ISE Industrial and Systems Engineering

Progressively Sampled Equality-Constrained Optimization

Frank E. Curtis, Lingjun Guo, and Daniel P. Robinson

Department of Industrial and Systems Engineering, Lehigh University

COR@L Technical Report 25T-017

LEHIGH UNIVERSITY

COR@L COMPUTATIONAL OPTIMIZATION RESEARCH AT LEHIGH

Progressively Sampled Equality-Constrained Optimization

Frank E. Curtis[*1], Lingjun Guo[†1], and Daniel P. Robinson[‡1]

[1]Department of Industrial and Systems Engineering, Lehigh University

October 1, 2025

Abstract

An algorithm is proposed, analyzed, and tested for solving continuous nonlinear-equality-constrained optimization problems where the constraints are defined by an expectation or an average over a large (finite) number of terms. The main idea of the algorithm is to solve a sequence of equality-constrained problems, each involving a finite sample of constraint-function terms, over which the sample set grows progressively. Under assumptions about the constraint functions and their first- and second-order derivatives that are reasonable in some real-world settings of interest, it is shown that—with a sufficiently large initial sample—solving a sequence of problems defined through progressive sampling yields a better worst-case sample complexity bound compared to solving a single problem with a full set of samples. The results of numerical experiments with a set of test problems demonstrate that the proposed approach can be effective in practice.

1 Introduction

We propose, analyze, and test an algorithm for solving *constrained* continuous optimization problems where the aim is to minimize a nonconvex objective function subject to nonlinear equality constraints. Our particular setting of interest is when the equality constraints are defined by an expectation or an average over a large finite number of terms. A broad setting in which such a problem arises is the context of least-squares regression when one knows, in advance, a target residual error, at least for a subset of regression terms. Another class of problems that fit our setting include some arising from physics-informed learning where the constraints state that the expected residual of a differential equation should be zero.

The inspiration for our proposed algorithm is the method for *unconstrained* continuous optimization problems proposed in [6]. In that work, the aim is to minimize a nonconvex objective function when the *objective function* is defined by an expectation or an average over a large finite number of terms. The proposed algorithm involves solving a sequence of subproblems, where in each subproblem the objective is defined by an average over a sample of objective terms. It is shown that, under certain conditions about the original objective, one can obtain an improved worst-case gradient-sample-complexity bound by solving a sequence of subproblems over which the sample sets grow progressively. That is, each sample set is a superset of the previous one. The main idea is to exploit the fact that, under their stated conditions, a minimizer of a sampled objective is close to a minimizer of an objective corresponding to an augmented sample set.

Our aim in this paper is to consider a similar algorithmic approach, except that instead of sampling objective function terms, we sample terms defining the *constraint function.* Our algorithm employs some similar ideas as proposed in [6], but our setting is quite distinct for multiple reasons. First, our setting

[*]E-mail: frank.e.curtis@lehigh.edu
[†]E-mail: lig423@lehigh.edu
[‡]E-mail: daniel.p.robinson@lehigh.edu

necessitates consideration of constrained subproblems and worst-case complexity bounds for finding second-order stationary points of such problems. With respect to this, we rely on recent complexity bounds developed in [4]. Second, the presence of constraints means that our method needs to rely on approximate stationarity conditions for constrained optimization when determining when one subproblem has been solved to sufficient accuracy and the next in the sequence should be considered. With respect to this, we rely on properties of least-squares multipliers for ensuring our overall sample complexity guarantees. Finally, due to fundamental changes in an optimization problem that result when a constraint function is perturbed, our overall sample complexity bound relies on theory pertaining to *acute perturbations*; see [8, 9, 10]. For the unconstrained setting in [6], a fundamental role is played by an assumption that the objective function is *strongly Morse*. Our analysis also involves a similar assumption (with respect to a Lagrangian function), but this alone is not sufficient. In the context of equality constraints, the sample sizes need to be sufficiently large in order to ensure that the constraint Jacobians are acute perturbations of each other, as revealed in our analysis.

We also include the results of numerical experiments to demonstrate that computational gains can be achieved through the use of our proposed algorithm.

1.1 Notation

We use $\mathbb{R}$ to denote the set of real numbers, $\mathbb{R}_{\geq r}$ (resp., $\mathbb{R}_{>r}$) to denote the set of real numbers greater than or equal to (resp., greater than) $r \in \mathbb{R}$, $\mathbb{R}^n$ to denote the set of n-dimensional real vectors, and $\mathbb{R}^{m\times n}$ to denote the set of m-by-n-dimensional real matrices. We define $\mathbb{N} := \{0, 1, 2, \dots\}$, and, for any integer $N \geq 1$, we use $[N]$ to denote the set $\{1, \dots, N\}$. For any finite set $\mathcal{S}$, we use $|\mathcal{S}|$ to denote its cardinality. For vectors we define $\|\cdot\| := \|\cdot\|_2$, or else specify the norm explicitly. We use $\|\cdot\|$ to denote the spectral norm of any matrix.

For any matrix $A \in \mathbb{R}^{m\times n}$, we use $\sigma_i(A)$ for each $i \in [\min\{m, n\}]$ to denote its ith largest singular value. Given any such A, we use $\text{null}(A)$ to denote its null space, i.e., $\{d \in \mathbb{R}^n : Ad = 0\}$. Assuming that $B \in \mathbb{R}^{n\times m}$ has full-column rank, its Moore-Penrose pseudoinverse is $B^\dagger := (B^TB)^{-1}B^T$. Note that in this case of B having full-column rank, the pseudoinverse $B^\dagger$ is a left inverse of B in the sense that $B^\dagger B = I$. For any subspace $\mathcal{X} \subseteq \mathbb{R}^n$ and point $x \in \mathbb{R}^n$, we denote the Euclidean projection of x onto $\mathcal{X}$ as $\text{proj}_{\mathcal{X}}(x) := \arg\min_{\bar{x}\in\mathcal{X}} \|\bar{x} - x\|$. Given $B \in \mathbb{R}^{n\times m}$ with full column rank, we use $\mathcal{R}(B) := BB^\dagger$ and $\mathcal{N}(B^T) = I - \mathcal{R}(B)$ to denote orthogonal projection matrices onto the span of the columns of B and the null space of B^T, respectively. That such $\mathcal{R}(B)$ and $\mathcal{N}(B^T)$ are orthogonal projection matrices follows since each of them are both symmetric and idempotent.

1.2 Organization

We state our problem class of interest and our proposed algorithm in §2. The analysis of our proposed algorithm is contained in §3. The results of our numerical experiments are provided in §4 and a conclusion is provided in §5.

2 Algorithm

Our proposed algorithm is designed to solve a sample average approximation (SAA) of the continuous nonlinear-equality-constrained problem

$$\min_{x\in\mathbb{R}^n} \; f(x) \;\text{ s.t. }\; \bar{c}(x) = 0, \tag{1}$$

where the objective and constraint functions, i.e., $f : \mathbb{R}^n \to \mathbb{R}$ and $\bar{c} : \mathbb{R}^n \to \mathbb{R}^m$, respectively, are twice-continuously differentiable, $m \leq n$, and the constraint function $\bar{c}$ is defined by an expectation. Formally, with respect to a random variable ω defined by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, expectation $\mathbb{E}$ defined by $\mathbb{P}$, and $\overline{C} : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$, the function $\bar{c}$ is defined by $\bar{c}(x) = \mathbb{E}[\overline{C}(x, \omega)]$ for all $x \in \mathbb{R}^n$.

The SAA of problem (1) that our algorithm is designed to solve is defined with respect to a (large) sample of $N \in \mathbb{N}$ realizations of the random variable ω, say, $\{\omega_i\}_{i\in[N]}$. Defining the SAA constraint function $c : \mathbb{R}^n \to \mathbb{R}^m$ for all $x \in \mathbb{R}^n$ by

$$c(x) = \tfrac{1}{N} \sum_{i\in[N]} c_i(x), \quad \text{where} \quad c_i(x) \equiv \overline{C}(x, \omega_i) \quad \text{for all} \quad i \in [N],$$

the problem that our algorithm is designed to solve is that given by

$$\min_{x\in\mathbb{R}^n} \; f(x) \;\; \text{s.t.} \;\; c(x) = 0. \tag{2}$$

Under reasonable assumptions about $\bar{c}$ and an assumption that N is sufficiently large, differences between values of $\bar{c}$ in (1) and c in (2) as well as values of their first- and second-order derivatives at any $x \in \mathbb{R}^n$ can be bounded with high probability [2, 5, 11], which in turn can be used to relate approximate stationary points of (2) with those of (1), again with high probability. Thus, for our purposes, we focus on our proposed algorithm and our analysis of it for solving problem (2).

The main idea of our proposed algorithm for solving problem (2) is to generate a sequence of iterates, each of which is a stationary point (at least approximately) with respect to a subsampled problem involving only a nonempty subset $\mathcal{S} \subseteq [N]$ of constraint function terms. For any such $\mathcal{S}$, we denote the approximate constraint function as $c_{\mathcal{S}} : \mathbb{R}^n \to \mathbb{R}^m$ and state the approximation of problem (2) as

$$\min_{x\in\mathbb{R}^n} \; f(x) \;\; \text{s.t.} \;\; c_{\mathcal{S}}(x) = 0, \quad \text{where} \quad c_{\mathcal{S}}(x) = \tfrac{1}{|\mathcal{S}|} \sum_{i\in\mathcal{S}} c_i(x). \tag{3}$$

Observe that the constraint function in (2) is $c_{[N]} \equiv c$. The primary benefit of considering (3) for $\mathcal{S} \subseteq [N]$, rather than (2) directly, is that any evaluation of the constraint function or Jacobian requires a sum of only $|\mathcal{S}| \leq N$ terms, as opposed to N terms. Also, under reasonable assumptions about the constraint functions, we show in this paper that, by starting with an approximate stationary point for (3) and aiming to solve a subsequent instance of (3) with respect to a superset $\overline{\mathcal{S}} \supseteq \mathcal{S}$, our proposed algorithm can obtain an approximate stationary point for the subsequent instance with lower sample complexity than if the problem with the larger sample set were solved directly from an arbitrary starting point. Overall, we show that for a sufficiently large sample set relative to N, a sufficiently approximate stationary point of (2) can be obtained more efficiently through progressive sampling than by solving the full-sample problem directly.

Now let us introduce stationarity conditions for (3) for any nonempty sample set $\mathcal{S} \subseteq [N]$, which in particular also represent stationarity conditions for (2) when one considers the case that $\mathcal{S} = [N]$. The objective f has already been assumed to be twice-continuously differentiable. Let us also assume that $c_{\mathcal{S}}$ is twice-continuously differentiable for all nonempty $\mathcal{S} \subseteq [N]$. Let the Lagrangian of problem (3) be denoted by $L_{\mathcal{S}} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, defined for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ by

$$L_{\mathcal{S}}(x, y) = f(x) + c_{\mathcal{S}}(x)^T y = f(x) + \tfrac{1}{|\mathcal{S}|} \sum_{i\in\mathcal{S}} c_i(x)^T y,$$

where $y \in \mathbb{R}^m$ is a vector of Lagrange multiplier estimates (also known as dual variables). When $\mathcal{S} = [N]$, we denote the Lagrangian simply as $L \equiv L_{[N]}$. Second-order necessary conditions for optimality for (3) can then be stated as

$$\nabla L_{\mathcal{S}}(x, y) \equiv \begin{bmatrix} \nabla_x L_{\mathcal{S}}(x, y) \\ \nabla_y L_{\mathcal{S}}(x, y) \end{bmatrix} \equiv \begin{bmatrix} \nabla f(x) + \nabla c_{\mathcal{S}}(x) y \\ c_{\mathcal{S}}(x) \end{bmatrix} = 0 \tag{4}$$

and, with $[c_{\mathcal{S}}]_j$ denoting the jth component of the constraint function $c_{\mathcal{S}}$,

$$d^T \nabla^2_{xx} L_{\mathcal{S}}(x, y) d \equiv d^T \left(\nabla^2 f(x) + \sum_{j\in[m]} \nabla^2 [c_{\mathcal{S}}]_j(x) y_j \right) d \geq 0 \quad \text{for all} \quad d \in \text{null}(\nabla c_{\mathcal{S}}(x)^T). \tag{5}$$

We refer to any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ satisfying (4) as a first-order stationary point with respect to (3), and we refer to any such point satisfying both (4) and (5) as a second-order stationary point. Also, consistent with the literature on worst-case complexity for nonconvex smooth optimization, we say for some $(\epsilon, \zeta) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ that $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ is (ϵ, ζ)-stationary with respect to (3) if and only if

$$\|\nabla L_{\mathcal{S}}(x, y)\| \leq \epsilon \tag{6a}$$

$$\text{and} \quad d^T \nabla^2_{xx} L_{\mathcal{S}}(x, y) d \geq -\zeta \|d\|_2^2 \quad \text{for all} \quad d \in \text{null}(\nabla c_{\mathcal{S}}(x)^T). \tag{6b}$$

Generally speaking, an algorithm for solving (3) can be a *primal* method that only generates a sequence of primal iterates $\{x_k\}$, or it can be a *primal-dual* method that generates a sequence of primal and dual iterate pairs $\{(x_k, y_k)\}$. For an application of our proposed algorithm, either type of method can be employed, but for certain results in our analysis, we use properties of *least-square multipliers* corresponding to a given primal point $x \in \mathbb{R}^n$. Assuming that the Jacobian of $c_{\mathcal{S}}$ at x (i.e., $\nabla c_{\mathcal{S}}(x)^T$) has full-row rank, the least-squares multipliers with respect to x are given by $y_{\mathcal{S}}(x) \in \mathbb{R}^m$ that minimizes $\|\nabla_x L_{\mathcal{S}}(x, \cdot)\|^2$, which is denoted by

$$y_{\mathcal{S}}(x) = -(\nabla c_{\mathcal{S}}(x)^T \nabla c_{\mathcal{S}}(x))^{-1} \nabla c_{\mathcal{S}}(x)^T \nabla f(x) = -\nabla c_{\mathcal{S}}(x)^{\dagger} \nabla f(x). \tag{7}$$

Our proposed method is stated as Algorithm 1 below. Our analysis in the next section formalizes assumptions under which Algorithm 1 is well defined and yields our claimed worst-case sample complexity guarantees.

Algorithm 1 Progressive Constraint-Sampling Method (PCSM) for (2)

Require: Initial sample size $p_1 \in [N]$, initial point $x_0 \in \mathbb{R}^n$, iteration limit $K = \lceil \log_2 \frac{N}{p_1} \rceil$, and subproblem tolerances $\{(\epsilon_k, \zeta_k)\}_{k=1}^K \subset \mathbb{R}_{>0} \times \mathbb{R}_{>0}$

set $\mathcal{S}_0 \leftarrow \emptyset$
for $k \in [K]$ **do**
  choose $\mathcal{S}_k \supseteq \mathcal{S}_{k-1}$ such that $|\mathcal{S}_k| = p_k$
  using x_{k-1} as a starting point, employ an algorithm to solve (3), terminating once a primal iterate x_k has been obtained such that $(x_k, y_{\mathcal{S}_k}(x_k))$ (see (7)) is (ϵ_k, ζ_k)-stationary with respect to subproblem (3) for $\mathcal{S} = \mathcal{S}_k$
  set $p_{k+1} \leftarrow \min\{2p_k, N\}$
end for
return $(x_K, y(x_K))$, which is (ϵ_K, ζ_K)-stationary with respect to (2)

3 Analysis

Our analysis is split into two main parts. First, we prove generic worst-case iteration and sample complexity bounds for Algorithm 1 under generic assumptions about the objective and constraint functions and the worst-case iteration complexity properties of the algorithm that is employed for solving the arising subproblems. We state these bounds in generic terms since they hold for any subproblem solver that possesses the stated worst-case iteration complexity properties. Second, for concreteness, we prove that a specific subproblem solver based on Fletcher's augmented Lagrangian function possesses the worst-case iteration complexity properties that lead to a desirable worst-case sample complexity bound.

3.1 Generic Worst-case Sample Complexity Bounds for Algorithm 1

We begin our analysis by showing conditions under which Algorithm 1 can offer an improved worst-case sample complexity bound compared to an algorithm that solves (2) directly. The main work in our analysis in this section is to show that if (2) is strongly Morse (in a manner defined in our assumptions in this section), then the sample problem (3) is strongly Morse as long as $|\mathcal{S}|$ is sufficiently large. At the core of this analysis is a result showing that $|\mathcal{S}|$ being sufficiently large ensures that the sample constraint Jacobian and the full-sample constraint Jacobian are acute perturbations of each other in a sense defined in [8, 9, 10].

Let us begin by stating the basic assumptions under which we prove our theoretical guarantees in this subsection. Assumption 3.1 below ensures that any minimizer of each subproblem is a second-order stationary point and that one can expect an algorithm that is employed to solve each subproblem to find a sufficiently approximate second-order stationary point. It would be possible to prove reasonable convergence guarantees for Algorithm 1 under looser assumptions. For example, if an algorithm employed to solve (3) for some nonempty sample set $\mathcal{S}$ were to encounter an (approximate) infeasible stationary point, then it would be reasonable to terminate the subproblem solver and either terminate Algorithm 1 in its entirety or move on to solve the next subproblem (with a larger sample set). However, since consideration of such scenarios would distract from the essential properties of our algorithm when each subproblem solve is successful, we make Assumption 3.1. Our remarks after the assumption justify it further.

Assumption 3.1. *The functions $f : \mathbb{R}^n \to \mathbb{R}$ and $c_i : \mathbb{R}^n \to \mathbb{R}^m$ for each $i \in [N]$ are thrice-continuously differentiable. In addition, the following hold for problem* (2)*, instances of subproblem* (3)*, and the algorithm employed to solve them.*

(a) There exists $\sigma_{\min} \in \mathbb{R}_{>0}$ such that, for all $x \in \mathbb{R}^n$ and $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$, the constraint Jacobian has $\sigma_m(\nabla c_{\mathcal{S}}(x)^T) \geq \sigma_{\min}$.

(b) For all $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$ and with any initial point, the algorithm employed to solve subproblem (3) *is guaranteed to generate a sequence of iterates for which a limit point is a second-order stationary point, i.e., one satisfying* (4) *and* (5)*.*

(c) For some $(\alpha, \beta) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}$, problem (2) *is (α, β)-strongly Morse in the sense that, for any $x \in \mathbb{R}^n$, if $(x, y(x))$ satisfies $\|\nabla L(x, y(x))\| \leq \alpha$, then*

$$|d^T \nabla^2_{xx} L(x, y(x)) d| \geq \beta \|d\|_2^2 \quad \text{for all} \quad d \in \text{null}(\nabla c(x)^T).$$

Algorithm 1 does not generally require the problem functions to be thrice-continuously differentiable, but since a later part of our analysis refers to third-order derivatives, we make the first part of Assumption 3.1. Part (a) of Assumption 3.1 guarantees that the algorithm employed to solve subproblem (3) will not, e.g., generate an infeasible stationary point. In addition to this assurance, part (b) of Assumption 3.1 implicitly requires one of various types of conditions in the literature on equality-constrained optimization that guarantee convergence of an algorithm to a second-order stationary point. (This assumption will be strengthened in our final theorem of this subsection.) Part (c) of Assumption 3.1 is essential for proving our desired worst-case sample complexity properties of Algorithm 1. Related to this, the following comment is important, so we emphasize it as a formal remark.

Remark 3.1. *If one were to change the requirement for the subproblem solver in Algorithm 1 and only require that the solver produce an approximate first-order stationary point, rather than an approximate second-order stationary point, then under Assumption 3.1 (with "second-order" replaced by "first-order" in part (b) and part (c) of the assumption removed) the algorithm would be well defined and, by its construction, would guarantee convergence to an approximate first-order stationary point of problem* (2)*. Therefore, in practice, Algorithm 1 might be run with only approximate first-order stationarity requirements. However, such a set-up would not allow us to prove our specific desired strong worst-case sample complexity guarantee, which focuses on attainment of approximate second-order stationarity. Therefore, for the purposes of our analysis, we state Algorithm 1 and Assumption 3.1 as they are given, and note that our analysis leverages approximate second-order stationarity in order to ensure better worst-case sample complexity compared to a method that solves the full-sample problem directly.*

Our next assumption articulates bounds on derivatives of the objective and constraint functions corresponding to the full-sample problem (2). An assumption of this type is typical in the literature on continuous constrained optimization.

Assumption 3.2. *There exists* $(\kappa_{\nabla f}, \kappa_{\nabla c}, \kappa_{\nabla^2 f}, \kappa_{\nabla^2 c}) \in \mathbb{R}^4_{>0}$ *such that, for all* $x \in \mathbb{R}^n$ *and* $j \in [m]$*, one has* $\|\nabla f(x)\| \leq \kappa_{\nabla f}$*,* $\|\nabla c(x)\| \leq \kappa_{\nabla c}$*,* $\|\nabla^2 f(x)\| \leq \kappa_{\nabla^2 f}$*, and* $\|\nabla^2 [c]_j(x)\| \leq \kappa_{\nabla^2 c}$*, where* $[c]_j$ *denotes the* j*th component of* c*.*

Our final assumption introduces constants that bound discrepancies between constraint Jacobians corresponding to individual samples and those corresponding to the full set of samples, and introduces constants that similarly bound discrepancies between individual-sample and the full-sample constraint Hessian matrices. We remark that the assumption is quite loose. Indeed, since $c(x) = 0$ at any feasible point for problem (2), having $\|c(x)\|$ on the right-hand side of the first inequality in Assumption 3.3 would make the assumption restrictive. For our analysis, only a constant bound is needed on the right-hand side. On the other hand, under a constraint qualification such as the linear independence constraint qualification (LICQ) at x, one has $\|\nabla c(x)\| > 0$, so the assumed bound on the differences in Jacobian values is reasonable. We also remark that Assumptions 3.2 and 3.3 complement each other in the sense that, together, they allow us to prove similar bounds for $c_{\mathcal{S}}$ for nonempty $\mathcal{S} \subseteq [N]$; see upcoming Lemma 3.2.

Assumption 3.3. *There exists* $(\gamma_c, \gamma_{\nabla c}, \gamma_{\nabla^2 c}) \in \mathbb{R}^3_{>0}$ *such that, for all* $(x, j) \in \mathbb{R}^n \times [m]$*,*

$$\begin{aligned} \tfrac{1}{N} \sum_{i \in [N]} \|c_i(x) - c(x)\|^2 &\leq \gamma_c, \\ \tfrac{1}{N} \sum_{i \in [N]} \|\nabla c_i(x) - \nabla c(x)\|^2 &\leq \gamma_{\nabla c} \|\nabla c(x)\|^2, \\ \textit{and} \ \ \tfrac{1}{N} \sum_{i \in [N]} \|\nabla^2 [c_i]_j(x) - \nabla^2 [c]_j(x)\|^2 &\leq \gamma_{\nabla^2 c} \|\nabla^2 [c]_j(x)\|^2. \end{aligned}$$

Our first lemma employs bounds introduced in Assumptions 3.1 and 3.2 to offer bounds on the pseudoinverse of the constraint Jacobian, the least-squares multipliers, and the Hessian of the Lagrangian at all points in the domain of problem (2).

