\nabla m_k^f(n_k) + J(x_k, s_k)^T y_k)\|_2 \quad \text{and} \quad \chi_k^f := \chi_k^f(y_k) := \frac{\nabla m_k^f(n_k)^T r_k(y_k)}{\pi_k^f(y_k)} \tag{2.25}$$

associated with minimizing the barrier function. In this computation, we require that y_k and the resulting r_k, π_k^f, and χ_k^f satisfy at least one of the following three sets of conditions:

$$\pi_k^f \leq \epsilon_\pi \text{ and } v_k \leq \epsilon_v; \tag{2.26a}$$
$$\pi_k^f \leq \omega_t(\pi_k^v); \tag{2.26b}$$
$$\chi_k^f \geq \kappa_\chi \pi_k^f. \tag{2.26c}$$

Here, $\{\epsilon_\pi, \epsilon_v\} > 0$ and $\kappa_\chi \in (0, 1)$ are constants and ω_t is a forcing function. For technical reasons (in the proof of Lemma 2.6(vii)), we require that the functions ω_n and ω_t (see (2.10) and (2.26b)) satisfy

$$\omega_t(\omega_n(\tau)) \leq \kappa_\omega \tau \text{ for all } \tau \geq 0 \text{ and for some } \kappa_\omega \in (0, 1). \tag{2.27}$$

The presence of P_k in (2.23) forces components of the approximate projected gradient in (2.24) to be large when the corresponding components of s_k are small. Thus, this scaling matrix helps prevent slack variables from approaching zero, just as it did in the formulation of the normal step subproblem (2.7). Later, Lemma 2.10 shows that we can always satisfy one of the three sets of conditions in (2.26), and thus this strategy for computing y_k (and the related quantities r_k, π_k^f, and χ_k^f) is well-posed.

If (2.26a) is satisfied, then (x_k, s_k, y_k) is an approximate first-order KKT point for the barrier subproblem for the tolerances $\{\epsilon_\pi, \epsilon_v\} > 0$, so we terminate the algorithm for solving (1.1). However, if (2.26a) is not satisfied, but (2.26b) holds, then the f-criticality measure π_k^f is insubstantial compared to the v-criticality measure π_k^v. In this case, the computation of a tangential step is skipped, i.e., we simply set $t_k \leftarrow 0$. Otherwise, when (2.26a) and (2.26b) do not hold (and necessarily (2.26c) holds), we proceed to compute a tangential step. In this case, it follows from the definition (2.25), the condition (2.26c) and the fact that $\pi_k^f > 0$ (since otherwise (2.26b) would have held) that r_k is a direction of strict ascent for $m_k^f(\cdot)$ at n_k. This property allows us to compute a tangential step t_k satisfying one of two sets of conditions as outlined in the following two subsections. Our choice of which set of conditions to satisfy depends on whether a normal step is computed. Specifically, if $n_k \neq 0$, then we require the computation of what we call a relaxed SQP tangential step. Otherwise, if $n_k = 0$, then we are still free to attempt to compute a relaxed SQP tangential step, but we may instead compute what we call a very relaxed SQP tangential step. In such a case, this latter option may be preferable as it involves a weaker restriction on linearized infeasibility of the step.

2.3.1 A relaxed SQP tangential step

Given constants $\kappa_{\text{fbt}} \in (0, 1)$ and $\kappa_{\text{tg}} \in (0, 1)$, a relaxed SQP tangential step is defined as follows.

Definition 2.1 (Relaxed SQP tangential step)**.** *Define the Cauchy point*

$$t_k^{\text{C}} := t_k^{\text{C}}(\alpha_{\text{T}}^{\text{C}}), \quad \textit{where} \quad t_k^{\text{C}}(\alpha) := \begin{pmatrix} t_k^{\text{C}x}(\alpha) \\ t_k^{\text{C}s}(\alpha) \end{pmatrix} := -\alpha \begin{pmatrix} r_k^x \\ r_k^s \end{pmatrix} = -\alpha r_k \tag{2.28}$$

and $\alpha_{\mathrm{T}}^{\mathrm{C}}$ *is the minimizer of*

$$\begin{array}{ll}\underset{\alpha\geq 0}{\text{minimize}} & m_k^f\big(n_k + t_k^{\mathrm{C}}(\alpha)\big) \\ \text{subject to} & \|P_k^{-1}\big(n_k + t_k^{\mathrm{C}}(\alpha)\big)\|_2 \leq \min\{\delta_k^v, \delta_k^f\}, \quad s_k + n_k^s + t_k^{\mathrm{C}s}(\alpha) \geq \kappa_{\mathrm{fbt}}(s_k + n_k^s).\end{array} \tag{2.29}$$

Then, t_k *is a relaxed SQP tangential step if*

$$\Delta m_k^{f,t} := m_k^f(n_k) - m_k^f(n_k + t_k) \geq m_k^f(n_k) - m_k^f(n_k + t_k^{\mathrm{C}}), \tag{2.30a}$$

$$s_k + n_k^s + t_k^s \geq \kappa_{\mathrm{fbt}}(s_k + n_k^s), \tag{2.30b}$$

$$\|P_k^{-1}(n_k + t_k)\|_2 \leq \min\{\delta_k^v, \delta_k^f\}, \quad \textit{and} \tag{2.30c}$$

$$m_k^v(n_k + t_k) \leq \kappa_{\mathrm{tg}} m_k^v(0) + (1 - \kappa_{\mathrm{tg}}) m_k^v(n_k). \tag{2.30d}$$

Condition (2.30a) ensures that the model of the barrier function is decreased at least as much as by the Cauchy point t_k^{C}, (2.30b) is a fraction-to-the-boundary constraint, (2.30c) is a trust-region constraint, and (2.30d) is a relaxation of the traditional SQP constraint that $c(x_k, s_k) + J(x_k, s_k)(n_k + t_k) = 0$ that ensures that linearized constraint infeasibility is sufficiently reduced.

If a relaxed SQP tangential step satisfying (2.30) is computed, then we must evaluate its usefulness in the sense that we must ensure that a relatively large tangential step results in a sufficient decrease in the model m_k^f of the barrier function. With this in mind, we check whether the conditions

$$\|P_k^{-1} t_k\|_2 > \kappa_{\mathcal{VS}} \|P_k^{-1} n_k\|_2 \ \text{ for some } \kappa_{\mathcal{VS}} > 1 \tag{2.31}$$

and

$$\Delta m_k^{f,d} := \Delta m_k^{f,n} + \Delta m_k^{f,t} \geq \kappa_\delta \Delta m_k^{f,t} \ \text{ for some } \kappa_\delta \in (0,1) \text{ with } \Delta m_k^{f,n} := m_k^f(0) - m_k^f(n_k) \tag{2.32}$$

are satisfied. The inequality (2.32) indicates that the predicted decrease in the barrier function obtained from the tangential step is substantial when compared to the possible increase resulting from the normal step. If the step t_k satisfies (2.31) but violates (2.32), it does not serve its role so we reset it to zero.

2.3.2 A very relaxed SQP tangential step

Condition (2.30) may be too restrictive in certain cases. Specifically, if $v_k = 0$, then the algorithm will set $n_k = 0$, from which it follows that (2.30d) requires t_k to be in the null space of $J(x_k, s_k)$. This is an unreasonable requirement in matrix-free settings; indeed (2.30d) may be unreasonable in any situation when $n_k = 0$. Thus, to avoid such a requirement, we allow for the computation of an alternative tangential step. Given the constant $\kappa_{\mathrm{fbt}} \in (0,1)$ employed in (2.30b), a constant $\kappa_{\mathrm{v}} \in (1, \infty)$, and a constant $\kappa_{\mathrm{tt}} \in (\kappa_{\mathrm{vv}}, 1)$ (with $\kappa_{\mathrm{vv}} \in (0,1)$ defined for (2.10)), the salient feature of our alternative is that it involves a relaxed condition on the linearized infeasibility of the step. We emphasize that we are only allowed to compute a tangential step of this type when $n_k = 0$, though we incorporate n_k into the conditions in the following definition so that one may more easily compare them to the conditions in Definition 2.1.

Definition 2.2 (Very relaxed SQP tangential step)**.** *Define the Cauchy point*

$$t_k^{\mathrm{C}} = t_k^{\mathrm{C}}(\alpha_{\mathrm{T}}^{\mathrm{C}}), \quad \textit{where} \quad t_k^{\mathrm{C}}(\alpha) := \begin{pmatrix} t_k^{\mathrm{C}x}(\alpha) \\ t_k^{\mathrm{C}s}(\alpha) \end{pmatrix} := -\alpha \begin{pmatrix} r_k^x \\ r_k^s \end{pmatrix} = -\alpha r_k \tag{2.33}$$

and $\alpha_{\mathrm{T}}^{\mathrm{C}}$ *is the minimizer of*

$$\begin{array}{ll}\underset{\alpha\geq 0}{\text{minimize}} & m_k^f\big(n_k + t_k^{\mathrm{C}}(\alpha)\big) \\ \text{subject to} & \|P_k^{-1}\big(n_k + t_k^{\mathrm{C}}(\alpha)\big)\|_2 \leq \min\{\delta_k^v, \delta_k^f, \kappa_{\mathrm{v}} v_k^{\max}\}, \quad s_k + n_k^s + t_k^{\mathrm{C}s}(\alpha) \geq \kappa_{\mathrm{fbt}}(s_k + n_k^s).\end{array} \tag{2.34}$$

Then, t_k is a very relaxed SQP tangential step if

$$\Delta m_k^{f,t} := m_k^f(n_k) - m_k^f(n_k + t_k) \geq m_k^f(n_k) - m_k^f(n_k + t_k^{\mathrm{C}}), \tag{2.35a}$$

$$s_k + n_k^s + t_k^s \geq \kappa_{\mathrm{fbt}}(s_k + n_k^s), \tag{2.35b}$$

$$\|P_k^{-1}(n_k + t_k)\|_2 \leq \min\{\delta_k^v, \delta_k^f, \kappa_{\mathrm{v}} v_k^{\max}\}, \quad \textit{and} \tag{2.35c}$$

$$m_k^v(n_k + t_k) \leq \kappa_{\mathrm{tt}} v_k^{\max}. \tag{2.35d}$$

Conditions (2.35a)–(2.35c) resemble and play the same role as conditions (2.30a)–(2.30c). However, we emphasize that since the Cauchy point defined by (2.33)–(2.34) involves a potentially smaller trust-region radius than that defined in (2.29), the bound imposed in (2.35a) may be different from that imposed in (2.30a), and this difference in the trust-region radii is matched in (2.35c) (c.f., (2.30c)). The name "very relaxed SQP tangential step" has been chosen because of condition (2.35d), which merely requires that the predicted constraint violation be sufficiently less than a fraction of the upper bound $v_k^{\max}$ rather than a fraction of the current violation (c.f., (2.30d)). In fact, the smaller trust-region radii in (2.34) and (2.35c) (as compared to those in (2.29) and (2.30c)) have been chosen to compensate for this relaxation.

2.4 Iteration type, step acceptance, and updating strategy

As in other trust-region methods, once we have computed the trial step $d_k := n_k + t_k$ and the trial point

$$(x_k^+, s_k^+) := (x_k, s_k) + d_k,$$

we are left with the task of accepting or rejecting them. Our proposal for making this choice is based on the distinction between *y-iterations*, *f-iterations* and *v-iterations* in the spirit of [9–11]. This characterization is made based on model values computed with the trial step, and the type of iteration influences the updates performed for various algorithmic quantities.

2.4.1 A *y*-iteration

A *y*-iteration is any iteration satisfying the following definition.

Definition 2.3 (*y*-iteration)**.** *The kth iteration is a y-iteration if $d_k = 0$.*

Note that a *y*-iteration will occur when n_k and t_k are both set to zero, but could (in theory) occur if $n_k = -t_k$ and some components are nonzero. (This latter case is ruled out by Lemma 2.6(vi).) During a *y*-iteration, we perform the updates

$$(x_{k+1}, s_{k+1}) \leftarrow (x_k, s_k), \quad \delta_{k+1}^f \leftarrow \delta_k^f, \quad \delta_{k+1}^v \leftarrow \delta_k^v, \quad \text{and} \quad v_{k+1}^{\max} \leftarrow v_k^{\max}. \tag{2.36}$$

Since a *y*-iteration is defined by a zero primal step, the only computation of interest is that of a new vector of Lagrange multiplier estimates. Therefore, the updates in (2.36) leave the trust-region radii and bound on the maximum allowed infeasibility unchanged for the subsequent iteration.

2.4.2 An *f*-iteration

The primary goal of an *f*-iteration is to reduce the barrier function.

Definition 2.4 (*f*-iteration)**.** *The kth iteration is an f-iteration if $t_k \neq 0$,* (2.32) *holds, and*

$$v(x_k^+, s_k^+) \leq v_k^{\max}. \tag{2.37}$$

Condition (2.37) ensures that, at the trial point (x_k^+, s_k^+), the constraint violation remains within the upper bound imposed by $v_k^{\max}$. Combining this with the fact that (2.32) holds, it follows that the main achievement of interest is a predicted decrease in the value of the barrier function (1.2).

Our updating strategy for f-iterations is based on the quantity

$$\rho_k^f := \frac{f(x_k, s_k) - f(x_k^+, s_k^+)}{\Delta m_k^{f,d}}, \tag{2.38}$$

that measures the ratio of actual-to-predicted decrease in the barrier function. Specifically, if $\rho_k^f \geq \eta_1$, we set

$$(x_{k+1}, s_{k+1}) \leftarrow (x_k^+, s_k^+) \tag{2.39}$$

$$[s_{k+1}]_i \leftarrow \begin{cases} [s_{k+1}]_i & \text{if } [c(x_{k+1}, s_{k+1})]_i \geq 0, \\ -[c(x_{k+1})]_i & \text{otherwise,} \end{cases} \tag{2.40}$$

$$\delta_{k+1}^f \begin{cases} \in [\delta_k^f, \infty) & \text{if } \rho_k^f \geq \eta_2, \\ \in [\gamma_2 \delta_k^f, \delta_k^f] & \text{otherwise,} \end{cases} \tag{2.41}$$

$$\delta_{k+1}^v \geq \max\{\kappa_{\delta vv} \pi^v(x_{k+1}, s_{k+1}), \delta_k^v\} \tag{2.42}$$

Otherwise (i.e., if $\rho_k^f < \eta_1$), we set

$$(x_{k+1}, s_{k+1}) \leftarrow (x_k, s_k), \tag{2.43}$$

$$\delta_{k+1}^f \in [\gamma_1 \delta_k^f, \gamma_2 \delta_k^f], \tag{2.44}$$

$$\delta_{k+1}^v \leftarrow \delta_k^v. \tag{2.45}$$

In both cases, we set

$$v_{k+1}^{\max} \leftarrow v_k^{\max}. \tag{2.46}$$

In (2.39)–(2.46), the constants should be chosen to satisfy $0 < \eta_1 \leq \eta_2 < 1$, $0 < \gamma_1 \leq \gamma_2 < 1$, and $\kappa_{\delta vv} \in (0, \infty)$. Overall, we accept the trial point (x_k^+, s_k^+) if the achieved decrease in the barrier function is comparable to the predicted decrease (and reject it otherwise), update δ_{k+1}^f using a typical trust-region updating strategy, possibly increase the normal step trust-region radius, and leave the infeasibility limit unchanged (since the success or failure of an f-iteration depends only on whether the barrier function was substantially reduced).

2.4.3 A v-iteration

When the conditions that define a y- and an f-iteration are not satisfied, the iteration type defaults to that of a v-iteration. As we shall see in the convergence analysis of our algorithm, the main achievement of interest in such an iteration is a reduction in constraint violation.

Definition 2.5 (v-iteration)**.** *The kth iteration is a v-iteration if it is not a y- or an f-iteration, i.e., if* $d_k \neq 0$ *and either* $t_k = 0$, (2.32) *does not hold, or* (2.37) *does not hold.*

A measure of decrease one might expect in v from the trial step d_k is

$$\Delta m_k^{v,d} := m_k^v(0) - m_k^v(d_k). \tag{2.47}$$

Indeed, our updating strategy in a v-iteration is based on the conditions

$$n_k \neq 0 \quad \text{and} \quad \Delta m_k^{v,d} \geq \kappa_{\text{cd}} \Delta m_k^{v,n} \tag{2.48}$$

for some constant $\kappa_{\text{cd}} \in (0, 1 - \kappa_{\text{tg}}] \subset (0, 1)$ with κ_{tg} defined in (2.30d), and the quantity

$$\rho_k^v := \frac{v_k - v(x_k^+, s_k^+)}{\Delta m_k^{v,d}} \tag{2.49}$$

that measures the ratio of actual-to-predicted decrease in the constraint violation. Using these conditions and quantities, if (2.48) holds and $\rho_k^v \geq \eta_1$, we set

$$(x_{k+1}, s_{k+1}) \leftarrow (x_k^+, s_k^+) \tag{2.50}$$

$$[s_{k+1}]_i \leftarrow \begin{cases} [s_{k+1}]_i & \text{if } [c(x_{k+1}, s_{k+1})]_i \geq 0, \\ -[c(x_{k+1})]_i & \text{otherwise,} \end{cases} \tag{2.51}$$

$$\delta_{k+1}^v \begin{cases} \geq \max\{\kappa_{\delta vv}\pi^v(x_{k+1}, s_{k+1}), \delta_k^v\} & \text{if } \rho_k^v \geq \eta_2, \\ = \max\{\kappa_{\delta vv}\pi^v(x_{k+1}, s_{k+1}), \delta_k^v\} & \text{otherwise,} \end{cases} \tag{2.52}$$

$$v_{k+1}^{\max} \leftarrow \max\{\kappa_{\text{t1}} v_k^{\max}, v(x_{k+1}, s_{k+1}) + \kappa_{\text{t2}}\big(v_k - v(x_{k+1}, s_{k+1})\big)\}. \tag{2.53}$$

Otherwise (i.e., if (2.48) does not hold or $\rho_k^v < \eta_1$), we set

$$(x_{k+1}, s_{k+1}) \leftarrow (x_k, s_k), \tag{2.54}$$

$$\delta_{k+1}^v \in [\gamma_1 \delta_k^v, \gamma_2 \delta_k^v], \tag{2.55}$$

$$v_{k+1}^{\max} \leftarrow v_k^{\max}. \tag{2.56}$$

In both cases, we set

$$\delta_{k+1}^f \leftarrow \delta_k^f. \tag{2.57}$$

In (2.50)–(2.57), the constants should be chosen to satisfy $\{\kappa_{\text{t1}}, \kappa_{\text{t2}}\} \subset (0, 1)$, and we recall that $\kappa_{\delta vv}$ is defined in (2.42). In this manner, the trial point is accepted if the normal step is nonzero and the improvement in linearized feasibility is comparable to its predicted value, which is itself comparable to the improvement yielded by the normal step. Moreover, the radius δ_{k+1}^v is set by a standard trust-region radius updating strategy, but the radius δ_{k+1}^f is left unchanged. Finally, we decrease the upper bound $v_k^{\max}$ when the trial step is accepted. It will be shown in our convergence analysis that the amount that this bound is decreased is nontrivial, but it is modest enough so that the funnel does not contract too quickly.

2.5 The trust-funnel algorithm

We formally state our trust-funnel method as Algorithm 1 on page 43. For convenience in our convergence analysis, we define several sets that classify each iteration, as well as the types of computations performed in them. The first group of sets distinguishes between y-, f-, and v-iterations, respectively:

$$\mathcal{Y} := \{k \in \mathbb{N} : d_k = 0\}, \quad \mathcal{F} := \{k \in \mathbb{N} : t_k \neq 0 \text{ and (2.32) and (2.37) hold}\}, \quad \text{and} \quad \mathcal{V} := \mathbb{N} \setminus (\mathcal{Y} \cup \mathcal{F}).$$

As can be seen by the results in Lemma 2.6 below, these sets are mutually exclusive and exhaustive.

Our next collection of index sets distinguishes iterations for which the normal and/or tangential

steps satisfy various conditions, and whether the tangential step was reset to zero:

$$\begin{aligned}
\mathcal{N} &:= \{k \in \mathbb{N} : n_k \text{ was computed to satisfy (2.13) and (2.14)}\};\\
\mathcal{T} &:= \{k \in \mathbb{N} : t_k \text{ was computed to satisfy either (2.30) or (2.35)}\};\\
\mathcal{T}_{\mathcal{D}} &:= \{k \in \mathcal{T} : \text{the computed } t_k \text{ satisfied (2.30)}\};\\
\mathcal{T}_0 &:= \{k \in \mathcal{T}_{\mathcal{D}} : \text{the computed } t_k \text{ satisfied (2.30) and (2.31), but not (2.32), and was reset to zero}\}.
\end{aligned}$$

Furthermore the set of iterations for which d_k satisfies the linearized constraint contraction condition (2.30d) plays an important role in our analysis. Thus, in addition to the sets above, we define

$$\mathcal{D} := \{k \in \mathbb{N} : \text{the step } d_k = n_k + t_k \text{ satisfies (2.30d)}\}.$$

Our last collection of sets distinguishes iterations that produce a change in the primal space. In particular, if $\rho_k^f \geq \eta_1$ holds during an f-iteration, or if (2.48) holds and $\rho_k^v \geq \eta_1$ during a v-iteration, then iteration k is called *successful*. The following sets capture these types of iterations:

$$\mathcal{S}_f := \{k \in \mathcal{F} : \rho_k^f \geq \eta_1\}; \quad \mathcal{S}_v := \{k \in \mathcal{V} : (2.48) \text{ holds and } \rho_k^v \geq \eta_1\}; \quad \mathcal{S} := \mathcal{S}_f \cup \mathcal{S}_v.$$

When a tangential step is computed, the size of the step is restricted by a trust-region radius (see (2.30c) and (2.35c)). For convenience, we capture these radii by defining $\delta_{-1}^t := 1$ and, for $k \geq 0$,

$$\delta_k^t := \begin{cases} \delta_{k-1}^t & \text{if } k \notin \mathcal{T}, \\ \min\{\delta_k^v, \delta_k^f\} & \text{if } k \in \mathcal{T} \cap \mathcal{T}_{\mathcal{D}}, \\ \min\{\delta_k^v, \delta_k^f, \kappa_{\text{v}} v_k^{\max}\} & \text{if } k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}. \end{cases} \tag{2.58}$$

As a guide to the salient properties of the various types of iterations we have defined, we provide the following lemma regarding basic facts that may be deduced from the design of our algorithm. Unless stated otherwise, reference to the tangential step t_k corresponds to the value used in Step 35 of Algorithm 1, i.e., the value after the possible reset in Step 28. For the purposes of this lemma, we assume that if the algorithm does not terminate during iteration k, then all steps of the algorithm during the iteration are well-defined. We prove this fact formally in the next subsection.

Lemma 2.6. *If Algorithm* 1 *does not terminate during the kth iteration, then the following hold.*

(i) *If* $k \in \mathcal{N}$, *then* $\chi_k^v > 0$, $\pi_k^v > 0$, $m_k^v(0) - m_k^v(n_k^{\text{C}}) > 0$, $m_k^v(0) - m_k^v(n_k) > 0$, *and* $n_k \neq 0$.

(ii) *If* $n_k \neq 0$, *then* $k \in \mathcal{N}$.

(iii) *If* $k \in \mathcal{T}$, *then* $\chi_k^f \geq \kappa_\chi \pi_k^f > 0$ *and* $m_k^f(n_k) - m_k^f(n_k + t_k^{\text{C}}) > 0$.

(iv) *If* $k \in \mathcal{T} \setminus \mathcal{T}_0$, *then* $t_k \neq 0$ *and* $m_k^f(n_k) - m_k^f(n_k + t_k) > 0$, *while if* $k \in \mathcal{T}_0$, *then* $t_k = 0$ *and* (2.21) *holds.*

(v) *If* $t_k \neq 0$, *then* $k \in \mathcal{T} \setminus \mathcal{T}_0$.

(vi) $k \in \mathcal{Y}$ *if and only if* $n_k = t_k = 0$.

(vii) *If* $k \in \mathcal{Y}$, *then* $k \in \mathcal{D}$ *and* $\pi_k^f \leq \kappa_\omega \pi_{k-1}^f$ *with* $\kappa_\omega \in (0, 1)$ *defined as in* (2.27).

(viii) *If* $k \notin \mathcal{D}$, *then* $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$ *and* (2.35) *holds.*

(ix) *If* $k \in \mathcal{D}$, *then the inequality in* (2.48) *holds.*

(x) $\mathcal{T}_{\mathcal{D}} \subseteq \mathcal{D}$.

(xi) *If* $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$, *then* $n_k = 0$ *and* $k \notin \mathcal{N}$.

