Extended convergence analysis of the Scholtes-type regularization for cardinality-constrained optimization problems

Sebastian Lämmel[1] [ORCID] · **Vladimir Shikhman[1]**

Abstract

We extend the convergence analysis of the Scholtes-type regularization method for cardinality-constrained optimization problems. Its behavior is clarified in the vicinity of saddle points, and not just of minimizers as it has been done in the literature before. This becomes possible by using as an intermediate step the recently introduced regularized continuous reformulation of a cardinality-constrained optimization problem. We show that the Scholtes-type regularization method is well-defined locally around a nondegenerate T-stationary point of this regularized continuous reformulation. Moreover, the nondegenerate Karush–Kuhn–Tucker points of the corresponding Scholtes-type regularization converge to a T-stationary point having the same index, i.e. its topological type persists. As consequence, we conclude that the global structure of the Scholtes-type regularization essentially coincides with that of CCOP.

1 Introduction

In nonconvex optimization Scholtes-type regularization methods became popular since the seminal paper [1]. Typically, nonsmooth constraints are relaxed by means of a parameter. Then, Karush–Kuhn–Tucker points of the induced nonlinear programs

✉ Sebastian Lämmel
sebastian.laemmel@mathematik.tu-chemnitz.de

Vladimir Shikhman
vladimir.shikhman@mathematik.tu-chemnitz.de

1   Faculty of Mathematics, Chemnitz University of Technology, Reichenhainer Straße 41, Chemnitz 09126, Germany

         ⚛ Springer

need to be computed. They are shown to converge towards some suitably defined stationary points of the original optimization problem as the regularization parameter tends to zero. Scholtes-type regularization methods for mathematical programs with complementarity (MPCC), vanishing (MPVC), switching (MPSC), and orthogonality type constrains (MPOC) were examined along these lines in the literature so far, see e.g. [1–4] for further details, respectively.

In this paper, we study the Scholtes-type regularization method for the class of cardinality-constrained optimization problems:

$$\text{CCOP}: \quad \min_{x} \ f(x) \quad \text{s.t.} \quad h(x) = 0, \quad g(x) \geq 0, \quad \|x\|_0 \leq s$$

with the feasible set given by equality, inequality, and cardinality constraints, where the so-called zero "norm" $\|x\|_0 = |\{i \in \{1, \ldots, n\} \mid x_i \neq 0\}|$ is counting non-zero entries of x. Here, we assume that the objective function f, as well as the equality and inequality constraints $h = (h_p, p \in P)$, $g = (g_q, q \in Q)$ are twice continuously differentiable, and $s \in \{1, \ldots, n-1\}$ is an integer. In order to arrive at the Scholtes-type regularization, the so-called continuous reformulation of CCOP from [5] is helpful:

$$\min_{x,y} \ f(x) \quad \text{s.t.} \quad h(x) = 0, \quad g(x) \geq 0, \quad \sum_{i=1}^{n} y_i \geq n - s,$$
$$x_i y_i = 0, \quad 0 \leq y_i \leq 1, \quad i = 1, \ldots, n. \tag{1}$$

As pointed out there, $\bar{x}$ solves CCOP if and only if there exists a vector $\bar{y}$ such that $(\bar{x}, \bar{y})$ solves (1). In order to tackle (1) numerically, [6] suggests to regularize the orthogonality type constraints by using the Scholtes' idea, cf. [1]:

$$\min_{x,y} \ f(x) \quad \text{s.t.} \ h(x) = 0, \quad g(x) \geq 0, \quad \sum_{i=1}^{n} y_i \geq n - s,$$
$$-t \leq x_i y_i \leq t, \quad 0 \leq y_i \leq 1, \quad i = 1, \ldots, n, \tag{2}$$

where $t > 0$. Further in [7], the authors prove that—under some suitable constraint qualification and second-order sufficient condition—the Scholtes-type regularization method is well-defined locally around a minimizer of (1). Moreover, the Karush–Kuhn–Tucker points of (2) converge to an S-stationary point of (1) whenever $t \to 0$.

Our goal is to extend the convergence analysis of the Scholtes-type regularization method beyond the case of minimizers of (1), but also for all kinds of its saddle points. By doing so, we intend to relate the indices of nondegenerate Karush–Kuhn–Tucker points of the Scholtes-type regularization with those of T-stationary points of the regularized continuous reformulation. Here, nondegeneracy refers to some tailored versions of linear independence constraint qualification, strict complementarity and second-order regularity. Assuming nondegeneracy, Karush–Kuhn–Tucker points and T-stationary points can be classified according to their quadratic and T-index, respectively. The index encodes the local structure of the optimization problem under consideration in algebraic terms and its global structure in the sense of Morse theory,

see [8, 9]. We note that for our purpose we need to preliminarily regularize the continuous reformulation (1). The reason is that all T-stationary points of (1)—considered as an MPOC instance—turn out to be degenerate, cf. [4]. To overcome this obstacle, it has been suggested in [9] not only to linearly perturb the objective function in (1) with respect to y-variables, but also to additionally relax the upper bounds on them. As for our main results, the Scholtes-type regularization method proves to be well-defined locally around a nondegenerate T-stationary point of the regularized continuous reformulation. Moreover, the nondegenerate Karush–Kuhn–Tucker points of its Scholtes-type regularization converge to a T-stationary point having the same index. These results allow us to relate the x-variables of the Karush–Kuhn–Tucker points of the Scholtes-type regularization to the M-stationary points of CCOP directly.

We emphasize that the study of saddle points for the Scholtes-type regularization is not only valuable from the global optimization perspective, but also from the practical point of view. Indeed, since the Scholtes-type regularization falls into the scope of nonlinear programming, we can only hope to efficiently compute its Karush–Kuhn–Tucker points. This can be done e.g. by using Newton-type methods, which—as well known—do not in general converge towards minimizers. These Karush–Kuhn–Tucker points of the Scholtes-type regularization will thus appear to be saddle points of different kinds. Their convergence to the saddle points of the regularized continuous reformulation of CCOP and of CCOP itself has then to be addressed.

The article is organized as follows. In Sect. 2 we discuss some preliminary results on CCOP and its regularized continuous reformulation. Sect. 3 is devoted to the extended convergence analysis of its Scholtes-type regularization.

Our notation is standard. The cardinality of a finite set A is denoted by $|A|$. The n-dimensional Euclidean space is denoted by $\mathbb{R}^n$ with the coordinate vectors $e_i, i = 1, \ldots, n$. The vector consisting of ones is denoted by e. Given a twice continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, ∇f denotes its gradient, and $D^2 f$ stands for its Hessian.

2 Preliminaries

We start with the notion of nondegenerate stationarity for CCOP as described in [10]. For that, we use the index set of active inequality constraints and the index set of vanishing x-variables, i.e.

$$Q_0(\bar{x}) = \left\{ q \in Q \mid g_q(\bar{x}) = 0 \right\}, \quad I_0(\bar{x}) = \{ i \in \{1, \ldots, n\} \mid \bar{x}_i = 0 \}.$$

Let us introduce the CCOP-tailored linear independence constraint qualification.

Definition 1 *(CC-LICQ, see [11])* We say that a feasible point $\bar{x}$ of CCOP satisfies the cardinality-constrained linear independence constraint qualification (CC-LICQ) if the following gradients are linearly independent:

$$\nabla h_p(\bar{x}), \ p \in P, \quad \nabla g_q(\bar{x}), \ q \in Q_0(\bar{x}), \quad e_i, \ i \in I_0(\bar{x}).$$

It was shown in [10] that the topologically relevant stationary concept for CCOP is M-stationarity, namely in the sense of the Morse theory.

Definition 2 *(M-stationarity, see* [5]*)* A CCOP feasible point $\bar{x}$ is called M-stationary if there exist multipliers

$$\bar{\lambda}_p, \ p \in P, \quad \bar{\mu}_q, \ q \in Q_0(\bar{x}), \quad \bar{\gamma}_i, \ i \in I_0(\bar{x}),$$

such that the following conditions hold:

$$\nabla f(\bar{x}) = \sum_{p \in P} \bar{\lambda}_p \nabla h_p(\bar{x}) + \sum_{q \in Q_0(\bar{x})} \bar{\mu}_q \nabla g_q(\bar{x}) + \sum_{i \in I_0(\bar{x})} \bar{\gamma}_i e_i, \tag{3}$$

$$\bar{\mu}_q \geq 0 \text{ for all } q \in Q_0(\bar{x}). \tag{4}$$

It is convenient to define the Lagrange function, since the multipliers are unique under CC-LICQ, cf. [6],

$$L(x) = f(x) - \sum_{p \in P} \bar{\lambda}_p h_p(x) - \sum_{q \in Q_0(\bar{x})} \bar{\mu}_q g_q(x) - \sum_{i \in I_0(\bar{x})} \bar{\gamma}_i x_i.$$

We also use the corresponding tangent space

$$\mathcal{T}_{\bar{x}} = \left\{ \xi \in \mathbb{R}^n \ \middle| \ Dh_p(\bar{x})\xi = 0, \ p \in P, \ Dg_q(\bar{x})\xi = 0, \ q \in Q_0(\bar{x}), \ \xi_i = 0, \ i \in I_0(\bar{x}) \right\}.$$

We now focus on the definition of nondegeneracy for M-stationary points, which was introduced in [10]. It is justified there by showing that all M-stationary points of CCOP are generically nondegenerate.

Definition 3 *(Nondegenerate M-stationarity, see* [10]*)* An M-stationary point $\bar{x}$ of CCOP is called nondegenerate if

NDM1: CC-LICQ holds at $\bar{x}$,
NDM2: $\bar{\mu}_q > 0$ for all $q \in Q_0(\bar{x})$,
NDM3: if $\|\bar{x}\|_0 < s$ then $\bar{\gamma}_i \neq 0$ for all $i \in I_0(\bar{x})$,
NDM4: the matrix $D^2 L(\bar{x}) \restriction_{\mathcal{T}_{\bar{x}}}$ is nonsingular.

For a nondegenerate M-stationary point we eventually use an additional condition:

NDM5: $\gamma_i \neq 0$ holds for all $i \in I_0(\bar{x})$.

With a nondegenerate M-stationary point $\bar{x}$ an M-index can be associated. The M-index captures the structure of CCOP locally around $\bar{x}$ and defines the type of an M-stationary point, see [10] for details. In particular, nondegenerate minimizers of CCOP are characterized by a vanishing M-index. If the M-index does not vanish, we get all kinds of saddle points.

Definition 4 *(M-index, see* [10]*)* Let $\bar{x}$ be a nondegenerate M-stationary point of CCOP. The number of negative eigenvalues of the matrix $D^2 L(\bar{x}) \restriction_{\mathcal{T}_{\bar{x}}}$ is called its quadratic index (QI). The number $s - \|\bar{x}\|_0$ is called the sparsity index (SI) of $\bar{x}$. We define the M-index (MI) as the sum of both, i. e. $MI = SI + QI$.

Now, we are ready to associate with CCOP the regularized continuous reformulation as suggested in [9]:

$$\mathcal{R}(c, \varepsilon): \quad \min_{x,y} \ f(x) + c^T y \quad \text{s.t.} \ h(x) = 0, \quad g(x) \geq 0, \quad \sum_{i=1}^{n} y_i \geq n - s,$$

$$x_i y_i = 0, \quad 0 \leq y_i \leq 1 + \varepsilon, \quad i = 1, \ldots, n,$$

where the components of $c \in \mathbb{R}^n$ are positive and pairwise different, and $0 < \varepsilon \leq \frac{1}{n-s}$. Given a feasible point $(\bar{x}, \bar{y})$ of $\mathcal{R}$, we define the index sets which correspond to the orthogonality type constraints $x_i y_i = 0$, $y_i \geq 0$:

$$a_{01}(\bar{x}, \bar{y}) = \{i \mid \bar{x}_i = 0, \bar{y}_i > 0\}, \quad a_{10}(\bar{x}, \bar{y}) = \{i \mid \bar{x}_i \neq 0, \bar{y}_i = 0\},$$
$$a_{00}(\bar{x}, \bar{y}) = \{i \mid \bar{x}_i = 0, \bar{y}_i = 0\}.$$

The index sets of the active inequality constraints will be denoted by

$$Q_0(\bar{x}) = \{q \in Q \mid g_q(\bar{x}) = 0\}, \quad \mathcal{E}(\bar{y}) = \{i \mid \bar{y}_i = 1 + \varepsilon\}.$$

The regularized continuous reformulation $\mathcal{R}$ is a special case of MPOC. The latter class was examined in [4], where the MPOC-tailored linear independence constraint qualification and the notion of (nondegenerate) T-stationary points with the corresponding T-index were introduced. It has been shown there that T-stationarity is the topologically relevant stationarity notion for MPOC, again in the sense of Morse theory. We note that the alternative concept of S-stationarity has been defined for the original continuous reformulation (1). It has been shown in [4] that S-stationarity implies T-stationarity for (1), but not vice versa. These both facts motivated us in [9] to apply T-, rather than S-stationarity to the regularization $\mathcal{R}$.