Lemma 3.1. *For all* $x \in \mathbb{R}^n$*, one has (recall (7)) that*

$$\|\nabla c(x)^\dagger\| \leq \sigma_{\min}^{-1}, \tag{8a}$$

$$\|y(x)\| \leq \kappa_{\nabla f} \sigma_{\min}^{-1}, \tag{8b}$$

$$\textit{and} \ \ \|\nabla^2_{xx} L(x, y(x))\| \leq \kappa_{\nabla^2 f} + \sqrt{m} \kappa_{\nabla f} \kappa_{\nabla^2 c} \sigma_{\min}^{-1}. \tag{8c}$$

Proof. First, consider (8a). As is well known, the spectral norm of the pseudoinverse of a matrix is equal to the reciprocal of its smallest singular value; see, e.g., [3, Chapter 21]. Therefore, since under Assumption 3.1 the smallest singular value of $\nabla c(x)^T$ is bounded below by $\sigma_{\min}$ for all $x \in \mathbb{R}^n$, the bound in (8a) follows. Second, consider (8b). By (7), submultiplicity of the matrix 2-norm, Assumption 3.2, and (8a) the desired conclusion follows since, for all $x \in \mathbb{R}^n$, one finds

$$\|y(x)\| = \|\nabla c(x)^\dagger \nabla f(x)\| \leq \|\nabla c(x)^\dagger\| \|\nabla f(x)\| \leq \kappa_{\nabla f} \sigma_{\min}^{-1}.$$

Finally, consider (8c). By the triangle inequality and absolute homogeneity of matrix norms, Assumption 3.2, the fact that for any vector $y \in \mathbb{R}^m$ one has $\|y\|_1 \leq \sqrt{m}\|y\|$, and (8b), one has for all $x \in \mathbb{R}^n$ that

$$\begin{aligned} \|\nabla^2_{xx} L(x, y(x))\| &= \left\| \nabla^2 f(x) + \sum_{j \in [m]} \nabla^2 [c]_j(x) [y(x)]_j \right\| \\ &\leq \|\nabla^2 f(x)\| + \sum_{j \in [m]} \|\nabla^2 [c]_j(x)\| \cdot |[y(x)]_j| \\ &\leq \kappa_{\nabla^2 f} + \kappa_{\nabla^2 c} \|y(x)\|_1 \\ &\leq \kappa_{\nabla^2 f} + \sqrt{m} \kappa_{\nabla^2 c} \|y(x)\| \end{aligned}$$

$$\leq \kappa_{\nabla^2 f} + \sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}\sigma_{\min}^{-1},$$

which is the desired conclusion. □ □

Our second lemma leverages Assumption 3.3 in order to provide bounds that are similar to those in the assumption, except that they are with respect to $c_{\mathcal{S}}$ and its derivatives for nonempty $\mathcal{S} \subseteq [N]$. The resulting bounds depend on the sample size $|\mathcal{S}|$. For convenience here and throughout the rest of the paper, let us define

$$\xi_{\mathcal{S}} := \sqrt{\tfrac{N(N-|\mathcal{S}|)}{|\mathcal{S}|^2}} \in [0, \sqrt{N(N-1)}] \ \text{ for all nonempty } \mathcal{S} \subseteq [N]. \tag{9}$$

Lemma 3.2. *For all $x \in \mathbb{R}^n$, nonempty $\mathcal{S} \subseteq [N]$, and $j \in [m]$, one has*

$$\|c_{\mathcal{S}}(x) - c(x)\|^2 \leq \xi_{\mathcal{S}}^2 \gamma_c, \tag{10a}$$

$$\|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\|^2 \leq \xi_{\mathcal{S}}^2 \gamma_{\nabla c} \|\nabla c(x)\|^2, \tag{10b}$$

$$\text{and} \ \ \|\nabla^2 [c_{\mathcal{S}}]_j(x) - \nabla^2 [c]_j(x)\|^2 \leq \xi_{\mathcal{S}}^2 \gamma_{\nabla^2 c} \|\nabla^2 [c]_j(x)\|^2. \tag{10c}$$

Proof. Each of the desired bounds can be proved in a similar manner. We prove the second bound, namely, (10b), with respect to the constraint derivatives. The other two bounds follow in a similar manner. First, observe that

$$\begin{aligned}\nabla c_{\mathcal{S}}(x) = \tfrac{1}{|\mathcal{S}|}\sum_{i \in \mathcal{S}} \nabla c_i(x) &= \tfrac{1}{|\mathcal{S}|}\sum_{i \in [N]} \nabla c_i(x) - \tfrac{1}{|\mathcal{S}|}\sum_{i \in [N]\setminus\mathcal{S}} \nabla c_i(x) \\ &= \tfrac{N}{|\mathcal{S}|}\nabla c(x) - \tfrac{1}{|\mathcal{S}|}\sum_{i \in [N]\setminus\mathcal{S}} \nabla c_i(x).\end{aligned}$$

Second, observe that for any vector in $v \in \mathbb{R}^{N-|\mathcal{S}|}$ one has from the Cauchy-Schwarz inequality that $(\mathbb{1}^T v)^2 \leq \|\mathbb{1}\|^2 \|v\|_2^2 = (N - |\mathcal{S}|)\|v\|_2^2$. Consequently, with the triangle inequality and Assumption 3.3, one finds that

$$\begin{aligned}\|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\|^2 &= \left\|\tfrac{N}{|\mathcal{S}|}\nabla c(x) - \nabla c(x) - \tfrac{1}{|\mathcal{S}|}\sum_{i \in [N]\setminus\mathcal{S}} \nabla c_i(x)\right\|^2 \\ &= \tfrac{1}{|\mathcal{S}|^2}\left\|\sum_{i \in [N]\setminus\mathcal{S}} (\nabla c(x) - \nabla c_i(x))\right\|^2 \\ &\leq \tfrac{1}{|\mathcal{S}|^2}\left(\sum_{i \in [N]\setminus\mathcal{S}} \|\nabla c(x) - \nabla c_i(x)\|\right)^2 \\ &\leq \left(\tfrac{N-|\mathcal{S}|}{|\mathcal{S}|^2}\right)\sum_{i \in [N]\setminus\mathcal{S}} \|\nabla c(x) - \nabla c_i(x)\|^2 \\ &\leq \left(\tfrac{N-|\mathcal{S}|}{|\mathcal{S}|^2}\right)\sum_{i \in [N]} \|\nabla c(x) - \nabla c_i(x)\|^2 \\ &\leq \left(\tfrac{N-|\mathcal{S}|}{|\mathcal{S}|^2}\right) N\gamma_{\nabla c}\|\nabla c(x)\|^2,\end{aligned}$$

which gives the desired conclusion. □ □

Our third lemma refers to the concept of an *acute perturbation* between two real-valued rectangular matrices. We provide the following definition, then state our lemma, which shows that if $\mathcal{S}$ has sufficiently large cardinality, then $\nabla c_{\mathcal{S}}(x)$ and $\nabla c(x)$ are acute perturbations of each other for all $x \in \mathbb{R}^n$.

Definition 3.1. Two matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{n\times m}$ are acute perturbations of each other if and only if $\text{rank}(AA^\dagger BA^\dagger A) = \text{rank}(A) = \text{rank}(B)$. In particular, if A and B are full-column-rank matrices, then $A^\dagger A = I$ and $\text{rank}(A) = \text{rank}(B) = m$, so A and B are acute perturbations of each other if and only if $\text{rank}(AA^\dagger B) = m$.

Lemma 3.3. *If nonempty $\mathcal{S} \subseteq [N]$ satisfies $|\mathcal{S}| \geq p_1$ and*

$$|\mathcal{S}| > \frac{2N}{1+\sqrt{1+\frac{4\sigma_{\min}^2}{\gamma_{\nabla c}\kappa_{\nabla c}^2}}} \in (0, N], \tag{11}$$

then $\nabla c(x)$ and $\nabla c_{\mathcal{S}}(x)$ are acute perturbations of each other for all $x \in \mathbb{R}^n$.

Proof. By Assumption 3.1, for all $x \in \mathbb{R}^n$, the matrices $\nabla c_{\mathcal{S}}(x) \in \mathbb{R}^{n\times m}$ and $\nabla c(x) \in \mathbb{R}^{n\times m}$ have full-column rank. All that remains is to show that the rank condition in Definition 3.1 holds for all $x \in \mathbb{R}^n$. Observe that, for any $x \in \mathbb{R}^n$, one finds

$$\begin{aligned}\nabla c(x)\nabla c(x)^\dagger \nabla c_{\mathcal{S}}(x) &= \nabla c(x)\nabla c(x)^\dagger(\nabla c(x) + \nabla c_{\mathcal{S}}(x) - \nabla c(x)) \\ &= \nabla c(x)(I + \nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x) - \nabla c(x))).\end{aligned} \tag{12}$$

Next observe that, for all $x \in \mathbb{R}^n$, one has from submultiplicity of the matrix 2-norm, Lemma 3.1, Lemma 3.2, and Assumption 3.2 that

$$\begin{aligned}\|\nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x) - \nabla c(x))\| &\leq \|\nabla c(x)^\dagger\|\|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\| \\ &\leq \tfrac{\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}}{\sigma_{\min}}\|\nabla c(x)\| \leq \tfrac{\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}}.\end{aligned} \tag{13}$$

Let us now show that (11) implies that the right-hand side of (13) is strictly less than one. Define $t := \frac{N}{|\mathcal{S}|}$ and recall the definition of $\xi_{\mathcal{S}}$ in (9). Then, (13) being strictly less than one is equivalent to $\frac{\kappa_{\nabla c}}{\sigma_{\min}}\sqrt{\gamma_{\nabla c}t(t-1)} < 1$, which in turn is equivalent to $t(t-1) < \frac{\sigma_{\min}^2}{\kappa_{\nabla c}^2\gamma_{\nabla c}}$. The left-hand side of this latter inequality is a quadratic function of t, so it follows that it is satisfied for all positive $t < \frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{4\sigma_{\min}^2}{\kappa_{\nabla c}^2\gamma_{\nabla c}}}$. Substituting $\frac{N}{|\mathcal{S}|}$ for t, one finds after rearrangement that this is equivalent to (11). Hence, by (11), one has that (13) yields $\|\nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x) - \nabla c(x))\| < 1$. Now observe that from [8, Theorem 6.6] (or see [10, Eq. (3)]) that

$$\begin{aligned}&\sigma_m(I) - \sigma_m(I + \nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x) - \nabla c(x))) \\ &\leq |\sigma_m(I) - \sigma_m(I + \nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x) - \nabla c(x)))| \\ &\leq \|\nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x) - \nabla c(x))\| < 1,\end{aligned}$$

from which it follows that $\sigma_m(I+\nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x)-\nabla c(x))) > 0$. Thus, the matrix $I+\nabla c(x)^\dagger(\nabla c_{\mathcal{S}}(x)-\nabla c(x))$ is nonsingular, which along with the fact that $\nabla c(x)$ has full-column rank means that the matrix in (12) has full rank, namely, m. □ □

Our next lemma shows that with a similar lower bound on the sample size one can bound differences between sampled and full-sample least-squares multipliers.

Lemma 3.4. *If nonempty $\mathcal{S} \subseteq [N]$ satisfies $|\mathcal{S}| \geq p_1$ and*

$$|\mathcal{S}| \geq \frac{2N}{1+\sqrt{1+\frac{4\sigma_{\min}^2}{9\gamma_{\nabla c}\kappa_{\nabla c}^2}}} \in (0, N], \tag{14}$$

then for all $x \in \mathbb{R}^n$ one has (recall (7)) that

$$\|y(x) - y_{\mathcal{S}}(x)\| \leq \frac{9\kappa_{\nabla f}\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{2\sigma_{\min}^2}.$$

Proof. First, let us remark that (14) implies that, for all $x \in \mathbb{R}^n$, the matrices $\nabla c(x)$ and $\nabla c_{\mathcal{S}}(x)$ are acute perturbations of each other, which is useful since it allows us to employ a theorem from [9] to bound $\|y(x) - y_{\mathcal{S}}(x)\|$ for all $x \in \mathbb{R}^n$. Indeed, this is straightforward to see since the right-hand side of (14) is larger than that of (11). Hence, when (14) holds, (11) also holds, in which case one has from Lemma 3.3 that $\nabla c(x)$ and $\nabla c_{\mathcal{S}}(x)$ are acute perturbations of each other for all $x \in \mathbb{R}^n$, as desired. At the same time, using a similar argument as in Lemma 3.3, it follows that (14) implies that the right-hand side of (13) is less than or equal to 1/3, which shows

$$\|\nabla c(x)^\dagger\|\|(\nabla c_{\mathcal{S}}(x) - \nabla c(x))\| \leq \tfrac{\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}} \leq \tfrac{1}{3}. \tag{15}$$

Our aim now is to state, for any $x \in \mathbb{R}^n$, the upper bound for $\|y(x) - y_{\mathcal{S}}(x)\|$ that is offered by [9, Theorem 5.2]. Note that this theorem applies to our setting since, as previously mentioned, under (14) the matrices $\nabla c(x)$ and $\nabla c_{\mathcal{S}}(x)$ are acute perturbations of each other for all $x \in \mathbb{R}^n$. Consider arbitrary $x \in \mathbb{R}^n$. To state the bound from [9, Theorem 5.2], the quantities that it involves need to be derived. To this end, observe that a singular value decomposition of $\nabla c(x)$ yields

$$\nabla c(x) = U \begin{bmatrix} S \\ 0 \end{bmatrix} V^T \implies \begin{bmatrix} S \\ 0 \end{bmatrix} = U^T \nabla c(x) V,$$

where $U \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{m\times m}$ are orthogonal, and where under Assumption 3.1 the diagonal matrix $S \in \mathbb{R}^{m\times m}$ is positive definite. Correspondingly, let us define

$$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} := U^T(\nabla c_{\mathcal{S}}(x) - \nabla c(x))V,\ \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} := U^T \nabla c_{\mathcal{S}}(x) V, \text{ and } \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} := U^T(-\nabla f(x)),$$

where $E_1 \in \mathbb{R}^{m\times m}$, $E_2 \in \mathbb{R}^{(n-m)\times m}$, $B_1 \in \mathbb{R}^{m\times m}$, $B_2 \in \mathbb{R}^{(n-m)\times m}$, $b_1 \in \mathbb{R}^m$, and $b_2 \in \mathbb{R}^{n-m}$. Note that our tuple of matrices (S, E_1, E_2, B_1, B_2) correspond to the tuple of matrices $(A_{11}, E_{11}, E_{21}, B_{11}, B_{21})$ that are introduced in [9, page 636], and note that since our Assumption 3.1 ensures that $\nabla c(x)$ has full-column rank, the matrices $(E_{12}, E_{22}, B_{12}, B_{22})$ that are introduced in [9, page 636] are not present in our setting. One can now state that [9, Theorem 5.2] in our setting yields

$$\tfrac{\|y(x)-y_{\mathcal{S}}(x)\|}{\|y(x)\|} \leq \bar\kappa \tfrac{\|E_1\|}{\|\nabla c(x)\|} + \bar\kappa^2 \tfrac{\|E_2\|}{\|\nabla c(x)\|}\left(\eta\tfrac{\|b_2\|}{\|b_1\|} + \tfrac{\|E_2\|}{\|\nabla c(x)\|}\right), \tag{16}$$

where, as in [9, Theorem 3.8] and the displayed equation prior to [9, Theorem 5.1],

$$\bar\kappa := \|\nabla c(x)\|\|B_1^{-1}\| = \|\nabla c(x)\|\|(S + E_1)^{-1}\| \text{ and } \eta := \tfrac{\|b_1\|}{\|\nabla c(x)\|\|y(x)\|}.$$

We remark in passing that our statement of (16) corrects a typo in the statement of [9, Theorem 5.2]. In particular, in the statement of [9, Theorem 5.2], the latter term on the right-hand side is stated with $\|E_{12}\|$ in the numerator outside of the parentheses. That is a typo. One can see through the proof of [9, Theorem 5.2] that our statement is correct, where in fact $\|E_{21}\|$ belongs in this numerator. In our setting, this corresponds to $\|E_2\|$ in this numerator, as we have stated in (16).

Our next goal is to prove upper bounds for the terms on the right-hand side of (16). First, with respect to $\|b_2\|$, one has from submultiplicity of the matrix 2-norm, $\|U\| = 1$ due to U being orthogonal, and Assumption 3.2 that

$$\|b_2\| \leq \|U^T(-\nabla f(x))\| \leq \|U\|\|\nabla f(x)\| \leq \kappa_{\nabla f}. \tag{17}$$

Second, with respect to both $\|E_1\|$ and $\|E_2\|$, one has from submultiplicity of the matrix 2-norm and $\|U\| = \|V\| = 1$ due to both U and V being orthogonal that

$$\max\{\|E_1\|, \|E_2\|\} \leq \left\|\begin{bmatrix} E_1 \\ E_2 \end{bmatrix}\right\| = \|U^T(\nabla c_{\mathcal{S}}(x) - \nabla c(x))V\| \leq \|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\|. \tag{18}$$

Third, let us bound $\|(S+E_1)^{-1}\|^2$. For this, let us employ [9, Theorem 2.2], where now our tuple of matrices (S, E_1, B_1) play the role of (A, E, B) in that theorem. In particular, let us employ the second part of [9, Theorem 2.2], which requires (in our notation) S to be nonsingular and $\|S^{-1}\|\|E_1\| < 1$. The fact that S is nonsingular has been established earlier through the fact that S is positive definite. On the other hand, let us now show $\|S^{-1}\|\|E_1\| < 1$. Toward this end, observe that

$$\begin{aligned}
\nabla c(x)^\dagger &= (\nabla c(x)^T \nabla c(x))^{-1} \nabla c(x)^T \\
&= \left(V \begin{bmatrix} S \\ 0 \end{bmatrix}^T U^T U \begin{bmatrix} S \\ 0 \end{bmatrix} V^T \right)^{-1} V \begin{bmatrix} S \\ 0 \end{bmatrix}^T U^T \\
&= (V S^2 V^T)^{-1} V \begin{bmatrix} S & 0 \end{bmatrix} U^T \\
&= (V S^{-2} V^T) V \begin{bmatrix} S & 0 \end{bmatrix} U^T \\
&= V \begin{bmatrix} S^{-1} & 0 \end{bmatrix} U^T.
\end{aligned}$$

That is, the right-hand side of the above equation is an SVD for $\nabla c(x)^\dagger$, from which it follows that $\|S^{-1}\| = \|\nabla c(x)^\dagger\|$. Hence, with (15) and (18), one finds that

$$\|S^{-1}\|\|E_1\| = \|\nabla c(x)^\dagger\|\|E_1\| \leq \|\nabla c(x)^\dagger\|\|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\| \leq \tfrac{1}{3}. \tag{19}$$

Since the requirements of [9, Theorem 2.2] thus hold, one obtains with (19) that

$$\|\,(S + E_1)^{-1}\,\| = \|B_1^{-1}\| \leq \tfrac{\|S^{-1}\|}{1 - \|S^{-1}\|\|E_1\|} \leq \tfrac{\|S^{-1}\|}{1 - \frac{1}{3}} = \tfrac{3}{2}\|\nabla c(x)^\dagger\|. \tag{20}$$

Combining (8a), (15), (16), (17), (18), and (20), one now obtains

$$\begin{aligned}
\tfrac{\|y(x) - y_{\mathcal{S}}(x)\|}{\|y(x)\|} &\leq \|E_1\|\|(S + E_1)^{-1}\| + \tfrac{\|(S+E_1)^{-1}\|^2\|E_2\|\|b_2\|}{\|y(x)\|} \\
&\quad + \|(S + E_1)^{-1}\|^2\|E_2\|^2 \\
&\leq \left(\tfrac{3}{2}\|\nabla c(x)^\dagger\| + \tfrac{9}{4}\tfrac{\|\nabla c(x)^\dagger\|^2 \kappa_{\nabla f}}{\|y(x)\|} \right) \|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\| \\
&\quad + \tfrac{9}{4}\|\nabla c(x)^\dagger\|^2\|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\|^2 \\
&\leq \left(\tfrac{9}{4}\|\nabla c(x)^\dagger\| + \tfrac{9}{4}\tfrac{\|\nabla c(x)^\dagger\|^2 \kappa_{\nabla f}}{\|y(x)\|} \right) \|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\| \\
&\leq \left(\tfrac{9}{4\sigma_{\min}} + \tfrac{9\kappa_{\nabla f}}{4\sigma_{\min}^2\|y(x)\|} \right) \|\nabla c(x) - \nabla c_{\mathcal{S}}(x)\|.
\end{aligned}$$

Multiplying $\|y(x)\|$ on both sides of the above inequality, one finds with (8b) that

$$\begin{aligned}
\|y(x) - y_{\mathcal{S}}(x)\| &\leq \left(\tfrac{9\|y(x)\|}{4\sigma_{\min}} + \tfrac{9\kappa_{\nabla f}}{4\sigma_{\min}^2} \right) \|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\| \\
&\leq \tfrac{9\kappa_{\nabla f}}{2\sigma_{\min}^2}\|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\|.
\end{aligned}$$

Further, using Assumptions 3.2 and Lemma 3.2, one finds that

$$\|y(x) - y_{\mathcal{S}}(x)\| \leq \tfrac{9\kappa_{\nabla f}\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}}{2\sigma_{\min}^2}\|\nabla c(x)\| \leq \tfrac{9\kappa_{\nabla f}\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{2\sigma_{\min}^2}.$$

This completes the proof since this bound is independent of $x \in \mathbb{R}^n$. □ □

Our next goal is to prove a bound on the difference of inner products involving the Hessians of the Lagrangians $L_{\mathcal{S}}$ and L when $|\mathcal{S}|$ is sufficiently large. Toward this end, we first present the following preliminary lemma, which for any $x \in \mathbb{R}^n$ shows a bound on the product between any element of the null space of the sample constraint Jacobian $\nabla c_{\mathcal{S}}(x)$ and a projection matrix that maps vectors onto the span of the columns of the full-sample constraint Jacobian $\nabla c(x)$.