Proof. To prove part (i), let $k \in \mathcal{N}$, in which case we have that the conditions in Step 9 held true. This could occur only if $\pi_k^v > 0$, or if in (2.10) we had $\pi_k^v > \omega_n(\pi_{k-1}^f) \geq 0$ or $v_k \geq \kappa_{\text{vv}} v_k^{\max}$. Thus, to prove that $k \in \mathcal{N}$ implies $\pi_k^v > 0$, all that remains is to investigate the case when $v_k \geq \kappa_{\text{vv}} v_k^{\max}$. Since $v_k^{\max} > 0$ by construction, this inequality implies $v_k > 0$. If $\pi_k^v = 0$ (which, since $v_k > 0$, implies $\chi_k^v = 0$), then the algorithm would have terminated in Step 8 with an infeasible stationary point. Thus, we may again conclude that $\pi_k^v > 0$, which establishes this strict inequality for all $k \in \mathcal{N}$. In turn, by (2.9) and the fact that $v_k > 0$ when $\pi_k^v > 0$, we must have $\chi_k^v > 0$ for all $k \in \mathcal{N}$. Now, since $\pi_k^v > 0$, it follows that $-P_k^2 J(x_k, s_k)^T c(x_k, s_k)$ is a direction of strict decrease for m_k^v at $n = 0$, from which it follows by (2.11) that $m_k^v(0) - m_k^v(n_k^{\text{C}}) > 0$. In turn, (2.14) implies the remainder of part (i).

Part (ii) follows since if $n_k \neq 0$, then the conditions in Step 9 must have held (or else the algorithm would have set $n_k \leftarrow 0$), in which case $k \in \mathcal{N}$.

Next, we prove part (iii). If $k \in \mathcal{T}$, then it follows from Steps 15–24 of the algorithm that after the computation of y_k (and all dependent quantities) both (2.26a) and (2.26b) did not hold (implying that $\pi_k^f > 0$), but (2.26c) did. Combining (2.26c) and the fact that $\pi_k^f > 0$ yields $\nabla m_k^f(n_k)^T r_k \geq \kappa_\chi(\pi_k^f)^2 > 0$ (as desired), which implies that r_k is a direction of strict ascent for m_k^f at n_k. Combining this fact with (2.28)/(2.33) and (2.29)/(2.34) yields $m_k^f(n_k) - m_k^f(n_k + t_k^{\text{C}}) > 0$, as desired.

Building on the proof of part (iii), we next prove part (iv). If we have $k \in \mathcal{T} \backslash \mathcal{T}_0$, then we may combine $m_k^f(n_k) - m_k^f(n_k + t_k^{\text{C}}) > 0$ with (2.30a)/(2.35a) to conclude that $t_k \neq 0$ and $m_k^f(n_k) - m_k^f(n_k + t_k) > 0$, as desired. (Since $k \notin \mathcal{T}_0$, this tangential step was not reset to zero, so we have maintained $t_k \neq 0$ in Step 35.) Finally, if $k \in \mathcal{T}_0$, it follows from Steps 14–28 that (2.21) holds, but that the algorithm reset $t_k \leftarrow 0$.

To prove part (v), we first note that if $t_k \neq 0$, then a tangential step was computed and thus $k \in \mathcal{T}$. Moreover, since $t_k \neq 0$, we know that $k \notin \mathcal{T}_0$, which means $k \in \mathcal{T} \setminus \mathcal{T}_0$, as desired.

We now prove part (vi). If $n_k = t_k = 0$, then $d_k = 0$ and we have $k \in \mathcal{Y}$ by the definition of $\mathcal{Y}$; this proves one direction. For the other direction, in order to derive a contradiction, suppose that $k \in \mathcal{Y}$ (so that $d_k = n_k + t_k = 0$), but that $n_k \neq 0$ and/or $t_k \neq 0$. Indeed, since $n_k + t_k = 0$, we must have $n_k \neq 0$ and $t_k \neq 0$. It then follows from parts (ii) and (v) that $k \in \mathcal{Y} \cap \mathcal{N} \cap (\mathcal{T} \setminus \mathcal{T}_0)$. Consequently, from part (i) we have that $m_k^v(0) > m_k^v(n_k)$. This fact and the equation $n_k + t_k = 0$ imply that (2.30d) must not be satisfied. However, according to Steps 21–22 of the algorithm, since $k \in \mathcal{N}$ we compute t_k to satisfy (2.30), a contradiction.

To prove part (vii), suppose $k \in \mathcal{Y}$. It follows from part (vi) that $n_k = t_k = 0$ so that (2.30d) holds (which means $k \in \mathcal{D}$, as desired), and then from part (i) that $k \notin \mathcal{N}$. Hence, from Step 9 of the algorithm, it follows that (2.10) must be violated. Moreover, since $n_k = 0$, we also know that (2.21) holds and thus an oblique projected gradient r_k was computed (as stipulated in Step 15) to satisfy at least one of (2.26a), (2.26b) and (2.26c). In fact, under the conditions of this lemma, it follows that (2.26a) must not have held, so we know that either (2.26b) or (2.26c) is satisfied as a result of this calculation. Suppose that (2.26c) holds so that the algorithm would have proceeded to compute a tangential step and $k \in \mathcal{T}$. If $k \notin \mathcal{T}_0$, then it would follow from part (iv) that $t_k \neq 0$, which by part (vi) contradicts the fact that $k \in \mathcal{Y}$. Thus, we must have $k \in \mathcal{T}_0$, i.e., we reset $t_k \leftarrow 0$ because the computed tangential step satisfied (2.31), but not (2.32). This is a contradiction because (2.32) would have been satisfied trivially since $n_k = 0$. Thus (2.26c) must not hold, which implies that (2.26b) must hold. Since we have shown that (2.26b) holds and (2.10) does not hold, we conclude that $\pi_k^f \leq \omega_t(\pi_k^v) \leq \omega_t(\omega_n(\pi_{k-1}^f)) \leq \kappa_\omega \pi_{k-1}^f$, where we have used the monotonicity of ω_t and (2.27).

To establish part (viii), let $k \notin \mathcal{D}$. It follows from part (vii) that $k \notin \mathcal{Y}$. Now, suppose that $t_k = 0$. Combining this with the fact that $k \notin \mathcal{Y}$ implies from part (vi) that $n_k \neq 0$, which may then be combined with part (ii) to deduce that $k \in \mathcal{N}$. This fact along with part (i) and the fact that $t_k = 0$ implies that $m_k^v(n_k + t_k) \leq \kappa_{\text{tg}} m_k^v(0) + (1 - \kappa_{\text{tg}}) m_k^v(n_k)$ (c.f., (2.30d)), and hence $k \in \mathcal{D}$, which is a contradiction. Therefore, we must have $t_k \neq 0$, which from part (v) implies that $k \in \mathcal{T} \setminus \mathcal{T}_0$ and that the computed tangential step was not reset to zero. Thus, t_k satisfies either (2.30) or (2.35). In fact, since $k \notin \mathcal{D}$ so that (2.30d) is not satisfied, we conclude that $k \notin \mathcal{T}_{\mathcal{D}}$ and (2.35) must be satisfied.

To prove part (ix), suppose $k \in \mathcal{D}$ so that (2.30d) holds. It follows that

$$\begin{aligned}\Delta m_k^{v,d} &= m_k^v(0) - m_k^v(d_k) \\ &\geq m_k^v(0) - \kappa_{\text{tg}} m_k^v(0) - (1-\kappa_{\text{tg}}) m_k^v(n_k) \\ &= (1-\kappa_{\text{tg}})\big(m_k^v(0) - m_k^v(n_k)\big) = (1-\kappa_{\text{tg}})\Delta m_k^{v,n},\end{aligned} \tag{2.59}$$

which, since $\kappa_{\text{cd}} \in (0, 1-\kappa_{\text{tg}}]$, means that the inequality in (2.48) holds, as desired.

To prove (x), let $k \in \mathcal{T}_{\mathcal{D}}$. It follows that a relaxed SQP tangential step t_k was computed to satisfy (2.30). Thus, if t_k is not reset to zero, we know that (2.30d) holds. However, if t_k was reset to zero, then (2.30d) holds trivially when $n_k = 0$ and from parts (i) and (ii) when $n_k \neq 0$. We have shown in all cases that (2.30d) holds, and therefore $k \in \mathcal{D}$.

Finally, to prove part (xi), let $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$. By Steps 21–28 of the algorithm, it follows that (2.35) holds and $k \notin \mathcal{N}$ for all $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$. It then follows from part (ii) that $n_k = 0$. □

2.6 Well-posedness

The purpose of this section is to prove that Algorithm 1 is well-posed in the sense that if iteration k is reached, then in a reasonable implementation of the algorithm, all computations within iteration k will terminate finitely. Our first result shows important consequences of the slack reset procedure.

Lemma 2.7. *The slack reset* (2.40) *and* (2.51) *in Steps 39 and 41 yields s_k such that (x_k, s_k) satisfies $s_k > 0$ and $c(x_k, s_k) \geq 0$.*

Proof. The fact that $s_k > 0$ follows from the choice $s_0 > 0$, the fact that the slack reset (2.40) and (2.51) only possibly increases the slack variables (as shown in (2.6)), and the fact that the fraction-to-the-boundary rules in (2.13) and (2.30b)/(2.35b) hold when normal and tangential steps are computed.

We now prove that $c(x_k, s_k) \geq 0$ holds. Prior to the slack reset performed in Steps 39 and 41, if $[c(x_k, s_k)]_i \geq 0$, then (2.40) and (2.51) leave $[s_k]_i$ unchanged so that $[c(x_k, s_k)]_i \geq 0$ still holds. Otherwise, if $[c(x_k, s_k)]_i < 0$, then after the slack reset (2.40) and (2.51) we have that $[c(x_k) + s_k]_i = 0$, which completes the proof. □

Next, we prove that the Cauchy step for the normal step subproblem is well-defined.

Lemma 2.8. *If $k \in \mathcal{N}$, then the Cauchy step n_k^{C} defined by* (2.11)–(2.12) *is computed and satisfies*

$$m_k^v(0) - m_k^v(n_k^{\text{C}}) \geq \kappa_k^{\text{cn}} \chi_k^v \min\{\pi_k^v, \delta_k^v, 1-\kappa_{\text{fbn}}\} > 0, \tag{2.60}$$

where

$$\kappa_k^{\text{cn}} := \frac{1}{\left(1 + \|J(x_k, s_k)P_k\|_2^2\right)} \in (0, 1].$$

Proof. Since $k \in \mathcal{N}$, we may observe from part (i) of Lemma 2.6 that $\pi_k^v > 0$ and $\chi_k^v > 0$, and hence $v_k > 0$. We now show that $n_k^{\text{C}}(\alpha)$ (recall (2.11)) is feasible for (2.12) during any iteration $k \in \mathcal{N}$ when

$$0 \leq \alpha \leq \frac{1}{\pi_k^v} \min\{\delta_k^v, \kappa_{\text{n}}\pi_k^v, (1-\kappa_{\text{fbn}})\} =: \alpha_{\text{B}}. \tag{2.61}$$

Indeed, consider any $\alpha \in [0, \alpha_{\text{B}}]$. It then follows from the definitions of $n_k^{\text{C}}(\alpha)$ and π_k^v that

$$\|P_k^{-1} n_k^{\text{C}}(\alpha)\|_2 = \|\alpha P_k J(x_k, s_k)^T c(x_k, s_k)\|_2 = \alpha\pi_k^v \leq \min\{\delta_k^v, \kappa_{\text{n}}\pi_k^v\}.$$

It also follows from the definitions of $n_k^{\text{C}s}(\alpha)$ and Lemma 2.7 that

$$\begin{aligned}[-n_k^{\text{C}s}(\alpha)]_i &= \alpha[S_k]_{ii}^2 [c(x_k, s_k)]_i \leq \alpha[s_k]_i \|P_k J(x_k, s_k)^T c(x_k, s_k)\|_2 \\ &= \alpha\pi_k^v [s_k]_i \leq (1-\kappa_{\text{fbn}})[s_k]_i \quad \text{for } i = 1, 2, \ldots M,\end{aligned}$$

which implies that $s_k + n_k^{\mathrm{C}s}(\alpha) \geq \kappa_{\mathrm{fbn}} s_k$. Overall, $n_k^{\mathrm{C}}(\alpha)$ is feasible for problem (2.7) for all $\alpha \in [0, \alpha_{\mathrm{B}}]$.

Now, observe that the minimizer $\alpha_{\mathrm{N}}^{\mathrm{C}}$ defined by (2.12) yields $m_k^v(n_k^{\mathrm{C}}) = m_k^v(n_k^{\mathrm{C}}(\alpha_{\mathrm{N}}^{\mathrm{C}})) \leq m_k^v(n_k^{\mathrm{C}}(\alpha))$ for all $\alpha \in [0, \alpha_{\mathrm{B}}]$. It then follows from [3, Lemma 1] with the quantities

$$\text{“}t\text{”} := \alpha_{\mathrm{B}}, \quad \text{“}a\text{”} := 2\|J(x_k, s_k)P_k^2 J(x_k, s_k)^T c(x_k, s_k)\|_2^2, \quad \text{“}b\text{”} := 2(\pi_k^v)^2 > 0,$$

the fact that

$$\text{“}a\text{”} \leq 2\|J(x_k, s_k)P_k\|_2^2 \|P_k J(x_k, s_k)^T c(x_k, s_k)\|_2^2 = 2\|J(x_k, s_k)P_k\|_2^2 (\pi_k^v)^2$$

and the definition of π_k^v that

$$\begin{aligned}
\big(m_k^v(0)\big)^2 - \big(m_k^v(n_k^{\mathrm{C}})\big)^2 &\geq \text{“}b\text{”} \min\left\{ \frac{\text{“}b\text{”}}{\text{“}a\text{”}}, \text{“}t\text{”} \right\} \\
&\geq 2(\pi_k^v)^2 \min\left\{ \frac{1}{\|J(x_k, s_k)P_k\|_2^2}, \frac{\delta_k^v}{\pi_k^v}, \kappa_{\mathrm{n}}, \frac{1-\kappa_{\mathrm{fbn}}}{\pi_k^v} \right\} \\
&= 2\pi_k^v \min\left\{ \frac{\pi_k^v}{1+\|J(x_k, s_k)P_k\|_2^2}, \delta_k^v, \kappa_{\mathrm{n}}\pi_k^v, 1-\kappa_{\mathrm{fbn}} \right\} \\
&= 2v_k\chi_k^v \min\left\{ \frac{\pi_k^v}{1+\|J(x_k, s_k)P_k\|_2^2}, \delta_k^v, 1-\kappa_{\mathrm{fbn}} \right\} > 0, \qquad (2.62)
\end{aligned}$$

where we have used the facts that $1 + \|J(x_k, s_k)P_k\|_2^2 \geq 1$ and $\kappa_{\mathrm{n}} \geq 1$ to derive the last equality. Hence, $m_k^v(n_k^{\mathrm{C}}) < m_k^v(0)$, and therefore

$$m_k^v(0) - m_k^v(n_k^{\mathrm{C}}) = \frac{(m_k^v(0))^2 - (m_k^v(n_k^{\mathrm{C}}))^2}{m_k^v(0) + m_k^v(n_k^{\mathrm{C}})} \geq \frac{(m_k^v(0))^2 - (m_k^v(n_k^{\mathrm{C}}))^2}{2m_k^v(0)} = \frac{(m_k^v(0))^2 - (m_k^v(n_k^{\mathrm{C}}))^2}{2v_k}.$$

The result follows from this inequality, (2.62), and the fact that $1 + \|J(x_k, s_k)P_k\|_2^2 \geq 1$. □

Next we establish the remaining claims made in (2.4). (We remark that certain bounds established in the proof of this lemma are specified in more detail in Lemma 3.12.)

Lemma 2.9. *The slack reset* (2.40) *and* (2.51) *in Steps 39 and 41 yields s_k such that the pair (x_k, s_k) satisfies $v_k \leq v_k^{\max}$ and, at the end of iteration k, we have $v_{k+1}^{\max} \leq v_k^{\max}$.*

Proof. Our proof is by induction. We have $v_0 \leq v_0^{\max}$ by the initialization of $v_0^{\max}$. Now suppose that $v_i \leq v_i^{\max}$ for $i = 0, \ldots, k-1$ for $k \geq 1$, and in particular that $v_{k-1} \leq v_{k-1}^{\max}$ at the start of the $(k-1)$-st iteration. The slack reset in Steps 39 and 41 cannot increase the constraint violation (recall (2.6)), so $v_{k-1} \leq v_{k-1}^{\max}$ holds following the slack reset. It is also clear from (2.36) and (2.37) that for $k-1 \in \mathcal{Y} \cup \mathcal{F}$ the inequality $v_k \leq v_k^{\max}$ continues to hold at the start of iteration k. Hence, it remains to consider $k-1 \in \mathcal{V}$. If $\rho_{k-1}^v < \eta_1$ or (2.48) (with k replaced by $k-1$) does not hold, then the step is rejected, so $v_k \leq v_k^{\max}$ holds at the start of iteration k as a consequence of (2.54) and (2.56). Otherwise, it follows from Lemma 2.8, (2.14) and (2.48) that $\Delta m_{k-1}^{v,d} > 0$ and thus $v_k < v_{k-1}$ from $\rho_{k-1}^v \geq \eta_1$, (2.49), and (2.50). Since $\kappa_{\mathrm{t2}} \in (0, 1)$ in (2.53), this implies

$$v_k < v_k + \kappa_{\mathrm{t2}}\big(v_{k-1} - v_k\big) < v_{k-1} \leq v_{k-1}^{\max} \qquad (2.63)$$

and hence from (2.53) we have $v_k^{\max} \leq v_{k-1}^{\max}$. Combining (2.53) and (2.63), we have that $v_k^{\max} \geq v_k + \kappa_{\mathrm{t2}}\big(v_{k-1} - v_k\big) > v_k$. Thus, in all cases, we have $v_k \leq v_k^{\max}$; the induction is complete.

To establish that $v_{k+1}^{\max} \leq v_k^{\max}$, note that if $k \notin \mathcal{V}$, then $v_{k+1}^{\max} \leftarrow v_k^{\max}$, so all that remains is to consider $k \in \mathcal{V}$. Observing (2.53), we see again that $v_{k+1}^{\max} \leftarrow v_k^{\max}$ if either (2.48) is violated or $\rho_k^v < \eta_1$. By contrast, if (2.48) holds and $\rho_k^v \geq \eta_1$, then we must have $n_k \neq 0$ and from part (ii) of Lemma 2.6 that $k \in \mathcal{N}$. Moreover, it follows from (2.50), (2.49), (2.48), (2.14) and Lemma 2.8 as above that $v_{k+1} < v_k$.

Thus, if the maximum value in (2.53) is the second term, it follows that $v_{k+1}^{\max} < v_k \leq v_k^{\max}$. Otherwise, if the maximum value in (2.53) is the first term, then $v_{k+1}^{\max} < v_k^{\max}$ trivially follows since $\kappa_{\text{t1}} \in (0,1)$. □

Next, we show that the computation of the least-squares multipliers y_k—along with the accompanying quantities r_k, π_k^f, and χ_k^f—is well-defined. We prove this result under the following reasonable assumption.

Assumption 2.1. *If* (2.21) *holds and the iterative solver employed to solve* (2.23) *is allowed to run for an infinite number of iterations, then it produces a bounded sequence* $\{y^{(i)}\}$ *with* $y^{(0)} = 0$ *such that*

$$\lim_{i \to \infty} \nabla m_k^{\mathcal{L}}(y^{(i)}) = 0. \tag{2.64}$$

Lemma 2.10. *If* (2.21) *holds and* $\{y^{(i)}\}$ *is produced by an iterative solver employed to solve* (2.23) *that satisfies Assumption 2.1, then for some (finite) index* i *the vector* $y_k \leftarrow y^{(i)}$ *yields* r_k, π_k^f, *and* χ_k^f *satisfying* (2.26a), (2.26b), *or* (2.26c).

Proof. For the purpose of deriving a contradiction, suppose that the iterative solver applied to solve (2.23) runs for an infinite number of iterations without satisfying (2.26a), (2.26b), or (2.26c). Under Assumption 2.1, the sequence $\{y^{(i)}\}$ is bounded, so with (2.64) we have that it has a limit point y^∞ satisfying

$$0 = \nabla m_k^{\mathcal{L}}(y^\infty) = J(x_k, s_k) r_k(y^\infty). \tag{2.65}$$

Suppose $\pi_k^f(y^\infty) = 0$. If $v_k \leq \epsilon_v$, then this implies that there exists some smallest index i such that with $y_k \leftarrow y^{(i)}$ condition (2.26a) will be satisfied, which is a contradiction. Otherwise, if $v_k > \epsilon_v$, then $\chi_k^v > 0$ or else Algorithm 1 would have terminated in Step 8. Since this implies that $\pi_k^v > 0$ (recall $v_k > \epsilon_v$), it follows from $\pi_k^f(y^\infty) = 0$ that there exists a smallest index i such that with $y_k \leftarrow y^{(i)}$ condition (2.26b) will be satisfied, which is a contradiction. We have shown that $\pi_k^f(y^\infty) > 0$, which combined with (2.65) and

$$\nabla m_k^f(n_k) = P_k^{-2} r_k(y^\infty) - J(x_k, s_k)^T y^\infty$$

shows that

$$\chi_k^f(y^\infty) = \frac{r_k(y^\infty)^T \nabla m_k^f(n_k)}{\pi_k^f(y^\infty)} = \frac{r_k(y^\infty)^T (P_k^{-2} r_k(y^\infty) - J(x_k, s_k)^T y^\infty)}{\pi_k^f(y^\infty)} = \frac{(\pi_k^f(y^\infty))^2}{\pi_k^f(y^\infty)} = \pi_k^f(y^\infty). \tag{2.66}$$

If $\nabla m_k^{\mathcal{L}}(0) = 0$, then we have with (2.66) that $y_k \leftarrow y^{(0)} = 0$ satisfies (2.26c), which is a contradiction. By contrast, if $\nabla m_k^{\mathcal{L}}(0) \neq 0$, then since $\kappa_\chi \in (0,1)$ we have from (2.66) and (2.64) that there is a smallest index i for which $y_k \leftarrow y^{(i)}$ satisfies condition (2.26c), which is another contradiction.

We have arrived at a contradiction in all cases, so the iterative solver must terminate finitely. □

We now give a bound on the decrease in our barrier model provided by the Cauchy step t_k^{C}.