Definition 5 *(MPOC-LICQ, [9])* We say that a feasible point $(\bar{x}, \bar{y})$ of $\mathcal{R}$ satisfies the MPOC-tailored linear independence constraint qualification (MPOC-LICQ) if the following vectors are linearly independent:

$$\begin{pmatrix} \nabla h_p(\bar{x}) \\ 0 \end{pmatrix}, \ p \in P, \quad \begin{pmatrix} \nabla g_q(\bar{x}) \\ 0 \end{pmatrix}, \ q \in Q_0(\bar{x}),$$

$$\begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in \mathcal{E}(\bar{y}), \quad \begin{pmatrix} 0 \\ e \end{pmatrix} \text{ if } \sum_{i=1}^{n} \bar{y}_i = n - s,$$

$$\begin{pmatrix} e_i \\ 0 \end{pmatrix}, \ i \in a_{01}(\bar{x}, \bar{y}) \cup a_{00}(\bar{x}, \bar{y}), \quad \begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in a_{10}(\bar{x}, \bar{y}) \cup a_{00}(\bar{x}, \bar{y}).$$

Definition 6 *(T-stationary point, [9])* A feasible point $(\bar{x}, \bar{y})$ of $\mathcal{R}$ is called T-stationary if there exist multipliers

$$\bar{\lambda}_p, \ p \in P, \quad \bar{\mu}_{1,q}, \ q \in Q_0(\bar{x}), \quad \bar{\mu}_{2,i}, \ i \in \mathcal{E}(\bar{y}), \quad \bar{\mu}_3,$$
$$\bar{\sigma}_{1,i}, \ i \in a_{01}(\bar{x}, \bar{y}), \quad \bar{\sigma}_{2,i}, \ i \in a_{10}(\bar{x}, \bar{y}), \quad \bar{\varrho}_{1,i}, \bar{\varrho}_{2,i}, \ i \in a_{00}(\bar{x}, \bar{y}),$$

such that the following conditions hold:

$$
\begin{pmatrix} \nabla f(\bar{x}) \\ c \end{pmatrix} = \sum_{p \in P} \bar{\lambda}_p \begin{pmatrix} \nabla h_p(\bar{x}) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(\bar{x})} \bar{\mu}_{1,q} \begin{pmatrix} \nabla g_q(\bar{x}) \\ 0 \end{pmatrix} - \sum_{i \in \mathcal{E}(\bar{y})} \bar{\mu}_{2,i} \begin{pmatrix} 0 \\ e_i \end{pmatrix}
$$

$$
+ \bar{\mu}_3 \begin{pmatrix} 0 \\ e \end{pmatrix} + \sum_{i \in a_{01}(\bar{x}, \bar{y})} \bar{\sigma}_{1,i} \begin{pmatrix} e_i \\ 0 \end{pmatrix} + \sum_{i \in a_{10}(\bar{x}, \bar{y})} \bar{\sigma}_{2,i} \begin{pmatrix} 0 \\ e_i \end{pmatrix}
$$

$$
+ \sum_{i \in a_{00}(\bar{x}, \bar{y})} \left(\bar{\varrho}_{1,i} \begin{pmatrix} e_i \\ 0 \end{pmatrix} + \bar{\varrho}_{2,i} \begin{pmatrix} 0 \\ e_i \end{pmatrix} \right), \tag{5}
$$

$$
\bar{\mu}_{1,q} \geq 0, \, q \in Q_0(\bar{x}), \quad \bar{\mu}_{2,i} \geq 0, \, i \in \mathcal{E}(\bar{y}),
$$

$$
\bar{\mu}_3 \geq 0, \, \bar{\mu}_3 \left(\sum_{i=1}^{n} \bar{y}_i - (n-s) \right) = 0, \tag{6}
$$

$$
\bar{\varrho}_{1,i} = 0 \text{ or } \bar{\varrho}_{2,i} \leq 0, \, i \in a_{00}(\bar{x}, \bar{y}). \tag{7}
$$

We define the appropriate Lagrange function:

$$
L^{\mathcal{R}}(x, y) = f(x) + c^T y - \sum_{p \in P} \bar{\lambda}_p h_p(x) - \sum_{q \in Q_0(\bar{x})} \bar{\mu}_{1,q} g_q(x)
$$

$$
+ \sum_{i \in \mathcal{E}(\bar{y})} \bar{\mu}_{2,i} (y_i - (1 + \varepsilon)) - \bar{\mu}_3 \left(\sum_{i=1}^{n} y_i - (n-s) \right)
$$

$$
- \sum_{i \in a_{01}(\bar{x}, \bar{y})} \bar{\sigma}_{1,i} x_i - \sum_{i \in a_{10}(\bar{x}, \bar{y})} \bar{\sigma}_{2,i} y_i - \sum_{i \in a_{00}(\bar{x}, \bar{y})} \left(\bar{\varrho}_{1,i} x_i + \bar{\varrho}_{2,i} y_i \right).
$$

Moreover, we set for the corresponding tangent space

$$
T^{\mathcal{R}}_{(\bar{x}, \bar{y})} = \left\{ \xi \in \mathbb{R}^{2n} \left| \begin{array}{l} \left(Dh_p(\bar{x}), 0 \right) \xi = 0, \, p \in P, \left(Dg_q(\bar{x}), 0 \right) \xi = 0, \, q \in Q_0(\bar{x}), \\ \left(0, e_i \right) \xi = 0, \, i \in \mathcal{E}(\bar{y}), \left(0, e \right) \xi = 0 \text{ if } \sum_{i=1}^{n} \bar{y}_i = n - s, \\ \left(e_i, 0 \right) \xi = 0, \, i \in a_{00}(\bar{x}, \bar{y}) \cup a_{01}(\bar{x}, \bar{y}), \\ \left(0, e_i \right) \xi = 0, \, i \in a_{00}(\bar{x}, \bar{y}) \cup a_{10}(\bar{x}, \bar{y}) \end{array} \right. \right\}.
$$

Definition 7 *(Nondegenerate T-stationary point,* [9]*)* A T-stationary point $(\bar{x}, \bar{y})$ of $\mathcal{R}$ with multipliers $(\bar{\lambda}, \bar{\mu}, \bar{\sigma}, \bar{\varrho})$ is called nondegenerate if

NDT1: MPOC-LICQ holds at $(\bar{x}, \bar{y})$,
NDT2: $\bar{\mu}_{1,q} > 0, q \in Q_0(\bar{x}), \bar{\mu}_{2,i} > 0, i \in \mathcal{E}(\bar{y})$, and $\bar{\mu}_3 > 0$ if $\sum_{i=1}^{n} \bar{y}_i = n - s$,
NDT3: $\bar{\varrho}_{1,i} \neq 0$ and $\bar{\varrho}_{2,i} < 0, i \in a_{00}(\bar{x}, \bar{y})$,
NDT4: the matrix $D^2 L^{\mathcal{R}}(\bar{x}, \bar{y}) \restriction_{T^{\mathcal{R}}_{(\bar{x}, \bar{y})}}$ is nonsingular.

For a nondegenerate T-stationary point we eventually use additional conditions:

NDT5: if $a_{00}(\bar{x}, \bar{y}) \neq \emptyset$, then $\bar{\sigma}_{1,i} \neq 0$ for all $i \in a_{01}(\bar{x}, \bar{y})$.
NDT6: $\bar{\sigma}_{1,i} \neq 0, i \in a_{01}(\bar{x}, \bar{y})$.

Definition 8 *(T-index, [9])* Let $(\bar{x}, \bar{y})$ be a nondegenerate T-stationary point of $\mathcal{R}$ with unique multipliers $(\bar{\lambda}, \bar{\mu}, \bar{\sigma}, \bar{\varrho})$. The number of negative eigenvalues of the matrix $D^2 L^{\mathcal{R}}(\bar{x}, \bar{y}) \restriction_{T^{\mathcal{R}}_{(\bar{x}, \bar{y})}}$ is called its quadratic index (QI). The cardinality of $a_{00}(\bar{x}, \bar{y})$ is called the biactive index (BI) of $(\bar{x}, \bar{y})$. We define the T-index (TI) as the sum of both, i.e. $TI = QI + BI$.

The nondegeneracy conditions NDT1-NDT4 are tailored for $\mathcal{R}$. Note that NDT2 corresponds to the strict complementarity and NDT4 to the second-order regularity as they are typically defined in the context of nonlinear programming. NDT1 substitutes the usual linear independence constraint qualification. NDT3 is new and says that the multipliers corresponding to biactive orthogonality type constraints must not vanish. With a nondegenerate T-stationary point $(\bar{x}, \bar{y})$ a T-index can be associated. The T-index captures the structure of $\mathcal{R}$ locally around $(\bar{x}, \bar{y})$ and defines the type of a T-stationary point, see [9] for details. In particular, nondegenerate minimizers of $\mathcal{R}$ are characterized by a vanishing T-index. If the T-index does not vanish, we get all kinds of saddle points.

Next Lemma 1 provides insights into the structure of auxiliary y-variables corresponding to a T-stationary point of $\mathcal{R}$.

Lemma 1 (Auxiliary y-variables in $\mathcal{R}$, [9]) *Let $(\bar{x}, \bar{y})$ be a T-stationary point of $\mathcal{R}$, then it holds:*

(a) *the summation inequality constraint is active, i.e. $\sum_{i=1}^{n} \bar{y}_i = n - s$,*
(b) *the index set $a_{01}(\bar{x}, \bar{y})$ consists of exactly $n - s$ elements,*
(c) *$n - s - 1$ components of $\bar{y}$ are equal to $1 + \varepsilon$, one component is equal to $1 - (n - s - 1)\varepsilon$, and s remaining components vanish.*

We note that nondegenerate M-stationary points of CCOP naturally correspond to nondegenerate T-stationary points of $\mathcal{R}$ and vice versa. As shown in [9], also their M- and T-indices coincide. Thus, the regularized continuous reformulation $\mathcal{R}$ can be likewise studied instead of (1).

Theorem 1 (Stationarity of $\mathcal{R}$ and CCOP, [9])

(a) *If $\bar{x}$ is an M-stationary point of CCOP, then there exist at least $\binom{n - \|\bar{x}\|_0 - 1}{n - s - 1}$ choices of $\bar{y}$ such that $(\bar{x}, \bar{y})$ is a T-stationary point of $\mathcal{R}$. If $\bar{x}$ is additionally nondegenerate with M-index m, then all corresponding T-stationary points $(\bar{x}, \bar{y})$ are also nondegenerate with T-index m. Moreover, their number is exactly $\binom{n - \|\bar{x}\|_0 - 1}{n - s - 1}$, and NDT5 holds at any of them.*

b) *If $(\bar{x}, \bar{y})$ is a T-stationary point of $\mathcal{R}$, then $\bar{x}$ is an M-stationary point of CCOP. If $(\bar{x}, \bar{y})$ is additionally nondegenerate with T-index m and satisfies NDT5, then $\bar{x}$ is also nondegenerate with M-index m.*

3 Scholtes-type regularization

Let us now regularize the orthogonality type constraints in $\mathcal{R}$ by using the Scholtes' idea, cf. [1]:

$$\mathcal{S}(t): \quad \min_{x,y} \ f(x) + c^T y \quad \text{s.t.} \ h(x) = 0, \quad g(x) \geq 0, \quad \sum_{i=1}^{n} y_i \geq n - s,$$

$$-t \leq x_i y_i \leq t, \quad 0 \leq y_i \leq 1 + \varepsilon, \quad i = 1, \ldots, n,$$

where $t > 0$. Note that $\mathcal{S}$ from above falls into the scope of nonlinear programming. The notation for the sets $Q_0(x)$ and $\mathcal{E}(y)$, which were used for $\mathcal{R}$, will be used here again. Furthermore, we define for a feasible point (x, y) of $\mathcal{S}$ the index set of vanishing y-components as well as the index sets of active relaxed orthogonality type constraints:

$$\mathcal{N}(y) = \{i \mid y_i = 0\}, \quad \mathcal{H}^{\geqq}(x, y) = \{i \mid x_i y_i = -t\}, \mathcal{H}^{\leqq}(x, y) = \{i \mid x_i y_i = t\}.$$

We also eventually use the following index sets:

$$\mathcal{H}(x, y) = \mathcal{H}^{\geqq}(x, y) \cup \mathcal{H}^{\leqq}(x, y), \quad \mathcal{O}(x, y) = (\mathcal{E}(y) \cup \mathcal{N}(y) \cup \mathcal{H}(x, y))^c.$$

For the sake of completeness we state the linear independence constraint qualification for the nonlinear programming problem $\mathcal{S}$.

Definition 9 *(LICQ)* We say that a feasible point (x, y) of $\mathcal{S}$ satisfies the linear independence constraint qualification (LICQ) if the following vectors are linearly independent:

$$\begin{pmatrix} \nabla h_p(x) \\ 0 \end{pmatrix}, \ p \in P, \quad \begin{pmatrix} \nabla g_q(x) \\ 0 \end{pmatrix}, \ q \in Q_0(x), \quad \begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in \mathcal{E}(y),$$

$$\begin{pmatrix} 0 \\ e \end{pmatrix} \text{ if } \sum_{i=1}^{n} y_i = n - s, \quad \begin{pmatrix} y_i e_i \\ x_i e_i \end{pmatrix}, \ i \in \mathcal{H}(x, y), \quad \begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in \mathcal{N}(y).$$

Let us relate MPOC-LICQ for $\mathcal{R}$ with LICQ for $\mathcal{S}$.