Lemma 3.5. *If nonempty $\mathcal{S} \subseteq [N]$ satisfies $|\mathcal{S}| \geq p_1$ and $\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c} < \sigma_{\min}$, then for all $x \in \mathbb{R}^n$ and for all $d_{\mathcal{S}} \in \text{null}(\nabla c_{\mathcal{S}}(x)^T) \setminus \{0\}$ one has that*

$$\tfrac{\|\mathcal{R}(\nabla c(x)) d_{\mathcal{S}}\|}{\|d_{\mathcal{S}}\|} \leq \tfrac{\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}} < 1.$$

Proof. Consider arbitrary $x \in \mathbb{R}^n$ and $d_{\mathcal{S}} \in \text{null}(\nabla c_{\mathcal{S}}(x)^T) \setminus \{0\}$. It follows from the definitions of $\mathcal{R}$ and $\mathcal{N}$ that $\mathcal{R}(\nabla c_{\mathcal{S}}(x))d_{\mathcal{S}} = 0$ and $\mathcal{N}(\nabla c_{\mathcal{S}}(x)^T)d_{\mathcal{S}} = d_{\mathcal{S}}$. Let us now consider $\mathcal{R}(\nabla c(x))d_{\mathcal{S}}$. By submultiplicity of the matrix 2-norm, one has

$$\|\mathcal{R}(\nabla c(x))d_{\mathcal{S}}\| = \|\mathcal{R}(\nabla c(x))\mathcal{N}(\nabla c_{\mathcal{S}}(x)^T)d_{\mathcal{S}}\| \leq \|\mathcal{R}(\nabla c(x))\mathcal{N}(\nabla c_{\mathcal{S}}(x)^T)\|\|d_{\mathcal{S}}\|.$$

Now, since by Assumption 3.1 the matrices $\nabla c_{\mathcal{S}}(x)$ and $\nabla c(x)$ have the same rank, it follows from [9, Theorem 2.4] that

$$\begin{aligned}\|\mathcal{R}(\nabla c(x))d_{\mathcal{S}}\| &\leq \|\mathcal{R}(\nabla c(x))\mathcal{N}(\nabla c_{\mathcal{S}}(x)^T)\|\|d_{\mathcal{S}}\| \\ &\leq \min\{\|\nabla c(x)^\dagger\|, \|\nabla c_{\mathcal{S}}(x)^\dagger\|\}\|\nabla c(x) - \nabla c_{\mathcal{S}}(x)\|\|d_{\mathcal{S}}\| \\ &\leq \|\nabla c(x)^\dagger\|\|\nabla c(x) - \nabla c_{\mathcal{S}}(x)\|\|d_{\mathcal{S}}\|. \end{aligned} \tag{21}$$

Since $d_{\mathcal{S}} \neq 0 \in \mathbb{R}^n$, by dividing the expression above by $\|d_{\mathcal{S}}\|$, the resulting expression can be bounded above as in (13), which yields the desired inequality. □ □

With the preceding lemma, we can now prove that if the sample set $\mathcal{S}$ is sufficiently large in cardinality, then (3) is $(\alpha_{\mathcal{S}}, \beta_{\mathcal{S}})$-strongly Morse with respect to a pair $(\alpha_{\mathcal{S}}, \beta_{\mathcal{S}}) \in \mathbb{R}^2_{>0}$ defined with respect to $(\alpha, \beta) \in \mathbb{R}^2_{>0}$ from Assumption 3.1.

Lemma 3.6. *Define the tuple $(\kappa_1, \kappa_2, \kappa_3) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ by*

$$\kappa_1 := \tfrac{\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}}, \quad \kappa_2 := \sqrt{\gamma_c + \kappa_{\nabla f}^2 \kappa_1^2}, \quad \textit{and} \tag{22}$$

$$\kappa_3 := 3\kappa_{\nabla^2 f}\kappa_1 + \sqrt{m}\left(\tfrac{5\kappa_{\nabla f}\kappa_{\nabla^2 c}}{2\sigma_{\min}}\right)(3\kappa_1 + \sqrt{\gamma_{\nabla^2 c}}). \tag{23}$$

Then, with $(\alpha, \beta) \in \mathbb{R}^2_{>0}$ from Assumption 3.1, if nonempty $\mathcal{S} \subseteq [N]$ has $|\mathcal{S}| \geq p_1$ and

$$\xi_{\mathcal{S}} \leq \min\left\{\tfrac{1}{3\kappa_1}, \tfrac{\alpha}{2\kappa_2}, \tfrac{7\beta}{18\kappa_3}\right\}, \tag{24}$$

then with respect to

$$\alpha_{\mathcal{S}} := \alpha - \xi_{\mathcal{S}}\kappa_2 \geq \tfrac{1}{2}\alpha > 0 \quad \textit{and} \quad \beta_{\mathcal{S}} := (1 - \tfrac{1}{3}\kappa_1\xi_{\mathcal{S}})\beta - \kappa_3\xi_{\mathcal{S}} \geq \tfrac{1}{2}\beta > 0 \tag{25}$$

subproblem (3) *is $(\alpha_{\mathcal{S}}, \beta_{\mathcal{S}})$-strongly Morse in the sense of Assumption 3.1; in other words, one has for any $x \in \mathbb{R}^n$ that if $(x, y_{\mathcal{S}}(x))$ satisfies $\|\nabla L(x, y_{\mathcal{S}}(x))\| \leq \alpha_{\mathcal{S}}$, then*

$$|d^T \nabla^2_{xx} L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d| \geq \beta_{\mathcal{S}}\|d\|_2^2 \quad \textit{for all} \quad d \in \text{null}(\nabla c_{\mathcal{S}}(x)^T).$$

Proof. Our first aim is to show that with (κ_1, κ_2) defined in (22) one finds for any point $x \in \mathbb{R}^n$ and nonempty $\mathcal{S} \subseteq [N]$ satisfying $|\mathcal{S}| \geq p_1$ that $\|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \alpha_{\mathcal{S}}$ implies that $\|\nabla L(x, y(x))\| \leq \alpha$ holds. Toward this end, consider arbitrary $x \in \mathbb{R}^n$ and nonempty $\mathcal{S} \subseteq [N]$ satisfying $|\mathcal{S}| \geq p_1$ with $\|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \alpha_{\mathcal{S}}$, where $\alpha_{\mathcal{S}}$ is defined as stated in the lemma and (κ_1, κ_2) is defined as in (22). Our first aim is to prove a bound on the norm of the difference between the gradient of the Lagrangian with respect to the sample set $\mathcal{S}$ and the gradient of the Lagrangian with respect to the full-sample set. First, observe that by (4), (7), submultiplicity of the matrix 2-norm, and Assumption 3.2, one finds

$$\begin{aligned}\|\nabla_x L(x, y(x)) - \nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| &= \|\nabla c(x)y(x) - \nabla c_{\mathcal{S}}(x)y_{\mathcal{S}}(x)\| \\ &= \|\nabla c_{\mathcal{S}}(x)\nabla c_{\mathcal{S}}(x)^\dagger \nabla f(x) - \nabla c(x)\nabla c(x)^\dagger \nabla f(x)\| \\ &\leq \|\mathcal{R}(\nabla c_{\mathcal{S}}(x)) - \mathcal{R}(\nabla c(x))\|\kappa_{\nabla f}. \end{aligned} \tag{26}$$

Since by Assumption 3.1 the Jacobians $\nabla c_{\mathcal{S}}(x)$ and $\nabla c(x)$ have full-column rank, it follows from (26), [9, Theorem 2.3], [9, Theorem 2.4], (8a), and (10b) that

$$\|\nabla_x L(x, y(x)) - \nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \|\mathcal{R}(\nabla c_{\mathcal{S}}(x)) - \mathcal{R}(\nabla c(x))\|\kappa_{\nabla f}$$

$$
\begin{aligned}
&= \|\mathcal{R}(\nabla c(x))\mathcal{N}(\nabla c_{\mathcal{S}}(x)^T)\|\kappa_{\nabla f} \\
&\leq \|\nabla c(x)^{\dagger}\|\|\nabla c_{\mathcal{S}}(x) - \nabla c(x)\|\kappa_{\nabla f} \\
&\leq \tfrac{\kappa_{\nabla f}\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}} = \kappa_{\nabla f}\xi_{\mathcal{S}}\kappa_1.
\end{aligned} \tag{27}
$$

Second, by (4) and (10a), one finds that

$$
\|\nabla_y L(x, y(x)) - \nabla_y L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| = \|c(x) - c_{\mathcal{S}}(x)\| \leq \xi_{\mathcal{S}}\sqrt{\gamma_c}. \tag{28}
$$

Combining (27) and (28), one finds that

$$
\|\nabla L(x, y(x)) - \nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \xi_{\mathcal{S}}\sqrt{\gamma_c + \kappa_{\nabla f}^2\kappa_1^2} = \xi_{\mathcal{S}}\kappa_2.
$$

Thus, since $\|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \alpha_{\mathcal{S}} = \alpha - \xi_{\mathcal{S}}\kappa_2$, as desired one obtains

$$
\begin{aligned}
\|\nabla L(x, y(x))\| &= \|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) + \nabla L(x, y(x)) - \nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \\
&\leq \|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| + \|\nabla L(x, y(x)) - \nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \\
&\leq \|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\|_2 + \xi_{\mathcal{S}}\kappa_2 \\
&\leq \alpha - \xi_{\mathcal{S}}\kappa_2 + \xi_{\mathcal{S}}\kappa_2 = \alpha.
\end{aligned}
$$

Finally, observe that $\alpha_{\mathcal{S}} := \alpha - \xi_{\mathcal{S}}\kappa_2 \geq \frac{1}{2}\alpha$ follows since (24) implies $\xi_{\mathcal{S}} \leq \alpha/(2\kappa_2)$.

We have shown that, for arbitrary $x \in \mathbb{R}^n$ and with (κ_1, κ_2), nonempty $\mathcal{S} \subseteq [N]$ satisfying $|\mathcal{S}| \geq p_1$, and $\alpha_{\mathcal{S}}$ satisfying the conditions of the lemma, having $\|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \alpha_{\mathcal{S}}$ yields $\|\nabla L(x, y(x))\| \leq \alpha$. Our next aim is to show that, for such x and $\mathcal{S}$, one has $|d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}}| \geq \beta_{\mathcal{S}}\|d_{\mathcal{S}}\|_2^2$ for all nonzero $d_{\mathcal{S}} \in \text{null}(\nabla c_{\mathcal{S}}(x)^T)$, where κ_3 and $\beta_{\mathcal{S}}$ are defined as in the lemma. (Observe that $|d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}}| \geq \beta_{\mathcal{S}}\|d_{\mathcal{S}}\|_2^2$ holds trivially for $d_{\mathcal{S}} = 0$. The value for κ_3 stated in the lemma is chosen in order to yield inequality (40) later in this proof.) Consider such x and $\mathcal{S}$, and note that by Assumption 3.1, namely, that (2) is (α, β)-strongly Morse, it follows that since $\|\nabla L(x, y(x))\| \leq \alpha$ one also has that $|d^T\nabla_{xx}^2 L(x, y(x))d| \geq \beta\|d\|^2$ for all $d \in \text{null}(\nabla c(x)^T)$. Let us now consider arbitrary nonzero $d_{\mathcal{S}} \in \text{null}(\nabla c_{\mathcal{S}}(x)^T)$. One first has that

$$
\begin{aligned}
&|d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}}| \\
&= |d_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))d_{\mathcal{S}} + d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}} - d_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))d_{\mathcal{S}}| \\
&\geq |d_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))d_{\mathcal{S}}| - |d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}} - d_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))d_{\mathcal{S}}|.
\end{aligned} \tag{29}
$$

Let us now derive lower bounds on the two terms on the right-hand side of (29). Note that $d_{\mathcal{S}}$ can be decomposed as the sum of two orthogonal vectors, namely, $d_{\mathcal{S}} = \bar{d}_{\mathcal{S}} + r_{\mathcal{S}}$, where $\bar{d}_{\mathcal{S}} = \mathcal{N}(\nabla c(x)^T)d_{\mathcal{S}}$ and $r_{\mathcal{S}} = \mathcal{R}(\nabla c(x))d_{\mathcal{S}}$. It follows that

$$
\|d_{\mathcal{S}}\|^2 = \|\bar{d}_{\mathcal{S}} + r_{\mathcal{S}}\|^2 = \|\bar{d}_{\mathcal{S}}\|^2 + 2\bar{d}_{\mathcal{S}}^T r_{\mathcal{S}} + \|r_{\mathcal{S}}\|^2 = \|\bar{d}_{\mathcal{S}}\|^2 + \|r_{\mathcal{S}}\|^2 \tag{30}
$$

along with $\max\{\|\bar{d}_{\mathcal{S}}\|, \|r_{\mathcal{S}}\|\} \leq \|d_{\mathcal{S}}\|$. Now, one finds with the triangle inequality, Cauchy-Schwarz inequality, submultiplicity of the matrix 2-norm, the fact that problem (2) is (α, β)-strongly Morse, and $\max\{\|\bar{d}_{\mathcal{S}}\|, \|r_{\mathcal{S}}\|\} \leq \|d_{\mathcal{S}}\|$ that

$$
\begin{aligned}
&|d_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))d_{\mathcal{S}}| \\
&= |(\bar{d}_{\mathcal{S}} + r_{\mathcal{S}})^T\nabla_{xx}^2 L(x, y(x))(\bar{d}_{\mathcal{S}} + r_{\mathcal{S}})| \\
&= |\bar{d}_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))\bar{d}_{\mathcal{S}} + 2\bar{d}_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))r_{\mathcal{S}} + r_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))r_{\mathcal{S}}| \\
&\geq |\bar{d}_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))\bar{d}_{\mathcal{S}}| - |2\bar{d}_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))r_{\mathcal{S}}| - |r_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))r_{\mathcal{S}}| \\
&\geq |\bar{d}_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))\bar{d}_{\mathcal{S}}| - 2\|\nabla_{xx}^2 L(x, y(x))\|\|\bar{d}_{\mathcal{S}}\|\|r_{\mathcal{S}}\| - \|\nabla_{xx}^2 L(x, y(x))\|\|r_{\mathcal{S}}\|^2
\end{aligned}
$$

$$\geq \beta\|\bar{d}_{\mathcal{S}}\|^2 - 3\|\nabla^2_{xx}L(x,y(x))\|\|d_{\mathcal{S}}\|\|r_{\mathcal{S}}\|. \tag{31}$$

Next, let us derive an upper bound for $\|r_{\mathcal{S}}\|$ with respect to $\|d_{\mathcal{S}}\|$. Recall that $r_{\mathcal{S}} = \mathcal{R}(\nabla c(x))d_{\mathcal{S}}$. Our aim is to employ Lemma 3.5, for which $\mathcal{S}$ needs to satisfy $\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c} < \sigma_{\min}$. To see that this holds, note that by (24) one has

$$\tfrac{\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}} = \kappa_1\xi_{\mathcal{S}} \leq \tfrac{1}{3} < 1.$$

Now applying the result of Lemma 3.5 and since $\|d_{\mathcal{S}}\| \neq 0$, one finds that

$$\tfrac{\|r_{\mathcal{S}}\|}{\|d_{\mathcal{S}}\|} = \tfrac{\|\mathcal{R}(\nabla c(x))d_{\mathcal{S}}\|}{\|d_{\mathcal{S}}\|} \leq \tfrac{\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}} = \kappa_1\xi_{\mathcal{S}}.$$

Multiplying by $\|d_{\mathcal{S}}\|$ on both sides yields $\|r_{\mathcal{S}}\| \leq \kappa_1\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|$. Combined with (30), this yields $\|d_{\mathcal{S}}\|^2 \leq \|\bar{d}_{\mathcal{S}}\|^2 + \kappa_1^2\xi_{\mathcal{S}}^2\|d_{\mathcal{S}}\|^2$, which with (24) implies

$$\|\bar{d}_{\mathcal{S}}\|^2 \geq \left(1 - \kappa_1^2\xi_{\mathcal{S}}^2\right)\|d_{\mathcal{S}}\|^2 \geq \left(1 - \tfrac{1}{3}\kappa_1\xi_{\mathcal{S}}\right)\|d_{\mathcal{S}}\|^2. \tag{32}$$

Combining (32), (8c), $\|r_{\mathcal{S}}\| \leq \kappa_1\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|$, and (31), it follows that

$$\begin{aligned} &|d_{\mathcal{S}}^T\nabla^2_{xx}L(x,y(x))d_{\mathcal{S}}| \\ \geq\ &\left(1 - \tfrac{1}{3}\kappa_1\xi_{\mathcal{S}}\right)\beta\|d_{\mathcal{S}}\|^2 - 3\left(\kappa_{\nabla^2 f} + \tfrac{\sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}}{\sigma_{\min}}\right)\kappa_1\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|^2, \end{aligned} \tag{33}$$

which gives us a useful lower bound for the first term on the right-hand side of (29). Now let us turn to proving a lower bound for the second term on the right-hand side of (29), which we derive by proving an upper bound for the absolute difference $|d_{\mathcal{S}}^T\nabla^2_{xx}L_{\mathcal{S}}(x,y_{\mathcal{S}}(x))d_{\mathcal{S}} - d_{\mathcal{S}}^T\nabla^2_{xx}L(x,y(x))d_{\mathcal{S}}|$. First, by the definitions of $L_{\mathcal{S}}$ and L,

$$\nabla^2_{xx}L_{\mathcal{S}}(x,y_{\mathcal{S}}(x)) - \nabla^2_{xx}L(x,y(x)) = \sum_{j\in[m]}(\nabla^2[c_{\mathcal{S}}]_j(x)[y_{\mathcal{S}}(x)]_j - \nabla^2[c]_j(x)[y(x)]_j),$$

from which it follows by submultiplicity of the matrix 2-norm that

$$\begin{aligned} &|d_{\mathcal{S}}^T\nabla^2_{xx}L_{\mathcal{S}}(x,y_{\mathcal{S}}(x))d_{\mathcal{S}} - d_{\mathcal{S}}^T\nabla^2_{xx}L(x,y(x))d_{\mathcal{S}}| \\ \leq\ &\left\|\sum_{j\in[m]}(\nabla^2[c_{\mathcal{S}}]_j(x)[y_{\mathcal{S}}(x)]_j - \nabla^2[c]_j(x)[y(x)]_j)\right\|\|d_{\mathcal{S}}\|^2. \end{aligned} \tag{34}$$

With respect to the first norm on the right of (34), one finds along with the triangle inequality, absolute homogeneity of norms, (10c), and Assumption 3.1 that

$$\begin{aligned} &\left\|\sum_{j\in[m]}(\nabla^2[c_{\mathcal{S}}]_j(x)[y_{\mathcal{S}}(x)]_j - \nabla^2[c]_j(x)[y(x)]_j)\right\| \\ =\ &\left\|\sum_{j\in[m]}((\nabla^2[c_{\mathcal{S}}]_j(x) - \nabla^2[c]_j(x)))[y_{\mathcal{S}}(x)]_j + \sum_{j\in[m]}\nabla^2[c]_j(x)([y_{\mathcal{S}}(x)]_j - [y(x)]_j)\right\| \\ \leq\ &\sum_{j\in[m]}\|\nabla^2[c_{\mathcal{S}}]_j(x) - \nabla^2[c]_j(x)\||[y_{\mathcal{S}}(x)]_j| + \sum_{j\in[m]}\|\nabla^2[c]_j(x)\||[y_{\mathcal{S}}(x)]_j - [y(x)]_j| \\ \leq\ &\sum_{j\in[m]}\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla^2 c}}\|\nabla^2[c]_j(x)\||[y_{\mathcal{S}}(x)]_j| + \sum_{j\in[m]}\|\nabla^2[c]_j(x)\||[y_{\mathcal{S}}(x)]_j - [y(x)]_j| \\ \leq\ &\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla^2 c}}\kappa_{\nabla^2 c}\sum_{j\in[m]}|[y_{\mathcal{S}}(x)]_j| + \kappa_{\nabla^2 c}\sum_{j\in[m]}|[y_{\mathcal{S}}(x)]_j - [y(x)]_j|. \end{aligned} \tag{35}$$

For the second sum in (35), it follows from the fact that $\|v\|_1 \leq \sqrt{m}\|v\|_2$ for any $v \in \mathbb{R}^m$, Lemma 3.4 and $\kappa_1 = \frac{\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{\sigma_{\min}}$ that

$$\sum_{j\in[m]} |[y_{\mathcal{S}}(x)]_j - [y(x)]_j| \leq \sqrt{m}\|y_{\mathcal{S}}(x) - y(x)\| \leq \frac{9\sqrt{m}\kappa_{\nabla f}\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{2\sigma_{\min}^2} = \frac{9\sqrt{m}\kappa_{\nabla f}\kappa_1\xi_{\mathcal{S}}}{2\sigma_{\min}}, \tag{36}$$

whereas for the first sum in (35), it follows from (36), (8b), and $\xi_{\mathcal{S}} \leq \frac{1}{3\kappa_1} = \frac{\sigma_{\min}}{3\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}$ that

$$\begin{aligned}\sum_{j\in[m]} |[y_{\mathcal{S}}(x)]_j| \leq \sqrt{m}\|y_{\mathcal{S}}(x)\| &\leq \sqrt{m}(\|y_{\mathcal{S}}(x) - y(x)\| + \|y(x)\|)\\ &\leq \frac{9\sqrt{m}\kappa_{\nabla f}\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla c}}\kappa_{\nabla c}}{2\sigma_{\min}^2} + \frac{\sqrt{m}\kappa_{\nabla f}}{\sigma_{\min}}\\ &= \frac{3\sqrt{m}\kappa_{\nabla f}}{2\sigma_{\min}} + \frac{\sqrt{m}\kappa_{\nabla f}}{\sigma_{\min}} = \frac{5\sqrt{m}\kappa_{\nabla f}}{2\sigma_{\min}}.\end{aligned} \tag{37}$$

Combining (36) and (37), one finds that the right-hand side of (35) has

$$\begin{aligned}&\xi_{\mathcal{S}}\sqrt{\gamma_{\nabla^2 c}}\kappa_{\nabla^2 c}\sum_{j\in[m]} |[y_{\mathcal{S}}(x)]_j| + \kappa_{\nabla^2 c}\sum_{j\in[m]} |[y_{\mathcal{S}}(x)]_j - [y(x)]_j|\\ &\leq \sqrt{m}\frac{\kappa_{\nabla f}\kappa_{\nabla^2 c}}{2\sigma_{\min}}\left(5\sqrt{\gamma_{\nabla^2 c}} + 9\kappa_1\right)\xi_{\mathcal{S}}.\end{aligned} \tag{38}$$

As a result, combining (34), (35), and (38) yields

$$\begin{aligned}&|d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}} - d_{\mathcal{S}}^T\nabla_{xx}^2 L(x, y(x))d_{\mathcal{S}}|\\ &\leq \sqrt{m}\frac{\kappa_{\nabla f}\kappa_{\nabla^2 c}}{2\sigma_{\min}}\left(5\sqrt{\gamma_{\nabla^2 c}} + 9\kappa_1\right)\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|^2.\end{aligned} \tag{39}$$

Thus, now combining (29), (33), and (39), one finds that

$$\begin{aligned}&|d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}}|\\ \geq\ &\left(1 - \tfrac{1}{3}\kappa_1\xi_{\mathcal{S}}\right)\beta\|d_{\mathcal{S}}\|^2 - 3\left(\kappa_{\nabla^2 f} + \frac{\sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}}{\sigma_{\min}}\right)\kappa_1\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|^2\\ &- \sqrt{m}\frac{\kappa_{\nabla f}\kappa_{\nabla^2 c}}{2\sigma_{\min}}\left(5\sqrt{\gamma_{\nabla^2 c}} + 9\kappa_1\right)\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|^2\\ =\ &\left(1 - \tfrac{1}{3}\kappa_1\xi_{\mathcal{S}}\right)\beta\|d_{\mathcal{S}}\|^2 - \left(3\kappa_{\nabla^2 f}\kappa_1 + \frac{5\sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}}{2\sigma_{\min}}\left(3\kappa_1 + \sqrt{\gamma_{\nabla^2 c}}\right)\right)\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|^2.\end{aligned}$$

Along with the definition of κ_3 given in (23), one finds that

$$|d_{\mathcal{S}}^T\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))d_{\mathcal{S}}| \geq \left(1 - \tfrac{1}{3}\kappa_1\xi_{\mathcal{S}}\right)\beta\|d_{\mathcal{S}}\|^2 - \kappa_3\xi_{\mathcal{S}}\|d_{\mathcal{S}}\|^2 = \beta_{\mathcal{S}}\|d_{\mathcal{S}}\|^2, \tag{40}$$

which concludes the proof. □ □

We now conclude this subsection with Theorem 3.1 below, which provides generic worst-case complexity bounds for Algorithm 1. The theorem states generic properties that a subproblem solver needs to have in order to leverage the strong-Morse properties of the subproblems such that, by solving a sequence of subproblems, the overall sample complexity can be less than that of solving the full-sample problem directly. (As an example, the theorem focuses on the behavior of a second-order-type method for solving the subproblems, although others may be employed in practice.) We note that a value of $p_1 \in [N]$ that satisfies the requirements of the theorem is not necessarily known in practice, but as shown in our numerical experiments one can obtain better performance by estimating such a value for p_1 and solving a sequence of subproblems rather than the full-sample problem directly.