Lemma 2.11. *If* $k \in \mathcal{T}$, *then the Cauchy step* t_k^{C} *defined by* (2.28)–(2.29) *or* (2.33)–(2.34) *is computed and satisfies*

$$m_k^f(n_k) - m_k^f(n_k + t_k^{\text{C}}) \geq \kappa_k^{\text{ct}} \pi_k^f \min\left\{\pi_k^f, (1-\kappa_{\text{B}})\delta_k^t, (1-\kappa_{\text{fbt}})\kappa_{\text{fbn}}\right\} > 0,$$

where

$$\kappa_k^{\text{ct}} := \frac{\kappa_\chi^2}{2(1 + \|P_k G_k P_k\|_2)} \in (0, 1/2).$$

Proof. We first consider the case when $k \in \mathcal{T}_{\mathcal{D}}$, i.e., when the Cauchy step t_k^{C} is computed from (2.28)–(2.29) with the trust region radius $\delta_k^t = \min\{\delta_k^v, \delta_k^f\}$ (see (2.58)). It follows from part (iii) of Lemma 2.6

that $\chi_k^f \geq \kappa_\chi \pi_k^f > 0$ so that $\nabla m_k^f(n_k)^T r_k \geq \kappa_\chi (\pi_k^f)^2 > 0$. We now show that $t_k^{\mathrm{C}}(\alpha)$ (recall (2.28)) is feasible for (2.29) during iteration $k \in \mathcal{T}_{\mathcal{D}}$ when

$$0 \leq \alpha \leq \frac{1}{\pi_k^f} \min\left\{(1-\kappa_{\mathrm{B}})\delta_k^t, (1-\kappa_{\mathrm{fbt}})\kappa_{\mathrm{fbn}}\right\} =: \alpha_{\mathrm{B}}.$$

Indeed, consider any $\alpha \in [0, \alpha_{\mathrm{B}}]$. It follows from the definitions of $t_k^{\mathrm{C}}(\alpha)$, r_k, and α_{B} that

$$\|P_k^{-1} t_k^{\mathrm{C}}(\alpha)\|_2 = \|P_k^{-1}\alpha r_k\|_2 = \alpha \|P_k^{-1} r_k\|_2 = \alpha \pi_k^f \leq (1-\kappa_{\mathrm{B}})\delta_k^t. \tag{2.67}$$

Using the triangle inequality, (2.21) (which must hold since $k \in \mathcal{T}_{\mathcal{D}} \subseteq \mathcal{T}$), (2.58), and (2.67), we then have

$$\|P_k^{-1}\big(n_k + t_k^{\mathrm{C}}(\alpha)\big)\|_2 \leq \|P_k^{-1} n_k\|_2 + \|P_k^{-1} t_k^{\mathrm{C}}(\alpha)\|_2 \leq \kappa_{\mathrm{B}}\delta_k^t + (1-\kappa_{\mathrm{B}})\delta_k^t \leq \delta_k^t = \min\{\delta_k^v, \delta_k^f\}, \tag{2.68}$$

which shows that $t_k^{\mathrm{C}}(\alpha)$ satisfies the first constraint in problem (2.29). To show that $t_k^{\mathrm{C}s}(\alpha)$ also satisfies the second constraint in problem (2.29), first observe that if $[t_k^{\mathrm{C}s}(\alpha)]_i = [-\alpha r_k^s]_i \geq 0$, then $[s_k + n_k^s + t_k^{\mathrm{C}s}(\alpha)]_i \geq [s_k + n_k^s]_i \geq \kappa_{\mathrm{fbt}}[s_k + n_k^s]_i \geq 0$ since $\kappa_{\mathrm{fbt}} \in (0,1)$. Thus it suffices to consider i such that $[r_k^s]_i > 0$. It follows from the definitions of α_{B} and π_k^f, (2.24), the fact that $[r_k^s]_i > 0$, Lemma 2.7, and (2.13) that

$$\begin{aligned}\alpha \leq \alpha_{\mathrm{B}} \leq \frac{(1-\kappa_{\mathrm{fbt}})\kappa_{\mathrm{fbn}}}{\pi_k^f} \leq \frac{(1-\kappa_{\mathrm{fbt}})\kappa_{\mathrm{fbn}}}{\|S_k^{-1} r_k^s\|_2} \\ \leq \frac{(1-\kappa_{\mathrm{fbt}})\kappa_{\mathrm{fbn}}}{|[r_k^s]_i / [S_k]_{ii}|} = \frac{(1-\kappa_{\mathrm{fbt}})\kappa_{\mathrm{fbn}}[s_k]_i}{[r_k^s]_i} \leq \frac{(1-\kappa_{\mathrm{fbt}})[s_k + n_k^s]_i}{[r_k^s]_i}.\end{aligned}$$

Using the definition of $t_k^{\mathrm{C}s}(\alpha)$ and the previous inequality leads to

$$[-t_k^{\mathrm{C}s}(\alpha)]_i = \alpha [r_k^s]_i \leq (1-\kappa_{\mathrm{fbt}})[s_k + n_k^s]_i$$

from which we may conclude overall that $[s_k + n_k^s + t_k^{\mathrm{C}s}(\alpha)]_i \geq \kappa_{\mathrm{fbt}}[s_k + n_k^s]_i$ for $i = 1, 2, \ldots, M$. This proves that $t_k^{\mathrm{C}s}(\alpha)$ satisfies the second constraint in problem (2.29), and completes the proof that $t_k^{\mathrm{C}}(\alpha)$ is feasible for problem (2.29) for all $\alpha \in [0, \alpha_{\mathrm{B}}]$.

We now observe that the minimizer $\alpha_{\mathrm{T}}^{\mathrm{C}}$ of (2.29) yields $m_k^f(n_k + t_k^{\mathrm{C}}) \equiv m_k^f(n_k + t_k^{\mathrm{C}}(\alpha_{\mathrm{T}}^{\mathrm{C}})) \leq m_k^f(n_k + t_k^{\mathrm{C}}(\alpha))$ for all $\alpha \in [0, \alpha_{\mathrm{B}}]$. We also have from the Cauchy-Schwarz and standard norm inequalities that

$$|r_k^T G_k r_k| = \left|\big(\nabla m_k^f(n_k) + J(x_k, s_k)^T y_k\big)^T P_k^2 G_k P_k^2 \big(\nabla m_k^f(n_k) + J(x_k, s_k)^T y_k\big)^T\right| \leq (\pi_k^f)^2 \|P_k G_k P_k\|_2.$$

It then follows from [3, Lemma 1] with the quantities

$$\text{“}t\text{”} := \alpha_{\mathrm{B}}, \quad \text{“}a\text{”} := |r_k^T G_k r_k|, \quad \text{“}b\text{”} := \nabla m_k^f(n_k)^T r_k > 0,$$

(the strict inequality was shown earlier in this proof) that

$$\begin{aligned}
m_k^f(n_k) - m_k^f(n_k + t_k^{\rm C}) &\geq \frac{\text{“}b\text{”}}{2} \min\left\{ \frac{\text{“}b\text{”}}{\text{“}a\text{”}}, \text{“}t\text{”} \right\} \\
&\geq \frac{\nabla m_k^f(n_k)^T r_k}{2} \min\left\{ \frac{\nabla m_k^f(n_k)^T r_k}{(\pi_k^f)^2 \|P_k G_k P_k\|_2}, \frac{(1-\kappa_{\rm B})\delta_k^t}{\pi_k^f}, \frac{(1-\kappa_{\rm fbt})\kappa_{\rm fbn}}{\pi_k^f} \right\} \\
&= \frac{\nabla m_k^f(n_k)^T r_k}{2\pi_k^f} \min\left\{ \frac{\nabla m_k^f(n_k)^T r_k}{\pi_k^f(1+\|P_k G_k P_k\|_2)}, (1-\kappa_{\rm B})\delta_k^t, (1-\kappa_{\rm fbt})\kappa_{\rm fbn} \right\} \\
&= \frac{\chi_k^f}{2} \min\left\{ \frac{\chi_k^f}{(1+\|P_k G_k P_k\|_2)}, (1-\kappa_{\rm B})\delta_k^t, (1-\kappa_{\rm fbt})\kappa_{\rm fbn} \right\} \\
&\geq \frac{\kappa_\chi^2 \pi_k^f}{2(1+\|P_k G_k P_k\|_2)} \min\left\{ \pi_k^f, (1-\kappa_{\rm B})\delta_k^t, (1-\kappa_{\rm fbt})\kappa_{\rm fbn} \right\},
\end{aligned}$$

where we have used $1 + \|P_k G_k P_k\|_2 \geq 1$ and $\chi_k^f \geq \kappa_\chi \pi_k^f$ with $\kappa_\chi \in (0,1)$ for the last inequality.

The proof for the case $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$ is similar, but uses $\delta_k^t = \min\{\delta_k^v, \delta_k^f, \kappa_{\rm v} v_k^{\max}\}$, (2.33) instead of (2.28), (2.34) instead of (2.29), and (by Lemma 2.6(xi)) the fact that $n_k = 0$ for $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$. □

Finally, we turn our attention to the tangential step computation. The following result shows one way to find a tangential step t_k that satisfies the required conditions described in Section 2.3.

Lemma 2.12. *If* (2.21) *holds and* $\{y^{(i)}\}$ *is produced by an iterative solver employed to solve* (2.23) *that satisfies Assumption 2.1, then for some (finite) index* i *the vector* $y_k \leftarrow y^{(i)}$ *yields* r_k *such that either*

(i) the Cauchy point $t_k^{\rm C}$ *defined by* (2.28)–(2.29) *satisfies* (2.30)*, or*

(ii) the Cauchy point $t_k^{\rm C}$ *defined by* (2.33)–(2.34) *satisfies* (2.35)*.*

Proof. As in the proof of Lemma 2.10, in order to derive contradictions, suppose that the iterative solver employed to solve (2.23) runs for an infinite number of iterations without yielding the desired result, in which case we have under Assumption 2.1 that the sequence $\{y^{(i)}\}$ has a limit point y^∞ satisfying (2.65). That is, as $i \to \infty$, we have $r_k(y^{(i)}) \to r_k(y^\infty) \in \text{Null}(J(x_k, s_k))$. We introduce the notation $t_k^{\rm Cr}(i) := t_k^{\rm C}$ when $t_k^{\rm C}$ is the Cauchy point defined by (2.28)–(2.29) with $r_k = r_k(y^{(i)})$ associated with the relaxed SQP tangential subproblem, and $t_k^{\rm Cv}(i) := t_k^{\rm C}$ when $t_k^{\rm C}$ is the Cauchy point defined by (2.33)–(2.34) with $r_k = r_k(y^{(i)})$ associated with the very relaxed SQP tangential subproblem. We observe from (2.28) and (2.33), the constraints of (2.29) and (2.34), and the fact that $r_k(y^\infty) \in \text{Null}(J(x_k, s_k))$ that there exist vectors $t_k^{\rm Cr}(\infty)$ and $t_k^{\rm Cv}(\infty)$ such that $t_k^{\rm Cr}(i) \to t_k^{\rm Cr}(\infty) \in \text{Null}(J(x_k, s_k))$ and $t_k^{\rm Cv}(i) \to t_k^{\rm Cv}(\infty) \in \text{Null}(J(x_k, s_k))$.

By definition, the Cauchy point $t_k^{\rm Cr}(i)$ satisfies (2.30a)–(2.30c) for all i. Similarly, the Cauchy point $t_k^{\rm Cv}(i)$ satisfies (2.35a)–(2.35c) for all i. Thus, to reach contradictions, we need only show that for sufficiently large i either $t_k^{\rm Cr}(i)$ satisfies (2.30d) or $t_k^{\rm Cv}$ satisfies (2.35d).

Suppose that $n_k \neq 0$, in which case part (ii) of Lemma 2.6 implies that $k \in \mathcal{N}$. It then follows from part (i) of Lemma 2.6 that $m_k^v(n_k) < m_k^v(0)$, and thus the right-hand side of (2.30d) is strictly greater than $m_k^v(n_k)$. Therefore, since $t_k^{\rm Cr}(\infty) \in \text{Null}(J(x_k, s_k))$, there exists some smallest index i such that $t_k^{\rm Cr}(i)$ satisfies (2.30d), which is to say that statement (i) holds, which is a contradiction.

Now suppose that $n_k = 0$, in which case part (i) of Lemma 2.6 implies that $k \notin \mathcal{N}$. By virtue of (2.10), this must mean that $v_k < \kappa_{\rm vv} v_k^{\max}$. It follows from the facts that $n_k = 0$, $v_k < \kappa_{\rm vv} v_k^{\max}$, $\kappa_{\rm tt} \in (\kappa_{\rm vv}, 1)$, and $t_k^{\rm Cv}(i) \to t_k^{\rm Cv}(\infty) \in \text{Null}(J(x_k, s_k))$ that $t_k^{\rm Cv}(i)$ satisfies (2.35d) for all sufficiently large i. We have reached the contradiction that statement (ii) holds. □

3 Convergence of the Trust-Funnel Algorithm for Solving the Barrier Subproblem

Our analysis requires the following assumption that is assumed to hold for the remainder of the paper.

Assumption 3.1. *The sequence of iterates $\{x_k\}$ is contained in a compact set.*

The following is an immediate consequence of Assumptions 1.1 and 3.1.

Lemma 3.1. *There exists a constant $\kappa_{\mathrm{H}} \geq 1$ such that we have*

$$\max_{k\geq 0}\left\{1, \|g(x_k)\|_2, \|c(x_k)\|_2, \|J(x_k)\|_2, \|\nabla_{xx}f(x_k)\|_2, \max_{1\leq i\leq M}\|\nabla_{xx}c_i(x_k)\|_2\right\} \leq \kappa_{\mathrm{H}}.$$

We may now prove that important sequences related to our method are uniformly bounded.

Lemma 3.2. *There exists a constant $\kappa_{\mathrm{ub}} \geq \kappa_{\mathrm{H}}$ such that we have*

$$\max_{k\geq 0}\left\{v_k, \|s_k\|_2, \|J(x_k,s_k)^Tc(x_k,s_k)\|_2, \pi_k^v, \|P_kJ(x_k,s_k)^T\|_2, \chi_k^v, \|P_kG_kP_k\|_2, \|P_k\nabla f(x_k,s_k)\|_2, \right\} \leq \kappa_{\mathrm{ub}}.$$

Proof. The result is clearly true if the algorithm terminates finitely. Otherwise, it follows from Lemma 2.9 that $v_k \leq v_k^{\max} \leq v_0^{\max}$ for all k, which proves that $\{v_k\}$ can be bounded as claimed. Combining this with the reverse triangle inequality yields

$$\|s_k\|_2 - \|c(x_k)\|_2 \leq \|c(x_k)+s_k\|_2 = \|c(x_k,s_k)\|_2 \leq v_0^{\max} \text{ for all } k.$$

We may deduce from this bound and Lemma 3.1 that $\{\|s_k\|_2\}$ can be bounded as claimed. It then follows from the triangle inequality that

$$\|J(x_k,s_k)^Tc(x_k,s_k)\|_2 \leq \left\|\begin{pmatrix} J(x_k)^Tc(x_k,s_k) \\ 0\end{pmatrix}\right\|_2 + \left\|\begin{pmatrix} 0 \\ c(x_k,s_k)\end{pmatrix}\right\|_2,$$

which may then be combined with the Cauchy-Schwarz inequality, Lemma 3.1, and the boundedness of $\{v_k\}$ to conclude that $\{\|J(x_k,s_k)^Tc(x_k,s_k)\|_2\}$ can be bounded as claimed. The boundedness of $\{\pi_k^v\}$ follows from that of $\{\|s_k\|_2\}$ and $\{\|J(x_k,s_k)^Tc(x_k,s_k)\|_2\}$. It then follows from the boundedness of $\{\|s_k\|_2\}$ and, by Lemma 3.1, that of $\{\|J(x_k)\|_2\}$ that $\{\|P_kJ(x_k,s_k)^T\|_2\}$ can be bounded as claimed. This, along with the Cauchy-Schwarz inequality, implies that $\{\chi_k^v\}$ can be bounded as claimed. The boundedness of $\|P_kG_kP_k\|_2$ follows from the boundedness of $\{\|s_k\|_2\}$, (2.18), (2.19), Assumptions 1.1 and 3.1, and (2.20). Finally, it follows from Lemma 3.1 and the fact that $P_k\nabla f(x_k,s_k) = (g(x_k), -\mu e)$ that $\{\|P_k\nabla f(x_k,s_k)\|_2\}$ can be bounded as claimed. □

Using Lemma 3.2, we may now improve the Cauchy decrease bounds provided in Lemmas 2.8 and 2.11 by making the leading constants independent of the iteration number k.

Lemma 3.3. *For all k, the following hold:*

(i) *If $k \in \mathcal{N}$, then the Cauchy step n_k^{C} defined by (2.11)–(2.12) is computed and satisfies*

$$m_k^v(0) - m_k^v(n_k^{\mathrm{C}}) \geq \kappa_{\mathrm{cn}}\chi_k^v \min\{\pi_k^v, \delta_k^v, 1-\kappa_{\mathrm{fbn}}\} > 0$$

for some constant $\kappa_{\mathrm{cn}} \in (0,1]$ independent of k.

(ii) *If $k \in \mathcal{T}$, then the Cauchy step t_k^{C} defined by (2.28)–(2.29) or (2.33)–(2.34) is computed and satisfies*

$$m_k^f(n_k) - m_k^f(n_k+t_k^{\mathrm{C}}) \geq \kappa_{\mathrm{ct}}\pi_k^f \min\{\pi_k^f, (1-\kappa_{\mathrm{B}})\delta_k^t, (1-\kappa_{\mathrm{fbt}})\kappa_{\mathrm{fbn}}\} > 0$$

for some constant $\kappa_{\mathrm{ct}} \in (0,1/2]$ independent of k.

Proof. The results follow from Lemmas 2.8 and 2.11 along with Lemma 3.2. □

We require the next lemma that bounds the size of the trial step in different scenarios.

Lemma 3.4. *If Algorithm 1 does not terminate during iteration k, then the following holds:*

$$\|P_k^{-1}d_k\|_2 \begin{cases} = \|P_k^{-1}n_k\|_2 \leq \min\{\delta_k^v, \kappa_{\rm n}\pi_k^v\} & \text{if } k \notin \mathcal{T}, \\ = \|P_k^{-1}n_k\|_2 \leq \min\{\delta_k^v, \delta_k^f, \kappa_{\rm n}\pi_k^v\} & \text{if } k \in \mathcal{T}_0, \\ \leq \delta_k^t & \text{if } k \in \mathcal{T} \setminus \mathcal{T}_0. \end{cases}$$

In particular, for all k, we have $\|P_k^{-1}d_k\|_2 \leq \delta_k^v$.

Proof. Let $k \notin \mathcal{T}$, from which we have that $t_k \leftarrow 0$ and $d_k = n_k$. If $n_k = 0$, then the result holds trivially. Conversely, if $n_k \neq 0$, then part (ii) of Lemma 2.6 implies that $k \in \mathcal{N}$ and the result follows from (2.13).

Next, let $k \in \mathcal{T}$. First, if $k \in \mathcal{T}_0$, then it follows from part (iv) of Lemma 2.6 that $t_k = 0$ and (2.21) holds. Combining this with $d_k = n_k + t_k = n_k$, (2.13), and the fact that $\kappa_{\rm B} \in (0,1)$ shows that

$$\|P_k^{-1}d_k\|_2 = \|P_k^{-1}n_k\|_2 \leq \min\{\kappa_{\rm B}\min\{\delta_k^v, \delta_k^f\}, \delta_k^v, \kappa_{\rm n}\pi_k^v\} \leq \min\{\delta_k^v, \delta_k^f, \kappa_{\rm n}\pi_k^v\},$$

as desired. Second, if $k \in \mathcal{T}_{\mathcal{D}} \setminus \mathcal{T}_0$, then the result follows from (2.30c) and the definition (2.58). Third, if $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$, then the result follows from (2.35c) and the definition (2.58). □

We now bound the discrepancies between the problem functions and their corresponding models.

Lemma 3.5. *The following hold:*

(i) *There exists a constant $\kappa_{\rm G} > 0$ independent of k such that*

$$|f(x_k + d_k^x, s_k + d_k^s) - m_k^f(d_k)| \leq \kappa_{\rm G}\|P_k^{-1}d_k\|_2^2 \quad \text{for all } k. \tag{3.69}$$

(ii) *There exists a constant $\kappa_{\rm C} > 0$ independent of k such that*

$$|v(x_k + d_k^x, s_k + d_k^s) - m_k^v(d_k)| \leq \kappa_{\rm C}\|P_k^{-1}d_k\|_2^2 \quad \text{for all } k. \tag{3.70}$$

Proof. We first prove part (i). By the triangle inequality, we have

$$\begin{aligned} & |f(x_k + d_k^x, s_k + d_k^s) - m_k^f(d_k)| \\ \leq\ & |f(x_k + d_k^x) - f(x_k) - \nabla f(x_k)^T d_k^x - \tfrac{1}{2}{d_k^x}^T \nabla_{xx}\mathcal{L}(x_k, y_k^{\rm B}) d_k^x| \\ +\ & \left| -\mu \sum_{i=1}^{M} \ln([s_k + d_k^s]_i) + \mu \sum_{i=1}^{M} \ln([s_k]_i) + \mu e^T S_k^{-1} d_k^s - \tfrac{1}{2}{d_k^s}^T D_k d_k^s \right|. \end{aligned} \tag{3.71}$$

Under Assumptions 1.1 and 3.1, and by (2.19), there exists a constant $\kappa_{\rm G1} > 0$ such that

$$|f(x_k + d_k^x) - f(x_k) - \nabla f(x_k)^T d_k^x - \tfrac{1}{2}{d_k^x}^T \nabla_{xx}\mathcal{L}(x_k, y_k^{\rm B}) d_k^x| \leq \kappa_{\rm G1}\|d_k^x\|_2^2. \tag{3.72}$$

Moreover, note that for each $i \in \{1, \ldots, M\}$, we have by (2.13) and (2.30b)/(2.35b) that $[s_k]_i + [d_k^s]_i \geq \kappa_{\rm fbt}\kappa_{\rm fbn}[s_k]_i > 0$ for all k regardless of whether a tangential step t_k was computed. The Mean Value Theorem yields $\ln([s_k]_i + [d_k^s]_i) - \ln[s_k]_i = [d_k^s]_i/\xi_i$, where ξ_i lies between $[s_k]_i$ and $[s_k]_i + [d_k^s]_i$. Hence

$$\begin{aligned} \left| \ln([s_k]_i + [d_k^s]_i) - \ln[s_k]_i - \frac{[d_k^s]_i}{[s_k]_i} \right| &\leq \sup_{\xi \in [[s_k]_i, [s_k]_i + [d_k^s]_i]} \left| \frac{[d_k^s]_i}{\xi} - \frac{[d_k^s]_i}{[s_k]_i} \right| \\ &= \frac{[s_k]_i}{[s_k]_i + [d_k^s]_i} \left(\frac{[d_k^s]_i}{[s_k]_i} \right)^2 \leq \frac{1}{\kappa_{\rm fbt}\kappa_{\rm fbn}} \left(\frac{[d_k^s]_i}{[s_k]_i} \right)^2, \end{aligned}$$

where in the middle equation we have used the fact that the sup occurs at $\xi = [s_k]_i + [d_k^s]_i$. Hence, by (2.20) and Lemma 3.2, we have that

$$\begin{aligned} &\left| -\mu \sum_{i=1}^{M} \ln([s_k + d_k^s]_i) + \mu \sum_{i=1}^{M} \ln([s_k]_i) + \mu e^T S_k^{-1} d_k^s - \tfrac{1}{2} {d_k^s}^T D_k d_k^s \right| \\ \leq\ & \frac{1}{\kappa_{\text{fbt}} \kappa_{\text{fbn}}} {d_k^s}^T (\mu S_k^{-2}) d_k^s + \tfrac{1}{2} |{d_k^s}^T D_k d_k^s| \leq \kappa_{\text{G2}} \|S_k^{-1} d_k^s\|_2^2. \end{aligned} \tag{3.73}$$

where $\kappa_{\text{G2}} = \mu/\kappa_{\text{fbt}}\kappa_{\text{fbn}} + \frac{1}{2}\kappa_{\text{ub}}^2 \kappa_{\text{D}} > 0$. The result now follows from (3.71)–(3.73), and Lemma 3.4 with $\kappa_{\text{G}} := \kappa_{\text{G1}} + \kappa_{\text{G2}}$.

We now prove part (ii). By Lemma 3.1, Taylor's expansion theorem yields

$$c(x_k + d_k^x, s_k + d_k^s) = c(x_k, s_k) + J(x_k, s_k) d_k + w_k \ \text{ where } \ [w_k]_i = \tfrac{1}{2} {d_k^x}^T \nabla_{xx} c_i(\xi_{ik}) d_k^x$$

for some scalars $\xi_{ik} \in [x_k, x_k + d_k^x]$. As a consequence, we obtain with the reverse triangle inequality that there exists a constant $\kappa_{\text{C}} > 0$ so that

$$\begin{aligned} |v(x_k + d_k^x, s_k + d_k^s) - m_k^v(d_k)| &= |\|c(x_k + d_k^x, s_k + d_k^s)\|_2 - \|c(x_k, s_k) + J(x_k, s_k) d_k\|_2| \\ &\leq \|w_k\|_2 \leq \kappa_{\text{C}} \|d_k^x\|_2^2 \leq \kappa_{\text{C}} \|P_k^{-1} d_k\|_2^2, \end{aligned}$$

where we have used Lemma 3.1 and the Cauchy-Schwarz inequality. The desired result follows. □

We now prove an important fact about v-iterations; namely, if $k \in \mathcal{V}$ and the trust region radii or $v_k^{\max}$ are sufficiently small, then $k \in \mathcal{D}$.

Lemma 3.6. *If $k \in \mathcal{V}$ and*

$$\min\{\delta_k^v, \delta_k^f, \kappa_{\text{v}} v_k^{\max}\} \leq \frac{(1 - \kappa_{\text{tt}})}{\kappa_{\text{C}} \kappa_{\text{v}}} =: \kappa_{\mathcal{V}}, \tag{3.74}$$

then $k \in \mathcal{D}$.

Proof. For a proof by contradiction, suppose that (3.74) holds while $k \in \mathcal{V} \setminus \mathcal{D}$. We show that all of the conditions of an f-iteration are satisfied, implying that $k \in \mathcal{F}$, contradicting the supposition that $k \in \mathcal{V}$.