Theorem 2 (MPOC-LICQ vs. LICQ) *Let a feasible point $(\bar{x}, \bar{y})$ of $\mathcal{R}$ fulfill MPOC-LICQ. Then, LICQ holds at all feasible points (x, y) of $\mathcal{S}$ for all sufficiently small t, whenever they are sufficiently close to $(\bar{x}, \bar{y})$.*

Proof Let us contrarily assume that there exists a sequence of feasible points (x^t, y^t) of $\mathcal{S}$ violating LICQ, which converges to $(\bar{x}, \bar{y})$ for $t \to 0$. Additionally, suppose that along some subsequence, which we index by t again, it holds $\sum_{i=1}^{n} y_i^t = n - s$. Then, we have $\sum_{i=1}^{n} \bar{y}_i = n - s$. Due to MPOC-LICQ at $(\bar{x}, \bar{y})$ as well as continuity of ∇h and ∇g, we have that for t sufficiently small all multipliers $\bar{\lambda}^t, \bar{\mu}^t, \bar{\sigma}^t, \bar{\varrho}^t$ in the following equation vanish:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{p \in P} \bar{\lambda}_p^t \begin{pmatrix} \nabla h_p(x^t) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(\bar{x})} \bar{\mu}_{1,q}^t \begin{pmatrix} \nabla g_q(x^t) \\ 0 \end{pmatrix} - \sum_{i \in \mathcal{E}(\bar{y})} \bar{\mu}_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}$$

$$
+ \bar{\mu}_3^t \begin{pmatrix} 0 \\ e \end{pmatrix} + \sum_{i \in a_{01}(\bar{x}, \bar{y})} \bar{\sigma}_{1,i}^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{i \in a_{10}(\bar{x}, \bar{y})} \bar{\sigma}_{2,i}^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}
$$
$$
+ \sum_{i \in a_{00}(\bar{x}, \bar{y})} \left(\bar{\varrho}_{1,i}^t \begin{pmatrix} e_i \\ 0 \end{pmatrix} + \bar{\varrho}_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix} \right). \tag{8}
$$

Moreover, due to the violation of LICQ at (x^t, y^t), there exist multipliers λ^t, μ^t, η^t, ν^t, not all vanishing, with

$$
\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{p \in P} \lambda_p^t \begin{pmatrix} \nabla h_p(x^t) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(x^t)} \mu_{1,q}^t \begin{pmatrix} \nabla g_q(x^t) \\ 0 \end{pmatrix} + \sum_{i \in \mathcal{E}(y^t)} \mu_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}
$$
$$
+ \mu_3^t \begin{pmatrix} 0 \\ e \end{pmatrix} + \sum_{i \in \mathcal{H}(x^t, y^t)} \eta_i^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{i \in \mathcal{N}(y^t)} \nu_i^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}.
$$

For t sufficiently small we have $Q_0(x^t) \subset Q_0(\bar{x})$ and $\mathcal{E}(y^t) \subset \mathcal{E}(\bar{y})$. In addition, it holds $\mathcal{H}(x^t, y^t) \subset a_{01}(\bar{x}, \bar{y}) \cup a_{10}(\bar{x}, \bar{y}) \cup a_{00}(\bar{x}, \bar{y})$ and $\mathcal{N}(y^t) \subset a_{10}(\bar{x}, \bar{y}) \cup a_{00}(\bar{x}, \bar{y})$. By setting some μ-multipliers to be zero if needed, we equivalently obtain:

$$
\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{p \in P} \lambda_p^t \begin{pmatrix} \nabla h_p(x^t) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(\bar{x})} \mu_{1,q}^t \begin{pmatrix} \nabla g_q(x^t) \\ 0 \end{pmatrix} + \sum_{i \in \mathcal{E}(\bar{y})} \mu_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}
$$
$$
+ \mu_3^t \begin{pmatrix} 0 \\ e \end{pmatrix} + \sum_{i \in \mathcal{H}(x^t, y^t) \cap a_{01}(\bar{x}, \bar{y})} \eta_i^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{i \in \mathcal{H}(x^t, y^t) \cap a_{10}(\bar{x}, \bar{y})} \eta_i^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}
$$
$$
+ \sum_{i \in \mathcal{H}(x^t, y^t) \cap a_{00}(\bar{x}, \bar{y})} \eta_i^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}
$$
$$
+ \sum_{i \in \mathcal{N}(y^t) \cap a_{10}(\bar{x}, \bar{y})} \nu_i^t \begin{pmatrix} 0 \\ e_i \end{pmatrix} + \sum_{i \in \mathcal{N}(y^t) \cap a_{00}(\bar{x}, \bar{y})} \nu_i^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}.
$$

This, however, implies that not all multipliers in the following equation vanish:

$$
\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{p \in P} \hat{\lambda}_p^t \begin{pmatrix} \nabla h_p(x^t) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(\bar{x})} \hat{\mu}_{1,q}^t \begin{pmatrix} \nabla g_q(x^t) \\ 0 \end{pmatrix} + \sum_{i \in \mathcal{E}(\bar{y})} \hat{\mu}_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}
$$
$$
+ \hat{\mu}_3^t \begin{pmatrix} 0 \\ e \end{pmatrix} + \sum_{i \in \mathcal{H}(x^t, y^t) \cap a_{01}(\bar{x}, \bar{y})} \hat{\eta}_i^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{i \in \mathcal{H}(x^t, y^t) \cap a_{10}(\bar{x}, \bar{y})} \hat{\eta}_i^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}
$$
$$
+ \sum_{i \in \mathcal{H}(x^t, y^t) \cap a_{00}(\bar{x}, \bar{y})} \hat{\eta}_{1,i}^t \begin{pmatrix} e_i \\ 0 \end{pmatrix} + \sum_{i \in \mathcal{H}(x^t, y^t) \cap a_{00}(\bar{x}, \bar{y})} \hat{\eta}_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}
$$
$$
+ \sum_{i \in \mathcal{N}(y^t) \cap a_{10}(\bar{x}, \bar{y})} \hat{\nu}_i^t \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{i \in \mathcal{N}(y^t) \cap a_{00}(\bar{x}, \bar{y})} \hat{\nu}_i^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}.
$$

A contradiction to (8) follows by taking into account that $\mathcal{H}(x^t, y^t) \cap \mathcal{N}(y^t) = \emptyset$. If instead we suppose that there is no subsequence with $\sum_{i=1}^{n} y_i^t = n - s$, then we can consider a subsequence with $\sum_{i=1}^{n} y_i^t > n - s$. By following a similar argumentation, we produce a contradiction to (8) again. $\qquad\square$

Next, we give the definitions of a (nondegenerate) Karush–Kuhn–Tucker point of $\mathcal{S}$ and of its quadratic index as it is meanwhile standard in nonlinear programming, see e.g. [8].

Definition 10 *(Karush–Kuhn–Tucker point)* A feasible point (x, y) of $\mathcal{S}$ is called Kurush–Kuhn–Tucker point if there exist multipliers

$$\lambda_p, \ p \in P, \quad \mu_{1,q}, \ q \in Q_0(x), \quad \mu_{2,i}, \ i \in \mathcal{E}(y), \quad \mu_3, \quad \eta_i^{\geq}, \eta_i^{\leq}, \nu_i, \ i \in \{1, \ldots, n\},$$

such that the following conditions hold:

$$\begin{pmatrix} \nabla f(x) \\ c \end{pmatrix} = \sum_{p \in P} \lambda_p \begin{pmatrix} \nabla h_p(x) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(x)} \mu_{1,q} \begin{pmatrix} \nabla g_q(x) \\ 0 \end{pmatrix}$$
$$- \sum_{i \in \mathcal{E}(y)} \mu_{2,i} \begin{pmatrix} 0 \\ e_i \end{pmatrix} + \mu_3 \begin{pmatrix} 0 \\ e \end{pmatrix}$$
$$+ \sum_{i \in \mathcal{H}^{\geq}(x,y)} \eta_i^{\geq} \begin{pmatrix} y_i e_i \\ x_i e_i \end{pmatrix} - \sum_{i \in \mathcal{H}^{\leq}(x,y)} \eta_i^{\leq} \begin{pmatrix} y_i e_i \\ x_i e_i \end{pmatrix} + \sum_{i \in \mathcal{N}(y)} \nu_i \begin{pmatrix} 0 \\ e_i \end{pmatrix},$$
$$\tag{9}$$

$$\mu_{1,q} \geq 0, \ q \in Q_0(x), \quad \mu_{2,i} \geq 0, \ i \in \mathcal{E}(y),$$
$$\mu_3 \geq 0, \ \mu_3 \left(\sum_{i=1}^{n} y_i - (n - s) \right) = 0, \tag{10}$$
$$\eta_i^{\geq} \geq 0, \ i \in \mathcal{H}^{\geq}(x, y), \quad \eta_i^{\leq} \geq 0, \ i \in \mathcal{H}^{\leq}(x, y), \quad \nu_i \geq 0, \ i \in \mathcal{N}(y). \tag{11}$$

We again define the Lagrange function as

$$L^{\mathcal{S}}(x, y) = f(x) + c^T y - \sum_{p \in P} \lambda_p h_p(x) - \sum_{q \in Q_0(x)} \mu_{1,q} g_q(x)$$
$$+ \sum_{i \in \mathcal{E}(y)} \mu_{2,i} (y_i - (1 + \varepsilon)) - \mu_3 \left(\sum_{i=1}^{n} y_i - (n - s) \right)$$
$$- \sum_{i \in \mathcal{H}^{\geq}(x,y)} \eta_i^{\geq} (x_i y_i + t) + \sum_{i \in \mathcal{H}^{\leq}(x,y)} \eta_i^{\leq} (x_i y_i - t) - \sum_{i \in \mathcal{N}(y)} \nu_i y_i.$$

The tangent space is given by

$$
\mathcal{T}^{\mathcal{S}}_{(x,y)} = \left\{ \xi \in \mathbb{R}^{2n} \;\middle|\; \begin{array}{l} \left(Dh_p(x), 0\right)\xi = 0, \, p \in P, \, \left(Dg_q(x), 0\right)\xi = 0, \, q \in Q_0(x), \\[4pt] \left(0, e_i\right)\xi = 0, \, i \in \mathcal{E}(y), \, \left(0, e\right)\xi = 0 \text{ if } \sum_{i=1}^{n} y_i = n - s, \\[4pt] \left(y_i e_i, x_i e_i\right)\xi = 0, \, i \in \mathcal{H}(x, y), \, \left(0, e_i\right)\xi = 0, \, i \in \mathcal{N}(y) \end{array} \right\}.
$$

Definition 11 *(Nondegenerate Karush–Kuhn–Tucker point)* A Karush–Kuhn–Tucker point (x, y) of $\mathcal{S}$ with multipliers $(\lambda, \mu, \eta, \nu)$ is called nondegenerate if

ND1: LICQ holds at (x, y),
ND2: $\mu_{1,q} > 0, q \in Q_0(x)$, $\mu_{2,i} > 0, i \in \mathcal{E}(y)$, $\eta_i^{\geq} > 0, i \in \mathcal{H}^{\geq}(x, y)$, $\eta_i^{\leq} > 0, i \in \mathcal{H}^{\leq}(x, y)$, $\nu_i > 0, i \in \mathcal{N}(y)$, and $\mu_3 > 0$ if $\sum_{i=1}^n y_i = n - s$,
ND3: the matrix $D^2 L^{\mathcal{S}}(x, y) \restriction_{\mathcal{T}^{\mathcal{S}}_{(x,y)}}$ is nonsingular.

Definition 12 *(Quadratic index)* Let (x, y) be a Karush–Kuhn–Tucker point of $\mathcal{S}$ with unique multipliers $(\lambda, \mu, \eta, \nu)$. The number of negative eigenvalues of the matrix $D^2 L^{\mathcal{S}}(x, y) \restriction_{\mathcal{T}^{\mathcal{S}}_{(x,y)}}$ is called its quadratic index (QI).

Note that ND1-ND3 are usual assumptions in nonlinear programming. ND1 refers to the linear independence constraint qualification, ND2 means the strict complementarity, and ND3 describes the second-order regularity. For the index of a nondegenerate Karush–Kuhn–Tucker point just the quadratic part is essential.

Lemma 2 examines the structure of y-components of a Karush–Kuhn–Tucker point of $\mathcal{S}$.