Theorem 3.1. *Suppose that Assumptions 3.1, 3.2, and 3.3 hold. In addition, suppose that $p_1 \in [N]$ is sufficiently large such that $\mathcal{S}_1$ satisfies $|\mathcal{S}_1| \geq p_1$ and (24) for $\mathcal{S} = \mathcal{S}_1$ (where the tuple $(\kappa_1, \kappa_2, \kappa_3)$ is defined by (22) and (23)), and that, for all $k \in [K]$, the pair of subproblem solver tolerances $(\epsilon_k, \zeta_k) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ is set as a fraction of $(\alpha_{\mathcal{S}_k}, \beta_{\mathcal{S}_k})$ (defined by (25) with $\mathcal{S} = \mathcal{S}_k$) such that any (ϵ_k, ζ_k)-stationary point with respect to (3) for $\mathcal{S} = \mathcal{S}_k$, call it $x_k \in \mathbb{R}^n$, satisfies*

$$\begin{aligned} &\|\nabla L_{\mathcal{S}_k}(x_k, y_{\mathcal{S}_k}(x_k))\| \leq \epsilon_k \\ \text{and} \quad &d^T \nabla^2_{xx} L_{\mathcal{S}_k}(x_k, y_{\mathcal{S}_k}(x_k)) d \geq -\zeta_k \|d\|_2^2 \quad \text{for all} \quad d \in \text{null}(\nabla c_{\mathcal{S}_k}(x_k)^T). \end{aligned}$$

Finally, suppose that the sequence of tolerances $\{(\epsilon_k, \zeta_k)\}$ is monotonically nonincreasing and the subproblem solver employed by Algorithm 1 guarantees that:

(a) The number of iterations required to reach an (ϵ_1, ζ_1)-stationary point with respect to subproblem (3) for $\mathcal{S} = \mathcal{S}_1$ is at most $\mathcal{O}(\max\{\epsilon_1^{-2}, \zeta_1^{-3}\})$.

(b) For any $k \geq 2$, the number of iterations required from x_{k-1} to reach an (ϵ_k, ζ_k)-stationary point with respect to subproblem (3) for $\mathcal{S} = \mathcal{S}_k$ is at most $\mathcal{O}(\log(\frac{\epsilon_{k-1}}{\epsilon_k}))$.

Then, the total number of constraint gradient evaluations required by a run of Algorithm 1 with an initial sample set size $p_1 < N$ is at most

$$\mathcal{O}\left(|\mathcal{S}_1| \max\{\epsilon_1^{-2}, \zeta_1^{-3}\} + \sum_{i=2}^{K} |\mathcal{S}_k| \log\left(\tfrac{\epsilon_{k-1}}{\epsilon_k}\right) \right),$$

whereas the total number of constraint gradient evaluations required by a run of Algorithm 1 with an initial sample set size $p_1 = N$ is at most

$$\mathcal{O}(N \max\{\epsilon_1^{-2}, \zeta_1^{-3}\}).$$

In our final theorem in the next subsection, namely, Theorem 3.2 on page 29, we show a more specific bound on the sample complexity of the algorithm when a particular subproblem solver is employed.

3.2 Subproblem Solver for Algorithm 1

Our goal in this subsection is to prove the existence of an algorithm that yields the subproblem requirements of our generic sample-complexity bound in Theorem 3.1. The algorithm that we consider is based on minimization of Fletcher's augmented Lagrangian function, and in particular the strategies proposed and analyzed in [4]. Much of our efforts in this subsection focus on showing that, by employing the algorithm—specifically, [4, Algorithm 1]—we can obtain our desired complexity bounds with respect to our termination conditions in (6). We emphasize that this is nontrivial since our termination conditions are different than those in [4].

We begin with a lemma showing that, if the initial sample size is sufficiently large and the subproblem tolerances are set appropriately, then the iterates computed by each subproblem solve are approximate second-order stationary points.

Lemma 3.7. *Suppose that $(\kappa_1, \kappa_2, \kappa_3)$ is defined by (22) and (23), and that with*

$$\tau_1 := \kappa_2 \quad \text{and} \quad \tau_2 := \left(\kappa_{\nabla^2 f} + \tfrac{\sqrt{m} \kappa_{\nabla f} \kappa_{\nabla^2 c}}{\sigma_{\min}} \right) \kappa_1 \tag{41}$$

the subproblem tolerances are set for all $k \in [K]$ (recall (9)) as

$$\epsilon_k := \tau_1 \xi_{\mathcal{S}_k} \quad \text{and} \quad \zeta_k := \tau_2 \xi_{\mathcal{S}_k}. \tag{42}$$

Then, as long as $p_1 \in [N]$ is sufficiently large such that

$$\sqrt{\tfrac{N(N-p_1)}{p_1^2}} \leq \min\left\{ \tfrac{1}{3\kappa_1}, \tfrac{\alpha}{4\kappa_2}, \tfrac{2\beta}{9\kappa_3} \right\}, \tag{43}$$

it follows for any $k \in [K]$ that if $x_{\mathcal{S}_k} \in \mathbb{R}^n$ is (ϵ_k, ζ_k)-stationary with respect to (3) *for $\mathcal{S} = \mathcal{S}_k$, then with respect to* (3) *with $\mathcal{S} = \mathcal{S}_{k+1}$ one has*

$$\begin{aligned} \|\nabla L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}))\| \leq 3\epsilon_k \leq \alpha_{\mathcal{S}_{k+1}} \quad \textit{and} \\ d^T \nabla^2_{xx} L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}))d \geq \beta_{\mathcal{S}_{k+1}} \|d\|^2 \textit{ for all } d \in \text{null}(\nabla c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})^T), \end{aligned} \tag{44}$$

where $\alpha_{\mathcal{S}_{k+1}} \geq \frac{1}{2}\alpha$ and $\beta_{\mathcal{S}_{k+1}} \geq \frac{1}{2}\beta$ are defined according to (25)*.*

Proof. Consider arbitrary $k \in [K]$. Our first aim is to bound the norm of the difference between the gradients of $L_{\mathcal{S}_k}$ and $L_{\mathcal{S}_{k+1}}$ with respect to the point $x_{\mathcal{S}_k}$. Toward this end, by (7), define $y_{\mathcal{S}_k} := y_{\mathcal{S}_k}(x_{\mathcal{S}_k})$ and $y_{\mathcal{S}_{k+1}} := y_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})$. Then, by similar arguments as led to (26) and the triangle inequality, one finds

$$\begin{aligned} &\|\nabla_x L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) - \nabla_x L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k})\| \\ \leq\ &\|\mathcal{R}(\nabla c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k}))\|\kappa_{\nabla f} \\ =\ &\|(\mathcal{R}(\nabla c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c(x_{\mathcal{S}_k}))) + (\mathcal{R}(\nabla c(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})))\|\kappa_{\nabla f} \\ \leq\ &(\|\mathcal{R}(\nabla c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c(x_{\mathcal{S}_k}))\| + \|\mathcal{R}(\nabla c(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k}))\|)\kappa_{\nabla f}. \end{aligned} \tag{45}$$

Next, let us bound the two terms in parentheses on the right-hand side of (45). Note that by Assumption 3.1 the matrices $\nabla c(x_{\mathcal{S}_k})$, $\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})$, and $\nabla c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})$ all have full-column rank, meaning that they have the same rank, so by [9, Theorems 2.3–2.4] and the same arguments as led to (27), one finds

$$\|\mathcal{R}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c(x_{\mathcal{S}_k}))\| \leq \kappa_1 \xi_{\mathcal{S}_k} \tag{46a}$$

$$\text{and} \quad \|\mathcal{R}(\nabla c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c(x_{\mathcal{S}_k}))\| \leq \kappa_1 \xi_{\mathcal{S}_{k+1}} \leq \kappa_1 \xi_{\mathcal{S}_k}, \tag{46b}$$

where the last inequality follows from (9) and the fact that the algorithm guarantees that $|\mathcal{S}_{k+1}| > |\mathcal{S}_k|$. Combining (46a) and (46b) with (45), one obtains

$$\|\nabla_x L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) - \nabla_x L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k})\| \leq 2\kappa_{\nabla f} \kappa_1 \xi_{\mathcal{S}_k}. \tag{47}$$

On the other hand, from the triangle inequality, (10a), and $\xi_{\mathcal{S}_{k+1}} \leq \xi_{\mathcal{S}_k}$, one has

$$\begin{aligned} &\|\nabla_y L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) - \nabla_y L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k})\| \\ =\ &\|c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}) - c_{\mathcal{S}_k}(x_{\mathcal{S}_k})\| \\ \leq\ &\|c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}) - c(x_{\mathcal{S}_k})\|_2 + \|c(x_{\mathcal{S}_k}) - c_{\mathcal{S}_k}(x_{\mathcal{S}_k})\| \\ \leq\ &\sqrt{\gamma_c}\xi_{\mathcal{S}_{k+1}} + \sqrt{\gamma_c}\xi_{\mathcal{S}_k} \leq 2\sqrt{\gamma_c}\xi_{\mathcal{S}_k}. \end{aligned} \tag{48}$$

Combining (47) and (48), one finds that

$$\|\nabla L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) - \nabla L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k})\|^2 \leq 4(\gamma_c + \kappa^2_{\nabla f}\kappa_1^2)\xi^2_{\mathcal{S}_k} = 4\kappa_2^2 \xi^2_{\mathcal{S}_k}, \tag{49}$$

which further gives under the conditions of the lemma that

$$\begin{aligned} &\|\nabla L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}})\| \\ \leq\ &\|\nabla L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) - \nabla L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k})\| + \|\nabla L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k})\| \qquad (50) \\ \leq\ &2\kappa_2 \xi_{\mathcal{S}_k} + \epsilon_k = 3\epsilon_k = 3\tau_1 \xi_{\mathcal{S}_k} \leq \tfrac{3}{4}\alpha. \qquad (51) \end{aligned}$$

At the same time, by (25), $\xi_{\mathcal{S}_{k+1}} \leq \xi_{\mathcal{S}_k}$, and $\kappa_2 \xi_{\mathcal{S}_k} \leq \frac{1}{4}\alpha$, one has

$$\alpha_{\mathcal{S}_{k+1}} = \alpha - \kappa_2 \xi_{\mathcal{S}_{k+1}} \geq \alpha - \kappa_2 \xi_{\mathcal{S}_k} \geq \alpha - \tfrac{1}{4}\alpha = \tfrac{3}{4}\alpha. \tag{52}$$

Combining (51) and (52), one obtains the first desired conclusion in (44) that

$$\|\nabla L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}})\| \leq 3\epsilon_k \leq \alpha_{\mathcal{S}_{k+1}}. \tag{53}$$

Our next goal is to prove the second inequality in (44). Toward this end, first note that Lemma 3.6 applies here since the right-hand side of (43) is less than or equal to the right-hand side of (24). Hence, by Lemma 3.6, it follows that (3) with $\mathcal{S} = \mathcal{S}_{k+1}$ is $(\alpha_{\mathcal{S}_{k+1}}, \beta_{\mathcal{S}_{k+1}})$-strongly Morse. Combined with the prior conclusion that $x_{\mathcal{S}_k}$ satisfies (53), one has for all $d_{\mathcal{S}_{k+1}} \in \text{null}(\nabla c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})^T)$ that

$$|d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}}| \geq \beta_{\mathcal{S}_{k+1}} \|d_{\mathcal{S}_{k+1}}\|^2.$$

To prove the second inequality in (44), it is necessary to show that, in fact, this inequality always holds without the absolute value on the left-hand side. Toward showing this, let us define $\bar{d}_{\mathcal{S}_{k+1}} := \mathcal{N}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})^T) d_{\mathcal{S}_{k+1}}$. Then, one finds

$$\begin{aligned}
&d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} \\
&= \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}} \\
&\quad + (d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}}) \\
&\geq \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}} \\
&\quad - |d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}}|.
\end{aligned} \tag{54}$$

Since $\mathcal{N}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})^T)$ is a projection matrix, $\|\mathcal{N}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})^T)\| \leq 1$ and

$$\|\bar{d}_{\mathcal{S}_{k+1}}\|^2 = \|\mathcal{N}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})^T) d_{\mathcal{S}_{k+1}}\|^2 \leq \|\mathcal{N}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})^T)\|^2 \|d_{\mathcal{S}_{k+1}}\|^2 \leq \|d_{\mathcal{S}_{k+1}}\|^2.$$

Along with the fact that $x_{\mathcal{S}_k}$ is (ϵ_k, ζ_k)-stationary, this means that the first-term on the right-hand side of (54) satisfies the inequalities

$$\bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}} \geq -\zeta_k \|\bar{d}_{\mathcal{S}_{k+1}}\|^2 \geq -\zeta_k \|d_{\mathcal{S}_{k+1}}\|^2. \tag{55}$$

On the other hand, with respect to the second term on the right-hand side of (54), observe that with $y(x_{\mathcal{S}_k})$ defined by (7) the triangle inequality yields

$$\begin{aligned}
&|d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}}| \\
&\leq |d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} - d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) d_{\mathcal{S}_{k+1}}| \\
&\quad + |d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) \bar{d}_{\mathcal{S}_{k+1}}| \\
&\quad + |\bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) \bar{d}_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}}|.
\end{aligned} \tag{56}$$

Next, one can observe that the first and third terms on the right-hand side of (56) can be bounded as in (39) with respect to $\mathcal{S} = \mathcal{S}_k$ and $\mathcal{S} = \mathcal{S}_{k+1}$, respectively. To see this, note that a bound of the form in (34) holds since (34) followed using general matrix-norm inequalities. Furthermore, a bound of the form in (35), which relies on (34) and (10c) holds, and bounds of the form in (36)–(37) hold since these rely on $\xi_{\mathcal{S}_k} \leq \frac{1}{3\kappa_1}$ and $\xi_{\mathcal{S}_{k+1}} \leq \frac{1}{3\kappa_1}$, which hold in the present setting. Thus, since (39) follows from a combination of (34), (35), (36), and (37), one obtains that with

$$\lambda := \sqrt{m} \tfrac{\kappa_{\nabla f} \kappa_{\nabla^2 c}}{2\sigma_{\min}} (5\sqrt{\gamma_{\nabla^2 c}} + 9\kappa_1)$$

the first and third terms on the right-hand side of (56) respectively satisfy

$$\begin{aligned}
&|d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, z_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} - d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) d_{\mathcal{S}_{k+1}}| \\
&\leq \lambda \xi_{\mathcal{S}_{k+1}} \|d_{\mathcal{S}_{k+1}}\|^2 \leq \lambda \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2 \\
\text{and} \quad &|\bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) \bar{d}_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_k}(x_{\mathcal{S}_k}, y_{\mathcal{S}_k}) \bar{d}_{\mathcal{S}_{k+1}}|
\end{aligned} \tag{57}$$

$$\le \lambda \xi_{\mathcal{S}_k} \|\bar{d}_{\mathcal{S}_{k+1}}\|^2 \le \lambda \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2, \tag{58}$$

As for the second term on the right of (56), it follows from the submultiplicity of the matrix 2-norm, the triangle inequality, $\|\bar{d}_{\mathcal{S}_{k+1}}\| \le \|d_{\mathcal{S}_{k+1}}\|$ and (8c) that

$$\begin{aligned}
&|d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k})) \bar{d}_{\mathcal{S}_{k+1}}| \\
&= |(d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}})^T \nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k}))(d_{\mathcal{S}_{k+1}} + \bar{d}_{\mathcal{S}_{k+1}})| \\
&\le \|\nabla_{xx}^2 L(x_{\mathcal{S}_k}, y(x_{\mathcal{S}_k}))\| \|d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}\| \|d_{\mathcal{S}_{k+1}} + \bar{d}_{\mathcal{S}_{k+1}}\| \\
&\le 2\left(\kappa_{\nabla^2 f} + \tfrac{\sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}}{\sigma_{\min}}\right) \|d_{\mathcal{S}_{k+1}}\| \|d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}\| \\
&= 2\tfrac{\tau_2}{\kappa_1} \|d_{\mathcal{S}_{k+1}}\| \|d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}\|.
\end{aligned} \tag{59}$$

Now with respect to $\|d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}\|$ one finds from the triangle inequality, submultiplicity of the matrix 2-norm, Lemma 3.5, and (46a) that

$$\begin{aligned}
&\|d_{\mathcal{S}_{k+1}} - \bar{d}_{\mathcal{S}_{k+1}}\| \\
&= \|\mathcal{R}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})) d_{\mathcal{S}_{k+1}}\| \\
&\le \|\mathcal{R}(\nabla c(x_{\mathcal{S}_k})) d_{\mathcal{S}_{k+1}}\| + \|(\mathcal{R}(\nabla c(x_{\mathcal{S}_k})) - \mathcal{R}(\nabla c_{\mathcal{S}_k}(x_{\mathcal{S}_k})))\| \|d_{\mathcal{S}_{k+1}}\| \\
&\le \kappa_1 \xi_{\mathcal{S}_{k+1}} \|d_{\mathcal{S}_{k+1}}\| + \kappa_1 \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\| \le 2\kappa_1 \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|.
\end{aligned} \tag{60}$$

Combining (54)–(60) with (23), (41), and $\zeta_k = \tau_2 \xi_{\mathcal{S}_k}$, one finds

$$\begin{aligned}
&d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} \\
\ge\ & -\tau_2 \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2 - 2\lambda \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2 - 4\tau_2 \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2 \\
=\ & \left(-5\left(\kappa_{\nabla^2 f} + \tfrac{\sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}}{\sigma_{\min}}\right)\kappa_1 - 2\tfrac{\sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}}{2\sigma_{\min}}\left(5\sqrt{\gamma_{\nabla^2 c}} + 9\kappa_1\right)\right) \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2 \\
\ge\ & -2\left(3\kappa_{\nabla^2 f}\kappa_1 + \left(\tfrac{5\sqrt{m}\kappa_{\nabla f}\kappa_{\nabla^2 c}}{2\sigma_{\min}}\right)\left(3\kappa_1 + \sqrt{\gamma_{\nabla^2 c}}\right)\right) \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2 \\
=\ & -2\kappa_3 \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2.
\end{aligned}$$

Since (43) requires $\xi_{\mathcal{S}_k} \le \frac{2\beta}{9\kappa_3}$ it follows that

$$d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} \ge -2\kappa_3 \xi_{\mathcal{S}_k} \|d_{\mathcal{S}_{k+1}}\|^2 \ge -\tfrac{4}{9}\beta \|d_{\mathcal{S}_{k+1}}\|^2. \tag{61}$$

On the other hand, recall from (25) that $\beta_{\mathcal{S}_{k+1}} = \left(1 - \frac{1}{3}\kappa_1 \xi_{\mathcal{S}_{k+1}}\right)\beta - \kappa_3 \xi_{\mathcal{S}_{k+1}}$, which along with $\xi_{\mathcal{S}_{k+1}} \le \xi_{\mathcal{S}_k}$ and (43) (specifically, $\xi_{\mathcal{S}_k} \le \min\{\frac{1}{3\kappa_1}, \frac{2\beta}{9\kappa_3}\}$) implies

$$\beta_{\mathcal{S}_{k+1}} \ge \left(1 - \tfrac{1}{3}\kappa_1 \xi_{\mathcal{S}_k}\right)\beta - \kappa_3 \xi_{\mathcal{S}_k} \ge \beta - \tfrac{1}{9}\beta - \tfrac{2}{9}\beta = \tfrac{2}{3}\beta.$$

Since subproblem (3) with $\mathcal{S} = \mathcal{S}_{k+1}$ is $(\alpha_{k+1}, \beta_{k+1})$-strongly morse, it follows that

$$|d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}}| \ge \beta_{\mathcal{S}_{k+1}} \|d_{\mathcal{S}_{k+1}}\|^2 \ge \tfrac{2}{3}\beta \|d_{\mathcal{S}_{k+1}}\|^2.$$

That said, since (61) holds and $-\frac{4}{9} > -\frac{2}{3}$, it must hold that

$$d_{\mathcal{S}_{k+1}}^T \nabla_{xx}^2 L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}) d_{\mathcal{S}_{k+1}} \ge \beta_{\mathcal{S}_{k+1}} \|d_{\mathcal{S}_{k+1}}\|^2,$$

which completes the proof. □ □

Our next aim is to employ [4, Theorem 3.5] to prove that [4, Algorithm 1] possesses a certain worst-case iteration complexity bound when employed as the subproblem solver in our Algorithm 1. For reference in our subsequent analysis, we introduce, for any $\mathcal{S} \subseteq [N]$, Fletcher's augmented Lagrangian function with respect to our subproblem (3), namely, $F_\mathcal{S} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$F_\mathcal{S}(x) = f(x) + c_\mathcal{S}(x)^T y_\mathcal{S}(x) + \rho_\mathcal{S} \|c_\mathcal{S}(x)\|^2, \tag{62}$$

where $y_\mathcal{S}(x)$ is defined as in (7) and $\rho_\mathcal{S} \in \mathbb{R}_{>0}$ is a penalty parameter.

Lemma 3.8. *Suppose there exists $R \in \mathbb{R}_{>0}$ such that for any $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$ the sublevel set $\mathcal{C}_{\mathcal{S},R} := \{x \in \mathbb{R}^n : \|c_\mathcal{S}(x)\| \leq R\}$ contains x_0 and is compact. Then, for any such $\mathcal{S}$, there exists $\hat{\rho}_\mathcal{S} \in \mathbb{R}_{>0}$ such that the requirements of [4, Theorem 3.5] hold for all $\rho_\mathcal{S} \geq \hat{\rho}_\mathcal{S}$. Thus, for any such $\rho_\mathcal{S}$ there exists $(u_{\mathcal{S},1}, u_{\mathcal{S},2}) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ such that for any $(\epsilon, \zeta) \in (0, \frac{\sqrt{5}}{2}R] \times (0, 1]$ one has that [4, Algorithm 1] with starting point x_0 locates an (ϵ, ζ)-stationary point (see (6)) in a number of iterations that is at most*

$$T_\mathcal{S} = \max\{u_{\mathcal{S},1}\epsilon^{-2}, u_{\mathcal{S},2}\zeta^{-3}\}. \tag{63}$$

Proof. To prove the lemma, it suffices to prove that the requirements of [4, Theorem 3.5] hold in our present setting and that the worst-case complexity bound from that theorem holds with respect to our termination condition (6), which is different from the termination condition for [4, Algorithm 1]. Toward these ends, let us begin by considering an arbitrary sample set $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$.