Since $k \notin \mathcal{D}$, we have from part (viii) of Lemma 2.6 that $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$ and (2.35) holds. Then, since $\mathcal{T}_0 \subseteq \mathcal{T}_{\mathcal{D}}$, it follows that $k \in \mathcal{T} \setminus \mathcal{T}_0$, so by part (iv) of Lemma 2.6 we have $t_k \neq 0$. Moreover, $k \in \mathcal{T} \setminus \mathcal{T}_0$ implies by Lemma 3.4 that $\|P_k^{-1} d_k\|_2 \leq \delta_k^t$, which along with the fact that $k \in \mathcal{T} \setminus \mathcal{T}_{\mathcal{D}}$ and (2.58) implies

$$\|P_k^{-1} d_k\|_2 \leq \min\{\delta_k^v, \delta_k^f, \kappa_{\text{v}} v_k^{\max}\} \leq \kappa_{\text{v}} v_k^{\max}. \tag{3.75}$$

Combining this fact with (3.70), the reverse triangle inequality, (2.35d), (3.75), and (3.74), we have that

$$v(x_k + d_k^x, s_k + d_k^s) \leq \kappa_{\text{tt}} v_k^{\max} + \kappa_{\text{C}} \|P_k^{-1} d_k\|_2^2 \leq \kappa_{\text{tt}} v_k^{\max} + \kappa_{\text{C}} \kappa_{\text{v}} v_k^{\max} \min\{\delta_k^v, \delta_k^f, \kappa_{\text{v}} v_k^{\max}\} \leq v_k^{\max}$$

so (2.37) holds. We have also argued (see the discussion after equation (2.37)) that (2.32) holds whenever (2.35) is satisfied. Thus, all of the conditions of an f-iteration are satisfied, so the result follows. □

Lemmas 3.4 and 3.6 have the following useful consequence.

Lemma 3.7. *There exists a constant $\kappa_{\text{n}\Delta 2} \in (0, 1]$ such that, if $k \in \mathcal{V}$ and*

$$\min\{\delta_k^v, \delta_k^f\} \leq \min\{1, \kappa_{\mathcal{V}}, \kappa_{\text{n}\Delta 2} \pi_k^v\}, \tag{3.76}$$

then $k \in \mathcal{N} \cap \mathcal{D}$.

Proof. We first note that, by Lemma 3.2, we have $\chi_k^v \leq \kappa_{\rm ub}$ for all k. Then, with

$$\kappa_{\rm n\Delta 2} := \min\left\{1, \frac{\kappa_{\rm v}}{\kappa_{\rm ub}}\right\} > 0, \tag{3.77}$$

we have with Lemma 2.9 that

$$\kappa_{\rm n\Delta 2}\pi_k^v = \kappa_{\rm n\Delta 2}\chi_k^v v_k \leq \kappa_{\rm n\Delta 2}\kappa_{\rm ub} v_k \leq \kappa_{\rm v} v_k \leq \kappa_{\rm v} v_k^{\max}. \tag{3.78}$$

Let $k \in \mathcal{V}$ and (3.76) hold. Then, along with (3.78) we have that

$$\min\{\delta_k^v, \delta_k^f, \kappa_{\rm v} v_k^{\max}\} = \min\{\delta_k^v, \delta_k^f\} \leq \kappa_{\mathcal{V}}.$$

Then, by Lemma 3.6, we have $k \in \mathcal{D}$ (as desired), so $k \in \mathcal{V} \cap \mathcal{D}$. Now, in order to derive a contradiction to the claim that $k \in \mathcal{N}$, suppose that $k \in (\mathcal{V} \cap \mathcal{D}) \backslash \mathcal{N}$. Since $k \notin \mathcal{N}$, we have from part (ii) of Lemma 2.6 that $n_k = 0$ (so that (2.32) holds). Then, since $k \in \mathcal{V}$, we must have $t_k \neq 0$ (since otherwise part (vi) of Lemma 2.6 would imply that $k \in \mathcal{Y}$, which is a contradiction). Thus, we have that $k \in \mathcal{T} \setminus \mathcal{T}_0$. At the same time, $k \notin \mathcal{N}$ implies that (2.10) does not hold, so $v_k < \kappa_{\rm vv} v_k^{\max} < \kappa_{\rm tt} v_k^{\max}$. This bound, (3.70), the reverse triangle inequality, (2.30d), the fact that $n_k = 0$, Lemma 2.9, the fact that $k \in \mathcal{T} \setminus \mathcal{T}_0$, Lemma 3.4, (2.58), (3.78) and (3.76) imply

$$v(x_k + d_k^x, s_k + d_k^s) < \kappa_{\rm tt} v_k^{\max} + \kappa_{\rm C}(\min\{\delta_k^v, \delta_k^f\})^2 \leq \kappa_{\rm tt} v_k^{\max} + \kappa_{\rm C}\kappa_{\rm v} v_k^{\max} \min\{\delta_k^v, \delta_k^f\},$$

which, when combined with (3.76) and (3.74), yields

$$v(x_k + d_k^x, s_k + d_k^s) \leq \kappa_{\rm tt} v_k^{\max} + (1 - \kappa_{\rm tt}) v_k^{\max} = v_k^{\max}$$

so that (2.37) holds. Combining this with the fact that $t_k \neq 0$ and the observation that (2.32) holds, shows that $k \in \mathcal{F}$, which is a contradiction. Thus, we must conclude that $k \in \mathcal{N}$. □

We now prove a relationship between the trust-region radii and a guarantee of a successful iteration.

Lemma 3.8. *The following hold:*

(i) *If $k \in \mathcal{F}$ and*

$$\delta_k^t \leq \min\left\{\frac{(1-\kappa_{\rm fbt})\kappa_{\rm fbn}}{1-\kappa_{\rm B}}, \frac{\pi_k^f}{1-\kappa_{\rm B}}, \frac{\kappa_\delta \kappa_{\rm ct}(1-\kappa_{\rm B})(1-\eta_2)\pi_k^f}{\kappa_{\rm G}}\right\} =: \min\{\kappa_{\Delta {\rm f}1}, \kappa_{\Delta {\rm f}2}\pi_k^f\} \tag{3.79}$$

then $\rho_k^f \geq \eta_2$, $k \in \mathcal{S}_f$, and $\delta_{k+1}^f \geq \delta_k^f$.

(ii) *If $k \in \mathcal{V}$ and*

$$\delta_k^v \leq \min\left\{\kappa_{\mathcal{V}}, 1-\kappa_{\rm fbn}, \kappa_{\rm n\Delta 2}\pi_k^v, \frac{\kappa_{\rm cd}\kappa_{\rm cn}\chi_k^v(1-\eta_2)}{\kappa_{\rm C}}, \right\} =: \min\{\kappa_{\Delta {\rm c}1}, \kappa_{\Delta {\rm c}2}\pi_k^v, \kappa_{\Delta {\rm c}3}\chi_k^v\}, \tag{3.80}$$

then $k \in \mathcal{N} \cap \mathcal{D} \cap \mathcal{S}_v$, $\rho_k^v \geq \eta_2$, and $\delta_{k+1}^v \geq \delta_k^v$.

Proof. For part (i), the proof that $\rho_k^f \geq \eta_2$, which implies that $k \in \mathcal{S}_f$, is the same as for [5, Theorem 6.4.2] and uses (2.38), (2.32) (which holds since $k \in \mathcal{F}$), (2.30a)/(2.35a), part (ii) of Lemma 3.3, (3.79), (3.69), the fact that $t_k \neq 0$, and Lemma 3.4. The fact that $\delta_{k+1}^f \geq \delta_k^f$ then follows from (2.41) and (2.44).

To prove part (ii), we first observe from (3.80) that $\pi_k^v > 0$ and $\chi_k^v > 0$ since $\delta_k^v > 0$ by construction in the algorithm. Moreover, (3.80) and Lemma 3.7 imply that $k \in \mathcal{N} \cap \mathcal{D}$. We now conclude from part

(ix) of Lemma 2.6 that (2.48) holds. Thus, using (3.70), Lemma 3.4, (2.48), (2.14), and Lemma 3.3(i), we have

$$|\rho_k^v - 1| = \left| \frac{v(x_k + d_k^x, s_k + d_k^s) - m_k^v(d_k)}{\Delta m_k^{v,d}} \right| \leq \left| \frac{\kappa_{\rm C}(\delta_k^v)^2}{\kappa_{\rm cd} \Delta m_k^{v,n}} \right| \leq \frac{\kappa_{\rm C}(\delta_k^v)^2}{\kappa_{\rm cd}\kappa_{\rm cn}\chi_k^v \min\{\pi_k^v, \delta_k^v, 1 - \kappa_{\rm fbn}\}}.$$

In fact, we have from (3.80) and the fact that $\kappa_{\rm n\Delta 2} \in (0, 1]$ that $\delta_k^v \leq \min\{\pi_k^v, \delta_k^v, 1 - \kappa_{\rm fbn}\}$ and

$$|\rho_k^v - 1| \leq \frac{\kappa_{\rm C}\delta_k^v}{\kappa_{\rm cd}\kappa_{\rm cn}\chi_k^v} \leq 1 - \eta_2.$$

Thus, $\rho_k^v \geq \eta_2 \geq \eta_1$, which means that $k \in \mathcal{S}_v$ and, by (2.52), $\delta_{k+1}^v \geq \delta_k^v$, as desired. □

We now provide a uniform lower bound on the tangential trust-region radius when our criticality measure π_k^f is bounded away from zero on f-iterates.

Lemma 3.9. *If there exists a constant $\epsilon_f > 0$ such that*

$$\pi_k^f \geq \epsilon_f \ \text{ for all } k \in \mathcal{F}, \tag{3.81}$$

then, for some constant $\epsilon_{\mathcal{F}} > 0$, we have

$$\delta_k^f \geq \epsilon_{\mathcal{F}} \ \text{ for all } k. \tag{3.82}$$

Proof. The statement follows from part (i) of Lemma 3.8, (2.58), the fact that $\mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, and the fact that $\delta_{k+1}^f \leftarrow \delta_k^f$ for $k \notin \mathcal{F}$. □

Lemma 3.10. *There exist constants $\{\kappa_{\Delta c4}, \kappa_{\Delta c5}, \kappa_{\Delta c6}\} \subset (0, 1)$ such that*

$$\delta_k^v \geq \min\{\kappa_{\Delta c4}, \kappa_{\Delta c5}\pi_k^v, \kappa_{\Delta c6}\chi_k^v\} \quad \text{for all } k. \tag{3.83}$$

Moreover, if there exists a constant $\epsilon_\theta > 0$ such that

$$\min\{v_k, \chi_k^v\} \geq \epsilon_\theta \ \text{for all } k \in \mathcal{V}, \tag{3.84}$$

then

$$\delta_k^v \geq \min\left\{\kappa_{\Delta c4}, \kappa_{\Delta c5}\epsilon_\theta^2, \kappa_{\Delta c6}\epsilon_\theta\right\} =: \epsilon_{\mathcal{C}} \ \text{ for all } k. \tag{3.85}$$

Proof. With $\gamma_1 \in (0, 1)$ defined for (2.44), we prove by induction that

$$\delta_k^v \geq \gamma_1 \min\{\delta_0^v, \kappa_{\Delta c1}, \kappa_{\delta vv}\pi_k^v, \kappa_{\Delta c2}\pi_k^v, \kappa_{\Delta c3}\chi_k^v\} \ \text{ for all } k. \tag{3.86}$$

This inequality holds trivially for $k = 0$, so supposing that it holds for iteration k, we prove that it holds for iteration $k+1$. First, suppose that $k \in \mathcal{Y} \cup (\mathcal{F} \setminus \mathcal{S}_f)$. Since $\delta_{k+1}^v \leftarrow \delta_k^v$ and $(x_{k+1}, s_{k+1}) \leftarrow (x_k, s_k)$ for such iterations, we conclude that (3.86) holds at iteration $k+1$. Second, if $k \in \mathcal{S}_f \cup \mathcal{S}_v$, then (2.42), (2.52), and the fact that $\gamma_1 \in (0, 1)$ ensure that (3.86) holds at iteration $k+1$. Finally, suppose that $k \in \mathcal{V} \setminus \mathcal{S}_v$. In this case, Lemma 3.8(ii) implies that $\delta_k^v > \min\{\kappa_{\Delta c1}, \kappa_{\Delta c2}\pi_k^v, \kappa_{\Delta c3}\chi_k^v\}$. This may then be combined with (2.55) and the fact that $(x_{k+1}, s_{k+1}) \leftarrow (x_k, s_k)$ to deduce that $\delta_{k+1}^v \geq \gamma_1 \min\{\kappa_{\Delta c1}, \kappa_{\Delta c2}\pi_k^v, \kappa_{\Delta c3}\chi_k^v\}$ so that (3.86) again holds at iteration $k + 1$. We therefore obtain that (3.83) holds for all k with $\kappa_{\Delta c4} := \gamma_1 \min\{\delta_0^v, \kappa_{\Delta c1}\}$, $\kappa_{\Delta c5} := \gamma_1 \min\{\kappa_{\delta vv}, \kappa_{\Delta c2}\}$ and $\kappa_{\Delta c6} := \gamma_1\kappa_{\Delta c3}$. The bound (3.85) then directly follows from (3.83), (3.84), (2.9), and the observation that δ_k^v is never decreased for $k \in \mathcal{Y} \cup \mathcal{F}$. □

We now give our first main result, namely that if there are finitely many successful iterations, then Algorithm 1 terminates finitely.

Theorem 3.11. *If $|\mathcal{S}| < \infty$, then Algorithm 1 terminates finitely.*

Proof. To derive a contradiction, suppose that Algorithm 1 does not terminate finitely. It then follows from the fact that $|\mathcal{S}| < \infty$, (2.36), (2.43), (2.46), (2.54), and (2.56) that for some $x_* \in \mathbb{R}^N$, $s_* \in \mathbb{R}^M$, and $\{v_*, v_*^{\max}, \pi_*^v, \chi_*^v\} \subset \mathbb{R}$ there exists an integer k_s such that

$$(x_k, s_k) = (x_*, s_*), \;\; v_k = v_*, \;\; v_k^{\max} = v_*^{\max}, \;\; \pi_k^v = \pi_*^v, \;\; \chi_k^v = \chi_*^v, \;\; \text{and } k \notin \mathcal{S} \text{ for all } k \geq k_s. \tag{3.87}$$

Also, the fact that $|\mathcal{S}| < \infty$ and Lemma 2.7 ensure that $s_* > 0$.

First, we prove that $|\mathcal{V}| < \infty$. In order to derive a contradiction, suppose that $|\mathcal{V}| = \infty$. Then, by (3.87) (in particular, the fact that $k \notin \mathcal{S}$ for $k \geq k_s$), it follows that (2.55) sets $\delta_{k+1}^v \leq \gamma_2 \delta_k^v$ for all $k \in \mathcal{V}$ with $k \geq k_s$. Combining this with the fact that (2.36) and (2.45) set $\delta_{k+1}^v \leftarrow \delta_k^v$ for all $k \in \mathcal{Y} \cup \mathcal{F}$ with $k \geq k_s$, it follows that $\{\delta_k^v\} \to 0$. We also have from part (ii) of Lemma 3.8 and the facts that $|\mathcal{V}| = \infty$ and $|\mathcal{S}| < \infty$ that we must have $0 = \lim_{k \in \mathcal{V}} \min\{\pi_k^v, \chi_k^v\} = \lim_{k \in \mathcal{V}} \min\{\chi_k^v v_k, \chi_k^v\} = \min\{\chi_*^v v_*, \chi_*^v\}$. If $v_* > 0$, then this implies that $\chi_*^v = 0$. However, this implies that for $k = k_s$ the algorithm would terminate finitely in Step 8, which contradicts the supposition of the proof. Thus, we must have that $v_* = 0$. Since $v_* = 0$, it follows from the conditions of Step 9 that $n_k = 0$ for all $k \geq k_s$. This implies that (2.21) will be satisfied for all $k \geq k_s$, which in turn implies by Step 14 of the algorithm that y_k, r_k, π_k^f, and χ_k^f will be computed to satisfy (2.26a), (2.26b), or (2.26c). If (2.26a) were to hold, then the algorithm would terminate finitely, which is a contradiction. Thus, we know that (2.26a) does not hold for all $k \geq k_s$, which combined with the fact that $v_* = 0$ implies that $\pi_k^f > \epsilon_\pi > 0$ for all $k \geq k_s$. It follows from this fact, part (i) of Lemma 3.8, the fact that $\{\delta_k^v\} \to 0$, (2.58), Lemma 3.4, and the fact that $|\mathcal{S}| < \infty$ that we must have $|\mathcal{F}| < \infty$. Next, it follows from the facts that $v_* = 0$ and $\{\delta_k^v\} \to 0$, Lemma 3.4, and (3.87) that (2.37) will be satisfied for all sufficiently large k. We may also deduce from the fact that $n_k = 0$ for all $k \geq k_s$ that (2.32) holds for all $k \geq k_s$. Since we have shown that $|\mathcal{F}| < \infty$ and that both (2.32) and (2.37) hold for sufficiently large k, we may conclude that $t_k = 0$ for all sufficiently large k. Therefore, since we have shown that $n_k = t_k = 0$ for all sufficiently large k, we have from part (vi) of Lemma 2.6 that $k \in \mathcal{Y}$ for all sufficiently large k, which combined with part (vii) of Lemma 2.6 implies that $\{\pi_k^f\} \to 0$. However, this contradicts our earlier conclusion that $\pi_k^f \geq \epsilon_\pi > 0$ for all $k \geq k_s$. Overall, we have contradicted the supposition that $|\mathcal{V}| = \infty$.

Next, suppose that $|\mathcal{F}| < \infty$. Combining this with the fact that $|\mathcal{V}| < \infty$ ensures that $k \in \mathcal{Y}$ for all sufficiently large k. It follows from this fact and part (vii) of Lemma 2.6 that $\{\pi_k^f\} \to 0$, and that y_k, r_k, π_k^f, and χ_k^f will be computed to satisfy (2.26a), (2.26b), or (2.26c) for all sufficiently large k. During the computation of these quantities, (2.26a) can never be satisfied, since in that case the algorithm would terminate finitely, which contradicts the supposition of the proof. Hence, since (2.26a) is never satisfied and $\{\pi_k^f\} \to 0$, we may deduce that $v_* > \epsilon_v > 0$. It then follows that $\chi_*^v > 0$ (and from (2.9) that $\pi_*^v > 0$), or else for $k = k_s$ the algorithm would terminate in Step 8, which is a contradiction. Thus, $\min\{\chi_*^v, \pi_*^v, v_*\} > 0$, which with (3.87), the fact that $\{\pi_k^f\} \to 0$, and (2.10) implies that $k \in \mathcal{N}$ for all sufficiently large k. Thus, by Lemma 2.6(i), we have $n_k \neq 0$, which by Lemma 2.6(vi) contradicts our earlier conclusion that $k \in \mathcal{Y}$. Overall, we have proven that we cannot have $|\mathcal{F}| < \infty$, so we must have $|\mathcal{F}| = \infty$.

Since $|\mathcal{F}| = \infty$, $|\mathcal{V}| < \infty$, and $|\mathcal{S}| < \infty$, we know from (2.36) and (2.44) that $\{\delta_k^f\} \to 0$, which when combined with (2.58), the fact that $\mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, and part (i) of Lemma 3.8 implies that $\{\pi_k^f\}_{k \in \mathcal{F}} \to 0$. Since (2.26a), (2.26b), or (2.26c) holds for $k \in \mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, and since the algorithm does not terminate finitely, we know that (2.26a) must not hold for all $k \in \mathcal{F}$. Combining this with the fact that $\{\pi_k^f\}_{k \in \mathcal{F}} \to 0$ implies that $v_k > \epsilon_v$ for all sufficiently large $k \in \mathcal{F}$. Hence, since $|\mathcal{F}| = \infty$, it follows from (3.87) that $v_* > \epsilon_v > 0$. We then must conclude that $\min\{v_*, \chi_*^v\} > 0$, or else for $k = k_s$ the algorithm would terminate finitely in Step 8, which is a contradiction. Also, from $\chi_*^v > 0$ and (2.9), it follows that $\pi_*^v > 0$. Since $\{\pi_k^f\}_{k \in \mathcal{F}} \to 0$, it follows that (2.26b) will be satisfied for all sufficiently large $k \in \mathcal{F}$, which implies that $t_k = 0$ and thus $k \notin \mathcal{F}$, which once again is a contradiction.

Overall, in all cases, we have reached contradictions of our supposition that Algorithm 1 does not terminate finitely, so the result is proved. □

We next prove a technical result about the violation decrease following a successful v-iteration.

Lemma 3.12. *There exist constants* $\{\kappa_{v\pi 1}, \kappa_{v\pi 2}, \kappa_{v\pi 3}\} \subset (0, \infty)$ *such that if* $k \in \mathcal{S}_v$, *then*

$$v_{k+1} \leq v_k - \chi_k^v \min\{\kappa_{v\pi 1}, \kappa_{v\pi 2}\pi_k^v, \kappa_{v\pi 3}\chi_k^v\}, \quad \textit{and} \tag{3.88a}$$

$$v_{k+1}^{\max} \leq \max\{\kappa_{\text{t}1} v_k^{\max}, v_k - (1-\kappa_{\text{t}2})\chi_k^v \min\{\kappa_{v\pi 1}, \kappa_{v\pi 2}\pi_k^v, \kappa_{v\pi 3}\chi_k^v\}\}, \tag{3.88b}$$

while (2.31) *does not hold.*

Proof. Let $k \in \mathcal{S}_v$, which by the definition of $\mathcal{S}_v$ means that (2.48) holds. In particular, we have $n_k \neq 0$. Combining this fact with part (ii) of Lemma 2.6 means that $k \in \mathcal{S}_v \cap \mathcal{N}$. It follows from this fact, (2.50), (2.49), (2.48), (2.14), Lemma 3.3(i), Lemma 3.10 (specifically (3.83)), and Lemma 3.2 that there exist constants $\{\kappa_{v\pi 1}, \kappa_{v\pi 2}, \kappa_{v\pi 3}\} \subset (0, \infty)$ such that (3.88a) holds, which in turn implies with (2.53) that (3.88b) holds. Note that (3.88a) and Lemma 2.9 imply that (2.37) holds.

We now prove that (2.31) does not hold. To reach a contradiction, suppose that (2.31) holds, which immediately implies that $t_k \neq 0$. Part (iv) of Lemma 2.6 then implies that $k \in \mathcal{T} \setminus \mathcal{T}_0$, which combined with the fact that (2.31) is assumed to hold shows that (2.32) holds. Thus all the conditions of an f-iteration are satisfied so that $k \in \mathcal{F}$, which, since $\mathcal{V} \cap \mathcal{F} = \emptyset$, contradicts the fact that $k \in \mathcal{S}_v \subseteq \mathcal{V}$. □

We now show that if there are infinitely many iterations, then the v-criticality measure χ_k^v converges to zero, at least along a subsequence of iterates.

Lemma 3.13. *If Algorithm 1 does not terminate finitely, then*

$$0 = \begin{cases} \lim\limits_{k \in \mathcal{S}_v} \min\{v_k, \chi_k^v\} & \text{if } |\mathcal{S}_v| = \infty, \\ \liminf\limits_{k \in \mathcal{S}_f} \min\{v_k, \chi_k^v\} & \text{if } |\mathcal{S}_v| < \infty. \end{cases} \tag{3.89}$$

Proof. Lemma 2.9 shows that $\{v_k^{\max}\}$ is monotonically decreasing and bounded below by zero. Therefore, if $|\mathcal{S}_v| = \infty$ and the update (2.53) sets $v_{k+1}^{\max} \leq \kappa_{\text{t}1} v_k^{\max}$ infinitely often, then $\{v_k^{\max}\} \to 0$, which implies by Lemma 2.9 that $\{v_k\} \to 0$, so (3.89) holds. Otherwise, if $|\mathcal{S}_v| = \infty$ and the update (2.53) sets $v_{k+1}^{\max} > \kappa_{\text{t}1} v_k^{\max}$ for all sufficiently large k, then by Lemma 3.12 we have $v_{k+1}^{\max} \leq v_k - (1-\kappa_{\text{t}2})\chi_k^v \min\{\kappa_{v\pi 1}, \kappa_{v\pi 2}\pi_k^v, \kappa_{v\pi 3}\chi_k^v\}$ for $k \in \mathcal{S}_v$, which implies that $\{\min\{\chi_k^v v_k, \chi_k^v\}\}_{k \in \mathcal{S}_v} = \{\min\{\pi_k^v, \chi_k^v\}\}_{k \in \mathcal{S}_v} \to 0$. If there is an infinite subsequence $\mathcal{K}_0 \subseteq \mathcal{S}_v$ over which $\{v_k\}_{\mathcal{K}_0}$ remains bounded away from zero, then this ensures that (3.89) holds over $\mathcal{K}_0$. Moreover, (3.89) clearly holds over each subsequence of $\mathcal{S}_v$ over which $\{v_k\}$ tends to zero. Thus, we deduce that (3.89) holds over the entire subsequence indexed by $\mathcal{S}_v$.