Lemma 2 *(Auxiliary y-variables in $\mathcal{S}$) Let (x, y) be a Karush–Kuhn–Tucker point of $\mathcal{S}$. Then, it holds:*

(a) *the summation inequality constraint is active, i.e. $\sum_{i=1}^n y_i = n - s$,*
(b) *the index set $\mathcal{E}(y) \cup \mathcal{H}(x, y)$ consists of at least $n - s - 1$ elements, and the index set $\mathcal{N}(y)$ consists of at most s elements. Additionally, there is at most one index, that does not belong to any of these sets, i.e. $|\mathcal{O}(x, y)| \leq 1$.*

Proof a) Let (x, y) be a Karush–Kuhn–Tucker point of $\mathcal{S}$ and $\sum_{i=1}^n y_i > n - s$. Then, there exist multipliers $(\lambda, \mu, \eta, \nu)$, such that (9)–(11) are fulfilled. Since $\mu_3 = 0$, we have that the $(n + i)$-th row of (9) reads as

$$
c_i = \begin{cases} -\mu_{2,i}, & \text{for } i \in \mathcal{E}(y)\backslash\mathcal{H}(x, y), \\ -\mu_{2,i} + \eta_i^{\geq}x_i, & \text{for } i \in \mathcal{E}(y) \cap \mathcal{H}^{\geq}(x, y), \\ -\mu_{2,i} - \eta_i^{\leq}x_i, & \text{for } i \in \mathcal{E}(y) \cap \mathcal{H}^{\leq}(x, y), \\ \eta_i^{\geq}x_i, & \text{for } i \in \mathcal{H}^{\geq}(x, y)\backslash\mathcal{E}(y), \\ -\eta_i^{\leq}x_i, & \text{for } i \in \mathcal{H}^{\leq}(x, y)\backslash\mathcal{E}(y), \\ \nu_i, & \text{for } i \in \mathcal{N}(y), \\ 0, & \text{else.} \end{cases}
$$

Due to (10), (11), and $c > 0$, it must hold that $i \in \mathcal{N}(y)$ for all $i \in \{1, \ldots, n\}$. This, however, contradicts $\sum_{i=1}^n y_i > n - s$.

b) As in the proof of statement a), we conclude that $\mu_3 > 0$ for a Karush–Kuhn–Tucker point (x, y) of $\mathcal{S}$. Hence, the $(n + i)$-th row now reads as

$$
c_i = \begin{cases}
-\mu_{2,i} + \mu_3, & \text{for } i \in \mathcal{E}(y)\backslash\mathcal{H}(x, y), \\
-\mu_{2,i} + \mu_3 + \eta_i^{\geq} x_i, & \text{for } i \in \mathcal{E}(y) \cap \mathcal{H}^{\geq}(x, y), \\
-\mu_{2,i} + \mu_3 - \eta_i^{\leq} x_i, & \text{for } i \in \mathcal{E}(y) \cap \mathcal{H}^{\leq}(x, y), \\
\mu_3 + \eta_i^{\geq} x_i, & \text{for } i \in \mathcal{H}^{\geq}(x, y)\backslash\mathcal{E}(y), \\
\mu_3 - \eta_i^{\leq} x_i, & \text{for } i \in \mathcal{H}^{\leq}(x, y)\backslash\mathcal{E}(y), \\
\mu_3 + v_i, & \text{for } i \in \mathcal{N}(y), \\
\mu_3, & \text{else.}
\end{cases}
\tag{12}
$$

It follows from (12) and the components of c being pairwise different that there can be at most one element $\bar{i} \in \mathcal{O}(x, y)$. If $\mathcal{E}(y) \cup \mathcal{H}(x, y)$ consists of fewer than $n - s - 1$ elements, we get:

$$
\sum_{i=1}^{n} y_i \leq (n - s - 2) \cdot (1 + \varepsilon) + y_{\bar{i}} < (n - s - 1) \cdot (1 + \varepsilon) < n - s,
$$

a contradiction. Finally, we assume that $\mathcal{N}(y)$ consists of more than s elements. In this case, there are at most $n-s-1$ nonvanishing components of y. Consequently,

$$
\sum_{i=1}^{n} y_i \leq (n - s - 1) \cdot (1 + \varepsilon) < n - s
$$

provides a contradiction. □

We apply the general result on the Scholtes-type regularization of MPOC in our context for the regularized continuous reformulation $\mathcal{R}$, see [4].

Theorem 3 (Convergence from $\mathcal{S}$ to $\mathcal{R}$, cf. [4]) *Suppose that a sequence of Karush–Kuhn–Tucker points (x^t, y^t) of $\mathcal{S}$ converges to $(\bar{x}, \bar{y})$ for $t \to 0$. If MPOC-LICQ holds at $(\bar{x}, \bar{y})$, then it is a T-stationary point of $\mathcal{R}$.*

From the proof of Theorem 3 in [4] also the convergence of the corresponding multipliers can be deduced.

Remark 1 (*Convergence of multipliers*) Let $\left(\lambda^t, \mu^t, \eta^t, v^t\right)$ be the multipliers of the Karush–Kuhn–Tucker points (x^t, y^t) of $\mathcal{S}$ and $\left(\bar{\lambda}, \bar{\mu}, \bar{\sigma}, \bar{\varrho}\right)$ of the T-stationary point $(\bar{x}, \bar{y})$ of $\mathcal{R}$ as in Theorem 3. Due to MPOC-LICQ at $(\bar{x}, \bar{y})$, we have:

a) $\lim_{t\to 0} \lambda^t = \bar{\lambda}$, $\lim_{t\to 0} \mu^t = \bar{\mu}$,

b) $\lim_{t\to 0} \left(\eta_i^{\geq,t} - \eta_i^{\leq,t}\right) y_i^t = \bar{\sigma}_{1,i}, i \in a_{01}(\bar{x}, \bar{y})$,

c) $\lim_{t\to 0} v_i^t + \left(\eta_i^{\geq,t} - \eta_i^{\leq,t}\right) x_i^t = \bar{\sigma}_{2,i}, i \in a_{10}(\bar{x}, \bar{y})$,

d) $\lim_{t\to 0} \left(\eta_i^{\geq,t} - \eta_i^{\leq,t}\right) y_i^t = \bar{\varrho}_{1,i}, \lim_{t\to 0} v_i^t + \left(\eta_i^{\geq,t} - \eta_i^{\leq,t}\right) x_i^t = \bar{\varrho}_{2,i}, i \in a_{00}(\bar{x}, \bar{y})$.

The convergence of nondegenerate Karush–Kuhn–Tucker points of $\mathcal{S}$ does not prevent the limiting T-stationary point of $\mathcal{S}$ from being degenerate. Let us present in Example 1 the failure of NDT2. Examples with the failure of NDT1, NDT3, or NDT4 are not difficult to construct analogously.

Example 1 *(Failure of NDT2)* We consider the following Scholtes-type regularization $\mathcal{S}$ with $n = 2$ and $s = 1$:

$$\mathcal{S}: \quad \min_{x,y} (x_1 - 1)^2 + (x_2 - 1)^2 + c_1 y_1 + (c_1 + \frac{5}{36})y_2$$
$$\text{s.t.} 1 + x_1 - x_2 \geq 0,$$
$$y_1 + y_2 \geq 1, \quad -t \leq x_i y_i \leq t, \quad 0 \leq y_i \leq 1 + \varepsilon, \quad i = 1, 2,$$

as well as the point $(x^t, y^t) = (t, 1, 1, 0)$. We claim that this point is a nondegenerate Karush–Kuhn–Tucker point for $t < \frac{1}{2} - \sqrt{\frac{13}{72}}$. Indeed, it holds:

$$\begin{pmatrix} 2t - 2 \\ 0 \\ c_1 \\ c_1 + \frac{5}{36} \end{pmatrix} = \mu_3^t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \eta_{\bar{1}}^{\leq,t} \begin{pmatrix} 1 \\ 0 \\ t \\ 0 \end{pmatrix} + v_2^t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

with the positive multipliers $\mu_3^t = c_1 + 2t - 2t^2$, $\eta_{\bar{1}}^{\leq,t} = 2 - 2t$, $v_2^t = \frac{5}{36} - 2t + 2t^2$. The tangent space is $\mathcal{T}_{(x^t, y^t)}^{\mathcal{S}} = \{\xi \in \mathbb{R}^4 \,|\, \xi_1 = \xi_3 = \xi_4 = 0\}$. The Hessian of the corresponding Lagrange function is

$$D^2 L^{\mathcal{S}}(x^t, y^t) = \begin{pmatrix} 2 & 0 & 2 - 2t & 0 \\ 0 & 2 & 0 & 0 \\ 2 - 2t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore, it is straightforward that $D^2 L^{\mathcal{S}}(x^t, y^t) \restriction_{\mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}}$ is nonsingular. We conclude that ND1-ND3 are fulfilled at (x^t, y^t). Moreover, (x^t, y^t) converges to $(\bar{x}, \bar{y}) = (0, 1, 1, 0)$ if $t \to 0$. This point is T-stationary for the corresponding regularized continuous reformulation $\mathcal{R}$ according to Theorem 3, since MPOC-LICQ is fulfilled. Indeed, we obtain the T-stationarity condition

$$\begin{pmatrix} -2 \\ 0 \\ c_1 \\ c_1 + \frac{5}{36} \end{pmatrix} = \bar{\mu}_1 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \bar{\mu}_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \bar{\sigma}_{1,1} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \bar{\sigma}_{2,2} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

with the unique multipliers $\bar{\mu}_1 = 0$, $\bar{\mu}_3 = c_1$, $\bar{\sigma}_{1,1} = -2$, $\bar{\sigma}_{2,2} = \frac{5}{36}$. However, NDT2 is violated at $(\bar{x}, \bar{y})$. $\qquad\square$

Due to Example 1, we cannot expect that a T-stationary point of $\mathcal{R}$, which is the limit of a sequence of nondegenerate Karush–Kuhn–Tucker points of $\mathcal{S}$, is also nondegenerate. Instead, we intend to examine its type if assuming nondegeneracy. Next Lemma 3 provides some valuable insights into the relations between active index sets while doing so.

Lemma 3 (Active index sets) *Suppose a sequence of Karush–Kuhn–Tucker points $\left(x^t, y^t\right)$ of $\mathcal{S}(t)$ converges to $(\bar{x}, \bar{y})$ for $t \to 0$. Moreover, let $(\bar{x}, \bar{y})$ be a nondegenerate T-stationary point of $\mathcal{R}$. Then, for all sufficiently small t it holds:*

(a) $Q_0(\bar{x}) = Q_0(x^t)$,
(b) $\mathcal{E}(\bar{y}) = \mathcal{E}(y^t)$,
(c) $a_{00}(\bar{x}, \bar{y}) \subset \mathcal{H}(x^t, y^t)$,
(d) $\mathcal{N}(y^t) \subset a_{10}(\bar{x}, \bar{y}) \subset \mathcal{N}(y^t) \cup \mathcal{H}(x^t, y^t)$.

Proof a) We start by proving $Q_0(\bar{x}) = Q_0(x^t)$. Due to continuity arguments, we have $Q_0(x^t) \subset Q_0(\bar{x})$ for all sufficiently small t. Let us now assume that there exists $\bar{i} \in Q_0(\bar{x}) \setminus Q_0(x^t)$ along a subsequence. Hence, for the corresponding multipliers it holds $\mu_{\bar{i}}^t = 0$. NDT1 allows us to apply Remark 1, and we thus have $\bar{\mu}_{\bar{i}} = \lim_{t \to \infty} \mu_{\bar{i}}^t = 0$, a contradiction to NDT2. Consequently, $Q_0(\bar{x}) = Q_0(x^t)$ holds for all sufficiently small t.

b) Next, we prove $\mathcal{E}(\bar{y}) = \mathcal{E}(y^t)$. Again, continuity arguments provide $\mathcal{E}(y^t) \subset \mathcal{E}(\bar{y})$ for all sufficiently small t. Similar to the first part of the proof, we now assume there exists $\bar{i} \in \mathcal{E}(\bar{y}) \setminus \mathcal{E}(y^t)$ along a subsequence. As we have seen in Lemma 1, T-stationarity of $(\bar{x}, \bar{y})$ implies in particular $c_{\bar{i}} = -\bar{\mu}_{2,\bar{i}} + \bar{\mu}_{3,\bar{i}}$. Moreover, NDT1 and Remark 1 provide $\lim_{t \to 0} \mu_3^t = \bar{\mu}_3$. Since $\bar{i} \notin \mathcal{N}(y^t)$, we distinguish the following cases:

(i) $\bar{i} \in \mathcal{H}^{\geq}(x^t, y^t) \setminus \mathcal{E}(y^t)$. Karush–Kuhn–Tucker conditions for (x^t, y^t) imply $c_{\bar{i}} = \mu_{3,\bar{i}}^t + \eta_{\bar{i}}^{\geq,t} x_{\bar{i}}$, cf. (12). It follows $-\bar{\mu}_{2,\bar{i}} + \bar{\mu}_{3,\bar{i}} = \mu_{3,\bar{i}}^t + \eta_{\bar{i}}^{\geq,t} x_{\bar{i}}$. By taking the limit, we can cancel out $\bar{\mu}_{3,\bar{i}}$ and $\mu_{3,\bar{i}}^t$. This leads to a contradiction because the left-hand side of the equation is strictly negative due to NDT2 and the right-hand side is nonnegative since $\eta_{\bar{i}}^{\geq,t}$ is nonnegative and $x_{\bar{i}}$ is positive.

(ii) $\bar{i} \in \mathcal{H}^{\leq}(x^t, y^t) \setminus \mathcal{E}(y^t)$. By using (12), we get $c_{\bar{i}} = \mu_{3,\bar{i}}^t - \eta_{\bar{i}}^{\leq,t} x_{\bar{i}}$. This leads to a contradiction just as in the previous case.

(iii) $\bar{i} \in \mathcal{O}(x^t, y^t)$. Analogously, we obtain $c_{\bar{i}} = \mu_{3,\bar{i}}^t$ from (12). It follows $-\bar{\mu}_{2,\bar{i}} + \bar{\mu}_{3,\bar{i}} = \mu_{3,\bar{i}}^t$. Taking the limits leads to $\bar{\mu}_{2,\bar{i}} = 0$, a contradiction with NDT2.

Altogether, $\mathcal{E}(\bar{y}) \setminus \mathcal{E}(y^t) = \emptyset$ for all sufficiently small t, and the assertion follows.

c) Clearly, $a_{00}(\bar{x}, \bar{y}) \cap \mathcal{E}(y^t) = \emptyset$ for sufficiently small t.