Assumption A1 in [4] requires the existence of $\hat{R} \in \mathbb{R}_{>0}$ and $\hat{\sigma} \in \mathbb{R}_{>0}$ such that for all $x \in \mathbb{R}^n$ with $\|c_\mathcal{S}(x)\| \leq \hat{R}$ one has $\sigma_{\min}(\nabla c_\mathcal{S}(x)^T) \geq \hat{\sigma}$. This holds in our present setting under the conditions of this lemma and Assumption 3.1. Assumption A2 in [4] requires the existence of R as stated in the conditions of this lemma. Assumption A3 in [4] requires the existence of $\eta_\mathcal{S} \in \mathbb{R}_{>0}$ such that

$$\|c_\mathcal{S}(x + d) - c_\mathcal{S}(x) - \nabla c_\mathcal{S}(x)^T d\| \leq \eta_\mathcal{S} \|d\|^2 \ \text{ for all } \ (x, d) \in \mathcal{C}_{\mathcal{S},R} \times \mathbb{R}^n. \tag{64}$$

To show that this holds in the present setting, observe that for all $(x, j) \in \mathbb{R}^n \times [m]$ it follows from Assumption 3.3, (10c), and the triangle inequality that

$$\begin{aligned} \|\nabla^2 [c_\mathcal{S}]_j(x)\| &= \|\nabla^2 [c]_j(x) + \nabla^2 [c_\mathcal{S}]_j(x) - \nabla^2 [c]_j(x)\| \\ &\leq \|\nabla^2 [c]_j(x)\| + \|\nabla^2 [c_\mathcal{S}]_j(x) - \nabla^2 [c]_j(x)\| \\ &\leq \|\nabla^2 [c]_j(x)\| + \xi_\mathcal{S} \sqrt{\gamma_{\nabla^2 c}} \|\nabla^2 [c]_j(x)\| \\ &\leq \left(1 + \sqrt{\tfrac{N(N-p_1)\gamma_{\nabla^2 c}}{p_1^2}}\right) \kappa_{\nabla^2 c} =: \kappa_{p_1}. \end{aligned} \tag{65}$$

Now consider arbitrary $(x, d) \in \mathcal{C}_{\mathcal{S},R} \times \mathbb{R}^n$ and observe that, by Taylor's theorem and (65), it follows that for all $j \in [m]$ there exists a point $\tilde{x}_j \in \mathbb{R}^n$ such that

$$\begin{aligned} &|[c_\mathcal{S}]_j(x + d) - [c_\mathcal{S}]_j(x) - \nabla [c_\mathcal{S}]_j(x)^T d| \\ &= |\tfrac{1}{2} d^T \nabla^2 [c_\mathcal{S}]_j(\tilde{x}_j) d| \leq \tfrac{1}{2} \|\nabla^2 [c_\mathcal{S}]_j(\tilde{x}_j)\| \|d\|^2 \leq \tfrac{1}{2} \kappa_{p_1} \|d\|^2. \end{aligned} \tag{66}$$

Consequently, one finds that

$$\|c_\mathcal{S}(x + d) - c_\mathcal{S}(x) - \nabla c_\mathcal{S}(x)^T d\| \leq \sqrt{\sum_{j \in [m]} \tfrac{1}{4} \kappa_{p_1}^2 \|d\|^4} = \tfrac{1}{2} \sqrt{m} \kappa_{p_1} \|d\|^2,$$

from which it follows that (64) holds for any $\eta_\mathcal{S} = \frac{1}{2}\sqrt{m}\kappa_{p_1}$. Assumption A4 in [4] requires that $x_0 \in \mathcal{C}_{\mathcal{S},R}$, as is required in the conditions of this lemma. Finally, Assumption A5 in [4] requires that the penalty parameter is chosen greater than

$$\max_{x \in \mathcal{C}_{\mathcal{S},R}} \max \left\{ \tfrac{\sigma_{\max}(\nabla c_\mathcal{S}(x)) \sigma_{\max}(\nabla y_\mathcal{S}(x))}{2\sigma_{\min}(\nabla c_\mathcal{S}(x))^2}, \tfrac{\sigma_{\max}(\nabla y_\mathcal{S}(x))}{\sigma_{\min}(\nabla c_\mathcal{S}(x))}, \tfrac{1}{\sigma_{\min}(\nabla c_\mathcal{S}(x))} \right\}.$$

Since $\mathcal{C}_{\mathcal{S},R}$ is compact, it follows under Assumption 3.1 and the extreme value theorem that there exists $\rho_{\mathcal{S}}$ as stated in the lemma. All together, one can conclude that the requirements of [4, Theorem 3.5] hold in our present setting, as desired.

All that remains is to prove that the worst-case iteration complexity bound from [4, Theorem 3.5] yields the desired conclusion of the lemma for our setting. Toward this end, let us introduce $\tau_{\mathcal{S}} \in \mathbb{R}_{>0}$ as equal to the positive real number "C" introduced in [4, Corollary 2.7] (dependent on second-order derivatives of the element functions $\{[c_{\mathcal{S}}]_j\}_{j\in[m]}$ and $\{[y_{\mathcal{S}}]_j\}_{j\in[m]}$ over $\mathcal{C}_{\mathcal{S},R}$). One can now state by [4, Theorem 3.5] that with respect to [4, Algorithm 1] employed to solve subproblem (3) with $x_0 \in \mathcal{C}_{\mathcal{S},R}$ that there exists $(\overline{u}_{\mathcal{S},1}, \overline{u}_{\mathcal{S},2}, \overline{u}_{\mathcal{S},3}) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ such that for any $(\epsilon, \zeta) \in (0, \frac{\sqrt{5}}{2}R] \times (0,1]$ and corresponding $\epsilon_{F_{\mathcal{S}}} := \min\{\frac{\sqrt{5}\epsilon}{5}, \frac{\zeta}{100\tau_{\mathcal{S}}}\}$ and $\zeta_{F_{\mathcal{S}}} := \frac{99\zeta}{100} + \tau_{\mathcal{S}}\epsilon_{F_{\mathcal{S}}}$ one finds that in a number of iterations that is at most

$$\overline{T}_{\mathcal{S}} := \max\left\{\overline{u}_{\mathcal{S},1}\epsilon_{F_{\mathcal{S}}}^{-2}, \overline{u}_{\mathcal{S},2}\zeta_{F_{\mathcal{S}}}^{-1}, \overline{u}_{\mathcal{S},3}\zeta_{F_{\mathcal{S}}}^{-3}\right\} \tag{67}$$

the algorithm produces a point $\overline{x} \in \mathbb{R}^n$ satisfying

$$\|\nabla F_{\mathcal{S}}(\overline{x})\| \leq \epsilon_{F_{\mathcal{S}}} \;\text{ and }\; d^T\nabla^2 F_{\mathcal{S}}(\overline{x})d \geq -\zeta_{F_{\mathcal{S}}}\|d\|^2 \;\text{ for all }\; d \in \mathbb{R}^n. \tag{68}$$

(Note that to apply [4, Theorem 3.5] it has been observed that the former tolerance satisfies $\epsilon_{F_{\mathcal{S}}} \leq \frac{\sqrt{5}}{5}\epsilon \leq \frac{1}{2}R$.) In addition, it also follows from [4, Theorem 3.5] that the point $\overline{x}$ satisfies $\|\nabla_x L(\overline{x}, y(\overline{x}))\| \leq 2\epsilon_{F_{\mathcal{S}}}$, $\|\nabla_y L(\overline{x}, y(\overline{x}))\| \leq \epsilon_{F_{\mathcal{S}}}$, and

$$d^T\nabla^2_{xx}L(\overline{x}, y(\overline{x}))d \geq -\zeta_{F_{\mathcal{S}}}\|d\|^2 \;\text{ for all }\; d \in \text{null}(\nabla c_{\mathcal{S}}(\overline{x})^T).$$

Observing that this means $\|\nabla L(\overline{x}, y(\overline{x}))\| \leq \sqrt{5}\epsilon_{F_{\mathcal{S}}} \leq \epsilon$, and observing that $\zeta_{F_{\mathcal{S}}} = \frac{99\zeta}{100} + \tau_{\mathcal{S}}\epsilon_{F_{\mathcal{S}}} \leq \zeta$, it follows that (6) holds, as desired. Finally, to show that this is achieved in a number of iterations that is at most $T_{\mathcal{S}}$ of the form in (63), one only needs to plug in the definitions of $\epsilon_{F_{\mathcal{S}}}$ and $\zeta_{F_{\mathcal{S}}}$ into $\overline{T}_{\mathcal{S}}$ in (67). This yields

$$\begin{aligned}\overline{T}_{\mathcal{S}} &\leq \max\{5\overline{u}_{\mathcal{S},1}\epsilon^{-2}, (100\tau_{\mathcal{S}})^2\overline{u}_{\mathcal{S},1}\zeta^{-2}, \overline{u}_{\mathcal{S},2}(\tfrac{99\zeta}{100} + \tau_{\mathcal{S}}\tfrac{\sqrt{5}\epsilon}{5})^{-1}, \overline{u}_{\mathcal{S},2}\zeta^{-1}, \\ &\qquad \overline{u}_{\mathcal{S},3}(\tfrac{99\zeta}{100} + \tau_{\mathcal{S}}\tfrac{\sqrt{5}\epsilon}{5})^{-3}, \overline{u}_{\mathcal{S},3}\zeta^{-3}\} \\ &\leq \max\{5\overline{u}_{\mathcal{S},1}\epsilon^{-2}, (100\tau_{\mathcal{S}})^2\overline{u}_{\mathcal{S},1}\zeta^{-2}, \overline{u}_{\mathcal{S},2}(\tfrac{99\zeta}{100})^{-1}, \overline{u}_{\mathcal{S},3}(\tfrac{99\zeta}{100})^{-3}\}.\end{aligned}$$

Thus, since $\zeta \in (0,1]$ implies $\max\{\zeta^{-1}, \zeta^{-2}\} \leq \zeta^{-3}$, the conclusion holds with $u_{\mathcal{S},1} = 5\overline{u}_{\mathcal{S},1}$ and $u_{\mathcal{S},2} = \max\{(100\tau_{\mathcal{S}})^2\overline{u}_{\mathcal{S},1}, \frac{100}{99}\overline{u}_{\mathcal{S},2}, (\frac{100}{99})^3\overline{u}_{\mathcal{S},3}\}$. □ □

Since [4] is based on Fletcher's Augmented Lagrangian function, we now show that (44) offers bounds on first- and second-order derivatives of (62).

Lemma 3.9. *Suppose there exists $R \in \mathbb{R}_{>0}$ such that for any $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$ the sublevel set $\mathcal{C}_{\mathcal{S},R} := \{x \in \mathbb{R}^n : \|c_{\mathcal{S}}(x)\| \leq R\}$ contains x_0 and is compact. In addition, for any such $\mathcal{S}$ and any given $\beta_{\mathcal{S}} \in \mathbb{R}_{>0}$, define*

$$\begin{aligned}\kappa_{\mathcal{S}} &:= \max_{x\in\mathcal{C}_{\mathcal{S},R}} \left\{\|\nabla f(x)\|, \|y_{\mathcal{S}}(x)\|, \|\nabla y_{\mathcal{S}}(x)\|, \|\nabla c_{\mathcal{S}}(x)\|, \right. \\ &\qquad\qquad \left. \max_{j\in[m]}\left\{\|\nabla^2[y_{\mathcal{S}}]_j(x)\|, \|\nabla^2[c_{\mathcal{S}}]_j(x)\|\right\}\right\}, \\ \eta_{\mathcal{S},1} &:= \left(\frac{2\sqrt{m}}{\sigma_{\min}} + m\right)\kappa_{\mathcal{S}}, \\ \eta_{\mathcal{S},2} &:= 2m\kappa_{\mathcal{S}}, \\ \eta_{\mathcal{S},3} &:= \frac{\frac{1}{2}\beta_{\mathcal{S}}\sigma_{\min}^2}{\eta_{\mathcal{S},1}\sigma_{\min}^2 + \frac{1}{2}\beta_{\mathcal{S}}\eta_{\mathcal{S},2}}, \\ \eta_{\mathcal{S},4} &:= \eta_{\mathcal{S},1}\eta_{\mathcal{S},3} + (1 + m\kappa_{\mathcal{S}})\kappa_{\mathcal{S}} \\ \textit{and}\;\; \overline{\epsilon}_{\mathcal{S}} &:= \frac{\frac{1}{2}\beta_{\mathcal{S}}\sigma_{\min}^2}{\eta_{\mathcal{S},1}\sigma_{\min}^2 + \frac{1}{2}\beta_{\mathcal{S}}\eta_{\mathcal{S},2} + \eta_{\mathcal{S},2}\eta_{\mathcal{S},4}}\end{aligned} \tag{69}$$

along with the functions $\underline{\rho_{\mathcal{S}}} : (0, \bar{\epsilon}_{\mathcal{S}}) \to \mathbb{R}_{>0}$ *and* $\overline{\rho_{\mathcal{S}}} : (0, \bar{\epsilon}_{\mathcal{S}}) \to \mathbb{R}_{>0}$ *defined by*

$$\underline{\rho_{\mathcal{S}}}(\epsilon) = \frac{\frac{1}{2}\beta_{\mathcal{S}} + \eta_{\mathcal{S},4}}{2\sigma_{\min}^2 - \eta_{\mathcal{S},2}\epsilon} \quad \text{and} \quad \overline{\rho_{\mathcal{S}}}(\epsilon) = \frac{\frac{1}{2}\beta_{\mathcal{S}} - \eta_{\mathcal{S},1}\epsilon}{\eta_{\mathcal{S},2}\epsilon}. \tag{70}$$

Then, for any such $\mathcal{S}$ *and any* $\beta_{\mathcal{S}} \in \mathbb{R}_{>0}$*, the following hold.*

(a) for any $\epsilon \in (0, \bar{\epsilon}_{\mathcal{S}})$ *one has* $0 < \underline{\rho_{\mathcal{S}}}(\epsilon) < \overline{\rho_{\mathcal{S}}}(\epsilon)$*;*

(b) for any $\epsilon \in (0, \min\{R, \bar{\epsilon}_{\mathcal{S}}\})$ *and* $\rho_{\mathcal{S}} \in (\underline{\rho_{\mathcal{S}}}(\epsilon), \overline{\rho_{\mathcal{S}}}(\epsilon))$*, if* x *yields*

$$\begin{aligned} &\|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \epsilon \\ \text{and} \quad &d^T \nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) d \geq \beta_{\mathcal{S}} \|d\|^2 \quad \text{for all} \quad d \in \operatorname{null}(\nabla c_{\mathcal{S}}(x)^T) \end{aligned} \tag{71}$$

then

$$\|c_{\mathcal{S}}(x)\| \leq R, \quad \|\nabla F_{\mathcal{S}}(x)\| \leq (1 + \kappa_{\mathcal{S}} + 2\rho_{\mathcal{S}}\kappa_{\mathcal{S}})\epsilon, \quad \text{and} \quad \nabla^2 F_{\mathcal{S}}(x) \succeq \tfrac{1}{2}\beta_{\mathcal{S}} I. \tag{72}$$

Proof. Consider arbitrary $\beta_{\mathcal{S}} \in \mathbb{R}_{>0}$ and $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$. Part (a) follows from the fact that, for any $\epsilon \in (0, \bar{\epsilon}_{\mathcal{S}})$, one finds $\epsilon < \bar{\epsilon}_{\mathcal{S}} \leq \frac{\sigma_{\min}^2}{\eta_{\mathcal{S},2}}$ and

$$\begin{aligned} &\left\{ \bar{\epsilon}_{\mathcal{S}} \leq \frac{\frac{1}{2}\beta_{\mathcal{S}}\sigma_{\min}^2}{\eta_{\mathcal{S},1}\sigma_{\min}^2 + \frac{1}{2}\beta_{\mathcal{S}}\eta_{\mathcal{S},2} + \eta_{\mathcal{S},2}\eta_{\mathcal{S},4}} \right\} \iff \left\{ \frac{\frac{1}{2}\beta_{\mathcal{S}} + \eta_{\mathcal{S},4}}{\sigma_{\min}^2} \leq \frac{\frac{1}{2}\beta_{\mathcal{S}} - \eta_{\mathcal{S},1}\bar{\epsilon}_{\mathcal{S}}}{\eta_{\mathcal{S},2}\bar{\epsilon}_{\mathcal{S}}} \right\} \\ &\implies \left\{ \frac{\frac{1}{2}\beta_{\mathcal{S}} + \eta_{\mathcal{S},4}}{2\sigma_{\min}^2 - \eta_{\mathcal{S},2}\epsilon} < \frac{\frac{1}{2}\beta_{\mathcal{S}} - \eta_{\mathcal{S},1}\epsilon}{\eta_{\mathcal{S},2}\epsilon} \right\} \iff \left\{ \underline{\rho_{\mathcal{S}}}(\epsilon) < \overline{\rho_{\mathcal{S}}}(\epsilon) \right\}. \end{aligned}$$

Let us now prove part (b). Toward this end, consider arbitrary $\epsilon \in (0, \min\{R, \bar{\epsilon}_{\mathcal{S}}\})$, $\rho_{\mathcal{S}} \in (\underline{\rho_{\mathcal{S}}}(\epsilon), \overline{\rho_{\mathcal{S}}}(\epsilon))$, and x satisfying (71). Thus,

$$\max\{\|\nabla f(x) + \nabla c_{\mathcal{S}}(x) y_{\mathcal{S}}(x)\|, \|c_{\mathcal{S}}(x)\|\} \leq \|\nabla L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \leq \epsilon \leq R, \tag{73}$$

which among other things gives the first inequality in (72). Consequently, $x \in \mathcal{C}_{\mathcal{S},R}$, so it follows from (62), the chain rule, the triangle inequality, and submultiplicity of the matrix 2-norm that

$$\begin{aligned} \|\nabla F_{\mathcal{S}}(x)\| &\leq \|\nabla f(x) + \nabla c_{\mathcal{S}}(x) y_{\mathcal{S}}(x)\| + (\|\nabla y_{\mathcal{S}}(x)\| + 2\rho_{\mathcal{S}}\|\nabla c_{\mathcal{S}}(x)\|)\|c_{\mathcal{S}}(x)\| \\ &\leq \|\nabla f(x) + \nabla c_{\mathcal{S}}(x) y_{\mathcal{S}}(x)\| + (1 + 2\rho_{\mathcal{S}})\kappa_{\mathcal{S}}\|c_{\mathcal{S}}(x)\|. \end{aligned} \tag{74}$$

Combining (74) with (73), one obtains the second inequality in (72). Our final aim is to prove the third inequality in (72). Toward this end, first note that

$$\begin{aligned} &\nabla^2 F_{\mathcal{S}}(x) \\ &= \nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) + \nabla y_{\mathcal{S}}(x) \nabla c_{\mathcal{S}}(x)^T + \nabla c_{\mathcal{S}}(x) \nabla y_{\mathcal{S}}(x)^T \\ &\quad + 2\rho_{\mathcal{S}} \nabla c_{\mathcal{S}}(x)^T \nabla c_{\mathcal{S}}(x) + \sum_{j \in [m]} (\nabla^2 [y_{\mathcal{S}}]_j(x) + 2\rho_{\mathcal{S}} \nabla^2 [c_{\mathcal{S}}]_j(x))[c_{\mathcal{S}}]_j(x). \end{aligned} \tag{75}$$

Let us now express this Hessian in an equivalent form involving a decomposition into orthogonal spaces defined by the constraint Jacobian at x. Let us first derive an expression for $\nabla y_{\mathcal{S}}(x) \nabla c_{\mathcal{S}}(x)^T$ by differentiating the linear system that defines $y_{\mathcal{S}}(x)$ through (7). Specifically, by (7), one finds that

$$\nabla c_{\mathcal{S}}(x)^T \nabla c_{\mathcal{S}}(x) y_{\mathcal{S}}(x) = -\nabla c_{\mathcal{S}}(x)^T \nabla f(x). \tag{76}$$

Abbreviating notation and differentiating the left-hand side yields

$$\nabla(\nabla c_{\mathcal{S}}^T \nabla c_{\mathcal{S}} y_{\mathcal{S}})|_x$$

$$
\begin{aligned}
&= \nabla((\nabla c_{\mathcal{S}}^T|_x \nabla c_{\mathcal{S}}|_x) y_{\mathcal{S}})|_x + \nabla(\nabla c_{\mathcal{S}}^T \nabla c_{\mathcal{S}}|_x y_{\mathcal{S}}|_x)|_x + \nabla(\nabla c_{\mathcal{S}}^T|_x \nabla c_{\mathcal{S}} y_{\mathcal{S}}|_x)|_x \\
&= \nabla y_{\mathcal{S}}|_x \nabla c_{\mathcal{S}}^T|_x \nabla c_{\mathcal{S}}|_x + \begin{bmatrix} \nabla^2 [c_{\mathcal{S}}]_1|_x \nabla c_{\mathcal{S}}|_x y_{\mathcal{S}}|_x & \cdots & \nabla^2 [c_{\mathcal{S}}]_m|_x \nabla c_{\mathcal{S}}|_x y_{\mathcal{S}}|_x \end{bmatrix} \\
&\quad + \left(\sum_{j \in [m]} \nabla^2 [c_{\mathcal{S}}]_j|_x [y_{\mathcal{S}}]_j|_x \right) \nabla c_{\mathcal{S}}|_x,
\end{aligned}
$$

while at the same time differentiating the right-hand side yields

$$
\begin{aligned}
\nabla(-\nabla c_{\mathcal{S}}^T \nabla f)|_x &= -\nabla(\nabla c_{\mathcal{S}}^T|_x \nabla f)|_x - \nabla(\nabla c_{\mathcal{S}}^T \nabla f|_x)|_x \\
&= -\nabla^2 f|_x \nabla c_{\mathcal{S}}|_x - \begin{bmatrix} \nabla^2 [c_{\mathcal{S}}]_1|_x \nabla f|_x & \cdots & \nabla^2 [c_{\mathcal{S}}]_m|_x \nabla f|_x \end{bmatrix}.
\end{aligned}
$$

Combining these derivations and rearranging yields

$$
\begin{aligned}
&\nabla y_{\mathcal{S}}(x) \nabla c_{\mathcal{S}}^T(x) \nabla c_{\mathcal{S}}(x) \\
&= -\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) \nabla c_{\mathcal{S}}(x) \\
&\quad - \underbrace{\begin{bmatrix} \nabla^2 [c_{\mathcal{S}}]_1(x) \nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) & \cdots & \nabla^2 [c_{\mathcal{S}}]_m(x) \nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) \end{bmatrix}}_{=:\mathcal{E}(x)}. \qquad (77)
\end{aligned}
$$

Multiplying (77) on the right by $(\nabla c_{\mathcal{S}}(x)^T \nabla c_{\mathcal{S}}(x))^{-1} \nabla c_{\mathcal{S}}(x)^T$ yields

$$
\begin{aligned}
&\nabla y_{\mathcal{S}}(x) \nabla c_{\mathcal{S}}(x)^T \\
&= -(\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) \nabla c_{\mathcal{S}}(x) + \mathcal{E}(x))(\nabla c_{\mathcal{S}}(x)^T \nabla c_{\mathcal{S}}(x))^{-1} \nabla c_{\mathcal{S}}(x)^T \\
&= -\nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) \mathcal{R}(\nabla c_{\mathcal{S}}(x)) - \mathcal{E}(x) \nabla c_{\mathcal{S}}(x)^\dagger. \qquad (78)
\end{aligned}
$$

On the other hand, denoting $\mathcal{R}(x) := \mathcal{R}(\nabla c_{\mathcal{S}}(x))$, $\mathcal{N}(x) := \mathcal{N}(\nabla c_{\mathcal{S}}(x)^T)$, and $H_{xx} := \nabla_{xx}^2 L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))$ for the sake of notational simplicity and using the fact that $\mathcal{R}(x) + \mathcal{N}(x) = I$, one finds that

$$
\begin{aligned}
H_{xx} - H_{xx}\mathcal{R}(x) - \mathcal{R}(x) H_{xx} &= H_{xx} \mathcal{N}(x) - \mathcal{R}(x) H_{xx} \\
&= (\mathcal{N}(x) + \mathcal{R}(x)) H_{xx} \mathcal{N}(x) - \mathcal{R}(x) H_{xx} \\
&= \mathcal{N}(x) H_{xx} \mathcal{N}(x) - \mathcal{R}(x) H_{xx} \mathcal{R}(x). \qquad (79)
\end{aligned}
$$