It remains to consider when $|\mathcal{S}_v| < \infty$, in which case, by the fact that $v_{k+1}^{\max} < v_k^{\max}$ only when $k \in \mathcal{S}_v$, there exists a constant $v_\infty^{\max} > 0$ such that $v_k^{\max} = v_\infty^{\max}$ for all sufficiently large k. By Theorem 3.11, the conditions of this lemma, and the fact that $|\mathcal{S}_v| < \infty$, it follows that $|\mathcal{S}_f| = \infty$. Now, to derive a contradiction, suppose that there exists a constant $\phi_{\min} > 0$ such that

$$\min\{v_k, \chi_k^v\} \geq \phi_{\min} > 0 \ \text{ for all sufficiently large } k. \tag{3.90}$$

Since $|\mathcal{S}_v| < \infty$, we know from (2.36) for $k \in \mathcal{Y}$, from (2.39) and (2.43) for $k \in \mathcal{F}$, from (2.54) for $k \in \mathcal{V} \setminus \mathcal{S}_v$, and the fact that the slack reset only possibly decreases the barrier function that $\{f(x_k, s_k)\}$ is monotonically decreasing. Moreover, it follows from Assumptions 1.1 and 3.1 and Lemma 3.2 that $\{f(x_k, s_k)\}$ is bounded below, so overall we have that $\{f(x_k, s_k)\} \to f_{\text{low}}$ for some $f_{\text{low}} > -\infty$. It follows from this fact, the fact that $|\mathcal{S}_f| = \infty$, (2.38), (2.39), (2.32) (which holds for $k \in \mathcal{F}$), (2.30a)/(2.35a), and part (ii) of Lemma 3.3 that $\lim_{k \in \mathcal{S}_f} \min\{\pi_k^f, \delta_k^t\} = 0$. Suppose that for some infinite index set $\mathcal{K}_1 \subseteq \mathcal{S}_f$ and scalar $\pi_{\min}^f > 0$ we have $\pi_k^f \geq \pi_{\min}^f$ for all $k \in \mathcal{K}_1$. It follows that $\{\delta_k^t\}_{k \in \mathcal{K}_1} \to 0$. However, from part (ii) of Lemma 3.8, the fact that $|\mathcal{S}_v| < \infty$, and (3.90), it follows that $\{\delta_k^v\}_{k \in \mathcal{V}}$ is bounded away from zero. In fact, since $\delta_{k+1}^v \leftarrow \delta_k^v$ for $k \notin \mathcal{V}$, we conclude that $\{\delta_k^v\}$ (defined over all k) is

bounded away from zero. Combining this with the facts that $\{\delta_k^t\}_{k\in\mathcal{K}_1} \to 0$ and $v_k^{\max} = v_\infty^{\max} > 0$ for all sufficiently large k implies that $\{\delta_k^f\}_{k\in\mathcal{K}_1} \to 0$. It then follows from Lemma 3.9 that there exists an infinite index set $\mathcal{K}_2 \subseteq \mathcal{F}$ such that $\{\pi_k^f\}_{k\in\mathcal{K}_2} \to 0$. Since $\mathcal{K}_2 \subseteq \mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, we know that (2.26a), (2.26b), or (2.26c) is satisfied for all $k \in \mathcal{K}_2$. However, we also know that (2.26a) cannot be satisfied since Algorithm 1 is assumed not to terminate finitely. It does, however, follow from $\{\pi_k^f\}_{k\in\mathcal{K}_2} \to 0$ and (3.90) that (2.26b) will be satisfied for all sufficiently large $k \in \mathcal{K}_2$ so that $t_k = 0$ for all sufficiently large $k \in \mathcal{K}_2 \subseteq \mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, which is a contradiction. Thus, we conclude that the set $\mathcal{K}_1$ cannot exist, so that $\lim_{k\in\mathcal{S}_f} \pi_k^f = 0$. It follows from this fact, (3.90), the definition of χ_k^v given in (2.9), the fact that (2.26a), (2.26b), or (2.26c) is satisfied for all $k \in \mathcal{F} \subseteq \mathcal{T}\setminus\mathcal{T}_0$, and since the algorithm does not terminate finitely that (2.26b) will be satisfied (and hence $t_k = 0$) for all sufficiently large $k \in \mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, which again is a contradiction. Thus, our supposition that (3.90) held must be incorrect and therefore there is a subsequence $\mathcal{K}$ such that $\lim_{k\in\mathcal{K}} \min\{v_k, \chi_k^v\} = 0$. Moreover, since $|\mathcal{S}_v| < \infty$ and $|\mathcal{S}_f| = \infty$, we conclude that (3.89) holds. □

To proceed further, we define the active and inactive slack variable sets

$$\mathcal{A}(s) := \{i \in \{1, 2, \ldots, M\} : [s]_i = 0\} \quad \text{and} \quad \mathcal{I}(s) := \{1, 2, \ldots M\} \setminus \mathcal{A}(s) \tag{3.91}$$

at $s \in \mathbb{R}^M$, and denote these sets at a point x_* by

$$\mathcal{A}_* := \mathcal{A}(s_*) \text{ and } \mathcal{I}_* := \mathcal{I}(s_*).$$

Lemma 3.14. *If Algorithm 1 does not terminate finitely and there exists an infinite index set $\mathcal{K}$ such that $\lim_{k\in\mathcal{K}} \min\{v_k, \chi_k^v\} \to 0$, then for an arbitrary limit point (x_*, s_*) of $\{(x_k, s_k)\}_{k\in\mathcal{K}}$ it follows that either $v(x_*, s_*) = 0$ so that (x_*, s_*) is feasible for problem* (NPs)*, or $\chi^v(x_*, s_*) = 0$ and x_* is an infeasible point at which the Jacobian of active constraints $J_{\mathcal{A}_*}(x_*)$ has linearly dependent rows.*

Proof. We first partition $\mathcal{K}$ into two disjoint index sets, call them $\mathcal{K}_1$ and $\mathcal{K}_2$, such that

$$\lim_{k\in\mathcal{K}_1} v_k = 0 \text{ and } \lim_{k\in\mathcal{K}_2} \chi_k^v = 0, \tag{3.92}$$

and such that v_k is bounded away from zero on $\mathcal{K}_2$. Any limit point (x_*, s_*) of the sequence $\{(x_k, s_k)\}_{k\in\mathcal{K}_1}$ yields $v(x_*, s_*) = 0$ so that (x_*, s_*) is feasible for problem (NPs), as desired.

Consider now a limit point of the sequence $\{(x_k, s_k)\}_{k\in\mathcal{K}_2}$, call it (x_*, s_*). By our definition of $\mathcal{K}_2$ and slack reset procedure (c.f., (2.5)), it follows that (x_*, s_*) is infeasible for problem (NPs). Moreover, since v_k is bounded away from zero on $\mathcal{K}_2$, the second limit in (3.92) implies

$$0 = \lim_{k\in\mathcal{K}_2} \chi_k^v = \lim_{k\in\mathcal{K}_2} \frac{\pi_k^v}{v_k} = \lim_{k\in\mathcal{K}_2} \frac{\|P_k J(x_k, s_k)^T c(x_k, s_k)\|_2}{\|c(x_k, s_k)\|_2} \geq \lim_{k\in\mathcal{K}_2} \sigma_{\min}\left(\begin{pmatrix} J(x_k)^T \\ S_k \end{pmatrix}\right) = \sigma_{\min}\left(\begin{pmatrix} J(x_*)^T \\ S_* \end{pmatrix}\right),$$

where $\sigma_{\min}(B)$ denotes the minimum singular value of a matrix B. Thus, we deduce that $(J(x_*)\ S_*)$ must have a subset of linearly dependent rows. Due to the structure of this matrix, we may assume without loss of generality that the subset does not contain row i when $[s_*]_i > 0$; it only contains rows indexed by $\mathcal{A}_*$, and thus $J_{\mathcal{A}_*}(x_*)$ has linearly dependent rows. □

We now make the following assumption throughout the rest of the paper.

Assumption 3.2. *If Algorithm 1 does not terminate finitely and $\mathcal{K}$ is an infinite index set such that $\{\pi_k^v\}_{k\in\mathcal{K}} \to 0$, then for an arbitrary limit point (x_*, s_*) of $\{(x_k, s_k)\}_{k\in\mathcal{K}}$ it follows that $\mathcal{A}_* = \emptyset$ or $J_{\mathcal{A}_*}(x_*)$ has full row rank (i.e., $J(x_*, s_*)P_*$ with $P_* := \operatorname{diag}(I, S_*)$ has full row rank), which implies that (x_*, s_*) is not an infeasible stationary point for problem* (NPs).

Our claim in this assumption—i.e., that due to full row rank of the scaled constraint Jacobian, the property $\{\pi_k^v\}_{k\in\mathcal{K}} \to 0$ implies that the algorithm avoids infeasible stationary points over $\mathcal{K}$—is

formally proved in the following lemma. This lemma represents a strengthening of Lemma 3.14.

Lemma 3.15. *If Algorithm 1 does not terminate finitely and $\mathcal{K}$ is an infinite index set such that $\{\pi_k^v\}_{k\in\mathcal{K}} \to 0$, then for an arbitrary limit point (x_*, s_*) of $\{(x_k, s_k)\}_{k\in\mathcal{K}}$ it follows that $v(x_*, s_*) = 0$ so that (x_*, s_*) is feasible for problem* (NPs).

Proof. Let us define the feasibility problem

$$\underset{x,s}{\text{minimize}} \ \tfrac{1}{2}v(x,s)^2 \ \text{ subject to } \ s \geq 0$$

for which we have the first-order KKT conditions

$$\min\{s, c(x,s)\} = 0 \quad \text{and} \quad J(x)^T c(x,s) = 0. \tag{3.93}$$

For an arbitrary limit point (x_*, s_*) of $\{(x_k, s_k)\}_{k\in\mathcal{K}}$, it follows from Lemma 2.7 and $\{\pi_k^v\}_{k\in\mathcal{K}} \to 0$ that

$$s_* \geq 0, \quad c(x_*, s_*) \geq 0, \quad S_* c(x_*, s_*) = 0, \ \text{ and } \ J(x_*)^T c(x_*, s_*) = 0. \tag{3.94}$$

In particular, using the definitions in (3.91) and (3.94), we have

$$[s_*]_{\mathcal{I}_*} > 0 \quad \text{and} \quad c_{\mathcal{I}_*}(x_*) < c_{\mathcal{I}_*}(x_*, s_*) = 0; \tag{3.95a}$$

$$[s_*]_{\mathcal{A}_*} = 0 \quad \text{and} \quad c_{\mathcal{A}_*}(x_*) = c_{\mathcal{A}_*}(x_*, s_*) \geq 0. \tag{3.95b}$$

Hence, from (3.94) and (3.95), we have that (x_*, s_*) satisfies (3.93). Now, if $\mathcal{A}_* = 0$, then by (3.95a) we have that $v(x_*, s_*) = 0$ and $c(x_*) \leq 0$, as desired. Otherwise, by (3.94) and (3.95a), we have

$$0 = J(x_*)^T c(x_*, s_*) = J_{\mathcal{A}}(x_*)^T c_{\mathcal{A}}(x_*, s_*) = J_{\mathcal{A}}(x_*)^T c_{\mathcal{A}}(x_*).$$

Under Assumption 3.2, we have that $J_{\mathcal{A}}(x_*)$ has full row rank, so the above implies that $0 = c_{\mathcal{A}}(x_*) = c_{\mathcal{A}}(x_*, s_*)$. Combining this with (3.95a) again yields $v(x_*, s_*) = 0$ and $c(x_*) \leq 0$, as desired. □

We now prove a useful fact about our employed infeasibility measures.

Lemma 3.16. *For any infinite index set $\mathcal{K}$, we have*

$$\lim_{k\in\mathcal{K}} \min\{v_k, \chi_k^v\} = 0 \quad \textit{if and only if} \quad \lim_{k\in\mathcal{K}} \pi_k^v = 0. \tag{3.96}$$

Proof. First, suppose that $\{\min\{v_k, \chi_k^v\}\}_{k\in\mathcal{K}} \to 0$. Then, as in the proof of Lemma 3.14, we can partition $\mathcal{K}$ into disjoint subsets $\mathcal{K}_1$ and $\mathcal{K}_2$ such that (3.92) holds and v_k is bounded away from zero on $\mathcal{K}_2$. By Lemma 3.2, it then follows that $\{\pi_k^v\}_{k\in\mathcal{K}_1} \to 0$, and by (2.9) we must also have $\{\pi_k^v\}_{k\in\mathcal{K}_2} \to 0$. Consequently, $\{\pi_k^v\}_{k\in\mathcal{K}} \to 0$, as desired. Second, suppose that $\{\pi_k^v\}_{k\in\mathcal{K}} \to 0$ and, to obtain a contradiction, that there exists some $\epsilon > 0$ such that the set $\mathcal{K}_\epsilon := \{k \in \mathcal{K} : \min\{v_k, \chi_k^v\} \geq \epsilon\}$ is infinite. It then follows from the definition of χ_k^v in (2.9) that the infinite sequence $\{\pi_k^v\}_{k\in\mathcal{K}_\epsilon}$ is bounded away from zero, which is a contradiction. Hence, $\{\min\{v_k, \chi_k^v\}\}_{k\in\mathcal{K}} \to 0$, as desired. □

We now prove that if there are an infinite number of successful v-iterations, then, amongst other things, feasibility is achieved at all limit points of the sequence of iterates computed by the algorithm.

Lemma 3.17. *If $|\mathcal{S}_v| = \infty$, then $\{v_k^{\max}\} \to 0$, $\{v_k\} \to 0$, $\{\pi_k^v\} \to 0$, and $\{n_k\} \to 0$.*

Proof. Since $|\mathcal{S}_v| = \infty$, it must be true that Algorithm 1 does not terminate finitely. This implies, for one thing, that the result of Lemma 3.13 holds true. Moreover, Lemma 2.9 shows that $\{v_k^{\max}\}$ is monotonically decreasing and bounded below by zero. Then, as in the proof of Lemma 3.13, we have that if the update (2.53) sets $v_{k+1}^{\max} \leq \kappa_{\text{t1}} v_k^{\max}$ infinitely often, then $\{v_k^{\max}\} \to 0$ and $\{v_k\} \to 0$, from which it follows by Lemma 3.2 that $\{\pi_k^v\} \to 0$. It then follows from these facts and (2.13) that $\{n_k\} \to 0$.

All that remains is to consider the case when the update (2.53) sets $v_{k+1}^{\max} > \kappa_{\text{t1}} v_k^{\max}$ for all large k. From Lemma 3.13 we have that $\{\min\{v_k, \chi_k^v\}\}_{k\in\mathcal{S}_v} \to 0$, which in turn by Lemma 3.16 implies that $\{\pi_k^v\}_{k\in\mathcal{S}_v} \to 0$. Then, by Lemma 3.15, Assumption 3.1, and the boundedness of $\{s_k\}$ stated in Lemma 3.2, there exists an infinite index set $\mathcal{K} \subseteq \mathcal{S}_v$ such that $\{v_k\}_{k\in\mathcal{K}} \to 0$. We then have from Lemma 3.12 (in particular, (3.88b)) that $\{v_k^{\max}\}_{k\in\mathcal{K}} \to 0$, which means that $\{v_k^{\max}\} \to 0$ and hence $\{v_k\} \to 0$ because of Lemma 2.9. Combining this with Assumptions 1.1 and 3.1 and Lemma 3.2, we thus have $\{\pi_k^v\} \to 0$. It follows from this fact, (2.13), and Lemma 3.2 that $\{n_k\} \to 0$. □

We next prove a result illustrating the importance of the sequence $\{\pi_k^f\}$. In particular, the result establishes that π_k^f is a valid criticality measure for (1.1).

Lemma 3.18. *If $\mathcal{K}$ is any subsequence and (x_*, s_*) is any point such that $\lim_{k\in\mathcal{K}}(x_k, s_k) = (x_*, s_*)$ with $v(x_*, s_*) = 0$ and $\lim_{k\in\mathcal{K}} \pi_k^f = 0$, then $\lim_{k\in\mathcal{K}} y_k = y_*$ where (x_*, s_*, y_*) is a KKT point for problem (1.1).*

Proof. Since $v(x^*, s^*) = 0$, it follows that $\lim_{k\in\mathcal{K}} c(x_k, s_k) = c(x_*, s_*) = 0$, which, when combined with Lemma 3.2, proves that $\lim_{k\in\mathcal{K}} \pi_k^v = 0$. Thus, it follows from (2.13) and Lemma 3.2 that

$$\lim_{k\in\mathcal{K}} n_k = 0. \tag{3.97}$$

Next, observe that

$$\begin{aligned} 0 = \lim_{k\in\mathcal{K}} \pi_k^f &= \lim_{k\in\mathcal{K}} \left\| P_k\Big(\nabla m_k^f(n_k) + J(x_k, s_k)^T y_k\Big)\right\|_2 \\ &= \lim_{k\in\mathcal{K}} \left\| \begin{pmatrix} g(x_k) + \nabla_{xx}\mathcal{L}(x_k, y_k^{\text{B}})n_k^x + J(x_k)^T y_k \\ -\mu e + S_k D_k n_k^s + S_k y_k \end{pmatrix}\right\|_2 \qquad (3.98) \\ &= \lim_{k\in\mathcal{K}} \left\| \begin{pmatrix} g(x_k) + \nabla_{xx}\mathcal{L}(x_k, y_k^{\text{B}})n_k^x + J(x_k)^T y_k \\ [-\mu e + S_k D_k n_k^s + S_k y_k]_{\mathcal{A}_*} \\ [-\mu e + S_k D_k n_k^s + S_k y_k]_{\mathcal{I}_*} \end{pmatrix}\right\|_2. \qquad (3.99) \end{aligned}$$

Using (3.99) (specifically the third row of the matrix inside the norm) with $\lim_{k\in\mathcal{K}}(x_k, s_k) = (x_*, s_*)$, the fact that $[s_*]_{\mathcal{I}_*} > 0$, (2.20), Lemma 3.2 and (3.97) shows that

$$\lim_{k\in\mathcal{K}} [y_k]_{\mathcal{I}_*} = [\mu S_*^{-1} e]_{\mathcal{I}_*} =: [y_*]_{\mathcal{I}_*}.$$

It then follows from (3.99) (specifically the first row of the matrix inside the norm), the fact that $\lim_{k\in\mathcal{K}}(x_k, s_k) = (x_*, s_*)$, (2.20), (2.19), Lemma 3.1, (3.97), the fact that $\lim_{k\in\mathcal{K}} \pi_k^v = 0$, and hence the full row rank of $J_{\mathcal{A}_*}(x_*)$ (stated in Assumption 3.2) that

$$\lim_{k\in\mathcal{K}} [y_k]_{\mathcal{A}_*} = -\big[J_{\mathcal{A}_*}(x_*)J_{\mathcal{A}_*}(x_*)^T\big]^{-1} J_{\mathcal{A}_*}(x_*)\Big(g(x_*) + J_{\mathcal{I}_*}(x_*)^T [y_*]_{\mathcal{I}_*}\Big) =: [y_*]_{\mathcal{A}_*}.$$

We have shown that the multiplier sequence converges on $\mathcal{K}$, i.e., that $\lim_{k\in\mathcal{K}} y_k = y_*$ for some $y_* \in \mathbb{R}^M$. Combining this with (3.98), the fact that $\lim_{k\in\mathcal{K}}(x_k, s_k) = (x_*, s_*)$, (2.20), (2.19), Lemma 3.1, and (3.97) proves that

$$g(x_*) + J(x_*)^T y_* = 0 \quad \text{and} \quad S_* y_* = \mu e. \tag{3.100}$$

Note that it follows from (3.100), Lemma 2.7, and the fact that $\mu > 0$ that $(s_*, y_*) > 0$. Combining this with (3.100) and $v(x_*, s_*) = 0$ proves that (x_*, y_*, s_*) is a KKT point for problem (1.1). □

Lemmas 3.15 and 3.18 prove that, under Assumption 3.2, we may obtain a first-order KKT point for the barrier subproblem (1.1) with any subsequence $\mathcal{K}$ over which $\{v_k\}_{k\in\mathcal{K}} \to 0$ and $\{\pi_k^f\}_{k\in\mathcal{K}} \to 0$. Now, to prove that such a sequence will exist, we make the following assumption—which is, at nearly feasible points, stronger than Assumption 3.2—for the remainder of our analysis. The assumption states that

at any nearly feasible point, the singular values of a scaled constraint Jacobian are bounded away from zero.

Assumption 3.3. *There exists a constant $\kappa_{\mathrm{c}} > 0$ independent of k such that if $v_k \leq \kappa_{\mathrm{c}}$, then the smallest singular value of $J(x_k, s_k)P_k$ is greater than κ_{J} for some constant $\kappa_{\mathrm{J}} > 0$ independent of k. Observe that this implies that if $v_k \leq \kappa_{\mathrm{c}}$, then $\chi_k^v \geq \kappa_{\mathrm{J}}$.*

We also define the following projection operator. Note that this operator is used for theoretical purposes only, i.e., computing such projections is unnecessary in an implementation of our algorithm.

Definition 3.19. *Let $\mathrm{Proj}_k(d)$ denote the orthogonal projection of d onto the range space of $P_k J(x_k, s_k)^T$.*

Lemma 3.20. *If $k \in \mathcal{N} \cap \mathcal{D}$ and $v_k \leq \kappa_{\mathrm{c}}$, then there exist constants $\{\kappa_{\mathrm{R1}}, \kappa_{\mathrm{R2}}\} \subset (0, \infty)$ such that*

$$\|\mathrm{Proj}_k(P_k^{-1}d_k)\|_2 \leq \frac{2}{\kappa_{\mathrm{J}}^2}\pi_k^v \quad \textit{and} \quad \Delta m_k^{v,d} \geq \kappa_{\mathrm{J}} \min\{\kappa_{\mathrm{R1}}, \kappa_{\mathrm{R2}}\|\mathrm{Proj}_k(P_k^{-1}d_k)\|_2\}. \tag{3.101}$$

Proof. Let $k \in \mathcal{N} \cap \mathcal{D}$ and define $m_k^{v,P}(d) := \|c(x_k, s_k) + J(x_k, s_k)P_k d\|_2$ and $d_k^P := P_k^{-1}d_k$. Then, it follows from the fact that $J(x_k, s_k)P_k\mathrm{Proj}_k(d_k^P) = J(x_k, s_k)P_k d_k^P$, part (i) of Lemma 2.6, and (2.30d) that

$$\begin{aligned} m_k^{v,P}(\mathrm{Proj}_k(d_k^P)) &= \|c(x_k, s_k) + J(x_k, s_k)P_k\mathrm{Proj}_k(d_k^P)\|_2 = \|c(x_k, s_k) + J(x_k, s_k)P_k d_k^P\|_2 \\ &= \|c(x_k, s_k) + J(x_k, s_k)d_k\|_2 < \|c(x_k, s_k)\|_2 = m_k^{v,P}(0). \end{aligned} \tag{3.102}$$

We may also note that since $v_k \leq \kappa_{\mathrm{c}}$, we have under Assumption 3.3 that the smallest eigenvalue of $\nabla_{xx}(m_k^{v,P}(0))^2 = P_k^T J(x_k, s_k)^T J(x_k, s_k)P_k$ is bounded below by $\kappa_{\mathrm{J}}^2 > 0$. We may now use this fact, (3.102), and [5, Lemma 6.5.1] applied to $(m_k^{v,P})^2$ to conclude that

$$\|\mathrm{Proj}_k(P_k^{-1}d_k)\|_2 = \|\mathrm{Proj}_k(d_k^P)\|_2 \leq \frac{2}{\kappa_{\mathrm{J}}^2}\nabla_x(m_k^{v,P}(0))^2 = \frac{2}{\kappa_{\mathrm{J}}^2}\pi_k^v,$$

which proves the first inequality in (3.101). It also follows from Lemma 3.4 and the fact that the orthogonal projection operator is nonexpansive that

$$\delta_k^v \geq \|P_k^{-1}d_k\|_2 \geq \|\mathrm{Proj}_k(P_k^{-1}d_k)\|_2.$$

Combining this with the fact that $k \in \mathcal{D}$, Lemma 2.6(ix), the inequality in (2.48), (2.14), Lemma 3.3(i), Assumption 3.3 and the first inequality in (3.101), we have that there exist constants $\{\kappa_{\mathrm{R1}}, \kappa_{\mathrm{R2}}\} \subset (0, \infty)$ such that the second inequality in (3.101) holds. □

We now bound the size of the normal step along a certain subsequence of unsuccessful v-iterations.