Let us assume there exists an $\bar{i} \in a_{00}(\bar{x}, \bar{y}) \cap \mathcal{N}(y^t)$. In view of (12), we have $c_{\bar{i}} = \mu_3^t + \nu_{\bar{i}}^t$. Due to the T-stationarity of $(\bar{x}, \bar{y})$, the $(n+i)$-th row of (5) reads as

$$
c_i = \begin{cases}
-\bar{\mu}_{2,i} + \bar{\mu}_3, & \text{for } i \in \mathcal{E}(\bar{y}), \\
\bar{\sigma}_{2,i} + \bar{\mu}_3, & \text{for } i \in a_{10}(\bar{x}, \bar{y}), \\
\bar{\varrho}_{2,i} + \bar{\mu}_3, & \text{for } i \in a_{00}(\bar{x}, \bar{y}), \\
\bar{\mu}_3, & \text{else.}
\end{cases}
\tag{13}
$$

This provides $c_{\bar{i}} = \bar{\varrho}_{2,\bar{i}} + \bar{\mu}_3$. According to Remark 1, we have $\lim_{t \to 0} \mu_3^t = \bar{\mu}_3$. Consequently, it must hold $\lim_{t \to 0} \nu_{\bar{i}}^t = \bar{\varrho}_{2,\bar{i}}$. This, however, cannot be true since $\nu_{\bar{i}}^t \geq 0$, while $\bar{\varrho}_{2,\bar{i}} < 0$ due to NDT3 from the nondegeneracy of $(\bar{x}, \bar{y})$, a contradiction. Let us assume now that there exists an $\bar{i} \in a_{00}(\bar{x}, \bar{y}) \cap \mathcal{O}(x^t, y^t)$. Analogously, we get $\bar{\varrho}_{2,\bar{i}} = 0$, again a contradiction to NDT3. Overall, we get the assertion.

d) Clearly, $a_{01}(\bar{x}, \bar{y}) \cap \mathcal{N}(y^t) = \emptyset$ for sufficiently small t. From c) we also know that $a_{00}(\bar{x}, \bar{y}) \cap \mathcal{N}(y^t) = \emptyset$. Altogether, the first inclusion of the assertion follows immediately. Further, it also holds $a_{10}(\bar{x}, \bar{y}) \cap \mathcal{E}(y^t) = \emptyset$ for sufficiently small t. Let us assume there exists an $\bar{i} \in a_{10}(\bar{x}, \bar{y}) \cap \mathcal{O}(x^t, y^t)$. Due to (12), we have $c_{\bar{i}} = \mu_3^t$. In view of Lemma 1c), there exists an index $\tilde{i} \in a_{01}(\bar{x}, \bar{y}) \backslash \mathcal{E}(\bar{y})$. Thus, T-stationarity of $(\bar{x}, \bar{y})$ implies via (13) that $c_{\tilde{i}} = \bar{\mu}_3$. By taking the limit and Remark 1, we obtain $c_{\bar{i}} = c_{\tilde{i}}$, but $\bar{i} \neq \tilde{i}$, a contradiction to the choice of c. $\qquad \square$

Theorem 4 highlights the convergence properties of the Scholtes-type regularization method. Its proof can be found in the Appendix below.

Theorem 4 (Convergence from $\mathcal{S}$ to $\mathcal{R}$ again) *Suppose that a sequence of nondegenerate Karush–Kuhn–Tucker points (x^t, y^t) of $\mathcal{S}$ with quadratic index m converges to $(\bar{x}, \bar{y})$ for $t \to 0$. If $(\bar{x}, \bar{y})$ is a nondegenerate T-stationary point of $\mathcal{R}$, then we have for its T-index:*

$$\max \left\{ m - \left| \{ i \in a_{01}(\bar{x}, \bar{y}) \,|\, \bar{\sigma}_{1,i} = 0 \} \right|, 0 \right\} \leq TI \leq m.$$

If additionally NDT6 holds at $(\bar{x}, \bar{y})$, then the indices coincide, i.e. $TI = m$.

Let us illustrate the necessity of NDT6 for the validity of Theorem 4.

Example 2 (*Necessity of NDT6*) We consider the following Scholtes-type regularization $\mathcal{S}$ with $n = 2$, $s = 1$ and $0 < c_1 < c_2$:

$$\mathcal{S}: \quad \min_{x,y} (1 + x_1)^2 + (3 - 2x_2)^2 + c_1 y_1 + c_2 y_2$$
$$\text{s.t.} x_1 + x_2 - 1 \geq 0,$$
$$y_1 + y_2 \geq 1, \quad -t \leq x_i y_i \leq t, \quad 0 \leq y_i \leq 1 + \varepsilon, \quad i = 1, 2,$$

as well as the point $(x^t, y^t) = (0, 1, 1, 0)$. We claim that this point is a nondegenerate Karush–Kuhn–Tucker point. Indeed, it holds:

$$\begin{pmatrix} 2 \\ 2 \\ c_1 \\ c_2 \end{pmatrix} = \mu_{1,1}^t \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mu_3^t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \nu_2^t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

with the positive multipliers $\mu_{1,1}^t = 2$, $\mu_3^t = c_1$, $\nu_2^t = c_2 - c_1$. Obviously, LICQ and strict complementarity, i.e. ND1 and ND2, respectively, are fulfilled. We show that $D^2 L^{\mathcal{S}}(x^t, y^t) \restriction_{T_{(x^t,y^t)}^{\mathcal{S}}}$ is nonsingular and calculate the number of its negative

eigenvalues. The tangent space is $\mathcal{T}^{\mathcal{S}}_{(x^t, y^t)} = \left\{ \xi \in \mathbb{R}^4 \mid \xi_1 + \xi_2 = 0, \xi_3 = \xi_4 = 0 \right\}$. For the Hessian of the corresponding Lagrange function we have:

$$D^2 L^{\mathcal{S}}(x^t, y^t) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, for $\xi \in \mathcal{T}^{\mathcal{S}}_{(x^t, y^t)}$ it holds:

$$\xi^T D^2 L^{\mathcal{S}}(x^t, y^t)\xi = 2\xi_1^2 - 4\xi_2^2 = -2\xi_1^2.$$

Hence, ND3 is also fulfilled, the Karush–Kuhn–Tucker point (x^t, y^t) is nondegenerate and its quadratic index equals one, i.e. $m = 1$ in Theorem 4. The limiting point is $(\bar{x}, \bar{y}) = (0, 1, 1, 0)$. This point is T-stationary for the corresponding regularized continuous reformulation $\mathcal{R}$ according to Theorem 3, since MPOC-LICQ is fulfilled. Indeed, we have:

$$\begin{pmatrix} 2 \\ 2 \\ c_1 \\ c_2 \end{pmatrix} = \bar{\mu}_{1,1} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \bar{\mu}_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \bar{\sigma}_{1,1} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \bar{\sigma}_{2,2} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

with the unique multipliers $\bar{\mu}_{1,1} = 2, \bar{\mu}_3 = c_1, \bar{\sigma}_{1,1} = 0, \bar{\sigma}_{2,2} = c_2 - c_1$. It is easy to see that this point is nondegenerate with vanishing T-index, i.e. $TI = 0$, since $a_{00}(\bar{x}, \bar{y}) = \emptyset$ and $\mathcal{T}^{\mathcal{R}}_{(\bar{x}, \bar{y})} = \{0\}$. Note that additionally $\left\{ i \in a_{01}(\bar{x}, \bar{y}) \mid \bar{\sigma}_{1,i} = 0 \right\} = \{1\}$. Although all assumptions of Theorem 4 are fulfilled, we have here:

$$TI = \max \left\{ m - \left| \left\{ i \in a_{01}(\bar{x}, \bar{y}) \mid \bar{\sigma}_{1,i} = 0 \right\} \right|, 0 \right\}.$$

With other words, the saddle points of the Scholtes-type regularization $\mathcal{S}$ approximate a minimizer of the regularized continuous reformulation $\mathcal{R}$. The reason is that the σ-multipliers corresponding to zero x- and nonzero y-variables vanish. The lower bound given in Theorem 4 is attained. $\qquad\qquad\square$

Next, we point out that the assumption NDT6 is not restrictive at all.

Remark 2 (*Genericity for NDT6*) Let us briefly sketch why condition NDT6 must be generically fulfilled at the T-stationary points of $\mathcal{R}$. First, we note that all T-stationary points of $\mathcal{R}$ are generically nondegenerate, see [9]. Now, let us count the losses of freedom induced by the definition of a T-stationary point. For feasibility we have $|P|$ equality constraints, $|Q_0|$ active inequality constraints, $|\mathcal{E}|$ bounding constraints on the y-variables, potentially one summation constraint, and $|a_{01}| + |a_{10}| + 2 |a_{00}|$ orthogonality type constraints. Additional losses of freedom come from the T-stationarity condition. They amount to $2n - |P| - |Q_0| - |\mathcal{E}| - 1 - |a_{01}| - |a_{10}| - 2 |a_{00}|$ if the summation constraint is active, and to $2n - |P| - |Q_0| - |\mathcal{E}| - |a_{01}| - |a_{10}| - 2 |a_{00}|$

otherwise. In both cases, the losses of freedom are equal to the number of variables $2n$. The violation of NDT6 would produce an additional loss of freedom, which would imply that the total available degrees of freedom $2n$ are exceeded. By virtue of the structured jet transversality theorem from [12], this cannot happen generically. $\qquad\square$

Let us examine the set of multipliers from Theorem 4 in terms of CCOP.

Lemma 4 *Let $\bar{x}$ be a nondegenerate M-stationary point of CCOP. Then, for any $\bar{y}$ such that $(\bar{x}, \bar{y})$ is a T-stationary point of $\mathcal{R}$ we have*

$$\{i \in I_0 (\bar{x}) \mid \bar{\gamma}_i = 0\} = \{i \in a_{01} (\bar{x}, \bar{y}) \mid \bar{\sigma}_i = 0\}.$$

Proof We refer to the proof of Theorem 3.7 from [9]. There, it was shown how any T-stationary point $(\bar{x}, \bar{y})$ of $\mathcal{R}$ can be constructed by means of a nondegenerate M-stationary point $\bar{x}$ of CCOP. Specifically, the corresponding multipliers were set as

$$\bar{\sigma}_{1,i} = \bar{\gamma}_i \text{ for all } i \in a_{01}(\bar{x}, \bar{y}).$$

We conclude

$$\{i \in a_{01} (\bar{x}, \bar{y}) \mid \bar{\gamma}_i = 0\} = \{i \in a_{01} (\bar{x}, \bar{y}) \mid \bar{\sigma}_i = 0\}.$$

Let us assume that $\|\bar{x}\|_0 < s$ additionally holds. Hence, in virtue of NDM3 we have

$$\{i \in I_0 (\bar{x}) \mid \bar{\gamma}_i = 0\} = \emptyset.$$

By recalling $a_{01} (\bar{x}, \bar{y}) \subset I_0 (\bar{x})$, the assertion follows.
Suppose now $\|\bar{x}\|_0 = s$ instead. Due to Lemma 1b), we have

$$|I_0 (\bar{x})| = n - s = |a_{01} (\bar{x}, \bar{y})|.$$

Since $I_0 (\bar{x}) = a_{00} (\bar{x}, \bar{y}) \cup a_{01} (\bar{x}, \bar{y})$, we conclude $a_{00} (\bar{x}, \bar{y}) = \emptyset$. Thus, $I_0 (\bar{x}) = a_{01} (\bar{x}, \bar{y})$ and the assertion follows. $\qquad\square$

In view of Theorem 1 we get the following convergence properties of the proposed Scholtes-type regularization with respect to the underlying CCOP.

Corollary 1 (Convergence from $\mathcal{S}$ to CCOP)

(a) *Suppose that a sequence of Karush–Kuhn–Tucker points (x^t, y^t) of $\mathcal{S}$ converges to $(\bar{x}, \bar{y})$ for $t \to 0$. If CC-LICQ holds at $\bar{x}$, then $\bar{x}$ is an M-stationary point of CCOP.*
(b) *Suppose that a sequence of nondegenerate Karush–Kuhn–Tucker points (x^t, y^t) of $\mathcal{S}$ with quadratic index m converges to $(\bar{x}, \bar{y})$ for $t \to 0$. If $\bar{x}$ is a nondegenerate M-stationary point of CCOP, then we have for its M-index MI:*

$$\max \{m - |\{i \in I_0 (\bar{x}) \mid \bar{\gamma}_i = 0\}|, 0\} \leq MI \leq m.$$

If additionally NDM5 holds, then the indices coincide, i.e. $MI = m$.

Proof a) Due to continuity arguments, $(\bar{x}, \bar{y})$ is feasible for $\mathcal{R}$. Let us show that the latter implies feasibility of $\bar{x}$ for CCOP. For this purpose we assume instead $\|\bar{x}\|_0 > s$. Consequently, we have

$$n - s > |I_0(\bar{x})| \geq |a_{01}(\bar{x}, \bar{y})|.$$

Thus, it holds for $(\bar{x}, \bar{y})$:

$$\sum_{i=1}^{n} \bar{y}_i = \sum_{i \in a_{01}(\bar{x}, \bar{y})} \bar{y}_i \leq (n - s - 1)(1 + \varepsilon) < n - s,$$

a contradiction to its feasibility. Overall, $\bar{x}$ has to be feasible for CCOP and, thus, we can apply Proposition 3.2a) from [9]. The latter states that if $\bar{x}$ is feasible for CCOP and satisfies CC-LICQ, then MPOC-LICQ holds at any $(\bar{x}, y)$ that is feasible for $\mathcal{R}$. Hence, in view of Theorem 3, $(\bar{x}, \bar{y})$ is a T-stationary point of $\mathcal{R}$. Therefore, $\bar{x}$ is M-stationary, due to Theorem 1b).

b) We deduce as above that $(\bar{x}, \bar{y})$ is a T-stationary point of $\mathcal{R}$. Using Theorem 1, we have that $(\bar{x}, \bar{y})$ is nondegenerate fulfilling NDT5. According to Theorem 4, for its T-index TI it holds

$$\max \left\{ m - \left| \left\{ i \in a_{01}(\bar{x}, \bar{y}) \,\middle|\, \bar{\sigma}_{1,i} = 0 \right\} \right|, 0 \right\} \leq TI \leq m.$$

However, we again use Theorem 1 to conclude $TI = MI$. In view of Lemma 4, the assertion follows. □

Let us briefly comment on condition NDM5. It ensures M-stationary points to have the same index as the approximating Karush–Kuhn–Tucker points of $\mathcal{S}$.