Combining (75), (78), and (79) now yields the expression for the Hessian as

$$
\begin{aligned}
\nabla^2 F_{\mathcal{S}}(x) = \mathcal{N}(x) H_{xx} \mathcal{N}(x) &- \mathcal{R}(x) H_{xx} \mathcal{R}(x) + 2\rho_{\mathcal{S}} \nabla c_{\mathcal{S}}(x)^T \nabla c_{\mathcal{S}}(x) \\
&- \mathcal{E}(x) \nabla c_{\mathcal{S}}(x)^\dagger - (\nabla c_{\mathcal{S}}(x)^\dagger)^T \mathcal{E}(x)^T \\
&+ \sum_{j \in [m]} (\nabla^2 [y_{\mathcal{S}}]_j(x) + 2\rho_{\mathcal{S}} \nabla^2 [c_{\mathcal{S}}]_j(x)) [c_{\mathcal{S}}]_j(x). \qquad (80)
\end{aligned}
$$

Let us now observe that from the definition of $\mathcal{E}(x)$, norm inequalities, submultiplicity of the matrix 2-norm, the fact that $x \in \mathcal{C}_{\mathcal{S},R}$, (69), and (73) that

$$
\begin{aligned}
\|\mathcal{E}(x)\| &= \|\mathcal{E}(x)^T\| \\
&= \max_{d \in \mathbb{R}^n \text{ s.t. } \|d\|=1} \|\mathcal{E}(x)^T d\| \\
&= \max_{d \in \mathbb{R}^n \text{ s.t. } \|d\|=1} \left\| \begin{bmatrix} \nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))^T \nabla^2 [c_{\mathcal{S}}]_1(x)^T d \\ \vdots \\ \nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))^T \nabla^2 [c_{\mathcal{S}}]_m(x)^T d \end{bmatrix} \right\| \\
&= \max_{d \in \mathbb{R}^n \text{ s.t. } \|d\|=1} \sqrt{\sum_{j \in [m]} (\nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))^T \nabla^2 [c_{\mathcal{S}}]_j(x)^T d)^2}
\end{aligned}
$$

$$
\begin{aligned}
&\le \sqrt{m} \max_{d\in\mathbb{R}^n \text{ s.t. } \|d\|=1} \left(\max_{j\in[m]} |\nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))^T \nabla^2 [c_{\mathcal{S}}]_j(x)^T d| \right) \\
&\le \sqrt{m} \max_{d\in\mathbb{R}^n \text{ s.t. } \|d\|=1} \left(\max_{j\in[m]} \|\nabla_x L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| \|\nabla^2 [c_{\mathcal{S}}]_j(x)\| \|d\| \right) \\
&\le \sqrt{m} \max_{d\in\mathbb{R}^n \text{ s.t. } \|d\|=1} (\epsilon \kappa_{\mathcal{S}} \|d\|) = \sqrt{m}\kappa_{\mathcal{S}}\epsilon. \qquad (81)
\end{aligned}
$$

Now consider arbitrary $d \in \mathbb{R}^n \setminus \{0\}$ decomposed as $d \equiv d_{\mathcal{N}} + d_{\mathcal{R}}$, where $d_{\mathcal{N}} \in \text{null}(\nabla c_{\mathcal{S}}(x)^T)$ and $d_{\mathcal{R}} \in \text{range}(\nabla c_{\mathcal{S}}(x))$. One has from (80) that

$$
\begin{aligned}
& d^T \nabla^2 F_{\mathcal{S}}(x) d \\
&= d^T \mathcal{N}(x) \nabla^2_{xx} L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) \mathcal{N}(x) d - d^T \mathcal{R}(x) \nabla^2_{xx} L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) \mathcal{R}(x) d \\
&\quad + 2\rho_{\mathcal{S}} d^T \nabla c_{\mathcal{S}}(x)^T \nabla c_{\mathcal{S}}(x) d - 2 d^T \mathcal{E}(x) \nabla c_{\mathcal{S}}(x)^\dagger d \\
&\quad + d^T \left(\sum_{j\in[m]} \left(\nabla^2 [y_{\mathcal{S}}]_j(x) + 2\rho_{\mathcal{S}} \nabla^2 [c_{\mathcal{S}}]_j(x) \right) [c_{\mathcal{S}}]_j(x) \right) d \\
&= d_{\mathcal{N}}^T \nabla^2_{xx} L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) d_{\mathcal{N}} - d_{\mathcal{R}}^T \nabla^2_{xx} L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) d_{\mathcal{R}} \\
&\quad + 2\rho_{\mathcal{S}} \|\nabla c_{\mathcal{S}}(x) d_{\mathcal{R}}\|^2 - 2 d^T \mathcal{E}(x) \nabla c_{\mathcal{S}}(x)^\dagger d \\
&\quad + d^T \left(\sum_{j\in[m]} \left(\nabla^2 [y_{\mathcal{S}}]_j(x) + 2\rho_{\mathcal{S}} \nabla^2 [c_{\mathcal{S}}]_j(x) \right) [c_{\mathcal{S}}]_j(x) \right) d, \qquad (82)
\end{aligned}
$$

where for the latter two terms one has from (69), (73), and (81) that

$$
\begin{aligned}
& - 2 d^T \mathcal{E}(x) \nabla c_{\mathcal{S}}(x)^\dagger d + d^T \left(\sum_{j\in[m]} \left(\nabla^2 [y_{\mathcal{S}}]_j(x) + 2\rho_{\mathcal{S}} \nabla^2 [c_{\mathcal{S}}]_j(x) \right) [c_{\mathcal{S}}]_j(x) \right) d \\
\ge\; & - 2 \|\mathcal{E}(x)\| \|\nabla c_{\mathcal{S}}(x)^\dagger\| \|d\|^2 - \left\| \sum_{j\in[m]} (\nabla^2 [y_{\mathcal{S}}]_j(x) + 2\rho_{\mathcal{S}} \nabla^2 [c_{\mathcal{S}}]_j(x)) [c_{\mathcal{S}}]_j(x) \right\| \|d\|^2 \\
\ge\; & - \tfrac{2\sqrt{m}\kappa_{\mathcal{S}}}{\sigma_{\min}} \|d\|^2 \epsilon - \sum_{j\in[m]} (\|\nabla^2 [y_{\mathcal{S}}]_j(x)\| + 2\rho_{\mathcal{S}} \|\nabla^2 [c_{\mathcal{S}}]_j(x)\|) |[c_{\mathcal{S}}]_j(x)| \|d\|^2 \\
\ge\; & - \tfrac{2\sqrt{m}\kappa_{\mathcal{S}}}{\sigma_{\min}} \|d\|^2 \epsilon - \sum_{j\in[m]} (\|\nabla^2 [y_{\mathcal{S}}]_j(x)\| + 2\rho_{\mathcal{S}} \|\nabla^2 [c_{\mathcal{S}}]_j(x)\|) \|c_{\mathcal{S}}(x)\| \|d\|^2 \\
\ge\; & - \tfrac{2\sqrt{m}\kappa_{\mathcal{S}}}{\sigma_{\min}} \|d\|^2 \epsilon - m\kappa_{\mathcal{S}} (1 + 2\rho_{\mathcal{S}}) \epsilon \|d\|^2. \qquad (83)
\end{aligned}
$$

Now since $d_{\mathcal{N}} \in \text{null}(\nabla c_{\mathcal{S}}(x)^T)$, (71) gives $d_{\mathcal{N}}^T \nabla^2_{xx} L_{\mathcal{S}}(x, y_{\mathcal{S}}(x)) d_{\mathcal{N}} \ge \beta_{\mathcal{S}} \|d_{\mathcal{N}}\|^2$. Along with (69), the triangle inequality, and submultiplicity of the matrix 2-norm,

$$
\begin{aligned}
\|\nabla^2_{xx} L_{\mathcal{S}}(x, y_{\mathcal{S}}(x))\| &= \left\| \nabla^2 f(x) + \sum_{j\in[m]} \nabla^2 [c_{\mathcal{S}}]_j(x) [y_{\mathcal{S}}]_j(x) \right\| \\
&\le \|\nabla^2 f(x)\| + \sum_{j\in[m]} \|\nabla^2 [c_{\mathcal{S}}]_j(x)\| \|y_{\mathcal{S}}(x)\| \le \kappa_{\mathcal{S}} + m\kappa_{\mathcal{S}}^2. \qquad (84)
\end{aligned}
$$

At the same time, let us derive a lower bound for $\|\nabla c_{\mathcal{S}}(x)^T d_{\mathcal{R}}\|$. Let $\nabla c_{\mathcal{S}}(x)$ have the singular value decomposition $\sum_{i\in[m]} u_i \sigma_i v_i^T$ where $(u_i, \sigma_i, v_i) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m$ for all $i \in [m]$ with $\{u_i\}_{i\in[m]}$ and $\{v_i\}_{i\in[m]}$ being sets of orthonormal vectors. Then, by Assumption 3.1 and the fact that $d_{\mathcal{R}} \in \text{range}(\nabla c_{\mathcal{S}}(x))$, one finds that

$$\|\nabla c_{\mathcal{S}}(x)^T d_{\mathcal{R}}\|^2 = \left\|\sum_{i\in[m]} v_i\sigma_i u_i^T d_{\mathcal{R}}\right\|^2 = \sum_{i\in[m]} \|v_i\sigma_i u_i^T d_{\mathcal{R}}\|^2 = \sum_{i\in[m]} \sigma_i^2 (u_i^T d_{\mathcal{R}})^2$$
$$\geq \sigma_{\min}^2 \sum_{i=1}^{m} (u_i^T d_{\mathcal{R}})^2 = \sigma_{\min}^2 \|d_{\mathcal{R}}\|^2. \tag{85}$$

Combining (82), (83), (84), (85), and the decomposition $d = d_{\mathcal{N}} + d_{\mathcal{R}}$, one has

$$\begin{aligned}
&d^T \nabla^2 F_{\mathcal{S}}(x) d \\
&\geq \beta_{\mathcal{S}} \|d_{\mathcal{N}}\|^2 + (2\rho_{\mathcal{S}}\sigma_{\min}^2 - (\kappa_{\mathcal{S}} + m\kappa_{\mathcal{S}}^2))\|d_{\mathcal{R}}\|^2 \\
&\quad - \left(\tfrac{2\sqrt{m}}{\sigma_{\min}} + m(1+2\rho_{\mathcal{S}})\right)\kappa_{\mathcal{S}}\|d\|^2\epsilon \\
&= \left(\beta_{\mathcal{S}} - \left(\tfrac{2\sqrt{m}}{\sigma_{\min}} + m(1+2\rho_{\mathcal{S}})\right)\kappa_{\mathcal{S}}\epsilon\right)\|d_{\mathcal{N}}\|^2 \\
&\quad + \left(2\rho_{\mathcal{S}}\sigma_{\min}^2 - (1+m\kappa_{\mathcal{S}})\kappa_{\mathcal{S}} - \left(\tfrac{2\sqrt{m}}{\sigma_{\min}} + m(1+2\rho_{\mathcal{S}})\right)\kappa_{\mathcal{S}}\epsilon\right)\|d_{\mathcal{R}}\|^2.
\end{aligned} \tag{86}$$

We now claim that our desired conclusion follows as long as $(\rho_{\mathcal{S}}, \epsilon)$ yields

$$\begin{aligned}
\beta_{\mathcal{S}} - \left(\tfrac{2\sqrt{m}}{\sigma_{\min}} + m(1+2\rho_{\mathcal{S}})\right)\kappa_{\mathcal{S}}\epsilon &\geq \tfrac{1}{2}\beta_{\mathcal{S}} \ \text{ and} \\
2\rho_{\mathcal{S}}\sigma_{\min}^2 - (1+m\kappa_{\mathcal{S}})\kappa_{\mathcal{S}} - \left(\tfrac{2\sqrt{m}}{\sigma_{\min}} + m(1+2\rho_{\mathcal{S}})\right)\kappa_{\mathcal{S}}\epsilon &\geq \tfrac{1}{2}\beta_{\mathcal{S}},
\end{aligned} \tag{87}$$

since these inequalities along with (86) would yield the desired fact that

$$d^T \nabla^2 F_{\mathcal{S}}(x) d \geq \tfrac{1}{2}\beta_{\mathcal{S}}\|d_{\mathcal{N}}\|^2 + \tfrac{1}{2}\beta_{\mathcal{S}}\|d_{\mathcal{R}}\|^2 = \tfrac{1}{2}\beta_{\mathcal{S}}\|d\|^2.$$

Indeed, since $\epsilon \in (0, \overline{\epsilon}_{\mathcal{S}})$ and $\rho_{\mathcal{S}} \in (\underline{\rho_{\mathcal{S}}}(\epsilon), \overline{\rho_{\mathcal{S}}}(\epsilon))$, one finds that

$$\begin{aligned}
\beta_{\mathcal{S}} - \left(\tfrac{2\sqrt{m}}{\sigma_{\min}} + m(1+2\rho_{\mathcal{S}})\right)\kappa_{\mathcal{S}}\epsilon &= \beta_{\mathcal{S}} - (\eta_{\mathcal{S},1}\epsilon + \eta_{\mathcal{S},2}\rho_{\mathcal{S}}\epsilon) \\
&\geq \beta_{\mathcal{S}} - (\eta_{\mathcal{S},1}\epsilon + \eta_{\mathcal{S},2}\overline{\rho_{\mathcal{S}}}(\epsilon)\epsilon) \\
&= \beta_{\mathcal{S}} - (\eta_{\mathcal{S},1}\epsilon + \tfrac{1}{2}\beta_{\mathcal{S}} - \eta_{\mathcal{S},1}\epsilon) = \tfrac{1}{2}\beta_{\mathcal{S}}
\end{aligned}$$

as well as

$$\begin{aligned}
&2\rho_{\mathcal{S}}\sigma_{\min}^2 - (1+m\kappa_{\mathcal{S}})\kappa_{\mathcal{S}} - \left(\tfrac{2\sqrt{m}}{\sigma_{\min}} + m(1+2\rho_{\mathcal{S}})\right)\kappa_{\mathcal{S}}\epsilon \\
&= 2\rho_{\mathcal{S}}\sigma_{\min}^2 - (\eta_{\mathcal{S},1}\epsilon + (1+m\kappa_{\mathcal{S}})\kappa_{\mathcal{S}}) - \eta_{\mathcal{S},2}\rho_{\mathcal{S}}\epsilon \\
&\geq 2\rho_{\mathcal{S}}\sigma_{\min}^2 - (\eta_{\mathcal{S},1}\overline{\epsilon}_{\mathcal{S}} + (1+m\kappa_{\mathcal{S}})\kappa_{\mathcal{S}}) - \eta_{\mathcal{S},2}\rho_{\mathcal{S}}\epsilon \\
&\geq 2\rho_{\mathcal{S}}\sigma_{\min}^2 - (\eta_{\mathcal{S},1}\eta_{\mathcal{S},3} + (1+m\kappa_{\mathcal{S}})\kappa_{\mathcal{S}}) - \eta_{\mathcal{S},2}\rho_{\mathcal{S}}\epsilon \\
&= 2\rho_{\mathcal{S}}\sigma_{\min}^2 - \eta_{\mathcal{S},4} - \eta_{\mathcal{S},2}\rho_{\mathcal{S}}\epsilon \\
&\geq 2\underline{\rho_{\mathcal{S}}}(\epsilon)\sigma_{\min}^2 - \eta_{\mathcal{S},4} - \eta_{\mathcal{S},2}\overline{\rho_{\mathcal{S}}}(\epsilon)\epsilon \\
&= \tfrac{(\beta_{\mathcal{S}}+2\eta_{\mathcal{S},4})\sigma_{\min}^2}{2\sigma_{\min}^2 - \eta_{\mathcal{S},2}\epsilon} - \eta_{\mathcal{S},4} - \left(\tfrac{1}{2}\beta_{\mathcal{S}} - \eta_{\mathcal{S},1}\epsilon\right) \\
&\geq \tfrac{(\beta_{\mathcal{S}}+2\eta_{\mathcal{S},4})\sigma_{\min}^2}{\sigma_{\min}^2} - 2\eta_{\mathcal{S},4} - \tfrac{1}{2}\beta_{\mathcal{S}} = \tfrac{1}{2}\beta_{\mathcal{S}}.
\end{aligned}$$

For the reasons stated previously, the proof is complete. □ □

To obtain our desired complexity result, we need one more assumption.

Assumption 3.4. *For any $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$ and any $\rho_{\mathcal{S}} \in \mathbb{R}_{>0}$, Fletcher's Augmented Lagrangian function has a Hessian function $\nabla^2 F_{\mathcal{S}} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ that is Lipschitz continuous in the sense that there exists $M_{\mathcal{S}} \in \mathbb{R}_{>0}$ such that*

$$\|\nabla^2 F_{\mathcal{S}}(x) - \nabla^2 F_{\mathcal{S}}(\overline{x})\| \leq M_{\mathcal{S}}\|x-\overline{x}\|^2 \ \ \textit{for all} \ \ (x,\overline{x}) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Lemma 3.10. *Suppose that the conditions of Lemma 3.7 hold in the sense that $(\kappa_1,\kappa_2,\kappa_3)$ is defined by (22) and (23), (τ_1,τ_2) is defined by (41), for all $k \in [K]$ the pair of subproblem tolerances (ϵ_k,ζ_k) is defined by (42), p_1 is sufficiently large such that (43) holds, and for all $k\in[K]$ the point $x_{\mathcal{S}_k} \in \mathbb{R}^n$ is (ϵ_k,ζ_k)-stationary with respect to (3) for $\mathcal{S} = \mathcal{S}_k$. In addition, suppose that the conditions of Lemma 3.9 hold in the sense that there exists $R \in \mathbb{R}_{>0}$ such that for any $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| \geq p_1$ the sublevel set $\mathcal{C}_{\mathcal{S},R} := \{x \in \mathbb{R}^n : \|c_{\mathcal{S}}(x)\| \leq R\}$ contains x_0 and is compact. Further, with the definitions in (69) and (70) for all $\beta_{\mathcal{S}} \geq \frac{1}{2}\beta$ and $\mathcal{S} \subset [N]$ with $|\mathcal{S}| \geq p_1$, suppose with*

$$\begin{aligned} \underline{\delta} &:= \min_{\mathcal{S}\subseteq[N]\,\text{s.t.}\,(43)} \left\{R, \overline{\epsilon}_{\mathcal{S}}, \frac{\frac{1}{2}\beta}{2\eta_{\mathcal{S},1} + 3\eta_{\mathcal{S},2}\frac{\max\{1,\kappa_{\mathcal{S}}\}}{\sigma_{\min}}}\right\}, \\ \overline{\omega} &:= \max_{\mathcal{S}\subseteq[N]\,\text{s.t.}\,(43)} \ 1 + \kappa_{\mathcal{S}} + 2\overline{\rho_{\mathcal{S}}}(\underline{\delta})\kappa_{\mathcal{S}}, \\ \textit{and} \ \ \overline{M} &:= \max_{\mathcal{S}\subseteq[N]\,\text{s.t.}\,(43)} \ M_{\mathcal{S}} \ \textit{(see Assumption 3.4)} \end{aligned} \tag{88}$$

that the sample set $\mathcal{S}_k$ for each $k \in [K]$ yields

$$\sqrt{\frac{N(N-|\mathcal{S}_k|)}{|\mathcal{S}_k|^2}} \leq \frac{1}{\kappa_2}\min\left\{\frac{\underline{\delta}}{3}, \frac{\beta}{18\overline{\omega}^2}, \frac{\beta^2}{432\overline{\omega}\overline{M}}, \frac{R}{3\overline{\omega}}\right\}. \tag{89}$$

Finally, suppose that with $\rho_{\mathcal{S}_{k+1}} \in \mathbb{R}_{>0}$ and $t_{\mathcal{S}_{k+1}} \in \mathbb{R}_{>0}$, gradient descent with constant step size $t_{\mathcal{S}_{k+1}}$ is employed to minimize $F_{\mathcal{S}_{k+1}}$ with initial point $x_{\mathcal{S}_k}$. Then, there exists a positive interval and a positive upper bound such that if $\rho_{\mathcal{S}_{k+1}}$ is within the positive interval and $t_{\mathcal{S}_{k+1}}$ is below the upper bound, in at most

$$\left\lceil \log_2 \frac{3\sqrt{5}\overline{\omega}\epsilon_k}{\epsilon_{k+1}} \right\rceil \ \textit{iterations}$$

the method gives $x_{\mathcal{S}_{k+1}}$ that is $(\epsilon_{k+1},\zeta_{k+1})$-stationary for (3) with $\mathcal{S} = \mathcal{S}_{k+1}$.