Lemma 3.21. *If $k \in (\mathcal{N} \cap \mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$ and*

$$v_k \leq \min\left\{\kappa_{\mathrm{c}}, \frac{\kappa_{\Delta\mathrm{c}1}}{\kappa_{\Delta\mathrm{c}2}\kappa_{\mathrm{J}}}, \frac{\kappa_{\Delta\mathrm{c}3}}{\kappa_{\Delta\mathrm{c}2}}, \frac{1-\kappa_{\mathrm{fbn}}}{\kappa_{\mathrm{J}}}, \frac{1-\kappa_{\mathrm{fbn}}}{\kappa_{\Delta\mathrm{c}2}\kappa_{\mathrm{J}}}\right\}, \tag{3.103}$$

then, for some constants $\{\kappa_{\mathrm{cld}}, \kappa_{\mathrm{sRn}}\} \subset (0, 1)$, we have

$$m_k^v(d_k) \leq \kappa_{\mathrm{cld}}v_k \quad \textit{and} \quad \|\mathrm{Proj}_k(P_k^{-1}d_k)\|_2 \geq \kappa_{\mathrm{sRn}}\|P_k^{-1}n_k\|_2. \tag{3.104}$$

Proof. Consider $k \in (\mathcal{N} \cap \mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$ such that (3.103) holds. It follows from the fact that $k \in \mathcal{N} \cap \mathcal{D}$, Lemma 2.6(ix), the inequality in (2.48), (2.14), Lemma 3.3(i), (3.103), Assumption 3.3, and (2.9) that

$$m_k^v(d_k) \leq m_k^v(0) - \kappa_{\mathrm{cd}}\kappa_{\mathrm{cn}}\chi_k^v \min\left\{\pi_k^v, \delta_k^v, 1-\kappa_{\mathrm{fbn}}\right\} \leq m_k^v(0) - \kappa_{\mathrm{cd}}\kappa_{\mathrm{cn}}\kappa_{\mathrm{J}} \min\left\{\kappa_{\mathrm{J}}v_k, \delta_k^v, 1-\kappa_{\mathrm{fbn}}\right\}. \tag{3.105}$$

It also follows from Lemma 3.8(ii), the fact that $k \in \mathcal{V} \setminus \mathcal{S}_v$, Assumption 3.3, (2.9), and (3.103) that

$$\delta_k^v > \min\left\{\kappa_{\Delta c1}, \kappa_{\Delta c2}\pi_k^v, \kappa_{\Delta c3}\chi_k^v\right\} \geq \min\left\{\kappa_{\Delta c1}, \kappa_{\Delta c2}\kappa_{\mathrm{J}}v_k, \kappa_{\Delta c3}\kappa_{\mathrm{J}}\right\} = \kappa_{\Delta c2}\kappa_{\mathrm{J}}v_k.$$

Substituting this into (3.105) ensures by (3.103) the existence of $\kappa_{\mathrm{cld}} \in (0,1)$ independent of k such that

$$m_k^v(d_k) \leq m_k^v(0) - \kappa_{\mathrm{cd}}\kappa_{\mathrm{cn}}\kappa_{\mathrm{J}} \min\left\{\kappa_{\mathrm{J}}v_k, \kappa_{\Delta c2}\kappa_{\mathrm{J}}v_k, 1-\kappa_{\mathrm{fbn}}\right\} = v_k - \kappa_{\mathrm{cd}}\kappa_{\mathrm{cn}}\kappa_{\mathrm{J}} \min\left\{\kappa_{\mathrm{J}}, \kappa_{\Delta c2}\kappa_{\mathrm{J}}\right\} v_k \leq \kappa_{\mathrm{cld}}v_k.$$

This is the first desired result. Next, defining $d_k^P := P_k^{-1}d_k$, we may use the inequality above, the reverse triangle inequality, and the fact that $J(x_k,s_k)P_kd_k^P = J(x_k,s_k)P_k\mathrm{Proj}_k(d_k^P)$ to have

$$\begin{aligned} v_k - \|J(x_k,s_k)P_k\mathrm{Proj}_k(d_k^P)\|_2 &\leq \|c(x_k,s_k) + J(x_k,s_k)P_k\mathrm{Proj}_k(d_k^P)\|_2 \\ &= \|c(x_k,s_k) + J(x_k,s_k)P_kd_k^P\|_2 \\ &= m_k^v(d_k) \leq \kappa_{\mathrm{cld}}v_k. \end{aligned}$$

Combining the above with the fact that $k \in \mathcal{N}$, (2.13), and standard norm inequalities then implies

$$\begin{aligned} \|P_k^{-1}n_k\|_2 \leq \kappa_{\mathrm{n}}\pi_k^v &\leq \kappa_{\mathrm{n}}\|P_kJ(x_k,s_k)^T\|_2v_k \\ &\leq \kappa_{\mathrm{n}}\|P_kJ(x_k,s_k)^T\|_2\frac{\|J(x_k,s_k)P_k\mathrm{Proj}_k(d_k^P)\|_2}{1-\kappa_{\mathrm{cld}}} \\ &\leq \kappa_{\mathrm{n}}\|P_kJ(x_k,s_k)^T\|_2\frac{\|J(x_k,s_k)P_k\|_2\|\mathrm{Proj}_k(d_k^P)\|_2}{1-\kappa_{\mathrm{cld}}}. \end{aligned}$$

It then follows from the definition of d_k^P, Lemma 3.2, and the fact that $\kappa_{\mathrm{cld}} \in (0,1)$ that for some $\kappa_{\mathrm{sRn}} \in (0,1)$ independent of k, we have

$$\|\mathrm{Proj}_k(P_k^{-1}d_k)\|_2 \geq \frac{1-\kappa_{\mathrm{cld}}}{\kappa_{\mathrm{n}}\|J(x_k,s_k)P_k\|_2^2}\|P_k^{-1}n_k\|_2 \geq \kappa_{\mathrm{sRn}}\|P_k^{-1}n_k\|_2,$$

which is the second desired result. □

For our next pair of results, we define the constants

$$\varsigma_{\mathrm{tn}} := \kappa_{\mathcal{VS}} \max\left\{1, \frac{2\kappa_{\mathrm{ub}}}{(1-\kappa_\delta)(\kappa_{\mathcal{VS}}-1)\kappa_{\mathrm{ct}}(1-\kappa_{\mathrm{B}})\epsilon_\pi}\right\} > 1 \text{ and} \tag{3.106a}$$

$$\varsigma_\delta := \min\left\{1, \frac{\epsilon_\pi}{1-\kappa_{\mathrm{B}}}, \frac{(1-\kappa_{\mathrm{fbt}})\kappa_{\mathrm{bfn}}}{1-\kappa_{\mathrm{B}}}\right\} \in (0,1]. \tag{3.106b}$$

Lemma 3.22. *If $k \notin \mathcal{Y}$,*

$$\pi_k^f \geq \epsilon_\pi > 0, \tag{3.107a}$$

$$\min\{\delta_k^v, \delta_k^f\} \leq \varsigma_\delta, \quad \textit{and} \tag{3.107b}$$

$$\|P_k^{-1}t_k\|_2 \geq \varsigma_{\mathrm{tn}}\|P_k^{-1}n_k\|_2, \tag{3.107c}$$

then $t_k \neq 0$ and (2.32) holds.

Proof. Let $k \notin \mathcal{Y}$ be such that (3.107) holds. If $k \in \mathcal{F}$, then the results follow by the definition of the index set $\mathcal{F}$. Thus, for the remainder of the proof, we may assume that $k \in \mathcal{V}$.

If $n_k = 0$, then $t_k \neq 0$ (since otherwise $k \in \mathcal{Y}$ by Lemma 2.6(vi)) and by (2.30a)/(2.35a), and Lemma 3.3(ii) we have $\Delta m_k^{f,d} = \Delta m_k^{f,t} \geq 0$, meaning that (2.32) holds, as desired. Otherwise, if $n_k \neq 0$, then since $s_k > 0$ and $P_k \succ 0$ for all k and (3.107c) holds, we have $t_k \neq 0$, which implies $k \in \mathcal{T} \setminus \mathcal{T}_0$ and (2.21) holds. It then follows from the reverse triangle inequality, (3.107c), and (3.106a)

that

$$\|P_k^{-1}d_k\|_2 \geq \|P_k^{-1}t_k\|_2 - \|P_k^{-1}n_k\|_2 = \left(1 - \frac{\|P_k^{-1}n_k\|_2}{\|P_k^{-1}t_k\|_2}\right)\|P_k^{-1}t_k\|_2 \geq \left(\frac{\kappa_{\mathcal{VS}} - 1}{\kappa_{\mathcal{VS}}}\right)\|P_k^{-1}t_k\|_2. \quad (3.108)$$

We also have that

$$-\Delta m_k^{f,n} = \nabla f(x_k, s_k)^T n_k + \tfrac{1}{2} n_k^T G_k n_k = \big(P_k \nabla f(x_k, s_k)\big)^T P_k^{-1} n_k + \tfrac{1}{2}(P_k^{-1}n_k)^T P_k G_k P_k (P_k^{-1}n_k). \quad (3.109)$$

Using the triangle and Cauchy-Schwarz inequalities, Lemma 3.2, and the fact that (2.21), (3.107b) and (3.106b) imply $\|P_k^{-1}n_k\|_2 \leq \min\{\delta_k^v, \delta_k^f\} \leq 1$, we then have

$$|\Delta m_k^{f,n}| \leq \kappa_{\text{ub}}(\|P_k^{-1}n_k\|_2 + \tfrac{1}{2}\|P_k^{-1}n_k\|_2^2) \leq 2\kappa_{\text{ub}}\|P_k^{-1}n_k\|_2. \quad (3.110)$$

Moreover, it follows from the fact that $k \in \mathcal{T} \setminus \mathcal{T}_0$, Lemma 3.3(ii), (3.107a), (2.58), (3.107b) and (3.106b) that

$$\Delta m_k^{f,t} \geq \kappa_{\text{ct}}\epsilon_\pi \min\{\epsilon_\pi, (1-\kappa_{\text{B}})\delta_k^t, (1-\kappa_{\text{fbt}})\kappa_{\text{fbn}}\} = \kappa_{\text{ct}}\epsilon_\pi(1-\kappa_{\text{B}})\delta_k^t.$$

Combining this with (3.110), the fact that $k \in \mathcal{T} \setminus \mathcal{T}_0$, Lemma 3.4, (3.108) and (3.107c) and (3.106a) yields

$$\frac{|\Delta m_k^{f,n}|}{\Delta m_k^{f,t}} \leq \frac{2\kappa_{\text{ub}}\|P_k^{-1}n_k\|_2}{\kappa_{\text{ct}}\epsilon_\pi(1-\kappa_{\text{B}})\delta_k^t} \leq \frac{2\kappa_{\text{ub}}\|P_k^{-1}n_k\|_2}{\kappa_{\text{ct}}\epsilon_\pi(1-\kappa_{\text{B}})\|P_k^{-1}d_k\|_2} \leq \frac{2\kappa_{\text{ub}}\kappa_{\mathcal{VS}}}{\kappa_{\text{ct}}\epsilon_\pi(1-\kappa_{\text{B}})(\kappa_{\mathcal{VS}}-1)}\frac{\|P_k^{-1}n_k\|_2}{\|P_k^{-1}t_k\|_2} \leq 1 - \kappa_\delta.$$

Hence, (2.32) holds, which completes the proof. □

We next prove that at nearly feasible points, certain v-iterates are guaranteed to be successful.

Lemma 3.23. *If $k \in \mathcal{V} \cap \mathcal{D}$,*

$$\|P_k^{-1}t_k\|_2 \leq \varsigma_{\text{tn}}\|P_k^{-1}n_k\|_2, \quad (3.111)$$

and

$$v_k \leq \min\left\{\kappa_{\text{c}}, \frac{\kappa_{\Delta\text{c1}}}{\kappa_{\Delta\text{c2}}\kappa_{\text{J}}}, \frac{\kappa_{\Delta\text{c3}}}{\kappa_{\Delta\text{c2}}}, \frac{1-\kappa_{\text{fbn}}}{\kappa_{\text{J}}}, \frac{1-\kappa_{\text{fbn}}}{\kappa_{\Delta\text{c2}}\kappa_{\text{J}}}, \frac{\kappa_{\text{R1}}}{\kappa_{\text{R2}}\kappa_{\text{sRn}}\kappa_{\text{n}}\kappa_{\text{ub}}}, \frac{\kappa_{\text{J}}\kappa_{\text{R2}}\kappa_{\text{sRn}}(1-\eta_1)}{\kappa_{\text{C}}(1+\varsigma_{\text{tn}})^2\kappa_{\text{n}}\kappa_{\text{ub}}}\right\}, \quad (3.112)$$

then $k \in \mathcal{S}_v$ and $\delta_{k+1}^v \geq \delta_k^v$.

Proof. Consider $k \in \mathcal{V} \cap \mathcal{D}$ such that (3.111) and (3.112) hold. If $n_k = 0$, then (3.111) implies that $t_k = 0$, which in turn implies by part (vi) of Lemma 2.6 that $k \in \mathcal{Y}$. However, this contradicts the supposition that $k \in \mathcal{V}$, so we must have $n_k \neq 0$. In this case, part (ii) of Lemma 2.6 ensures that $k \in \mathcal{N}$, so that overall we have $k \in \mathcal{N} \cap \mathcal{V} \cap \mathcal{D}$.

To obtain a contradiction, suppose that $k \notin \mathcal{S}_v$, so that overall we have $k \in (\mathcal{N} \cap \mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$. This and the bound (3.112) imply that the results of Lemmas 3.20 and 3.21 hold, i.e., that (3.101) and (3.104) hold. Moreover, the fact that $k \in \mathcal{D}$ and Lemma 2.6(ix) imply that (2.48) holds. Using this and the facts that $n_k \neq 0$ and $k \in \mathcal{V} \setminus \mathcal{S}_v$, it follows from (2.54) that $\rho_k^v < \eta_1$. However, since (3.101) and (3.104) hold,

$$\Delta m_k^{v,d} \geq \kappa_{\text{J}} \min\{\kappa_{\text{R1}}, \kappa_{\text{R2}}\|\text{Proj}_k(P_k^{-1}d_k)\|_2\} \geq \kappa_{\text{J}} \min\{\kappa_{\text{R1}}, \kappa_{\text{R2}}\kappa_{\text{sRn}}\|P_k^{-1}n_k\|_2\}.$$

In fact, it follows from (2.13), Lemma 3.2 and (3.112) that

$$\kappa_{\text{R2}}\kappa_{\text{sRn}}\|P_k^{-1}n_k\|_2 \leq \kappa_{\text{R2}}\kappa_{\text{sRn}}\kappa_{\text{n}}\pi_k^v \leq \kappa_{\text{R2}}\kappa_{\text{sRn}}\kappa_{\text{n}}\kappa_{\text{ub}}v_k \leq \kappa_{\text{R1}},$$

and thus

$$\Delta m_k^{v,d} \geq \kappa_{\text{J}}\kappa_{\text{R2}}\kappa_{\text{sRn}}\|P_k^{-1}n_k\|_2. \quad (3.113)$$

Furthermore, by (2.49), (3.70), (3.113), the triangle inequality, (3.111), (2.13), the Cauchy-Schwarz inequality, Lemma 3.2, and (3.112), we have that

$$|\rho_k^v - 1| = \left| \frac{v(x_k + d_k^x, s_k + d_k^s) - m_k^v(d_k)}{\Delta m_k^{v,d}} \right| \leq \frac{\kappa_{\rm C} \|P_k^{-1} d_k\|_2^2}{\kappa_{\rm J}\kappa_{\rm R2}\kappa_{\rm sRn}\|P_k^{-1} n_k\|_2}$$
$$\leq \frac{\kappa_{\rm C}(1+\varsigma_{\rm tn})^2\|P_k^{-1} n_k\|_2}{\kappa_{\rm J}\kappa_{\rm R2}\kappa_{\rm sRn}} \leq \frac{\kappa_{\rm C}(1+\varsigma_{\rm tn})^2\kappa_{\rm n}\kappa_{\rm ub}}{\kappa_{\rm J}\kappa_{\rm R2}\kappa_{\rm sRn}}\, v_k \leq 1 - \eta_1,$$

and hence $\rho_k^v \geq \eta_1$, which is a contradiction. Thus, we must conclude that $k \in \mathcal{S}_v$. The fact that $\delta_{k+1}^v \geq \delta_k^v$ now follows from the fact that $k \in \mathcal{S}_v$ and (2.52). □

We now prove that our algorithm terminates finitely if there are finitely many successful v-iterations.

Lemma 3.24. *If $|\mathcal{S}_v| < \infty$, Algorithm 1 terminates finitely.*

Proof. We prove the result by contradiction, and so suppose that $|\mathcal{S}_v| < \infty$, but that Algorithm 1 does not terminate finitely. It then follows from Theorem 3.11 that $|\mathcal{S}| = \infty$, which when combined with the fact that $|\mathcal{S}_v| < \infty$ implies that $|\mathcal{S}_f| = \infty$; i.e., it follows that there are an infinite number of successful iterations, and all belong to $\mathcal{S}_f$ for all sufficiently large k. We may also deduce from these facts—and since the barrier function is decreased for $k \in \mathcal{S}_f$ and the slack reset only possibly decreases the barrier function—that the sequence $\{f(x_k, s_k)\}$ is monotonically decreasing for sufficiently large k. Moreover, since $v_{k+1}^{\max} \leftarrow v_k^{\max}$ for all $k \notin \mathcal{S}_v$ and $|\mathcal{S}_v| < \infty$, we have that there exists a constant $v_\infty^{\max} > 0$ such that

$$v_k^{\max} = v_\infty^{\max} > 0 \text{ for all sufficiently large } k. \tag{3.114}$$

We complete the proof by considering two cases depending on whether, for some $\epsilon_f > 0$, (3.81) holds.
Case 1: Suppose that (3.81) holds for some $\epsilon_f > 0$. It then follows from Lemma 3.9 that (3.82) also holds, in which case we have from (2.30a)/(2.35a), the fact that $\mathcal{S}_f \subseteq \mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, Lemma 3.3(ii), (3.81), (3.82), (2.58), and (3.114) that

$$\begin{aligned}\Delta m_k^{f,t} &\geq \kappa_{\rm ct}\epsilon_f \min\{\epsilon_f, (1-\kappa_{\rm B})\delta_k^t, (1-\kappa_{\rm fbt})\kappa_{\rm fbn}\} \\ &\geq \kappa_{\rm ct}\epsilon_f \min\{\epsilon_f, (1-\kappa_{\rm B})\min\{\delta_k^v, \epsilon_{\mathcal{F}}, \kappa_{\rm v} v_\infty^{\max}\}, (1-\kappa_{\rm fbt})\kappa_{\rm fbn}\} \text{ for sufficiently large } k \in \mathcal{S}_f.\end{aligned} \tag{3.115}$$

We now consider two subcases, deriving contradictions in each, which will prove that the condition of this case (namely, that there exists $\epsilon_f > 0$ such that (3.81) holds) cannot occur.
Subcase 1.1: Suppose there exists an infinite subsequence $\mathcal{K}_f \subseteq \mathcal{S}_f$ such that $\{\delta_k^v\}_{k\in\mathcal{K}_f} \to 0$. Since $\delta_{k+1}^v < \delta_k^v$ only if $k \in \mathcal{V}\setminus\mathcal{S}_v$ and $\delta_{k+1}^v \leftarrow \delta_k^v$ otherwise, it follows that there exists an infinite subsequence $\mathcal{K}_v \subseteq \mathcal{V}\setminus\mathcal{S}_v$ such that $\{\delta_k^v\}_{k\in\mathcal{K}_v} \to 0$. Our goal in the remainder of this subcase is to prove that for all sufficiently large $k \in \mathcal{K}_v \subseteq \mathcal{V}$, we have that all of the conditions of an f-iteration are satisfied, which is a contradiction since $\mathcal{V} \cap \mathcal{F} = \emptyset$. This will prove that such a sequence $\mathcal{K}_f \subseteq \mathcal{S}_f$ cannot exist.

Using the fact that $\{\delta_k^v\}_{k\in\mathcal{K}_v} \to 0$ and Lemma 3.6, we may conclude that for all sufficiently large $k \in \mathcal{K}_v$ we have $k \in (\mathcal{V}\cap\mathcal{D})\setminus\mathcal{S}_v$. In addition, since $|\mathcal{S}_v| < \infty$ and $\{\delta_k^v\}_{k\in\mathcal{K}_v} \to 0$, we may conclude from part (ii) of Lemma 3.8 and Lemma 3.16 that $\{\pi_k^v\}_{k\in\mathcal{K}_v} \to 0$, which in turn implies with Lemma 3.15 that $\{v_k\}_{k\in\mathcal{K}_v} \to 0$. Now, suppose that there exists an infinite subsequence $\mathcal{K}_v' \subseteq \mathcal{K}_v$ such that $\mathcal{K}_v' \cap \mathcal{N} = \emptyset$. The following then hold for all sufficiently large $k \in \mathcal{K}_v' \subseteq \mathcal{K}_v \subseteq \mathcal{V}\setminus\mathcal{S}_v$:

(a) $n_k = 0$ by part (ii) of Lemma 2.6 (and thus (2.32) holds);

(b) $t_k \neq 0$ by (a), part (vi) of Lemma 2.6, and the fact that $k \in \mathcal{V}$; and

(c) $v_k < \kappa_{\rm vv} v_k^{\max} = \kappa_{\rm vv} v_\infty^{\max}$ by Step 9, (2.10), and (3.114).

It then follows from Assumption 1.1, Lemma 3.4, the fact that $\{\delta_k^v\}_{k\in\mathcal{K}'_v} \to 0$, statement (c) above, and the bound $\kappa_{\text{vv}} < 1$ that $v(x_k + d_k^x, s_k + d_k^s) \leq v_k^{\max}$ for all sufficiently large $k \in \mathcal{K}'_v$. Overall, this yields (2.37), and thus we have that all of the conditions of an f-iteration hold, so $k \in \mathcal{F}$. However, this is a contradiction since $k \in \mathcal{K}'_v \subseteq \mathcal{V}$ and $\mathcal{V} \cap \mathcal{F} = \emptyset$. Thus, such an infinite subsequence $\mathcal{K}'_v \subseteq \mathcal{K}_v$ cannot exist, so we may conclude that for all sufficiently large $k \in \mathcal{K}_v$ we have $k \in \mathcal{N}$. To summarize, at this point in this subcase, we may assume without loss of generality that there exists an infinite subsequence $\mathcal{K}_v \subseteq (\mathcal{N} \cap \mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$ over which $\{\delta_k^v\}_{k\in\mathcal{K}_v} \to 0$, $\{\pi_k^v\}_{k\in\mathcal{K}_v} \to 0$, and $\{v_k\}_{k\in\mathcal{K}_v} \to 0$.

It follows from Lemma 3.21 and the facts that $\mathcal{K}_v \subseteq (\mathcal{N} \cap \mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$ and $\{v_k\}_{k\in\mathcal{K}_v} \to 0$ that $m_k^v(d_k) \leq \kappa_{\text{cld}} v_k$ for all sufficiently large $k \in \mathcal{K}_v$. Using this fact, (3.70), the reverse triangle inequality, Lemma 3.4, Lemma 2.9, and (3.114), we have

$$v(x_k^+, s_k^+) \leq \kappa_{\text{cld}} v_\infty^{\max} + \kappa_{\text{C}} (\delta_k^v)^2 \text{ for all sufficiently large } k \in \mathcal{K}_v.$$

This relationship then implies that

$$v(x_k^+, s_k^+) \leq v_\infty^{\max} = v_k^{\max} \text{ for all sufficiently large } k \in \mathcal{K}_v \text{ such that } (\delta_k^v)^2 \leq \frac{(1-\kappa_{\text{cld}})}{\kappa_{\text{C}}} v_\infty^{\max}.$$

Thus, since $\{\delta_k^v\}_{k\in\mathcal{K}_v} \to 0$, we may conclude that (2.37) holds for all sufficiently large $k \in \mathcal{K}_v$.

Next, suppose that for $\varsigma_{\text{tn}} > 0$ defined in (3.106a), we have

$$\|P_k^{-1} t_k\|_2 \leq \varsigma_{\text{tn}} \|P_k^{-1} n_k\|_2 \text{ for all sufficiently large } k \in \mathcal{K}_v. \tag{3.116}$$

We may then use the facts that $\mathcal{K}_v \subseteq (\mathcal{N} \cap \mathcal{V} \cap \mathcal{D})$ and $\{v_k\}_{k\in\mathcal{K}_v} \to 0$, (3.116), and Lemma 3.23 to conclude that $|\mathcal{S}_v \cap \mathcal{K}_v| = \infty$, which contradicts the fact that $|\mathcal{S}_v| < \infty$. Therefore, there exists an infinite subsequence $\mathcal{K}''_v \subseteq \mathcal{K}_v$ such that if $k \in \mathcal{K}''_v$ then (3.116) fails.