Remark 3 *(On condition NDM5)* It follows from Lemma 4 that for a nondegenerate M-stationary point $\bar{x}$ of CCOP the following statements are equivalent:

a) NDM5 holds at $\bar{x}$,
b) NDT6 holds at a T-stationary point $(\bar{x}, \bar{y})$ of $\mathcal{R}$,
c) NDT6 holds at all T-stationary points $(\bar{x}, \bar{y})$ of $\mathcal{R}$.

Further, due to Theorem 1a), all M-stationary points are induced by at least one T-stationary point. Theorem 1b), Remark 2, and the equivalence above provide that all M-stationary points generically fulfill NDM5. As a consequence, we conclude that generically the bounds given in Corollary 1 are tight, i.e. $MI = m$. The latter holds in particular for $\|\bar{x}\|_0 < s$ regardless of NDM5, since NDM3 suffices. □

Now, we prove that the Scholtes-type regularization method is well-defined. For the proof see again the Appendix below.

Theorem 5 (*Well-posedness of $\mathcal{S}$ from $\mathcal{R}$*) *Let $(\bar{x}, \bar{y})$ be a nondegenerate T-stationary point of $\mathcal{R}$ with T-index m, additionally, fulfilling NDT6. Then, for all sufficiently small t there exists a nondegenerate Karush–Kuhn–Tucker point (x^t, y^t) of $\mathcal{S}$ within a neighborhood of $(\bar{x}, \bar{y})$, which has the same quadratic index m. Moreover, for any fixed t sufficiently small, such (x^t, y^t) is the unique Karush–Kuhn–Tucker point of $\mathcal{S}$ in a sufficiently small neighborhood of $(\bar{x}, \bar{y})$.*

Again, we state the result analogous to Theorem 5 in terms of CCOP.

Corollary 2 (*Well-posedness of $\mathcal{S}$ from CCOP*) *Let $\bar{x}$ be a nondegenerate M-stationary point of CCOP with M-index m, additionally, fulfilling NDM5. Then, for all sufficiently small t there exists a nondegenerate Karush–Kuhn–Tucker point (x^t, y^t) of S with x^t being within a neighborhood of $\bar{x}$, which has the same quadratic index m.*

Proof Due to Theorem 1a), there exists at least one nondegenerate T-stationary point $(\bar{x}, \bar{y})$ of $\mathcal{R}$. Moreover, Lemma 4 provides that it also fulfills NDT6. Thus, the assertion follows straightforward in view of Theorem 5. □

Let us compare our results with those for the initially proposed continuous reformulation (1) and the Scholtes-type relaxation (2) from [5] and [6], respectively. There, the concept of S-stationarity for (1) becomes crucial.

Definition 13 (*S-stationary, [5]*) A feasible point $(\bar{x}, \bar{y})$ of (1) is called S-stationary if there exist multipliers

$$\bar{\lambda}_p, \ p \in P, \quad \bar{\mu}_q, \ q \in Q_0(\bar{x}), \quad \bar{\gamma}_i, \ i \in a_{01}(\bar{x}, \bar{y}),$$

such that the following conditions hold:

$$\nabla f(\bar{x}) = \sum_{p \in P} \bar{\lambda}_p \nabla h_p(\bar{x}) + \sum_{q \in Q_0(\bar{x})} \bar{\mu}_q \nabla g_q(\bar{x}) + \sum_{i \in a_{01}(\bar{x}, \bar{y})} \bar{\gamma}_i e_i,$$
$$\bar{\mu}_q \geq 0, \ q \in Q_0(\bar{x}).$$

Example 3 We consider the following CCOP with $n = 3$ and $s = 1$:

$$\min_{x} \ (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 \quad \text{s.t.} \quad \|x\|_0 \leq 1.$$

It has minimizers at $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ as well as a saddle point at $(0, 0, 0)$. It is straightforward to check that all these points are nondegenerate M-stationary points, which additionally fulfill NDM5. For its continuous reformulation (1) we have

$$\min_{x, y} \ (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 \quad \text{s.t.} \quad y_1 + y_2 + y_3 \geq 2,$$
$$x_i y_i = 0, \quad 0 \leq y_i \leq 1, \quad i = 1, 2, 3.$$

We get as its S-stationary points:

$$(1, 0, 0, 0, 1, 1), \quad (0, 1, 0, 1, 0, 1), \quad (0, 0, 1, 1, 1, 0),$$

and

$$(0, 0, 0, y_1, y_2, y_3) \text{ with } y_1 + y_2 + y_3 \geq 2, 0 < y_i \leq 1, i = 1, 2, 3.$$

Hence, we have a continuum of saddle points. Moreover, it was shown in [4] that all S-stationary points of reformulation (1) are degenerate T-stationary points, i.e. violating at least one of the conditions NDT1-NDT4. Further, we turn our attention to the Scholtes-type regularization (2)

$$\min_{x,y} (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 \quad \text{s.t. } y_1 + y_2 + y_3 \geq 2,$$
$$-t \leq x_i y_i \leq t, \ 0 \leq y_i \leq 1, \ i = 1, 2, 3.$$

For t sufficiently small its Karush–Kuhn–Tucker points include

$$z^{1,t} = (\tilde{x}^t, t, \hat{x}^t, \tilde{y}^t, 1, \hat{y}^t), \qquad z^{2,t} = (\tilde{x}^t, \hat{x}^t, t, \tilde{y}^t, \hat{y}^t, 1),$$
$$z^{3,t} = (t, \tilde{x}^t, \hat{x}^t, 1, \tilde{y}^t, \hat{y}^t), \qquad z^{4,t} = (\hat{x}^t, \tilde{x}^t, t, \hat{y}^t, \tilde{y}^t, 1).$$
$$z^{5,t} = (t, \hat{x}^t, \tilde{x}^t, 1, \hat{y}^t, \tilde{y}^t), \qquad z^{6,t} = (\hat{x}^t, t, \tilde{x}^t, \hat{y}^t, 1, \tilde{y}^t),$$
$$z^{7,t} = (t, 2t, 2t, 1, 1/2, 1/2), \qquad z^{8,t} = (2t, t, 2t, 1/2, 1, 1/2),$$
$$z^{9,t} = (2t, 2t, t, 1/2, 1/2, 1), \quad z^{10,t} = (3t/2, 3t/2, 3t/2, 2/3, 2/3, 2/3),$$

where

$$\tilde{x}^t = \frac{t + 1 + \sqrt{1 - 2t - 3t^2}}{2}, \quad \hat{x}^t = \frac{\tilde{x}^t t}{\tilde{x}^t - t}, \quad \tilde{y}^t = \frac{t}{\tilde{x}^t}, \quad \hat{y}^t = \frac{t}{\hat{x}^t}.$$

For every minimizer of the underlying CCOP there exist at least two sequences of Karush–Kuhn–Tucker points of the Scholtes-type regularization (2), which approximate the corresponding S-stationary points of the continuous reformulation (1), i.e.

$$z^{1,t}, z^{2,t} \rightarrow (1, 0, 0, 0, 1, 1),$$
$$z^{3,t}, z^{4,t} \rightarrow (0, 1, 0, 1, 0, 1),$$
$$z^{5,t}, z^{6,t} \rightarrow (0, 0, 1, 1, 1, 0).$$

For the saddle point of CCOP there exist at least four sequences of Karush–Kuhn–Tucker points of the Scholtes-type regularization (2), which approximate the corresponding S-stationary points of the continuous reformulation (1), i.e.

$$z^{7,t} \rightarrow (0, 0, 0, 1, 1/2, 1/2), \qquad z^{8,t} \rightarrow (0, 0, 0, 1/2, 1, 1/2),$$
$$z^{9,t} \rightarrow (0, 0, 0, 1/2, 1/2, 1), \qquad z^{10,t} \rightarrow (0, 0, 0, 2/3, 2/3, 2/3).$$

We list M-, S-stationary, and Karush–Kuhn–Tucker points in Table 1.

Let us apply our results to the given CCOP. In view of Theorem 1a), we know that the regularized continuous reformulation $\mathcal{R}$ has in total five T-stationary points, all of them being nondegenerate. Three of them are minimizers and two of them are

Table 1 M-, S-stationary, and Karush–Kuhn–Tucker points

M-stationary for CCOP	S-stationary for (1)	Karush–Kuhn–Tucker points for (2)
(1, 0, 0)	(1, 0, 0, 0, 1, 1)	$z^{1,t}$, $z^{2,t}$
(0, 1, 0)	(0, 1, 0, 1, 0, 1)	$z^{3,t}$, $z^{4,t}$
(0, 0, 1)	(0, 0, 1, 1, 1, 0)	$z^{5,t}$, $z^{6,t}$
(0, 0, 0)	(0, 0, 0, 1, 1/2, 1/2)	$z^{7,t}$
	(0, 0, 0, 1/2, 1, 1/2)	$z^{8,t}$
	(0, 0, 0, 1/2, 1/2, 1)	$z^{9,t}$
	(0, 0, 0, 2/3, 2/3, 2/3)	$z^{10,t}$

saddle points of $\mathcal{R}$. Also, we know from Remark 3 that all of them fulfill NDT6. Due to Theorem 3, any convergent sequence of Karush–Kuhn–Tucker points of the Scholtes-type regularization $\mathcal{S}$ converges to one of these T-stationary points of $\mathcal{R}$. We apply Theorem 5 to conclude that for any fixed t sufficiently small there are exactly five Karush–Kuhn–Tucker points of $\mathcal{S}$. All of them are nondegenerate. Three of them are minimizers and two of them are saddle points of $\mathcal{S}$. Overall, not only the global structure of $\mathcal{R}$ is more accessible than that of (1), but also the global structure of $\mathcal{S}$ is more accessible than that of (2). This shows the advantage of our approach in comparison to the existing literature, at least for the presented example. □

Conclusions

In [9], the number of saddle points for the regularized continuous reformulation of CCOP has been estimated. Namely, each saddle point of CCOP generates exponentially many saddle points of $\mathcal{R}$, all of them having the same index. It has been concluded there that the introduction of auxiliary y-variables shifts the complexity of dealing with the cardinality constraint in CCOP into the appearance of multiple saddle points for its continuous reformulation. From our extended convergence analysis of the Scholtes-type regularization it follows that the number of its saddle points also grows exponentially as compared to that of CCOP. We emphasize that this issue is at the core of numerical difficulties if solving CCOP up to global optimality by means of the Scholtes-type regularization method. To the best of our knowledge this is the first paper studying convergence properties of the Scholtes-type regularization method in the vicinity of saddle points, rather than of minimizers. The ideas from our analysis can be potentially applied not only for classes of nonsmooth optimization problems, such as MPCC, MPVC, MPSC, and MPOC, but also for other regularization schemes known from the literature.

Declarations

Appendix

Proof of Theorem 4 The proof will be divided into 4 major steps.

Step 1a. We rewrite the tangent space corresponding to the Karush–Kuhn–Tucker point (x^t, y^t). For that, we use Lemma 2a) which provides that the summation constraint is active:

$$
\mathcal{T}^{\mathcal{S}}_{(x^t,y^t)} = \left\{ \xi \in \mathbb{R}^{2n} \left| \begin{array}{l} \left(Dh_p(x^t),0\right)\xi = 0,\, p \in P,\, \left(Dg_q(x^t),0\right)\xi = 0,\, q \in Q_0(x^t), \\ (0,e)\,\xi = 0,\, (0,e_i)\,\xi = 0,\, i \in \mathcal{E}(y^t), \\ \left(y_i^t e_i, x_i^t e_i\right)\xi = 0,\, i \in \mathcal{H}\left(x^t,y^t\right) \cap a_{00}\left(\bar{x},\bar{y}\right), \\ \left(y_i^t e_i, x_i^t e_i\right)\xi = 0,\, i \in \mathcal{H}\left(x^t,y^t\right) \cap a_{01}\left(\bar{x},\bar{y}\right), \\ \left(y_i^t e_i, x_i^t e_i\right)\xi = 0,\, i \in \mathcal{H}\left(x^t,y^t\right) \cap a_{10}\left(\bar{x},\bar{y}\right), \\ (0,e_i)\,\xi = 0,\, i \in \mathcal{N}\left(y^t\right) \end{array} \right. \right\}.
$$