Proof. Consider arbitrary $k \in [K]$. By Lemma 3.7, it follows that (44) holds, where $\alpha_{\mathcal{S}_{k+1}} \geq \frac{1}{2}\alpha$ and $\beta_{\mathcal{S}_{k+1}} \geq \frac{1}{2}\beta$ are defined according to (25). Consequently, one finds that the result of Lemma 3.9(b) holds with $\mathcal{S} = \mathcal{S}_{k+1}$, $\epsilon = 3\epsilon_k$, $\beta_{\mathcal{S}_{k+1}} \geq \frac{1}{2}\beta$ defined by (25), $x = x_{\mathcal{S}_k}$, and $\rho_{\mathcal{S}_{k+1}} = \overline{\rho_{\mathcal{S}_{k+1}}}(\underline{\delta})$. Indeed, with (41), (42), the first term of the minimum in (89), and (88), one finds that $3\epsilon_k = 3\kappa_2\xi_{\mathcal{S}_k} \leq \underline{\delta} \leq \min\{R, \overline{\epsilon}_{\mathcal{S}_{k+1}}\}$ and $\rho_{\mathcal{S}_{k+1}} \geq \underline{\rho_{\mathcal{S}_{k+1}}}(\underline{\delta}) \geq \underline{\rho_{\mathcal{S}_{k+1}}}(3\epsilon_k)$. Thus,

$$x_{\mathcal{S}_k} \in \mathcal{C}_{\mathcal{S}_{k+1},R},\ \|\nabla F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k})\| \leq 3\overline{\omega}\epsilon_k \text{ and } \nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}) \succeq \tfrac{1}{2}\beta_{\mathcal{S}_{k+1}} I. \tag{90}$$

Before proceeding, since $\mathcal{C}_{\mathcal{S}_{k+1},R} := \{x \in \mathbb{R}^n : \|c_{\mathcal{S}_{k+1}}(x)\| \leq R\}$ is compact, let

$$\sigma_{\nabla^2 F_{\mathcal{S}_{k+1}}} := \max_{x\in\mathcal{C}_{\mathcal{S}_{k+1},R}} \{\sigma_{\max}(\nabla^2 F_{\mathcal{S}_{k+1}}(x))\} < \infty.$$

Let us now consider the behavior of gradient descent employed with penalty parameter $\rho_{\mathcal{S}_{k+1}} \in \mathbb{R}_{>0}$ and step size $t_{\mathcal{S}_{k+1}} \in \mathbb{R}_{>0}$; that is, with $x^0_{\mathcal{S}_k} \leftarrow x_{\mathcal{S}_k}$, consider the iterative method defined for all $i = 0,1,2,\ldots$ by

$$x^{i+1}_{\mathcal{S}_k} \leftarrow x^i_{\mathcal{S}_k} - t_{\mathcal{S}_{k+1}}\nabla F_{\mathcal{S}_{k+1}}(x^i_{\mathcal{S}_k}) \tag{91}$$

$$\text{where} \ \ t_{\mathcal{S}_{k+1}} \leftarrow \min\left\{\frac{\beta_{\mathcal{S}_{k+1}}}{9\overline{\omega}\overline{M}\epsilon_k}, \frac{1}{\sigma_{\nabla^2 F_{\mathcal{S}_{k+1}}}}, \frac{3}{\beta_{\mathcal{S}_{k+1}}}\right\}. \tag{92}$$

Our next aim is to show that, for any such $i \geq 1$, one has that

$$x_{\mathcal{S}_k}^i \in \mathcal{C}_{\mathcal{S}_{k+1},R}, \quad \|\nabla F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)\| \leq \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)^i \|\nabla F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^0)\| \tag{93}$$

and $\sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^0)) \geq \frac{1}{2}\beta_{\mathcal{S}_{k+1}}$ while for $i \geq 1$ one has

$$\begin{aligned} &\sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)) \\ &\geq \tfrac{1}{2}\beta_{\mathcal{S}_{k+1}} - \overline{M}\|\nabla F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^0)\| t_{\mathcal{S}_{k+1}} \sum_{l=0}^{i-1} \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)^l \geq \tfrac{1}{3}\beta_{\mathcal{S}_{k+1}}. \end{aligned} \tag{94}$$

Toward this end, first observe that by defining $G_i := \nabla F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)$ for all i one has from Taylor's theorem and submultiplicity that

$$\begin{aligned} \|G_{i+1}\| &\leq \left\|G_i + \nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)(x_{\mathcal{S}_k}^{i+1} - x_{\mathcal{S}_k}^i)\right\| + \tfrac{1}{2}\overline{M}\|x_{\mathcal{S}_k}^{i+1} - x_{\mathcal{S}_k}^i\|^2 \\ &= \|G_i - t_{\mathcal{S}_{k+1}}\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i) G_i\| + \tfrac{1}{2}\overline{M} t_{\mathcal{S}_{k+1}}^2 \|G_i\|^2 \\ &\leq \|I - t_{\mathcal{S}_{k+1}}\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)\|\|G_i\| + \tfrac{1}{2}\overline{M} t_{\mathcal{S}_{k+1}}^2 \|G_i\|^2. \end{aligned} \tag{95}$$

Let us now employ the aforementioned induction. For $i = 0$, the base case holds since $x_{\mathcal{S}_k}^0 = x_{\mathcal{S}_k}$ satisfies (90). Now consider an arbitrary positive integer i and suppose that $x_{\mathcal{S}_k}^i$ satisfies (93)–(94). By (92), one has

$$t_{\mathcal{S}_{k+1}}\sigma_{\max}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)) \leq t_{\mathcal{S}_{k+1}}\sigma_{\nabla^2 F_{\mathcal{S}_{k+1}}} \leq 1,$$

which in turn along with (94) shows that

$$\begin{aligned} &\|I - t_{\mathcal{S}_{k+1}}\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)\| \\ &= \max_{v \in \mathbb{R}^n \text{ s.t. } \|v\|=1} \left|v^T v - t_{\mathcal{S}_{k+1}} v^T \nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i) v\right| \\ &= \max\left\{|1 - t_{\mathcal{S}_{k+1}}\sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i))|, |1 - t_{\mathcal{S}_{k+1}}\sigma_{\max}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i))|\right\} \\ &= 1 - t_{\mathcal{S}_{k+1}}\sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)) \\ &\leq 1 - \tfrac{1}{3}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}. \end{aligned} \tag{96}$$

At the same time, combining (90), (92), and (93), one finds that

$$\overline{M} t_{\mathcal{S}_{k+1}}\|G_i\| \leq \overline{M} t_{\mathcal{S}_{k+1}}\|G_0\| \leq 3\overline{M} t_{\mathcal{S}_{k+1}}\overline{\omega}\epsilon_k \leq \tfrac{1}{3}\beta_{\mathcal{S}_{k+1}}. \tag{97}$$

Now combining (95), (96), and (97), one obtains that

$$\begin{aligned} \|G_{i+1}\| &\leq \left(1 - \tfrac{1}{3}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)\|G_i\| + \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\|G_i\| \\ &= \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)\|G_i\| \leq \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)^{i+1}\|G_0\|, \end{aligned} \tag{98}$$

which shows the inequality in (93), and in fact shows that $\|\nabla F_{\mathcal{S}_{k+1}}(\cdot)\|$ decreases with each iteration. Also, from [8, Theorem 6.6] (or see [10, Eq. (3)]), one finds

$$\begin{aligned} &|\sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1})) - \sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i))| \\ &\leq \|\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1})) - \nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)\| \leq \overline{M}\|x_{\mathcal{S}_k}^{i+1} - x_{\mathcal{S}_k}^i\| = \overline{M} t_{\mathcal{S}_{k+1}}\|G_i\|. \end{aligned}$$

Combined with (92), (94), and (98), one now obtains that

$$\begin{aligned} &\sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1})) \\ &\geq \sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i)) - |\sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1})) - \sigma_{\min}(\nabla^2 F_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^i))| \end{aligned}$$

$$
\begin{aligned}
&\geq \tfrac{1}{2}\beta_{\mathcal{S}_{k+1}} - \overline{M}\|G_0\|t_{\mathcal{S}_{k+1}} \sum_{l=0}^{i-1} \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)^l - \overline{M} t_{\mathcal{S}_{k+1}} \|G_i\| \\
&\geq \tfrac{1}{2}\beta_{\mathcal{S}_{k+1}} - \overline{M}\|G_0\|t_{\mathcal{S}_{k+1}} \sum_{l=0}^{i} \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)^l \\
&\geq \tfrac{1}{2}\beta_{\mathcal{S}_{k+1}} - \overline{M}\|G_0\|t_{\mathcal{S}_{k+1}} \sum_{l=0}^{\infty} \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)^l \\
&= \tfrac{1}{2}\beta_{\mathcal{S}_{k+1}} - \overline{M}\|G_0\|t_{\mathcal{S}_{k+1}} \tfrac{6}{\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}} \\
&\geq \tfrac{1}{2}\beta_{\mathcal{S}_{k+1}} - \tfrac{18\overline{M}\overline{\omega}\epsilon_k}{\beta_{\mathcal{S}_{k+1}}} \geq \tfrac{1}{3}\beta_{\mathcal{S}_{k+1}}. \qquad (99)
\end{aligned}
$$

Here, for the last inequality, note that $\epsilon_k = \tau_1 \xi_{\mathcal{S}_k} = \kappa_2 \xi_{\mathcal{S}_k}$ by (42). Then, since $\mathcal{S}_k$ satisfies (89), one finds $\epsilon_k \leq \frac{\beta^2}{432\overline{M}\overline{\omega}}$, which further gives $\frac{18\overline{M}\overline{\omega}\epsilon_k}{\beta_{\mathcal{S}_{k+1}}} \leq \frac{\beta^2}{24\beta_{\mathcal{S}_{k+1}}}$. Now since $\beta_{\mathcal{S}_{k+1}} \geq \frac{1}{2}\beta$ by (25), it follows that $\frac{18\overline{M}\overline{\omega}\epsilon_{\mathcal{S}_k}}{\beta_{\mathcal{S}_{k+1}}} \leq \frac{\beta_{k+1}}{6}$. As a result, one has $\frac{1}{2}\beta_{\mathcal{S}_{k+1}} - \frac{18\overline{M}\overline{\omega}\epsilon_k}{\beta_{\mathcal{S}_{k+1}}} \geq \frac{1}{2}\beta_{\mathcal{S}_{k+1}} - \frac{1}{6}\beta_{\mathcal{S}_{k+1}} = \frac{1}{3}\beta_{\mathcal{S}_{k+1}}$, which gives the last inequality.

All that remains to complete the induction is to show that $x^{i+1}_{\mathcal{S}_k} \in \mathcal{C}_{\mathcal{S}_{k+1},R}$. Toward this end, let us employ [4, Corollary 2.7]. This requires that the singular values of $\nabla c_{\mathcal{S}_{k+1}}$ are bounded away from zero and that $\mathcal{C}_{\mathcal{S}_{k+1},R}$ is compact, both of which are assumed here. It also requires that

$$
\rho_{\mathcal{S}_{k+1}} > \max_{x \in \mathcal{C}_{\mathcal{S}_{k+1},R}} \frac{\max\{1, \|\nabla y_{\mathcal{S}_{k+1}}(x)\|\}}{\sigma_{\min}(\nabla c_{\mathcal{S}_{k+1}}(x))}. \qquad (100)
$$

Let us show that the choice $\rho_{\mathcal{S}_{k+1}} = \overline{\rho_{\mathcal{S}_{k+1}}}(\underline{\delta})$, which is reflected in (88), yields this lower bound. Indeed, since $\sigma_{\min}(\nabla c_{\mathcal{S}_{k+1}}(x)) \geq \sigma_{\min}$ by Assumption 3.1 and $\|\nabla y_{\mathcal{S}_{k+1}}(x)\| \leq \kappa_{\mathcal{S}_{k+1}}$ by the definition of $\kappa_{\mathcal{S}_{k+1}}$ in (69), one sees (100) holds if

$$
\rho_{\mathcal{S}_{k+1}} = \overline{\rho_{\mathcal{S}_{k+1}}}(\underline{\delta}) \geq \frac{3\max\{1, \kappa_{\mathcal{S}_{k+1}}\}}{2\sigma_{\min}}.
$$

For any $\delta > 0$, the definition $\overline{\rho_{\mathcal{S}_{k+1}}}(\delta) = \frac{\frac{1}{2}\beta_{\mathcal{S}_{k+1}} - \eta_{\mathcal{S}_{k+1},1}\delta}{\eta_{\mathcal{S}_{k+1},2}\delta}$ yields

$$
\begin{aligned}
&\left\{\overline{\rho_{\mathcal{S}_{k+1}}}(\delta) \geq \frac{3\max\{1,\kappa_{\mathcal{S}_{k+1}}\}}{2\sigma_{\min}}\right\} \iff \left\{\frac{\frac{1}{2}\beta_{\mathcal{S}_{k+1}} - \eta_{\mathcal{S}_{k+1},1}\delta}{\eta_{\mathcal{S}_{k+1},2}\delta} \geq \frac{3\max\{1,\kappa_{\mathcal{S}_{k+1}}\}}{2\sigma_{\min}}\right\} \\
&\iff \left\{\beta_{\mathcal{S}_{k+1}}\sigma_{\min} - 2\eta_{\mathcal{S}_{k+1},1}\delta\sigma_{\min} \geq 3\eta_{\mathcal{S}_{k+1},2}\max\{1,\kappa_{\mathcal{S}_{k+1}}\}\delta\right\} \\
&\iff \left\{\delta \leq \frac{\beta_{\mathcal{S}_{k+1}}\sigma_{\min}}{2\eta_{\mathcal{S}_{k+1},1}\sigma_{\min} + 3\eta_{\mathcal{S}_{k+1},2}\max\{1,\kappa_{\mathcal{S}_{k+1}}\}}\right\},
\end{aligned}
$$

where the last inequality is ensured to hold with $\underline{\delta}$ defined in (88). Now combining inequalities (98) and (99) to write

$$
\|G_{i+1}\| \leq \tilde{\epsilon} := \left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right)^{i+1} \|G_0\| \quad \text{and} \quad \nabla^2 F_{\mathcal{S}_{k+1}}(x^{i+1}_{\mathcal{S}_k}) \succeq \tfrac{1}{3}\beta_{\mathcal{S}_{k+1}} I,
$$

it follows from [4, Corollary 2.7] that the iterate $x^{i+1}_{\mathcal{S}_k}$ is an $(\tilde{\epsilon}, 2\tilde{\epsilon}, -\frac{\beta_{\mathcal{S}_{k+1}}}{3} + (1 + \frac{2}{\sigma_{\min}})\kappa_{\mathcal{S}_{k+1}}\tilde{\epsilon})$ stationary point of (3), in the sense that

$$
\begin{aligned}
&\|c_{\mathcal{S}_{k+1}}(x^{i+1}_{\mathcal{S}_k})\| \leq \tilde{\epsilon}, \quad \|\nabla_x L_{\mathcal{S}}(x^{i+1}_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}(x^{i+1}_{\mathcal{S}_k}))\| \leq 2\tilde{\epsilon}, \\
&\text{and } d^T \nabla^2_{xx} L_{\mathcal{S}_{k+1}}(x^{i+1}_{\mathcal{S}_k}, y_{\mathcal{S}_{k+1}}(x^{i+1}_{\mathcal{S}_k}))d \geq \left(\tfrac{1}{3}\beta_{\mathcal{S}_{k+1}} - \left(1 + \tfrac{2}{\sigma_{\min}}\right)\kappa_{\mathcal{S}_{k+1}}\tilde{\epsilon}\right)\|d\|^2 \\
&\text{for all } d \in \text{null}(\nabla c_{\mathcal{S}_{k+1}}(x^{i+1}_{\mathcal{S}_{k+1}})^T). \qquad (101)
\end{aligned}
$$

Note that $\left(1 - \frac{1}{6}\beta_{\mathcal{S}_{k+1}} t_{\mathcal{S}_{k+1}}\right) < 1$, so $\tilde{\epsilon} \leq \|G_0\|$. Combined with the fact that $\|G_0\| \leq 3\overline{\omega}\epsilon_{\mathcal{S}_k} \leq R$ by (89), one finds that $\|c_{\mathcal{S}_{k+1}}(x^{i+1}_{\mathcal{S}_k})\| \leq \tilde{\epsilon} \leq R$. Thus, the desired conclusion that $x^{i+1}_{\mathcal{S}_k} \in \mathcal{C}_{\mathcal{S}_{k+1},R}$ has been proved.

Through the induction it has been shown that (101) holds for all $i \geq 0$. Let us now show that that right-hand side of the second line of (101) is nonnegative for all $i \geq 0$. Consider arbitrary such i. From $\beta_{\mathcal{S}_{k+1}} \geq \frac{1}{2}\beta$ and $\tilde{\epsilon} \leq \|G_0\| \leq 3\overline{\omega}\epsilon_k$,

$$\tfrac{1}{3}\beta_{\mathcal{S}_{k+1}} - \left(1 + \tfrac{2}{\sigma_{\min}}\right)\kappa_{\mathcal{S}_{k+1}}\tilde{\epsilon} \geq \tfrac{1}{6}\beta - \left(1 + \tfrac{2}{\sigma_{\min}}\right)\kappa_{\mathcal{S}_{k+1}} 3\overline{\omega}\epsilon_k.$$

Next, from (88) and (100), it follows that

$$\begin{aligned}\overline{\omega} &= \max_{\mathcal{S}\subseteq[N]\,\text{s.t.}\,(43)} 1 + \kappa_{\mathcal{S}} + 2\overline{\rho_{\mathcal{S}}}(\underline{\delta})\kappa_{\mathcal{S}} \\ &\geq \kappa_{\mathcal{S}_{k+1}} + 2\overline{\rho_{\mathcal{S}_{k+1}}}(\underline{\delta})\kappa_{\mathcal{S}_{k+1}} \geq \kappa_{\mathcal{S}_{k+1}} + \tfrac{3\max\{1,\kappa_{\mathcal{S}_{k+1}}\}}{\sigma_{\min}}\kappa_{\mathcal{S}_{k+1}} \geq \left(1 + \tfrac{2}{\sigma_{\min}}\right)\kappa_{\mathcal{S}_{k+1}}.\end{aligned}$$

Combined with the previous displayed inequality, one finds that $\frac{1}{3}\beta_{\mathcal{S}_{k+1}} - (1 + \frac{2}{\sigma_{\min}})\kappa_{\mathcal{S}_{k+1}}\tilde{\epsilon} \geq \frac{1}{6}\beta - 3\overline{\omega}^2\epsilon_k$. Now combined with (89) and $\epsilon_k = \kappa_2\xi_{\mathcal{S}_k}$, one has

$$\tfrac{1}{3}\beta_{\mathcal{S}_{k+1}} - \left(1 + \tfrac{2}{\sigma_{\min}}\right)\kappa_{\mathcal{S}_{k+1}}\tilde{\epsilon} \geq \tfrac{1}{6}\beta - 3\overline{\omega}^2\epsilon_k \geq 0.$$

Thus, the right-hand side of the second line of (101) is nonnegative for all $i \geq 0$.

Let us now bound the number of iterations until gradient descent terminates, which due to the result in the previous paragraph can be determined by the number of iterations until the first-order condition is satisfied. Recall the fact that $\nabla_y L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1}, y_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1})) = c_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1})$. By (101), one finds for all $i \geq 0$ that

$$\begin{aligned}&\|\nabla L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1}, y_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1}))\| \\ &= \sqrt{\|\nabla_x L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1}, y_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1}))\|^2 + \|\nabla_y L_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1}, y_{\mathcal{S}_{k+1}}(x_{\mathcal{S}_k}^{i+1}))\|^2} \\ &\leq \sqrt{5}\tilde{\epsilon} = \sqrt{5}\left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}}t_{\mathcal{S}_{k+1}}\right)^{i+1}\|G_0\|.\end{aligned} \tag{102}$$

By the step-size rule in (92), the right-hand side of this expression is monotonically decreasing as i increases. Letting T denote the iteration index at which the first-order condition is satisfied, one finds from (102) that

$$\sqrt{5}\left(1 - \tfrac{1}{6}\beta_{\mathcal{S}_{k+1}}t_{\mathcal{S}_{k+1}}\right)^T\|G_0\| \leq \epsilon_{k+1} \Longrightarrow T \leq \left\lceil \log_{1-\frac{1}{6}\beta_{\mathcal{S}_{k+1}}t_{\mathcal{S}_{k+1}}} \tfrac{\epsilon_{k+1}}{\sqrt{5}\|G_0\|} \right\rceil.$$

Moreover, from (90) and (92), one finds for the log term that

$$\log_{1-\frac{1}{6}\beta_{\mathcal{S}_{k+1}}t_{\mathcal{S}_{k+1}}} \tfrac{\epsilon_{k+1}}{\sqrt{5}\|G_0\|} = \log_{\frac{1}{1-\frac{1}{6}\beta_{\mathcal{S}_{k+1}}t_{\mathcal{S}_{k+1}}}} \tfrac{\sqrt{5}\|G_0\|}{\epsilon_{k+1}} \leq \log_2 \tfrac{\sqrt{5}G_0}{\epsilon_{k+1}} \leq \log_2 \tfrac{3\sqrt{5}\overline{\omega}\epsilon_k}{\epsilon_{k+1}},$$

which completes the proof. □ □

Theorem 3.2. *Suppose that Assumptions 3.1, 3.2, 3.3, and 3.4 hold.*

(a) Suppose that with $p_1 = N$ the conditions of Lemma 3.8 hold and that [4, Algorithm 1] is employed to solve subproblem (3) for $\mathcal{S} = \mathcal{S}_1$. Then, there exists $(u_{[N],1}, u_{[N],2}) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ such that, for any $(\epsilon, \zeta) \in (0, \frac{\sqrt{5}}{2}] \times (0, 1]$, the number of constraint gradient evaluations that are required until Algorithm 1 terminates with an (ϵ, ζ) stationary point of (2) is

$$N\left\lceil \max\left\{u_{[N],1}\epsilon^{-2}, u_{[N],2}\zeta^{-3}\right\}\right\rceil. \tag{103}$$

(b) Suppose that with $p_1 < N$ the conditions of Lemma 3.8 hold and that [4, Algorithm 1] is employed to solve subproblem (3) for $\mathcal{S} = \mathcal{S}_1$, whereas gradient descent is employed to minimize Fletcher's augmented Lagrangian function for all subsequent $k \in [K]$ under the conditions of Lemma 3.10. Then,

there exists $(u_{\mathcal{S}_1,1}, u_{\mathcal{S}_1,2}) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ *such that, for any* $(\epsilon, \zeta) \in (0, \frac{\sqrt{5}}{2}] \times (0, 1]$ *and with* $(\tau_1, \kappa_1, \overline{\omega})$ *defined in Lemma 3.10, the number of constraint gradient evaluations that are required until Algorithm 1 terminates with an* (ϵ, ζ) *stationary point of* (2) *is*

$$|\mathcal{S}_1| \left\lceil \max\left\{u_{\mathcal{S}_1,1}\epsilon_1^{-2}, u_{\mathcal{S}_1,2}\zeta_1^{-3}\right\}\right\rceil + (N - 2|\mathcal{S}_1|) \left\lceil \tfrac{1}{2}\log_2\left(180\overline{\omega}^2\left(1 + \tfrac{N}{2}\right)\right)\right\rceil + N \left\lceil \log_2 \tfrac{\sqrt{5}\overline{\omega}\tau_1}{\kappa_1\epsilon}\right\rceil. \quad (104)$$

Proof. Part (a) of the theorem and the aspect of part (b) pertaining to the first subproblem follow from Lemma 3.8, [4, Theorem 3.5], and the fact that each iteration of [4, Algorithm 1] requires $\mathcal{O}(p_1)$ constraint gradient evaluations.

Let us now proceed to analyze the aspect of part (b) pertaining to the subproblem solves after the initial one. Consider two phases: (1) all subproblems in iterations prior to the last one, and (2) the final subproblem in iteration K.