We now show that with $k \in \mathcal{K}''_v \subseteq \mathcal{K}_v \subseteq \mathcal{V} \setminus \mathcal{S}_v$, the conditions of Lemma 3.22 hold. Consider $k \in \mathcal{K}''_v$. First, since $k \in \mathcal{K}''_v \subseteq \mathcal{V}$, we know that $k \notin \mathcal{Y}$. Second, since $k \in \mathcal{K}''_v$, we know from the previous paragraph that (3.116) does not hold, and therefore that $t_k \neq 0$ and r_k was computed to satisfy (2.26a), (2.26b), or (2.26c). Since we have supposed that the algorithm does not terminate finitely, we may use the fact that $\{v_k\}_{k\in\mathcal{K}_v} \to 0$ along with (2.26a) to conclude that (3.107a) holds for all sufficiently large $k \in \mathcal{K}''_v$. Third, since $\{\delta_k^v\}_{k\in\mathcal{K}_v} \to 0$, we have that (3.107b) holds for all sufficiently large $k \in \mathcal{K}''_v$. Fourth, we know from the definition of the set $\mathcal{K}''_v$ that (3.116) fails, which is to say that (3.107c) holds. We may now apply Lemma 3.22 to deduce that $t_k \neq 0$ and (2.32) holds for all sufficiently large $k \in \mathcal{K}''_v$. Thus, along with our previous conclusion that (2.37) holds for all sufficiently large $k \in \mathcal{K}_v$, we may conclude that for all sufficiently large $k \in \mathcal{K}''_v$ we have that all of the conditions of an f-iteration are satisfied. However, as previously mentioned, this is impossible since $\mathcal{K}''_v \subseteq \mathcal{K}_v \subseteq \mathcal{V}$ and $\mathcal{F} \cap \mathcal{V} = \emptyset$. Hence our stated supposition for Subcase 1.1, i.e., that there is infinite subsequence $\mathcal{K}_f \subseteq \mathcal{S}_f$ such that $\{\delta_k^v\}_{k\in\mathcal{K}_f} \to 0$, must be impossible.

Subcase 1.2: Suppose that there exists $\epsilon_* > 0$ such that $\delta_k^v \geq \epsilon_*$ for all $k \in \mathcal{S}_f$, and recall that $|\mathcal{S}_f| = \infty$. We may combine (3.115) and the bound $\delta_k^v \geq \epsilon_*$ for all $k \in \mathcal{S}_f$ to conclude that there exists k' such that

$$\Delta m_k^{f,t} \geq \kappa_{\text{ct}} \epsilon_f \min\{\epsilon_f, (1-\kappa_{\text{B}}) \min\{\epsilon_*, \epsilon_{\mathcal{F}}, \kappa_{\text{v}} v_\infty^{\max}\}, (1-\kappa_{\text{fbt}})\kappa_{\text{fbn}}\} > 0 \text{ for all } k \geq k' \text{ with } k \in \mathcal{S}_f. \tag{3.117}$$

Combining the facts that $|\mathcal{S}_v| < \infty$ and $|\mathcal{S}_f| = \infty$, (2.38), and (2.32) (which is required to hold for $k \in \mathcal{F}$), we have that

$$f(x_{k'}, s_{k'}) - f(x_k, s_k) = \sum_{j=k', j\in\mathcal{S}_f}^{k-1} [f(x_j, s_j) - f(x_{j+1}, s_{j+1})] \geq \eta_1 \kappa_\delta \sum_{j=k', j\in\mathcal{S}_f}^{k-1} \Delta m_j^{f,t}, \tag{3.118}$$

which in view of (3.117) proves that $\{f(x_k, s_k)\} \to -\infty$. However, this is a contradiction since the

barrier function is bounded below as a consequence of Lemma 3.2 and Assumptions 1.1 and 3.1.

Since we have proved that neither Subcase 1.1 nor 1.2 can occur, the premise of Case 1 cannot be true.

Case 2: Suppose that the condition of Case 1 does not hold, which is to say that there exists $\mathcal{K} \subseteq \mathcal{F}$ with

$$\lim_{k \in \mathcal{K}} \pi_k^f = 0. \tag{3.119}$$

For all $k \in \mathcal{K} \subseteq \mathcal{F} \subseteq \mathcal{T} \setminus \mathcal{T}_0$, we have that $t_k \neq 0$ was computed (and not reset to zero), in which case it must be true that (2.26b) does not hold. Combining this fact with (3.119) yields

$$0 = \lim_{k \in \mathcal{K}} \pi_k^f \geq \lim_{k \in \mathcal{K}} \omega_t(\pi_k^v) \geq 0, \quad \text{so that} \quad \lim_{k \in \mathcal{K}} \pi_k^v = 0.$$

It follows from this fact, Assumptions 3.1, 3.2, and 3.3, and Lemmas 3.2 and 3.15 that $\{v_k\}_{k \in \mathcal{K}} \to 0$, which when combined with (3.119) shows that (2.26a) will be satisfied for all sufficiently large $k \in \mathcal{K}$. However, this contradicts our supposition that the algorithm does not terminate finitely. □

The previous result proves that if the algorithm does not terminate finitely, then there are an infinite number of successful v-iterations. We now establish an important consequence of this fact.

Lemma 3.25. *If $|\mathcal{S}_v| = \infty$ and* (3.111) *holds for all sufficiently large $k \in \mathcal{V} \cap \mathcal{D}$, then*

$$\delta_k^v \geq \epsilon_* \text{ for some } \epsilon_* > 0 \text{ for all } k. \tag{3.120}$$

Proof. First, by Lemma 3.17, the fact that $|\mathcal{S}_v| = \infty$ implies that $\{v_k\} \to 0$. Hence, for sufficiently large $k \in \mathcal{V} \cap \mathcal{D}$, we have that (3.111) and (3.112) hold, which in turn implies by Lemma 3.23 that $\delta_{k+1}^v \geq \delta_k^v$. Second, if $k \in \mathcal{V} \setminus \mathcal{D}$, then it follows from Lemma 3.6 that $\delta_k^v \geq \min\{\delta_k^v, \delta_k^f, \kappa_v v_k^{\max}\} > \kappa_{\mathcal{V}}$. Third, if $k \in \mathcal{Y} \cup \mathcal{F}$, then by (2.36), (2.42) and (2.45) we have that $\delta_{k+1}^v \geq \delta_k^v$. The result follows by combining these facts. □

We next prove a result about certain v-iterations that are unsuccessful.

Lemma 3.26. *If $k \in \mathcal{V} \setminus \mathcal{S}_v$,* (3.103) *holds,*

$$v_k^{\max} \leq \min\left\{ \left(\frac{1 - \kappa_{\text{cld}}}{\kappa_{\text{C}}} \right)^2, \left(\frac{1 - \kappa_{\text{vv}}}{\kappa_{\text{C}}} \right)^2, \kappa_{\mathcal{V}}^{\frac{4}{3}} \right\}, \tag{3.121}$$

and

$$\delta_k^v \leq (v_k^{\max})^{\frac{3}{4}} \tag{3.122}$$

then $k \in \mathcal{D}$ and (2.37) *holds.*

Proof. Let $k \in \mathcal{V} \setminus \mathcal{S}_v$ and observe that (3.121) and (3.122) imply that $\delta_k^v \leq \kappa_{\mathcal{V}}$. Hence, by Lemma 3.6, we have that $k \in \mathcal{D}$. That is, $k \in (\mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$. We now consider two cases depending on whether or not $k \in \mathcal{N}$.

Suppose $k \in \mathcal{N}$ so that $k \in (\mathcal{N} \cap \mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$. It then follows from (3.70), the reverse triangle inequality, the fact that (3.103) holds, and Lemmas 3.4 and 3.21 that

$$v(x_k + d_k^x, s_k + d_k^s) \leq \kappa_{\text{cld}} v_k + \kappa_{\text{C}} (\delta_k^v)^2.$$

Then, from this inequality, Lemma 2.9, (3.122), and (3.121), we have that

$$v(x_k + d_k^x, s_k + d_k^s) \leq \kappa_{\text{cld}} v_k^{\max} + \kappa_{\text{C}} (v_k^{\max})^{\frac{3}{2}} = v_k^{\max} \left(\kappa_{\text{cld}} + \kappa_{\text{C}} \sqrt{v_k^{\max}} \right) \leq v_k^{\max},$$

which means that (2.37) holds, as desired.

Now suppose $k \notin \mathcal{N}$. It then follows from (3.70), the reverse triangle inequality, Lemmas 3.4 and 2.9, (2.30d) (which holds since $k \in \mathcal{D}$), and the fact that $v_k < \kappa_{\text{vv}} v_k^{\max}$ (which holds by (2.10) since $k \notin \mathcal{N}$), (3.121), and (3.122) that

$$v(x_k + d_k^x, s_k + d_k^s) \leq m_k^v(d_k) + \kappa_{\text{C}}(\delta_k^v)^2 \leq \kappa_{\text{vv}} v_k^{\max} + \kappa_{\text{C}} \left(v_k^{\max}\right)^{\frac{3}{2}} \leq v_k^{\max} \left(\kappa_{\text{vv}} + \kappa_{\text{C}} \sqrt{v_k^{\max}}\right) \leq v_k^{\max},$$

which again means that (2.37) holds, as desired. □

We now come to the conclusion that there are a finite number of successful v-iterations.

Theorem 3.27. *The set $\mathcal{S}_v$ is finite.*

Proof. We prove the result by contradiction, and so suppose that $|\mathcal{S}_v| = \infty$. It then follows from Lemma 3.17 that $\{v_k^{\max}\} \to 0$, $\{v_k\} \to 0$, $\{\pi_k^v\} \to 0$, and $\{n_k\} \to 0$. Moreover, from the fact that $|\mathcal{S}_v| = \infty$, we have that (2.26a) must not hold for all sufficiently large k, or else the algorithm would terminate finitely in Step 17 or 32, which is a contradiction. Thus, since $\{v_k\} \to 0$, we have

$$\pi_k^f \geq \epsilon_\pi > 0 \text{ for all sufficiently large } k. \tag{3.123}$$

It follows from this fact and Lemma 3.9 that (3.82) holds. Also it follows from the facts that $\{v_k\} \to 0$, $\{v_k^{\max}\} \to 0$, and $|\mathcal{S}_v| = \infty$ that there exists k_0 such that (3.103), (3.112), and (3.121) hold for all $k \geq k_0$.

We now prove a lower bound for δ_k^v that holds for all sufficiently large k, written as equation (3.127) below. We prove the bound by considering two cases, the latter of which is composed of two subcases.
Case 1: Suppose that (3.111) holds for all sufficiently large $k \geq k_0$ such that $k \in \mathcal{V} \cap \mathcal{D}$. Then, since $|\mathcal{S}_v| = \infty$, we may apply Lemma 3.25 to deduce that (3.120) holds for all sufficiently large k.
Case 2: Suppose that the situation in Case 1 does not occur in that there exists an infinite index set

$$\mathcal{K}_1 := \{k \geq k_0 : k \in \mathcal{V} \cap \mathcal{D} \text{ and } \|P_k^{-1} t_k\|_2 > \varsigma_{\text{tn}} \|P_k^{-1} n_k\|_2 \}.$$

Since δ_k^v ($v_k^{\max}$) is not decreased (increased) for $k \in \mathcal{S}_v \cup \mathcal{Y} \cup \mathcal{F}$, our goal is to provide a lower bound for δ_k^v over $k \in \mathcal{K}_1 \setminus \mathcal{S}_v$. We do this by considering two subcases depending on whether or not $k \in \mathcal{N}$.
Subcase 1: Consider k such that $k_0 \leq k \in \mathcal{K}_1 \setminus (\mathcal{S}_v \cup \mathcal{N})$. Since $k \notin \mathcal{N}$, it follows from part (ii) of Lemma 2.6 that $n_k = 0$. By part (vi) of Lemma 2.6, this means that $t_k \neq 0$ (since otherwise we would have $k \in \mathcal{Y}$), which in turn means by part (v) of Lemma 2.6 that $k \in \mathcal{T} \setminus \mathcal{T}_0$ and that (2.32) holds (since $n_k = 0$). We may then conclude from the fact that $k \in \mathcal{V} \setminus \mathcal{S}_v$, the choice of k_0 being large enough such that (3.103) and (3.121) hold for $k \geq k_0$, and Lemma 3.26 that if (3.122) holds, then (2.37) also holds. However, this would imply that $k \in \mathcal{F}$, which contradicts the definition of $\mathcal{K}_1$ since $\mathcal{V} \cap \mathcal{F} = \emptyset$. Thus, (3.122) must not hold and

$$\delta_k^v > (v_k^{\max})^{\frac{3}{4}} \text{ for all } k \text{ such that } k_0 \leq k \in \mathcal{K}_1 \setminus (\mathcal{S}_v \cup \mathcal{N}). \tag{3.124}$$

Subcase 2: Consider k such that $k_0 \leq k \in (\mathcal{K}_1 \cap \mathcal{N}) \setminus \mathcal{S}_v$. By (3.123), we have that (3.107a) holds. Similarly, by the definition of $\mathcal{K}_1$, we have that (3.107c) holds. Now suppose that (3.107b) and (3.122) both hold. Then, since $k \notin \mathcal{Y}$ and (3.107a), (3.107b), and (3.107c) all hold, we may apply Lemma 3.22 to conclude that $t_k \neq 0$ and (2.32) holds. Also, since $k \in \mathcal{V} \setminus \mathcal{S}_v$, we have shown that (3.103) and (3.121) hold, and we have supposed that (3.122) holds, we may apply Lemma 3.26 to conclude that (2.37) holds. Overall, we have shown that all of the conditions of an f-iteration are satisfied so that $k \in \mathcal{F}$. However, this contradicts the fact that $k \in \mathcal{K}_1 \subseteq \mathcal{V}$ and $\mathcal{V} \cap \mathcal{F} = \emptyset$. Therefore, we may deduce that at least one of (3.107b) or (3.122) must not hold, yielding

$$\delta_k^v > \min\left\{\varsigma_\delta, (v_k^{\max})^{\frac{3}{4}}\right\} \text{ for all } k \text{ such that } k_0 \leq k \in (\mathcal{K}_1 \cap \mathcal{N}) \setminus \mathcal{S}_v. \tag{3.125}$$

Combining (3.124) and (3.125) from Subcases 1 and 2 shows that, for Case 2, we have

$$\delta_k^v \geq \min\left\{\varsigma_\delta, (v_k^{\max})^{\frac{3}{4}}\right\} \text{ for all } k \text{ such that } k_0 \leq k \in \mathcal{K}_1 \setminus \mathcal{S}_v. \tag{3.126}$$

Moreover, the fact that $\{v_k\} \to 0$ and Lemma 3.23 implies that for any k with $k_0 \leq k \in (\mathcal{V} \cap \mathcal{D}) \setminus \mathcal{K}_1$, we have $k \in \mathcal{S}_v$. Thus, for all $k \geq k_0$ with $k \in (\mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$, we have $k \in \mathcal{K}_1 \setminus \mathcal{S}_v$. As a result, the inequality in (3.126) holds for all k with $k_0 \leq k \in (\mathcal{V} \cap \mathcal{D}) \setminus \mathcal{S}_v$. This conclusion, along with the deduction that $\delta_k^v > \kappa_\nu$ for all $k \in \mathcal{V} \setminus \mathcal{D}$ from Lemma 3.6 yields

$$\delta_k^v \geq \min\left\{\varsigma_\delta, (v_k^{\max})^{\frac{3}{4}}, \kappa_\nu\right\} \text{ for all } k \text{ with } k_0 \leq k \in \mathcal{V} \setminus \mathcal{S}_v,$$

which, when combined with the fact that δ_k^v ($v_k^{\max}$) is not decreased (increased) for $k \in \mathcal{S}_v \cup \mathcal{Y} \cup \mathcal{F}$, yields

$$\delta_k^v \geq \min\left\{\varsigma_\delta, (v_k^{\max})^{\frac{3}{4}}, \kappa_\nu\right\} \text{ for all } k \geq k_0.$$

Combining the results of Cases 1 and 2, we have that

$$\delta_k^v \geq \min\left\{\epsilon_*, \varsigma_\delta, (v_k^{\max})^{\frac{3}{4}}, \kappa_\nu\right\} \text{ for all sufficiently large } k. \tag{3.127}$$

Using this fact, (3.82), and the fact that $\{v_k^{\max}\} \to 0$ yields

$$\min\{\delta_k^v, \delta_k^f\} \geq \min\left\{\epsilon_*, \varsigma_\delta, (v_k^{\max})^{\frac{3}{4}}, \kappa_\nu, \epsilon_{\mathcal{F}}\right\} = (v_k^{\max})^{\frac{3}{4}} \text{ for all sufficiently large } k. \tag{3.128}$$

Under our supposition that the set $\mathcal{S}_v$ is infinite, at least one of the following two scenarios must occur. In both, we reach a contradiction to this supposition that $\mathcal{S}_v$ is infinite, which proves the theorem.

Scenario 1: Suppose that $\mathcal{S}_1 := \mathcal{S}_v \setminus \mathcal{T}$ is infinite. For $k \in \mathcal{S}_1$, we have that either (2.21) does not hold or (2.26b) holds. In fact, since (3.123) holds and $\{\pi_k^v\} \to 0$, condition (2.26b) cannot hold infinitely often for $k \in \mathcal{S}_1$, implying that for all sufficiently large $k \in \mathcal{S}_1$ we have that (2.21) does not hold. Then, since $t_k = 0$ for $k \in \mathcal{S}_1 \subseteq \mathcal{V}$, we have by Lemma 2.6(vi) that $n_k \neq 0$ (or else $k \in \mathcal{Y}$). We may now use the facts that $v_k^{\max} > 0$, $\delta_k^v > 0$, and $\delta_k^f > 0$ for all k, (2.13), (3.128), Lemmas 2.9 and 3.2, and the fact that $\{v_k\} \to 0$ to conclude that

$$\frac{\|P_k^{-1} n_k\|_2}{\min\{\delta_k^v, \delta_k^f\}} \leq \frac{\kappa_{\rm n}\pi_k^v}{(v_k^{\max})^{\frac{3}{4}}} \leq \frac{\kappa_{\rm n}\kappa_{\rm ub} v_k}{(v_k)^{\frac{3}{4}}} \leq \kappa_{\rm n}\kappa_{\rm ub} v_k^{\frac{1}{4}} \leq \kappa_{\rm B} \text{ for sufficiently large } k \in \mathcal{S}_1.$$

However, this means that (2.21) holds for all sufficiently large $k \in \mathcal{S}_1$, contradicting our earlier conclusion that it does not. This contradiction implies that this scenario cannot occur.

Scenario 2: Suppose that $\mathcal{S}_2 = \mathcal{S}_v \cap \mathcal{T}$ is infinite. Our goal is to show that for all sufficiently large $k \in \mathcal{S}_2$, we have that all of the conditions of an f-iteration are satisfied, which is impossible since $\mathcal{S}_2 \subseteq \mathcal{V}$ and $\mathcal{V} \cap \mathcal{F} = \emptyset$. We begin by showing that (2.32) holds for all sufficiently large $k \in \mathcal{S}_2$. Using (3.109), the triangle and Cauchy-Schwarz inequalities, Lemma 3.2, (2.13), and the fact that $\{\pi_k^v\} \to 0$ (and hence that $\kappa_{\rm n}\pi_k^v \leq 1$ for all sufficiently large k), it follows as in the proof of Lemma 3.22 (see (3.110)) that

$$|\Delta m_k^{f,n}| \leq \kappa_{\rm ub}(\|P_k^{-1} n_k\|_2 + \tfrac{1}{2}\|P_k^{-1} n_k\|_2^2) \leq 2\kappa_{\rm ub}\kappa_{\rm n}\pi_k^v \leq 2\kappa_{\rm ub}^2\kappa_{\rm n} v_k \text{ for all sufficiently large } k \in \mathcal{S}_2. \tag{3.129}$$

It also follows from the facts that $\{v_k^{\max}\} \to 0$ and $\mathcal{S}_2 \subseteq \mathcal{V}$ along with Lemma 3.6 that $k \in \mathcal{D}$ for all sufficiently large $k \in \mathcal{S}_2$. Moreover, since $\mathcal{S}_2 \subseteq \mathcal{T}$, it follows that for all $k \in \mathcal{S}_2$ a tangential step $t_k \neq 0$ was computed to satisfy either (2.30) or (2.35). However, for all $k \in \mathcal{S}_2$, it follows from (2.48) that $n_k \neq 0$, and then from Lemma 2.6(xi) that $k \in \mathcal{T}_{\mathcal{D}}$, i.e., that (2.30) holds. This implies by (2.58) that $\delta_k^t = \min\{\delta_k^v, \delta_k^f\}$ for sufficiently large $k \in \mathcal{S}_2$. It follows from this fact, the fact that $k \in \mathcal{T}_{\mathcal{D}}$, (2.30a),

part (ii) of Lemma 3.3, (3.123), (3.128), the fact that $\{v_k^{\max}\} \to 0$, and Lemma 2.9 that

$$\begin{aligned}\Delta m_k^{f,t} &\geq \kappa_{\rm ct}\epsilon_\pi \min\left\{\epsilon_\pi, (1-\kappa_{\rm B})\delta_k^t, (1-\kappa_{\rm fbt})\kappa_{\rm fbn}\right\} \\ &= \kappa_{\rm ct}\epsilon_\pi \min\left\{\epsilon_\pi, (1-\kappa_{\rm B})\min\{\delta_k^v, \delta_k^f\}, (1-\kappa_{\rm fbt})\kappa_{\rm fbn}\right\} \\ &\geq \kappa_{\rm ct}\epsilon_\pi(1-\kappa_{\rm B})(v_k^{\max})^{\frac{3}{4}} \geq \kappa_{\rm ct}\epsilon_\pi(1-\kappa_{\rm B})v_k^{\frac{3}{4}} \text{ for all sufficiently large } k \in \mathcal{S}_2.\end{aligned}$$

Combining this with (3.129) and the fact that $\{v_k\} \to 0$ shows that

$$\frac{|\Delta m_k^{f,n}|}{\Delta m_k^{f,t}} \leq \frac{2\kappa_{\rm ub}^2\kappa_{\rm n}v_k^{\frac{1}{4}}}{\kappa_{\rm ct}\epsilon_\pi(1-\kappa_{\rm B})} \leq 1-\kappa_\delta \text{ for all sufficiently large } k \in \mathcal{S}_2.$$

Hence, (2.32) holds for sufficiently large $k \in \mathcal{S}_2$, as desired. From here, it follows from Step 27 that the computed tangential step is not reset to zero, i.e., $k \in \mathcal{T}_{\mathcal{D}} \setminus \mathcal{T}_0$ for all sufficiently large $k \in \mathcal{S}_2$, from which it follows that $t_k \neq 0$ for all sufficiently large $k \in \mathcal{S}_2$. Moreover, since $k \in \mathcal{S}_v$ implies by Lemma 2.9 that (2.37) holds, we have from the fact that $\mathcal{S}_2 \subseteq \mathcal{S}_v$ that (2.37) holds for all $k \in \mathcal{S}_2$. To summarize, we have shown that for all sufficiently large $k \in \mathcal{S}_2$, all conditions of an f-iteration are satisfied, which is a contradiction. Thus, this scenario cannot occur.

Overall, we have shown that under our supposition that $|\mathcal{S}_v| = \infty$, neither Scenario 1 nor 2 may occur. However, since one of the two must occur in order to have $|\mathcal{S}_v| = \infty$, we have reached a contradiction to our supposition, meaning that the result is proved. □

We conclude by summarizing our convergence results.

Theorem 3.28. *The following hold for Algorithm 1:*

(i) *If Assumptions* 1.1, 2.1, *and* 3.1 *hold, then either Algorithm* 1 *terminates finitely or there exists an infinite subsequence* $\mathcal{K}$ *such that* $\lim_{k\in\mathcal{K}} \min\{v_k, \chi_k^v\} = \lim_{k\in\mathcal{K}} \pi_k^v = 0$. *In the latter case, any limit point* (x_*, s_*) *of* $\{(x_k, s_k)\}_{k\in\mathcal{K}}$ *satisfies* $\pi^v(x_*, s_*) = 0$ *and is therefore a critical point of* $\frac{1}{2}v(x,s)^2$ *subject to* $s \geq 0$.

(ii) *If Assumptions* 1.1, 2.1, 3.1, *and* 3.2 *hold, then either Algorithm* 1 *terminates finitely or there exists an infinite subsequence* $\mathcal{K}$ *such that* $\lim_{k\in\mathcal{K}} \min\{v_k, \chi_k^v\} = \lim_{k\in\mathcal{K}} \pi_k^v = 0$. *In the latter case, any limit point* (x_*, s_*) *of* $\{(x_k, s_k)\}_{k\in\mathcal{K}}$ *satisfies* $v(x_*, s_*) = 0$ *so that* (x_*, s_*) *is feasible for* (NPs).