In total there are, due to LICQ,

$$
\begin{aligned}
\alpha_t^{\mathcal{S}} = {}& |P| + \left|Q_0\left(x^t\right)\right| + 1 + \left|\mathcal{E}\left(y^t\right)\right| + \left|\mathcal{H}\left(x^t,y^t\right) \cap a_{00}\left(\bar{x},\bar{y}\right)\right| \\
& + \left|\mathcal{H}\left(x^t,y^t\right) \cap a_{01}\left(\bar{x},\bar{y}\right)\right| + \left|\mathcal{H}\left(x^t,y^t\right) \cap a_{10}\left(\bar{x},\bar{y}\right)\right| + \left|\mathcal{N}\left(y^t\right)\right|
\end{aligned}
$$

linearly independent vectors involved. We use Lemma 3a) and 3b) to substitute $\left|Q_0\left(x^t\right)\right|$ with $|Q_0(\bar{x})|$ and $\left|\mathcal{E}\left(y^t\right)\right|$ with $|\mathcal{E}(\bar{y})|$, respectively. The latter set has cardinality of $n - s - 1$ due to Lemma 1c). Additionally, we use Lemma 3c) and 3d) to conclude:

$$
\begin{aligned}
\alpha_t^{\mathcal{S}} = {}& |P| + |Q_0(\bar{x})| + 1 + n - s - 1 + |a_{00}(\bar{x},\bar{y})| \\
& + \left|\mathcal{H}\left(x^t,y^t\right) \cap a_{01}(\bar{x},\bar{y})\right| + |a_{10}(\bar{x},\bar{y})|.
\end{aligned}
$$

Finally, $|a_{00}(\bar{x},\bar{y})| + |a_{10}(\bar{x},\bar{y})| = s$, cf. Lemma 1b). Thus, we have:

$$
\alpha_t^{\mathcal{S}} = |P| + |Q_0(\bar{x})| + \left|\mathcal{H}\left(x^t,y^t\right) \cap a_{01}(\bar{x},\bar{y})\right| + n.
$$

Step 1b. We examine the tangent space corresponding to the T-stationary point $(\bar{x}, \bar{y})$. For this purpose, we consider the following vectors from its definition:

$$\begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in \mathcal{E}(\bar{y}), \quad \begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in a_{10}(\bar{x}, \bar{y}) \cup a_{00}(\bar{x}, \bar{y}), \quad \begin{pmatrix} 0 \\ e \end{pmatrix}.$$

The latter vector is involved due to Lemma 1a). The number of these vectors is due to Lemma 1c) equal to $(n - s - 1) + s + 1 = n$. Moreover, they are linearly independent due to MPOC-LICQ. Hence, we can write the respective tangent space as

$$\mathcal{T}^{\mathcal{R}}_{(\bar{x}, \bar{y})} = \left\{ \xi \in \mathbb{R}^{2n} \ \middle| \ \begin{array}{l} (Dh_p(\bar{x}), 0)\, \xi = 0, \ p \in P, \ (Dg_q(\bar{x}), 0)\, \xi = 0, \ q \in Q_0(\bar{x}), \\ (e_i, 0)\, \xi = 0, \ i \in a_{00}(\bar{x}, \bar{y}) \cup a_{01}(\bar{x}, \bar{y}), \ \xi_{n+1} = \ldots = \xi_{2n} = 0 \end{array} \right\}.$$

In total there are, due to MPOC-LICQ,

$$\alpha^{\mathcal{R}} = |P| + |Q_0(\bar{x})| + |a_{00}(\bar{x}, \bar{y})| + |a_{01}(\bar{x}, \bar{y})| + n$$

linearly independent vectors involved.

Step 2. Let $\mathcal{T} \subset \mathbb{R}^{2n}$ be a linear subspace. We denote the number of negative eigenvalues of $D^2 L^{\mathcal{S}}(x^t, y^t) \upharpoonright_{\mathcal{T}}$ by $QI^{\mathcal{S}}_{t, \mathcal{T}}$. Analogously, $QI^{\mathcal{R}}_{t, \mathcal{T}}$ stands for the number of negative eigenvalues of $D^2 L^{\mathcal{R}}(x^t, y^t) \upharpoonright_{\mathcal{T}}$ and $\overline{QI}^{\mathcal{R}}_{\mathcal{T}}$ stands for the number of negative eigenvalues of $D^2 L^{\mathcal{R}}(\bar{x}, \bar{y}) \upharpoonright_{\mathcal{T}}$. We have the following relation between the involved Hessians of the Lagrange functions by denoting $E(i) = e_i e_{n+i}^T + e_{n+i} e_i^T, \ i = 1, \ldots, n$:

$$D^2 L^{\mathcal{S}}(x^t, y^t) = D^2 L^{\mathcal{R}}(x^t, y^t) - \sum_{i \in \mathcal{H}^{\geq}(x^t, y^t)} \eta_i^{\geq, t} E(i) + \sum_{i \in \mathcal{H}^{\leq}(x^t, y^t)} \eta_i^{\leq, t} E(i). \tag{14}$$

Step 2a. It holds for t sufficiently small:

$$QI^{\mathcal{S}}_{t, \mathcal{T}^{\mathcal{R}}_{(\bar{x}, \bar{y})}} = \overline{QI}^{\mathcal{R}}_{\mathcal{T}^{\mathcal{R}}_{(\bar{x}, \bar{y})}}.$$

Indeed, by using (14), we derive for any $\xi \in \mathcal{T}^{\mathcal{R}}_{(\bar{x}, \bar{y})}$:

$$\xi^T D^2 L^{\mathcal{S}}(x^t, y^t)\, \xi = \xi^T D^2 L^{\mathcal{R}}(x^t, y^t)\, \xi - \sum_{i \in \mathcal{H}^{\geq}(x^t, y^t)} 2\eta_i^{\geq, t} \xi_i \xi_{n+i}$$

$$+ \sum_{i \in \mathcal{H}^{\leq}(x^t, y^t)} 2\eta_i^{\leq, t} \xi_i \xi_{n+i} = \xi^T D^2 L^{\mathcal{R}}(x^t, y^t)\, \xi, \tag{15}$$

since $\xi_{n+1} = \ldots = \xi_{2n} = 0$ as seen in Step 1b. Hence, we get $QI^{\mathcal{S}}_{t,\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}} = QI^{\mathcal{R}}_{t,\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}}$. Due to NDT4, continuity arguments provide $QI^{\mathcal{R}}_{t,\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}} = \overline{QI^{\mathcal{R}}_{\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}}}$.

Step 2b. We claim that the numbers of positive and negative eigenvalues of $D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \restriction_{\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}}$ and of $D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \restriction_{\mathcal{T}'}$, respectively, coincide, where

$$\mathcal{T}' = \left\{ \xi \in \mathbb{R}^{2n} \left| \begin{array}{l} \left(Dh_p(x^t), 0\right)\xi = 0, p \in P, \left(Dg_q(x^t), 0\right)\xi = 0, q \in Q_0(\bar{x}), \\ (0, e)\xi = 0, (0, e_i)\xi = 0, i \in \mathcal{E}(\bar{y}), \\ (e_i, 0)\xi = 0, i \in a_{00}(\bar{x}, \bar{y}), (0, e_i)\xi = 0, i \in a_{00}(\bar{x}, \bar{y}), \\ (y^t_i e_i, x^t_i e_i)\xi = 0, i \in a_{01}(\bar{x}, \bar{y}) \cup a_{10}(\bar{x}, \bar{y}) \end{array} \right. \right\}.$$

Let $\{\lambda^+_1, \ldots, \lambda^+_{k^+}\}$ be the positive eigenvalues of $D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \restriction_{\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}}$ with corresponding eigenvectors $\{\xi^+_1, \ldots, \xi^+_{k^+}\}$. Hence, for all $k = 1, \ldots, k^+$:

$$\xi^{+T}_k D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \xi^+_k > 0.$$

We rewrite the tangent space

$$\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})} = \left\{ \xi \in \mathbb{R}^{2n} \left| \begin{array}{l} \left(Dh_p(\bar{x}), 0\right)\xi = 0, p \in P, \left(Dg_q(\bar{x}), 0\right)\xi = 0, q \in Q_0(\bar{x}), \\ (0, e)\xi = 0, (0, e_i)\xi = 0, i \in \mathcal{E}(\bar{y}), \\ (e_i, 0)\xi = 0, i \in a_{00}(\bar{x}, \bar{y}), (0, e_i)\xi = 0, i \in a_{00}(\bar{x}, \bar{y}), \\ (\bar{y}_i e_i, \bar{x}_i e_i)\xi = 0, i \in a_{01}(\bar{x}, \bar{y}) \cup a_{10}(\bar{x}, \bar{y}) \end{array} \right. \right\}.$$

Due to MPOC-LICQ, the application of the implicit function theorem provides the existence of $\delta_2, \delta_3 > 0$ such that for all $k = 1, \ldots, k^+$ and $t < \delta_2$ there exists $\xi_{k,t}$ with $\|\xi_{k,t} - \xi^+_k\| < \delta_3$ and $\xi_{k,t} \in \mathcal{T}'$. We can choose t even smaller, such that $\xi_{1,t}, \ldots, \xi_{k^+,t}$ remain linearly independent and for all $k = 1, \ldots, k^+$ it holds:

$$\xi_{k,t}{}^T D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \xi_{k,t} > 0.$$

Hence, $D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \restriction_{\mathcal{T}'}$ has at least k^+ positive eigenvalues. If we repeat the above reasoning for negative eigenvalues, the matrix $D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \restriction_{\mathcal{T}'}$ has at least as many negative eigenvalues as $D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \restriction_{\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}}$. Additionally, we show that the dimensions of $\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}$ and $\mathcal{T}'$ coincide. By Step 1b, we have $2n - \alpha^{\mathcal{R}}$ for the dimension of $\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}$. Since MPOC-LICQ remains valid in the neighborhood of $(\bar{x}, \bar{y})$, we get again $2n - \alpha^{\mathcal{R}}$ for the dimension of $\mathcal{T}'$. By continuity arguments, NDT4 and (15) provide that $D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \restriction_{\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}}$ is nonsingular. Altogether, the assertion follows.

Step 2c. We claim that

$$QI^{\mathcal{S}}_{t,\mathcal{T}^{\mathcal{S}}_{(x^t,y^t)}} \leq QI^{\mathcal{S}}_{t,\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}} + \alpha^{\mathcal{R}} - \alpha^{\mathcal{S}}_t.$$

For that, we focus on the dimension of $\mathcal{T}^{\mathcal{S}}_{(x^t,y^t)}$. As a consequence of Step 1a it is $2n - \alpha^{\mathcal{S}}_t$. Due to continuity arguments, we can choose t small enough to ensure $x^t_i \neq 0$, $i \in a_{10}(\bar{x}, \bar{y})$ and $y^t_i \neq 0$, $i \in a_{01}(\bar{x}, \bar{y})$. Using this and Lemma 3a), 3b), and 3d), it follows that $\mathcal{T}' \subset \mathcal{T}^{\mathcal{S}}_{(x^t,y^t)}$. Therefore, using Step 2b, $QI^{\mathcal{S}}_{t,\mathcal{T}^{\mathcal{S}}_{(x^t,y^t)}} \leq 2n - \alpha^{\mathcal{S}}_t - k^+$. We observe in view of NDT4 and Step 1b that $QI^{\mathcal{S}}_{t,\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}} = 2n - \alpha^{\mathcal{R}} - k^+$. The assertion follows immediately.

Step 3. Let us show that

$$\max\left\{ m - \left| \{ i \in a_{01}(\bar{x}, \bar{y}) \mid \bar{\sigma}_{1,i} = 0 \} \right|, 0 \right\} \leq TI.$$

In view of Step 2a, Step 2c, and due to continuity, we have for t sufficiently small:

$$m = QI^{\mathcal{S}}_{t,\mathcal{T}^{\mathcal{S}}_{(x^t,y^t)}} \overset{Step\ 2c}{\leq} QI^{\mathcal{S}}_{t,\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}} + \alpha^{\mathcal{R}} - \alpha^{\mathcal{S}}_t \overset{Step\ 2a}{=} \overline{QI}^{\mathcal{R}}_{\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}} + \alpha^{\mathcal{R}} - \alpha^{\mathcal{S}}_t$$

$$\overset{Step\ 1}{=} \overline{QI}^{\mathcal{R}}_{\mathcal{T}^{\mathcal{R}}_{(\bar{x},\bar{y})}} + |a_{00}(\bar{x}, \bar{y})| + |a_{01}(\bar{x}, \bar{y})|$$

$$- \left| \mathcal{H}(x^t, y^t) \cap a_{01}(\bar{x}, \bar{y}) \right|$$

$$= TI + |a_{01}(\bar{x}, \bar{y})| - \left| \mathcal{H}(x^t, y^t) \cap a_{01}(\bar{x}, \bar{y}) \right|.$$

We show for t sufficiently small:

$$|a_{01}(\bar{x}, \bar{y})| - \left| \mathcal{H}(x^t, y^t) \cap a_{01}(\bar{x}, \bar{y}) \right| \leq \left| \{ i \in a_{01}(\bar{x}, \bar{y}) \mid \bar{\sigma}_{1,i} = 0 \} \right|,$$

and the assertion will follow immediately since $TI \geq 0$. Clearly,

$$|a_{01}(\bar{x}, \bar{y})| - \left| \mathcal{H}(x^t, y^t) \cap a_{01}(\bar{x}, \bar{y}) \right| = \left| a_{01}(\bar{x}, \bar{y}) \backslash \mathcal{H}(x^t, y^t) \right|.$$

Suppose $\bar{i} \in a_{01}(\bar{x}, \bar{y})$ with $\bar{\sigma}_{1,\bar{i}} \neq 0$. In view of Remark 1, the difference $\eta^{\geq,t}_{\bar{i}} - \eta^{\leq,t}_{\bar{i}}$ cannot vanish for all t sufficiently small. In particular, one of the multipliers $\eta^{\geq,t}_{\bar{i}}$ or $\eta^{\leq,t}_{\bar{i}}$ has to be not vanishing for all t sufficiently small. Hence, $\bar{i} \in \mathcal{H}(x^t, y^t)$. We therefore have:

$$a_{01}(\bar{x}, \bar{y}) \backslash \mathcal{H}(x^t, y^t) \subset \{ i \in a_{01}(\bar{x}, \bar{y}) \mid \bar{\sigma}_{1,i} = 0 \}.$$