Consider iterations $k \in \{2, \dots, K-1\}$ of Algorithm 1. The aim of each iteration is to find an (ϵ_k, ζ_k)-stationary point using the final iterate from the previous subproblem solve as a starting point. By Lemma 3.10, for all $k \in \{2, \dots, K-1\}$ it follows that the number of iterations required by gradient descent is at most

$$\left\lceil \log_2 \tfrac{3\sqrt{5}\overline{\omega}\epsilon_{k-1}}{\epsilon_k}\right\rceil.$$

Recalling that $|\mathcal{S}_k| = 2|\mathcal{S}_{k-1}| < N$ for all such k, one finds that

$$\begin{aligned}\tfrac{\epsilon_{k-1}}{\epsilon_k} &= \tfrac{\tau_1\xi_{\mathcal{S}_{k-1}}}{\tau_1\xi_{\mathcal{S}_k}} = \sqrt{\tfrac{(N-|\mathcal{S}_{k-1}|)|\mathcal{S}_k|^2}{|\mathcal{S}_{k-1}|^2(N-|\mathcal{S}_k|)}} \\ &= \sqrt{\tfrac{4(N-\frac{1}{2}|\mathcal{S}_k|)}{(N-|\mathcal{S}_k|)}} = 2\sqrt{1 + \tfrac{\frac{1}{2}|\mathcal{S}_k|}{(N-|\mathcal{S}_k|)}} \le 2\sqrt{1+\tfrac{N}{2}}.\end{aligned}$$

Here, the last inequality follows from the fact that $x/(N-x)$ increases as x increases over $(0, N)$, and since $|\mathcal{S}_k| \in \mathbb{N}$ and $|\mathcal{S}_k| < N$ imply $|\mathcal{S}_k| \le N-1$. Since each iteration of gradient descent requires $|\mathcal{S}_k|$ constraint gradient evaluations, it follows that the total number of constraint gradient evaluations in iteration k is at most

$$|\mathcal{S}_k| \left\lceil \log_2\left(6\sqrt{5}\overline{\omega}\sqrt{1+\tfrac{N}{2}}\right)\right\rceil = |\mathcal{S}_k| \left\lceil \tfrac{1}{2}\log_2\left(180\overline{\omega}^2\left(1+\tfrac{N}{2}\right)\right)\right\rceil.$$

Thus, the total number of constraint gradient evaluations required to solve all of the subproblems in iterations $k \in \{2, \dots, K-1\}$ is at most

$$\begin{aligned}&\sum_{k=2}^{K-1} |\mathcal{S}_k| \left\lceil \tfrac{1}{2}\log_2\left(180\overline{\omega}^2\left(1+\tfrac{N}{2}\right)\right)\right\rceil \\ &= \left\lceil \tfrac{1}{2}\log_2\left(180\overline{\omega}^2\left(1+\tfrac{N}{2}\right)\right)\right\rceil \sum_{i=1}^{\lceil \log_2 N/|\mathcal{S}_1|\rceil - 2} |\mathcal{S}_1| 2^i \\ &= \left\lceil \tfrac{1}{2}\log_2\left(180\overline{\omega}^2\left(1+\tfrac{N}{2}\right)\right)\right\rceil 2|\mathcal{S}_1|(2^{\lceil \log_2 N/|\mathcal{S}_1|\rceil - 2} - 1) \\ &\le \left\lceil \tfrac{1}{2}\log_2\left(180\overline{\omega}^2\left(1+\tfrac{N}{2}\right)\right)\right\rceil 2|\mathcal{S}_1|(2^{\log_2 N/|\mathcal{S}_1| - 1} - 1) \\ &= \left\lceil \tfrac{1}{2}\log_2\left(180\overline{\omega}^2\left(1+\tfrac{N}{2}\right)\right)\right\rceil (N - 2|\mathcal{S}_1|).\end{aligned}$$

Now consider iteration $k = K$, the aim of which is to find an (ϵ, ζ)-stationary point by gradient descent using the final iterate from the subproblem solve in iteration $K-1$ as the starting point. Combining Lemma 3.10, the inequality $\epsilon_{K-1} = \tau_1\xi_{K-1} \le \tau_1/(3\kappa_1)$ given by (43), and the fact that each iteration involves N individual constraint gradient evaluations, it follows that the number of constraint gradient evaluations that are required is at most

$$N\left\lceil \log_2 \tfrac{3\sqrt{5}\overline{\omega}\epsilon_{K-1}}{\epsilon}\right\rceil \le N\left\lceil \log_2 \tfrac{\sqrt{5}\overline{\omega}\tau_1}{\kappa_1\epsilon}\right\rceil.$$

Combining the bounds that have been proved, the proof is complete. □ □

4 Numerical Results

The purpose of our numerical experiments is to demonstrate the performance of Algorithm 1 with $p_1 \ll N$ versus a one-shot approach that solves the SAA problem (3) with $\mathcal{S} = [N]$ directly (i.e., $p_1 = N$). We present the results of two experiments. The first experiment involves an artificial two-dimensional problem, for which we present the results obtained using two (sub)problem solvers, namely, [4, Algorithm 1] based on Fletcher's augmented Lagrangian function and [1, Algorithm 2.2] based on a sequential quadratic optimization (SQP) approach. We also use this two-dimensional problem to illustrate experimentally that for a problem with form (3) that is strongly Morse when $\mathcal{S} = [N]$, a sampled problem with $\mathcal{S} \subset [N]$ is also strongly Morse when $|\mathcal{S}|$ is large enough relative to N. Our second experiment involves training a neural network to predict the solution of an ordinary differential equation. All software was written using Matlab R2024b.

4.1 An Artificial Problem

Consider problem (1) with $n = 2$, $m = 1$, and the objective function defined by $f(x) = x_1$. As for the constraint function $\bar{c} : \mathbb{R}^n \to \mathbb{R}$, let ω be a two-dimensional random vector with each component having a uniform distribution over $[-\pi, \pi]$. Then, for any given $(a, \phi) \in \mathbb{R}^2_{>0}$, let us define $\overline{C} : \mathbb{R}^2 \times [-\pi, \pi]^2 \to \mathbb{R}$ according to

$$\overline{C}(x, \omega) = x_1 - x_2^2 + a \sin(\phi x_1 + \omega_1) + a \cos(\phi x_2 + \omega_2).$$

That is, for any $x \in \mathbb{R}^2$, one finds that $\bar{c}(x) = \mathbb{E}[\overline{C}(x, \omega)] = x_1 - x_2^2$. On the other hand, for any positive integer N, the corresponding SAA problem has a constraint function that is the average of functions involving sin and cos terms.

For our experiments in this section, we set $a = 10^{-4}$, $\phi = 100$, and $N = 2048$. Our first aim is to demonstrate numerically that problem (2) (i.e., with $\mathcal{S} = [N]$) is strongly Morse over $[-1, 1]^2 \subset \mathbb{R}^2$ (i.e., a box around the origin, the optimal solution of problem (1)), and that problem (3) is also strongly Morse over this region when $|S|$ is sufficiently large relative to N. This is done through the contour plots in Figure 1. These contour plots can be understood as follows. First, for each sample size we computed the reduced Hessian of the Lagrangian at x—i.e., the Hessian of the Lagrangian over the null space of the constraint Jacobian—over the region, which in this setting is a real number at each point. Second, for each sample size we computed the norm of the gradient of the Lagrangian evaluated at $(x, y(x))$. Third, for reference, we plot the point $x^*_{[N]} \approx 0$. That each problem can be seen to be strongly Morse can be found by observing that λ is bounded away from zero when $\|\nabla L\|$ is small. Specifically, the dashed contour lines for the gradient of the Lagrangian show the boundaries of the region in which $\|\nabla L_{\mathcal{S}}\| \leq 0.6$, where over this region one always finds that $\lambda_{\mathcal{S}} \geq 0.8$, which is to say that each problem appears to be $(0.6, 0.8)$-strongly Morse in the region $[-1, 1]^2$.

Let us now present the results of Algorithm 1 with $p_1 \ll N$ vs. $p_1 = N$. We present the results with two subproblem solvers. In each case, our performance criterion is the number of constraint gradient evaluations—i.e., evaluations of ∇c_i for some i—required before an iterate satisfying (6) was obtained with tolerances $\epsilon = \zeta = 10^{-6}$. An initial point x_0 was chosen at random from $[-1, 1]^2$ and used for each run. For the subproblem tolerances, we set for all $k \in [K]$ the tolerances

$$\epsilon_k = \zeta_k = 10^{-6} \sqrt{\tfrac{N(N - |\mathcal{S}_k|)}{|\mathcal{S}_k|^2} + 1}. \tag{105}$$

This choice is motivated by (42), but with the addition of 1 inside the square root to ensure that the tolerance does not vanish at $k = K$. Rather, the final tolerances (for $k = K$) are $\epsilon = \zeta = 10^{-6}$, as stated previously.

Our first comparison considers [4, Algorithm 1] as the subproblem solver. This is a second-order method based on minimizing Fletcher's augmented Lagrangian function; recall $F_{\mathcal{S}_k} : \mathbb{R}^n \to \mathbb{R}$ defined by (62). In short, the algorithm employs gradient descent with a backtracking line search on the augmented Lagrangian function until the norm of gradient of the augmented Lagrangian is below a threshold. If the resulting point satisfies the second-order termination conditions, then the algorithm terminates; otherwise, a direction of negative curvature is obtained along which a backtracking line search is employed to determine the next

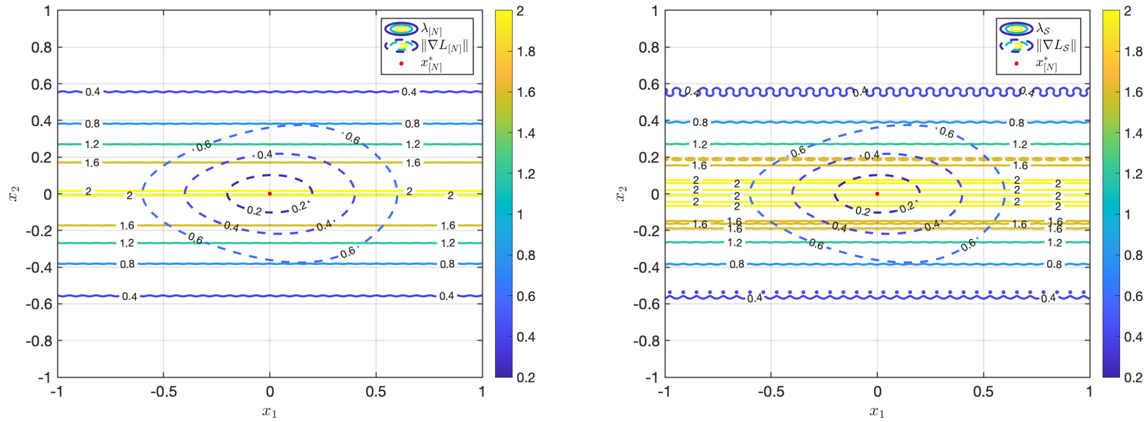

Figure 1: Contour plots for the problem defined in §4.1 with different constraint sample sizes. On the left and the right, λ represents the reduced Hessian at x (a real number), $\|\nabla L\|$ represents the norm of the gradient of the Lagrangian at $(x, y(x))$, and x^* is the optimal solution of (2). On the left, the contours correspond to function values over the full sample set with $N = 2048$. On the right, the contours correspond to values over a sample set with $|\mathcal{S}| = 64$.

iterate. The algorithm continues until the second-order termination conditions with tolerances in (105) are satisfied. In our experiment, we set the penalty parameter $\rho_{\mathcal{S}_k}$ to be 10 for all subproblems. With respect to the other parameters in [4, Algorithm 1], we chose $\alpha_{01} = \alpha_{02} = 1$, $c_1 = c_2 = 10^{-4}$, and $\tau_1 = \tau_2 = \frac{1}{2}$.

The results are shown in Figure 2. On the left in the figure, we show the norm of the Lagrangian corresponding to the full-sample problem as a function of the number of constraint gradient evaluations as the algorithm proceeds. One finds that despite the fact that $p_1 \ll N$, the progressive sampling approach yields a small norm of the Lagrangian after relatively few constraint gradient evaluations. By contrast, by solving the full-sample problem directly, many more individual constraint gradient evaluations are required before the norm of the gradient of the Lagrangian falls below the desired tolerance. On the right in the figure, we show the relative number of constraint gradient evaluations required when the final tolerance varies from 10^{-3} to 10^{-6}. In particular, one finds that the progressive sampling approach (with $p_1 \ll N$) requires only a small percentage of the constraint gradient evaluations that are required when solving the full-sample problem directly.

Our second comparison considers using [1, Algorithm 2.2] as a subproblem solver, which is an SQP-based method. This algorithm does not guarantee convergence to an approximate second-order stationary point, yet we still find that progressive sampling yields computational benefits. For the parameters required by [1, Algorithm 2.2], we chose $(\alpha, \nu, \sigma, \eta, \tau) = (1, 0.5, 0.5, 0.5, 1)$. The results are shown in Figure 3. As for the other subproblem solver, one finds that progressive sampling yields better results than the one-shot approach when considering individual constraint gradient evaluations as the performance measure.

4.2 Training a Physics-Informed Neural Network

Our second experiment involved training a neural network to predict the solution of an ordinary differential equation (ODE). In particular, our experiment trained a physics-informed neural network to predict the movement of a damped harmonic (mass-spring) oscillator under the influence of a restoring force and friction [7]. For the sake of simplicity when demonstrating the relative performance of our progressive sampling method, we trained the model for known ODE parameters and a single initial condition. Our algorithm is readily applicable to other such settings of training physics-informed neural networks as well.

The system that we considered is described by a linear, homogeneous, second-order ordinary differential

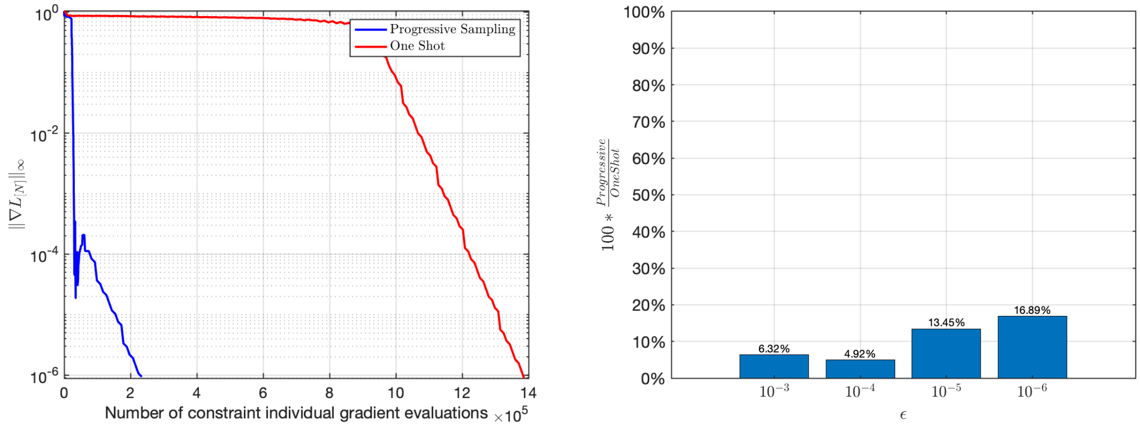

Figure 2: Performance of Algorithm 1 with $p_1 \ll N$ ("Progressive Sampling") versus $p_1 = N$ ("One Shot") when solving the problem defined in §4.1 using a subproblem solver based on minimizing Fletcher's augmented Lagrangian function. On the left, the norm of the gradient of the Lagrangian with respect to the full-sample problem as a function of the number of constraint gradient evaluations requested by the algorithm instances. On the right, the relative number of constraint gradient evaluations required by the two algorithm instances to obtain solutions satisfying the stated final tolerances.

equation with constant coefficients, namely,

$$m\frac{d^2u(t)}{dt^2} + \mu\frac{du(t)}{dt} + ku(t) = 0 \quad \text{over} \quad t \in [0, 10] \quad \text{where} \quad (m, \mu, k) = (1, 0.1, 1).$$

We let the initial conditions be $u(0) = 1$ and $du(t)/dt = -1$. To learn the solution of this ODE, we constructed a multilayer perceptron with one node in the input layer (representing time t), two hidden layers each with 128 nodes and tanh activation, and an output layer with a single node (representing the predicted height at time t, i.e., $u(t)$). The trainable parameters of the network are the edge weights as well as bias terms at each of the hidden layer nodes and the output node. Overall, the network defined a function $\mathcal{N} : \mathbb{R}^n \times \mathbb{R}_{>0} \to \mathbb{R}$, where the first argument is the vector of trainable parameters and the second argument is the input (time).

The supervised training process that we considered employed a set of known input-output pairs to define the objective function; namely, for $N_f = 128$, let the pairs be denoted $\{(t_i, u_i)\}_{i\in[N_f]}$, where u_i represents the true height of the spring at time t_i for all $i \in [N_f]$. The pairs were determined by solving the ODE with a finite-difference method over 128 evenly spaced points over the time interval $[0, 10]$. We defined the objective function to have two terms, a data-fitting term and a term to minimize the residual of the ODE at the given input times. Three constraints were included to aid in the training process. The first two were the initial conditions while the latter required that the average ODE residual at a set of $N = 512$ times equals zero. Overall, the training problem was

$$\begin{aligned}
\min_{x\in\mathbb{R}^n} \ & \frac{1}{N_f}\sum_{i\in[N_f]}\left(u_i - \mathcal{N}\left(x, \tfrac{10i}{N_f}\right)\right)^2 + \\
& \frac{1}{N_f}\sum_{i\in[N_f]}\left(\frac{\partial^2\mathcal{N}\left(x, \frac{10i}{N_f}\right)}{\partial t^2} + 0.1\frac{\partial\mathcal{N}\left(x, \frac{10i}{N_f}\right)}{\partial t} + \mathcal{N}\left(x, \tfrac{10i}{N_f}\right)\right)^2 \\
\text{s.t. } & \mathcal{N}(x, 0) = 1, \quad \frac{\partial\mathcal{N}(x,0)}{\partial t} = -1, \quad \text{and} \\
& \frac{1}{N}\sum_{i\in[N]}\left(\frac{\partial^2\mathcal{N}\left(x, \frac{10i}{N}\right)}{\partial t^2} + 0.1\frac{\partial\mathcal{N}\left(x, \frac{10i}{N}\right)}{\partial t} + \mathcal{N}\left(x, \tfrac{10i}{N}\right)\right) = 0.
\end{aligned} \tag{106}$$

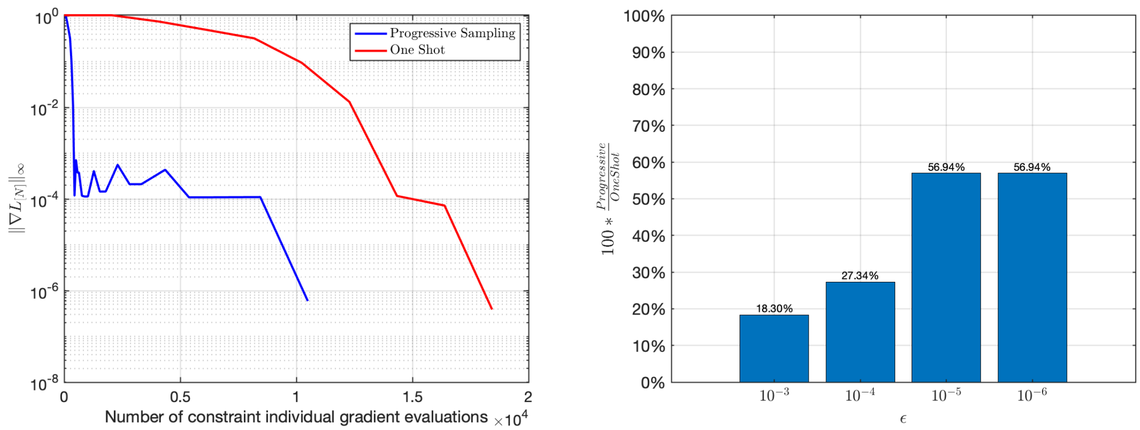

Figure 3: Performance of Algorithm 1 with $p_1 \ll N$ ("Progressive Sampling") versus $p_1 = N$ ("One Shot") when solving the problem defined in §4.1 using a subproblem solver based on a first-order SQP approach. On the left, the norm of the gradient of the Lagrangian with respect to the full-sample problem as a function of the number of constraint gradient evaluations requested by the algorithm instances. On the right, the relative number of constraint gradient evaluations required by the two algorithm instances to obtain solutions satisfying the stated final tolerances.

We ran Algorithm 1 twice with different p_1 values. For the subproblem solver, we employed an implementation of [1, Algorithm 2.2]. We emphasize that this is a first-order method, not a second-order method, yet our results still show a computational benefit of progressive sampling. First, we solved the full-sample problem directly by running Algorithm 1 with $p_1 = 512 = N$. To obtain the desired solution accuracy (specifically, (6a) with $\mathcal{S} = [N]$ below a threshold of 10^{-4}), the algorithm required approximately 4.3895e+07 individual constraint gradient evaluations. Second, we solved the problem using progressive sampling with $p_1 = 128$. To obtain the same desired solution accuracy, the algorithm required approximately 2.9022e+07 individual constraint gradients, meaning that it required approximately 33% fewer individual constraint gradient evaluations.

The true and predicted solutions (obtained after training) are shown in Figure 4. The prediction solutions with both $p_1 = 512$ and $p_1 = 128$ were qualitatively similar, so the figure only shows the latter. The final average ODE residuals with $\mathcal{S} = [N]$ for the two runs were both on the order of 10^{-7}. Let us highlight that when the progressive sampling approach with $|\mathcal{S}_1| = 128$ terminated with the first subproblem solution, the average ODE residual was only on the order of 10^{-2}. This means that the solution of (106) with $N_f = N = 128$ did not yield a very small ODE residual. However, using this solution as a starting point, a solution with a much lower ODE residual was obtained by adding further terms into the constraint with relatively modest additional computational effort. This demonstrates that our progressive sampling approach may be a viable option for enhancing the solutions obtained by solvers for training physics-informed neural networks.

5 Conclusion

We have proposed, analyzed, and tested an algorithm for solving a sample average approximation (2) of any problem of the form (1), where the equality-constraint function is defined by an expectation. The algorithm is based on a progressive sampling strategy, where in each iteration a deterministic algorithm can be employed to solve the resulting equality-constrained problem. We have shown that the algorithm can achieve improved worst-case constraint-gradient sample complexity compared to an approach that solves (2) directly. Our numerical experiments have shown that our algorithm can offer computational savings in

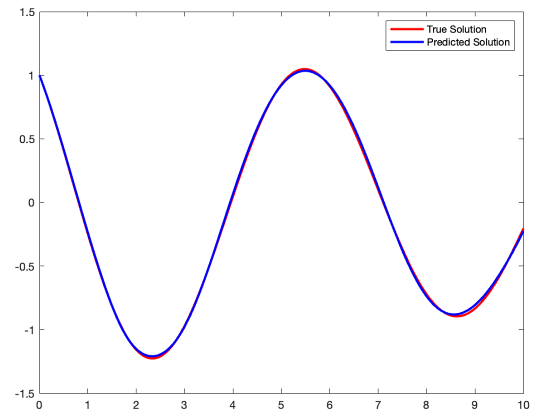

Figure 4: True and predicted ODE solutions.

practice.

6 Acknowledgments

This work was supported by Office of Naval Research award N00014-24-1-2703.

References

[1] Albert S Berahas, Frank E Curtis, Daniel Robinson, and Baoyu Zhou. Sequential quadratic optimization for nonlinear equality constrained stochastic optimization. *SIAM Journal on Optimization*, 31(2):1352–1379, 2021.

[2] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, February 2013.

[3] Jean Gallier and Jocelyn Quaintance. Linear algebra for computer vision, robotics, and machine learning. *University of Pennsylvania*, 2019.

[4] Florentin Goyens, Armin Eftekhari, and Nicolas Boumal. Computing Second-Order Points Under Equality Constraints: Revisiting Fletcher's Augmented Lagrangian. *Journal of Optimization Theory and Applications*, 201(3):1198–1228, June 2024.

[5] Song Mei, Yu Bai, and Andrea Montanari. The Landscape of Empirical Risk for Nonconvex Losses. *The Annals of Statistics*, 46(6A):2747–2774, 2018.

[6] Aryan Mokhtari, Asuman Ozdaglar, and Ali Jadbabaie. Efficient nonconvex empirical risk minimization via adaptive sample size methods. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 2485–2494. PMLR, 2019.

[7] Ben Moseley. So, what is a physics-informed neural network? `https://github.com/benmoseley/harmonic-oscillator-pinn/blob/main/Harmonic%20oscillator%20PINN.ipynb`, 2018. Published: 2021-08-28.

[8] G. W. Stewart. *Introduction to Matrix Computations*. Computer Science and Applied Mathematics. Academic Press, 1973.

[9] G. W. Stewart. On the perturbation of pseudo-inverses, projections and linear least squares problems. *SIAM Review*, 19(4):634–662, 1977.

[10] G. W. Stewart. A note on the perturbation of singular values. *Linear Algebra and its Applications*, 28:213–216, 1979.

[11] S. A. van de Geer. *Applications of Empirical Process Theory*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge ; New York, 2000.