(iii) *If Assumptions* 1.1, 2.1, 3.1, 3.2, *and* 3.3 *hold, then either Algorithm* 1 *terminates finitely in Step 8 with an infeasible stationary point* (x_k, s_k) *with* $v_k > \kappa_{\rm c}$ *or it terminates finitely in Step 17 or 32 with an approximate first-order KKT point* (x_k, s_k, y_k) *for the barrier problem* (1.1).

Proof. Part (i) follows from Lemmas 3.13 and 3.16, Assumption 1.1, and the criticality conditions (3.93) for minimizing $\frac{1}{2}v(x,s)^2$ subject to $s \geq 0$. Then, part (ii) follows from part (i) and Lemma 3.15. Finally, it follows from Theorems 3.27 and 3.24 that Algorithm 1 terminates finitely. Thus, part (iii) follows since, under Assumption 3.3, the algorithm does not converge to an infeasible stationary point with $v_k \leq \kappa_{\rm c}$. □

4 A Trust-Funnel Algorithm for Solving the Nonlinear Optimization Problem

The previous section considers the global convergence properties of our new trust-funnel algorithm when applied to the barrier subproblem (1.1). This section describes how a sequence of barrier subproblems with decreasing values for the barrier parameter may be solved to find an approximate first-order KKT point for problem (NP) (equivalently, problem (NPs)).

To achieve our stated goal, we require the constants ϵ_π and ϵ_v in Algorithm 1 to depend on μ. Moreover, for practical reasons, it is advisable to make other constants in Algorithm 1 depend on μ as well. In the previous section, for ease of exposition, we did not explicitly state these dependencies since μ was fixed. This does not pose a problem in this section since we use Algorithm 1 to solve a sequence of barrier problems where for each particular instance the penalty parameter is fixed and therefore our previous analysis still holds. A summary of the constants that depend on μ and precisely where they are used is given in Table 4.1. In addition to requiring them to be positive, it is practical to have them satisfy

$$\lim_{\mu\to 0}\epsilon_\pi(\mu) = \lim_{\mu\to 0}\epsilon_v(\mu) = \lim_{\mu\to 0}\kappa_{\rm fbn}(\mu) = \lim_{\mu\to 0}\kappa_{\rm fbt}(\mu) = 0 \quad \text{and} \tag{4.130}$$

$$\lim_{\mu\to 0}\kappa_{\rm y}(\mu) = \lim_{\mu\to 0}\kappa_{\rm D}(\mu) = \infty. \tag{4.131}$$

Moreover, the convergence result that we present in this section additionally assumes that

$$\epsilon_\pi(\mu_j) \le \zeta_1\mu_j \quad \text{and} \quad \epsilon_v(\mu_j) \le \zeta_2\mu_j \tag{4.132}$$

for some chosen constants $\zeta_1 \in (0,1)$ and $\zeta_2 \in (0,\infty)$, and that a particular choice for the positive-definite matrix D_k defined in (2.20) is used. Specifically, for each $1 \le i \le m$, we define

$$[d_k]_i := [D_k]_{ii} := \begin{cases} \kappa_{\rm D}(\mu_j) & \text{if } \mu_j[s_k]_i^{-2} > \kappa_{\rm D}(\mu_j), \\ \mu_j[s_k]_i^{-2} & \text{otherwise.} \end{cases} \tag{4.133}$$

Other choices are possible, e.g., based on the primal-dual update $D_k = Y_kS_k^{-1}$, and only require a small modification in the proof.

With these requirements, we may now state our algorithm for solving problem (NPs).

Table 4.1: Parameters from Algorithm 1 that depend on the barrier parameter.

Parameter	Used	Parameter	Used	Parameter	Used
$\kappa_{\rm y} = \kappa_{\rm y}(\mu)$	(2.19)	$\kappa_{\rm D} = \kappa_{\rm D}(\mu)$	(2.20)	$\epsilon_\pi = \epsilon_\pi(\mu)$	(2.26a)
$\kappa_{\rm fbt} = \kappa_{\rm fbt}(\mu)$	(2.30b)/(2.35b)	$\kappa_{\rm fbn} = \kappa_{\rm fbn}(\mu)$	(2.7)/(2.13)	$\epsilon_v = \epsilon_v(\mu)$	(2.26a)

Algorithm 2 Trust-funnel algorithm for solving (NPs).

Input: (x_0, s_0, y_0, μ_0) satisfying $(s_0, y_0, \mu_0) > 0$.
Choose a parameter $\gamma_\mu \in (0,1)$ and any two forcing functions $\epsilon_\pi(\cdot)$ and $\epsilon_v(\cdot)$.
Set $(x_0^{\rm start}, s_0^{\rm start}, y_0^{\rm start}) \leftarrow (x_0, s_0, y_0)$ and $j \leftarrow 0$.
loop
  Call Algorithm 1 with input $(x_j^{\rm start}, s_j^{\rm start}, y_j^{\rm start}, \mu_j)$ and $\big(\epsilon_\pi(\mu_j), \epsilon_v(\mu_j)\big)$ to compute $(x_{j+1}, s_{j+1}, y_{j+1})$.
  if Algorithm 1 terminated in Step 8 **then**
    Return the infeasible stationary point (x_{j+1}, s_{j+1}).
  Set $\mu_{j+1} \in (0, \gamma_\mu\mu_j]$.
  Use μ_j, μ_{j+1}, and $(x_{j+1}, s_{j+1}, y_{j+1})$ to compute the next starting point $(x_{j+1}^{\rm start}, s_{j+1}^{\rm start}, y_{j+1}^{\rm start})$.
  Set $j \leftarrow j+1$.

Theorem 4.1. *If Assumptions* 1.1, 2.1, 3.1, 3.2, *and* 3.3 *hold, and both* (4.132) *and* (4.133) *hold, then either*

(i) *Algorithm* 2 *returns an infeasible stationary point in Step* 7, *or*

(ii) *there exists a limit point* (x_*, s_*, y_*) *of the iterates* $\{(x_{j+1}, s_{j+1}, y_{j+1})\}$ *computed by Algorithm* 2 *such that* (x_*, s_*, y_*) *is a first-order KKT point for problem* (NPs).

Proof. If statement (i) occurs, then there is nothing left to prove. Therefore, suppose that statement (i) does not occur, in which case we have that Algorithm 1 never terminates in Step 8, which by (2.26a) and (4.132) means that for all $j \geq 0$ we have

$$\pi^f_{j+1}(y_{j+1}) \leq \epsilon_\pi(\mu_j) \leq \zeta_1 \mu_j \quad \text{and} \quad v_{j+1} \leq \epsilon_v(\mu_j) \leq \zeta_2 \mu_j. \tag{4.134}$$

In particular, we have that the sequence $\{(x_{j+1}, s_{j+1}, y_{j+1})\}$ is infinite, and from the second part of (4.134), the reverse triangle inequality, and Assumption 3.1, that $\{s_{j+1}\}$ is bounded. Combining this fact with Assumption 3.1 implies the existence of an infinite index set $\mathcal{J}$ and a point (x_*, s_*) with $s_* \geq 0$ such that

$$\lim_{j \in \mathcal{J}} (x_{j+1}, s_{j+1}) = (x_*, s_*). \tag{4.135}$$

It follows from this fact, (4.134), $\mu_j \to 0$, and Assumption 1.1 that

$$\lim_{j \in \mathcal{J}} v_{j+1} = v(x_*, s_*) = 0. \tag{4.136}$$

We comment that for the remainder of the proof, the quantities P_{j+1}, n_{j+1}, etc. are used to represent the final values of the relevant quantities computed in Algorithm 1 when it is called in line 5 during iteration j of Algorithm 2; they are the complementary quantities to $(x_{j+1}, s_{j+1}, y_{j+1})$.

It follows from norm inequalities, the definition of P_{j+1}, (2.13), (4.135), Assumption 1.1, and (4.134) that

$$\left| \frac{[n^s_{j+1}]_i}{[s_{j+1}]_i} \right| \leq \|S^{-1}_{j+1} n^s_{j+1}\|_\infty \leq \|S^{-1}_{j+1} n^s_{j+1}\|_2 \leq \|P^{-1}_{j+1} n_{j+1}\|_2 \leq \kappa_{\text{n}} \pi^v_{j+1} = \mathcal{O}(v_{j+1}) = \mathcal{O}(\mu_j) \text{ for } j \in \mathcal{J}.$$

Since we maintain positive slack vectors throughout Algorithm 1, we may then conclude that

$$|[n^s_{j+1}]_i| = \mathcal{O}(\mu_j [s_{j+1}]_i) \quad \text{for all } 1 \leq i \leq m \text{ and } j \in \mathcal{J}. \tag{4.137}$$

We now develop a crucial bound by considering two cases motivated by the definition of D_k. First, suppose that for a given i we have $\mu_j [s_{j+1}]_i^{-2} \leq \kappa_{\text{D}}(\mu_j)$, so that from (4.133) we have $[d_{j+1}]_i = \mu_j [s_{j+1}]_i^{-2}$. It then follows from this fact and (4.137) that

$$|[s_{j+1}]_i [d_{j+1}]_i [n^s_{j+1}]_i| = \mathcal{O}(\mu_j^2) \quad \text{for } j \in \mathcal{J}.$$

Second, suppose that for a given i we have $\mu_j [s_{j+1}]_i^{-2} > \kappa_{\text{D}}(\mu_j)$, so that from (4.133) we have $[d_{j+1}]_i = \kappa_{\text{D}}(\mu_j) < \mu_j [s_{j+1}]_i^{-2}$, and thus $[s_{j+1}]_i^2 [d_{j+1}]_i < \mu_j$. Combining this fact with (4.137) shows that

$$|[s_{j+1}]_i [d_{j+1}]_i [n^s_{j+1}]_i| = \mathcal{O}(\mu_j [s_{j+1}]_i^2 [d_{j+1}]_i) = \mathcal{O}(\mu_j^2) \quad \text{for } j \in \mathcal{J}. \tag{4.138}$$

Therefore, we have shown that (4.138) holds in both cases, i.e., (4.138) holds for all $1 \leq i \leq m$ and $j \in \mathcal{J}$. We may now use the same proof as for Lemma 3.18, combined with (4.136), (4.138), and the first part of (4.134) to deduce that $\lim_{j \in \mathcal{J}} y_{j+1} = y_*$ for some y_* satisfying $g(x_*) + J(x_*)^T y_* = 0$ and $S_* y_* = 0$. To prove that (x_*, s_*, y_*) is a first-order KKT point for problem (NPs), it only remains to prove that $y^* \geq 0$, which we now proceed to do.

From the first part of (4.134), we know that

$$\begin{aligned}\zeta_1\mu_j &\geq \left\| \begin{pmatrix} g(x_{j+1}) + \nabla_{xx}\mathcal{L}(x_{j+1}, y^{\text{B}}_{j+1})n^x_{j+1} + J(x_{j+1})^T y_{j+1} \\ -\mu_j e + S_{j+1}D_{j+1}n^s_{j+1} + S_{j+1}y_{j+1} \end{pmatrix} \right\|_2 \\ &\geq \left\| -\mu_j e + S_{j+1}D_{j+1}n^s_{j+1} + S_{j+1}y_{j+1} \right\|_2 \geq \left\| -\mu_j e + S_{j+1}D_{j+1}n^s_{j+1} + S_{j+1}y_{j+1} \right\|_\infty \\ &\geq | -\mu_j + [s_{j+1}]_i[d_{j+1}]_i[n^s_{j+1}]_i + [s_{j+1}]_i[y_{j+1}]_i| \text{ for all } 1 \leq i \leq m. \end{aligned} \tag{4.139}$$

We now consider two cases. First, suppose that i is such that $[s_*]_i > 0$. In this case it follows from (4.139), (4.138), the fact that $\mu_j \to 0$, and (4.135) that $\lim_{j\in\mathcal{J}}[y_{j+1}]_i = [y^*]_i = 0$, as desired. Second, suppose that i is such that $[s_*]_i = 0$. It may be observed from (4.139) that

$$-\zeta_1\mu_j \leq -\mu_j + [s_{j+1}]_i[d_{j+1}]_i[n^s_{j+1}]_i + [s_{j+1}]_i[y_{j+1}]_i,$$

and hence that

$$[y_{j+1}]_i \geq \frac{-\zeta_1\mu_j + \mu_j - [s_{j+1}]_i[d_{j+1}]_i[n^s_{j+1}]_i}{[s_{j+1}]_i}.$$

It follows from the previous inequality, the facts that $\zeta_1 \in (0, 1)$ and $\mu_j \to 0$, (4.138), and the fact that the slack vectors are maintained to be positive in Algorithm 1, that $[y_{j+1}]_i > 0$ for all sufficiently large $j \in \mathcal{J}$. Combining this with $\lim_{j\in\mathcal{J}} y_{j+1} = y_*$ shows that $[y_*]_i \geq 0$. This completes the proof. □

5 Conclusion and discussion

In this paper, we have presented a new algorithm for solving constrained nonlinear optimization problems. The algorithm is of the inexact barrier-SQP variety, i.e., it approximately solves a sequence of barrier subproblems using an inexact SQP method. In Sections 2 and 3, we proved that each barrier subproblem could be solved approximately using a new inexact-SQP method based on a trust-funnel mechanism (not requiring a filter or penalty function). The algorithm is extremely flexible in that, during each iteration, it automatically determines the types of steps and updates that are expected to be most productive, where potential productivity is determined by available criticality measures. In each iteration, each subproblem may be solved approximately using matrix-free iterative methods, which means that the algorithm is viable for solving large-scale barrier subproblems. We then proved in Section 4 that an approximate solution of the original nonlinear optimization problem may be obtained by approximately solving a sequence of barrier subproblems for a decreasing sequence of barrier parameters.

Although we have not considered them explicitly in this paper, we remark that equality constraints, call them $c_{\text{E}}(x) = 0$, may easily be included in our algorithm. To do this, one may simply redefine

$$c(x, s) := \begin{pmatrix} c(x) + s \\ c_{\text{E}}(x) \end{pmatrix}$$

and adjust the barrier problem (1.1), violation measure (2.3) and v-criticality measure (2.9) in obvious ways. Clearly, two-sided bounds on inequality constraints may also be incorporated in a similar fashion.

We are currently implementing our new algorithm. Once complete, it will be part of the GALAHAD [20] thread-safe library of Fortran 90 packages for the numerical solution of optimization problems.

References

[1] M. ARGÁEZ AND R. TAPIA, *On the global convergence of a modified augmented Lagrangian line-search interior-point Newton method for nonlinear programming*, Journal of Optimization Theory and Applications, 114 (2002), pp. 1–25.

[2] R. H. Byrd, F. E. Curtis, and J. Nocedal, *An inexact SQP method for equality constrained optimization*, SIAM Journal on Optimization, 19 (2008), pp. 351–369.

[3] R. H. Byrd, J. C. Gilbert, and J. Nocedal, *A trust region method based on interior point techniques for nonlinear programming*, Math. Program., 89 (2000), pp. 149–185.

[4] R. H. Byrd, M. E. Hribar, and J. Nocedal, *An interior point algorithm for large-scale nonlinear programming*, SIAM J. Optim., 9 (1999), pp. 877–900.

[5] A. R. Conn, N. I. M. Gould, and Ph. L. Toint, *Trust-Region Methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[6] F. E. Curtis, O. Schenk, and A. Wächter, *An interior-point algorithm for large-scale nonlinear optimization with inexact step computations*, SIAM Journal on Scientific Computing, 32 (2010), pp. 3447–3475.

[7] J. Czyzyk, R. Fourer, and S. Mehrotra, *Using a massively parallel processor to solve large sparse linear programs by an interior-point method*, SIAM J. Sci. Comput., 19 (1998), pp. 553–565.

[8] R. Fletcher, *Practical Methods of Optimization*, Wiley-Interscience [John Wiley & Sons], New York, 2001.

[9] R. Fletcher, N. I. M. Gould, S. Leyffer, Ph. L. Toint, and A. Wächter, *Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming*, SIAM J. Optim., 13 (2002), pp. 635–659 (electronic) (2003).

[10] R. Fletcher and S. Leyffer, *Nonlinear programming without a penalty function*, Math. Program., 91 (2002), pp. 239–269.

[11] R. Fletcher, S. Leyffer, and Ph. L. Toint, *On the global convergence of a filter-SQP algorithm*, SIAM J. Optim., 13 (2002), pp. 44–59 (electronic).

[12] R. Fourer and S. Mehrotra, *Performance of an augmented system approach for solving least-squares problems in an interior-point method for linear programming*, Math. Program., 19 (1991), pp. 26–31.

[13] R. Fourer and S. Mehrotra, *Solving symmetric indefinite systems in an interior-point method for linear programming*, Math. Program., 62 (1993), pp. 15–39.

[14] R. Fourer and S. Mehrotra, *Solving symmetric indefinite systems in an interior-point method for linear programming*, Math. Programming, 62 (1993), pp. 15–39.

[15] E. M. Gertz and P. E. Gill, *A primal-dual trust region algorithm for nonlinear optimization*, Mathematical Programming, Series B, 100 (2004), pp. 49–94.

[16] P. E. Gill, W. Murray, and M. A. Saunders, *SNOPT: An SQP algorithm for large-scale constrained optimization*, SIAM Rev., 47 (2005), pp. 99–131.

[17] P. E. Gill, W. Murray, and M. H. Wright, *Practical optimization*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1981.

[18] J. Gondzio, *Interior point methods 25 years later*, European Journal of Operational Research, 218 (2012), pp. 587–601.

[19] N. I. M. Gould, D. Orban and D. P. Robinson, *Trajectory-following methods for large-scale degenerate convex quadratic programming*, Mathematical Programming Computation, 5 (2013), pp. 113–142.

[20] N. I. M. Gould, D. Orban, and Ph. L. Toint, *GALAHAD, a library of thread-safe Fortran 90 packages for large-scale nonlinear optimization*, ACM Trans. Math. Software, 29 (2003), pp. 353–372.

[21] N. I. M. Gould and D. P. Robinson, *A second derivative SQP method: Global convergence*, SIAM J. Optim., 20 (2010), pp. 2023–2048.

[22] ——, *A second derivative SQP method: Local convergence and practical issues*, SIAM J. Optim., 20 (2010), pp. 2049–2079.

[23] ——, *A second derivative SQP method with a "trust-region-free" predictor step*, IMA J. Numer. Anal., 32 (2012), pp. 580–601.

[24] N. I. M. Gould, D. P. Robinson, and Ph. L. Toint, *Corrigendum: nonlinear programming without a penalty function or a filter*, Tech. Rep. RAL-TR-2011-006, Rutherford Appleton Laboratory, Chilton, England, 2011.

[25] N. I. M. Gould and Ph. L. Toint, *Nonlinear programming without a penalty function or a filter*, Math. Program., (2009).

[26] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.

[27] W. Karush, *Minima of functions of several variables with inequalities as side conditions*, Master's thesis, Department of Mathematics, University of Chicago, Illinois, USA, 1939.

[28] H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, in Proceedings of the second Berkeley symposium on mathematical statistics and probability, J. Neyman, ed., California, USA, 1951, University of Berkeley Press.

[29] M. Lalee, J. Nocedal, and T. Plantenga, *On the implementation of an algorithm for large-scale equality constrained optimization*, SIAM J. Optim., 8 (1998), pp. 682–706.

[30] S. Mehrotra, *On the implementation of a primal-dual interior point method*, SIAM J. Optim., 2 (1992), pp. 575–601.

[31] J. L. Morales, J. Nocedal, and Y. Wu, *A sequential quadratic programming algorithm with an additional equality constrained phase*, IMA J. Numer. Anal., 32 (2012), pp. 553–579.

[32] R. J. Vanderbei, *LOQO: an interior point code for quadratic programming*, Optimization Methods and Software, 11 (1999), pp. 451–484.

[33] A. Wächter and L. T. Biegler, *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming*, Mathematical Programming, Series A, 106 (2006), pp. 25–57.

[34] H. Yabe and H. Yamashita, *Q-superlinear convergence of primal-dual interior point quasi-Newton methods for constrained optimization*, J. Oper. Res. Soc. Japan, 40 (1997), pp. 415–436.

[35] H. Yamashita and H. Yabe, *Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization*, Math. Program., 75 (1996), pp. 377–397.

[36] ——, *An interior point method with a primal-dual quadratic barrier penalty function for nonlinear optimization*, SIAM Journal on Optimization, 14 (2003), pp. 479–499.

[37] H. Yamashita, H. Yabe, and T. Tanabe, *A globally and superlinearly convergent primal-dual interior point trust region method for large scale constrained optimization*, Mathematical Programming, Series A, 102 (2005), pp. 111–151.

Algorithm 1 Trust-funnel algorithm for minimizing the barrier problem (1.1).

Input: (x_0, s_0, y_{-1}, μ) and $(\epsilon_\pi, \epsilon_v)$ with $(s_0, y_{-1}, \mu) > 0$ and $(\epsilon_\pi, \epsilon_v) > 0$, respectively.
Choose $\{\delta_0^f, \delta_0^v, \kappa_{\rm ca}, \kappa_{\rm y}, \kappa_{\rm D}, \kappa_{\delta vv}\} \subset (0,\infty)$, $\{\kappa_{\rm cr}, \kappa_{\mathcal{VS}}, \kappa_{\rm v}\} \subset (1,\infty)$, $0 < \eta_1 \le \eta_2 < 1$, $0 < \gamma_1 \le \gamma_2 < 1$, $\kappa_{\rm n} \ge 1$, $\{\kappa_{\rm tt}, \kappa_\delta, \kappa_{\rm tg}, \kappa_\omega, \kappa_\chi, \kappa_{\rm B}, \kappa_{\rm vv}, \kappa_{\rm fbn}, \kappa_{\rm fbt}, \kappa_{\rm t1}, \kappa_{\rm t2}\} \subset (0,1)$, and $\kappa_{\rm cd} \in (0, 1-\kappa_{\rm tg}]$.
Perform a slack reset to s_0 as given by (2.5).
Set $v_0^{\max} \leftarrow \max\{\kappa_{\rm ca}, \kappa_{\rm cr} v(x_0, s_0)\}$, and $\pi_{-1}^f \leftarrow 0$.
for $k = 0, 1, \ldots$ **do**
    Compute v_k from (2.4) and π_k^v and χ_k^v from (2.9).
    if $\chi_k^v = 0$ and $v_k > 0$ **then**
        Return the infeasible stationary point (x_k, s_k).
    if (2.10) holds, or at least $\pi_k^v > 0$ **then**
        Compute n_k satisfying (2.13) and (2.14). $[k \in \mathcal{N}]$
    else
        Set $n_k \leftarrow 0$.
    Choose $y_k^{\rm B}$ satisfying (2.19) and D_k satisfying (2.20), and then set G_k by (2.18).
    if (2.21) holds **then**
        Compute y_k, r_k, π_k^f, and χ_k^f from (2.23)–(2.25) to satisfy (2.26a), (2.26b), or (2.26c).
        if (2.26a) holds **then**
            Return the (approximate) first-order KKT point (x_k, s_k, y_k) for the barrier problem (1.1).
        else if (2.26b) holds **then**
            Set $t_k \leftarrow 0$.
        else
            if $k \in \mathcal{N}$ **then**
                Compute t_k so that (2.30) is satisfied. $[k \in \mathcal{T}]$
            else
                Compute t_k so that either (2.30) or (2.35) is satisfied. $[k \in \mathcal{T}]$
            if (2.30) holds **then**
                Add iteration k to the set $\mathcal{T}_{\mathcal{D}}$. $[k \in \mathcal{T}_{\mathcal{D}}]$
                if (2.31) is satisfied but (2.32) fails **then**
                    Set $t_k \leftarrow 0$. $[k \in \mathcal{T}_0]$
    else
        Set $y_k \leftarrow y_{k-1}$ and $t_k \leftarrow 0$, and then set r_k, π_k^f, and χ_k^f by (2.24)–(2.25).
        if (2.26a) holds **then**
            Return the (approximate) first-order KKT point (x_k, s_k, y_k) for the barrier problem (1.1).
    if (2.30d) holds **then**
        Add iteration k to the set $\mathcal{D}$. $[k \in \mathcal{D}]$
    Set the trial step $d_k \leftarrow n_k + t_k$ and trial iterate $(x_k^+, s_k^+) \leftarrow (x_k, s_k) + d_k$.
    if $d_k = 0$ **then**
        Perform the y-iteration updates given by (2.36). $[k \in \mathcal{Y}]$
    else if $t_k \neq 0$ and both (2.32) and (2.37) hold **then**
        Perform the f-iteration updates given by (2.39)–(2.46). $[k \in \mathcal{F}]$
    else
        Perform the v-iteration updates given by (2.50)–(2.57). $[k \in \mathcal{V}]$