Step 4. Without loss of generality—considering subsequences if needed—we can assume that for any $i \in a_{00}(\bar{x}, \bar{y})$ at least one of the sequences $\frac{x^t_i}{y^t_i}$ or $\frac{y^t_i}{x^t_i}$

is convergent. First, we note that the quotients are well defined due to Lemma 3c). Moreover, if the former sequence does not contain a convergent subsequence, we find a subsequence that tends to plus or minus infinity. Consequently, the corresponding subsequence of the latter reciprocal sequence has to converge to zero. We define the following auxiliary sets:

$$a_{00}^x(\bar{x}, \bar{y}) = \left\{ i \in a_{00}(\bar{x}, \bar{y}) \,\middle|\, \frac{x_i^t}{y_i^t} \text{ converges for } t \to 0 \right\},$$
$$a_{00}^y(\bar{x}, \bar{y}) = a_{00}(\bar{x}, \bar{y}) \backslash a_{00}^x(\bar{x}, \bar{y}).$$

For $\bar{i} \in a_{00}^x(\bar{x}, \bar{y})$ we consider $\mathcal{T}_{(\bar{x}, \bar{y})}^{\mathcal{R}}$ and replace two of the involved equations, namely $(e_{\bar{i}}, 0)\,\xi = 0$ and $(0, e_{\bar{i}})\,\xi = 0$ by one equation $\left(e_{\bar{i}}, \lim\limits_{t \to 0} \frac{x_{\bar{i}}^t}{y_{\bar{i}}^t} e_{\bar{i}} \right) \xi = 0$. Clearly, the vectors involved in the definition of the newly generated linear space, i.e.

$$
\mathcal{T}^{\bar{i}} = \left\{ \xi \in \mathbb{R}^{2n} \,\middle|\,
\begin{array}{l}
(Dh_p(\bar{x}), 0)\,\xi = 0,\, p \in P,\, (Dg_q(\bar{x}), 0)\,\xi = 0,\, q \in Q_0(\bar{x}), \\
(0, e)\,\xi = 0,\, (0, e_i)\,\xi = 0,\, i \in \mathcal{E}(\bar{y}), \\
(e_i, 0)\,\xi = 0,\, i \in a_{00}(\bar{x}, \bar{y})\backslash\{\bar{i}\},\, (0, e_i)\,\xi = 0,\, i \in a_{00}(\bar{x}, \bar{y})\backslash\{\bar{i}\}, \\
\left(e_{\bar{i}}, \lim\limits_{t \to 0} \frac{x_{\bar{i}}^t}{y_{\bar{i}}^t} e_{\bar{i}} \right) \xi = 0, \\
(\bar{y}_i e_i, \bar{x}_i e_i)\,\xi = 0,\, i \in a_{01}(\bar{x}, \bar{y}) \cup a_{10}(\bar{x}, \bar{y}) \}
\end{array}
\right\},
$$

remain linearly independent. The dimension of $\mathcal{T}^{\bar{i}}$ is greater than the dimension of $\mathcal{T}_{(\bar{x}, \bar{y})}^{\mathcal{R}}$ by one. Moreover, there exists $\xi^{\bar{i}} \in \mathcal{T}^{\bar{i}}$ with $\xi_{n+\bar{i}}^{\bar{i}} \neq 0$. Indeed, assume that no such $\xi^{\bar{i}}$ exists, then we can add the equation $(0, e_{\bar{i}})\,\xi = 0$ to the defining equations of $\mathcal{T}^{\bar{i}}$ without changing it. The resulting space, however, is identical to $\mathcal{T}_{(\bar{x}, \bar{y})}^{\mathcal{R}}$, a contradiction. Without loss of generality, we assume $\xi_{n+\bar{i}}^{\bar{i}} = 1$. Further, by straightforward application of the implicit function theorem and due to Lemma 3a) and 3b), we find a sequence of vectors $\xi_t^{\bar{i}} \in \mathcal{T}_t^{\bar{i}}$ that converges to $\xi^{\bar{i}}$ for $t \to 0$. For this, we define

$$
\mathcal{T}_t^{\bar{i}} = \left\{ \xi \in \mathbb{R}^{2n} \,\middle|\,
\begin{array}{l}
(Dh_p(x^t), 0)\,\xi = 0,\, p \in P,\, (Dg_q(x^t), 0)\,\xi = 0,\, q \in Q_0(x^t), \\
(0, e)\,\xi = 0,\, (0, e_i)\,\xi = 0,\, i \in \mathcal{E}(y^t), \\
(e_i, 0)\,\xi = 0,\, i \in a_{00}(\bar{x}, \bar{y})\backslash\{\bar{i}\},\, (0, e_i)\,\xi = 0,\, i \in a_{00}(\bar{x}, \bar{y})\backslash\{\bar{i}\}, \\
\left(e_{\bar{i}}, \frac{x_{\bar{i}}^t}{y_{\bar{i}}^t} e_{\bar{i}} \right) \xi = 0,\, (y_i^t e_i, x_i^t e_i)\,\xi = 0,\, i \in a_{01}(\bar{x}, \bar{y}) \cup a_{10}(\bar{x}, \bar{y})
\end{array}
\right\}.
$$

We again have, due to continuity arguments, that $\xi_{t, n+\bar{i}}^{\bar{i}} \neq 0$. For $\bar{i} \in a_{00}^y(\bar{x}, \bar{y})$ we proceed analogously by considering $\mathcal{T}_{(\bar{x}, \bar{y})}^{\mathcal{R}}$ again and replace two of the involved equations $(e_{\bar{i}}, 0)\,\xi = 0$ and $(0, e_{\bar{i}})\,\xi = 0$ by the equation $\left(\lim\limits_{t \to 0} \frac{y_{\bar{i}}^t}{x_{\bar{i}}^t} e_{\bar{i}}, e_{\bar{i}} \right) \xi = 0$. By the same arguments as before, we find $\xi^{\bar{i}} \in \mathcal{T}^{\bar{i}}$

with $\xi_{\bar{i}}^{\bar{i}} \neq 0$. Again we will assume $\xi_{\bar{i}}^{\bar{i}} = 1$ and find a sequence of vectors $\xi_t^{\bar{i}} \in \mathcal{T}_t^{\bar{i}}$ that converges to $\xi^{\bar{i}}$ for $t \to 0$. Due to continuity, it holds then $\xi_{t,\bar{i}}^{\bar{i}} \neq 0$.

It is straightforward to verify the following observations for t sufficiently small:

a) Let $\{\xi'^{,1}, \ldots, \xi'^{,\ell}\}$ be a base of $\mathcal{T}'$, cf. Step 2b, then $\{\xi'^{,1}, \ldots, \xi'^{,\ell}\} \cup \{\xi_t^{\bar{i}} \mid \bar{i} \in a_{00}(\bar{x}, \bar{y})\}$ is a set of linear independent vectors. In fact, suppose for some coefficients $b_i \in \mathbb{R}, i \in a_{00}(\bar{x}, \bar{y}), \beta_i \in \mathbb{R}, i = 1, \ldots, \ell$ it holds:

$$\sum_{i \in a_{00}(\bar{x}, \bar{y})} b_i \xi_t^i + \sum_{i=1}^{\ell} \beta_i \xi'^{,i} = 0.$$

For $\bar{i} \in a_{00}^x(\bar{x}, \bar{y})$ we consider the $(n + \bar{i})$-th row of this sum

$$\underbrace{b_{\bar{i}} \xi_{t,n+\bar{i}}^{\bar{i}}}_{\neq 0} + \underbrace{\sum_{i \in a_{00}(\bar{x}, \bar{y}) \setminus \{\bar{i}\}} b_i \xi_{t,n+\bar{i}}^i}_{=0} + \underbrace{\sum_{i=1}^{\ell} \beta_i \xi_{n+\bar{i}}'^{,i}}_{=0} = 0.$$

If instead $\bar{i} \in a_{00}^y(\bar{x}, \bar{y})$ we consider the $\bar{i}$-th row of the sum

$$\underbrace{b_{\bar{i}} \xi_{t,\bar{i}}^{\bar{i}}}_{\neq 0} + \underbrace{\sum_{i \in a_{00}(\bar{x}, \bar{y}) \setminus \{\bar{i}\}} b_i \xi_{t,\bar{i}}^i}_{=0} + \underbrace{\sum_{i=1}^{\ell} \beta_i \xi_{\bar{i}}'^{,i}}_{=0} = 0.$$

Altogether, it must hold $b_i = 0, i \in a_{00}(\bar{x}, \bar{y})$. However, this implies

$$\sum_{i=1}^{\ell} \beta_i \xi'^{,i} = 0.$$

Hence, $\beta_i = 0$ for $i = 1, \ldots, \ell$.

b) It holds $\xi_t^{\bar{i}} \in \mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$, cf. Step 1a, for any $\bar{i} \in a_{00}(\bar{x}, \bar{y})$.

c) It holds $(\eta_i^{\leq,t} - \eta_i^{\geq,t}) \xi_{t,i}^{\bar{i}} \xi_{t,n+i}^{\bar{i}} \leq 0, i \in \mathcal{H}(x^t, y^t)$. Since $\xi_t^{\bar{i}} \in \mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$, we obtain:

$$(\eta_i^{\leq,t} - \eta_i^{\geq,t}) \xi_{t,i}^{\bar{i}} \xi_{t,n+i}^{\bar{i}} = (\eta_i^{\geq,t} - \eta_i^{\leq,t}) \left(\xi_{t,n+i}^{\bar{i}}\right)^2 \frac{x_i^t}{y_i^t}.$$

If $i \in \mathcal{H}^{\geq}(x^t, y^t)$, we have $x_i^t < 0, y_i^t > 0$ and $\eta_i^{\leq,t} = 0$. Moreover, due to ND2, we have $\eta_i^{\geq,t} > 0$. The assertion follows immediately. The other case $i \in \mathcal{H}^{\leq}(x^t, y^t)$ is completely analogous.

d) It holds $\lim_{t\to 0}\left(\eta_{\bar i}^{\le,t}-\eta_{\bar i}^{\ge,t}\right)\xi_{t,i}^{\bar i}\xi_{t,n+i}^{\bar i}=-\infty$. We calculate:

$$\left(\eta_{\bar i}^{\le,t}-\eta_{\bar i}^{\ge,t}\right)\xi_{t,i}^{\bar i}\xi_{t,n+i}^{\bar i}=\begin{cases}\left(\eta_{\bar i}^{\ge,t}-\eta_{\bar i}^{\le,t}\right)\left(\xi_{t,n+i}^{\bar i}\right)^2\dfrac{x_{\bar i}^t}{y_{\bar i}^t} & \text{for } \bar i\in a_{00}^x(\bar x,\bar y),\\[2ex]\left(\eta_{\bar i}^{\ge,t}-\eta_{\bar i}^{\le,t}\right)\left(\xi_{t,i}^{\bar i}\right)^2\dfrac{y_{\bar i}^t}{x_{\bar i}^t} & \text{for } \bar i\in a_{00}^y(\bar x,\bar y).\end{cases}$$

Let us suppose $\bar i\in a_{00}^x(\bar x,\bar y)$. We have $y_{\bar i}^t\neq 0$ and, thus, $v_{\bar i}^t=0$. We use Remark 1 and NDT3 to conclude that the sequence $\left(\eta_{\bar i}^{\ge,t}-\eta_{\bar i}^{\le,t}\right)x_{\bar i}^t$ converges to $\varrho_{2,\bar i}<0$ for $t\to 0$. Further, $\frac{1}{y_{\bar i}^t}>0$ tends to infinity for $t\to 0$. Finally, $\left(\xi_{t,n+\bar i}^{\bar i}\right)^2$ converges to 1 for $t\to 0$, due to the construction of $\xi_t^{\bar i}$. Thus, the assertion follows. Instead, let us suppose $\bar i\in a_{00}^y(\bar x,\bar y)$. This time, we have that $\left(\xi_{t,\bar i}^{\bar i}\right)^2$ converges to 1 for $t\to 0$. Due to Remark 1 and NDT3, $\left(\eta_{\bar i}^{\ge,t}-\eta_{\bar i}^{\le,t}\right)y_{\bar i}^t$ converges to $\varrho_{1,\bar i}\neq 0$ for $t\to 0$. If $\bar i\in\mathcal H^{\ge}\left(x^t,y^t\right)$, then $\varrho_{1,\bar i}$ is positive from here. Also, $\frac{1}{x_{\bar i}^t}<0$ tends to minus infinity for $t\to 0$. The other case $\bar i\in\mathcal H^{\le}\left(x^t,y^t\right)$ is completely analogous.

e) We notice that $\xi_t^{\bar i^T}D^2L^{\mathcal R}\left(x^t,y^t\right)\xi_t^{\bar i}$ converges for $t\to 0$ due to the construction above.

Finally, for $\bar i\in a_{00}(\bar x,\bar y)$ we estimate:

$$\xi_t^{\bar i^T}D^2L^{\mathcal S}\left(x^t,y^t\right)\xi_t^{\bar i}\overset{(14)}{=}\xi_t^{\bar i^T}D^2L^{\mathcal R}\left(x^t,y^t\right)\xi_t^{\bar i}$$
$$-\sum_{i\in\mathcal H^{\ge}(x^t,y^t)}2\eta_{\bar i}^{\ge,t}\xi_{t,i}^{\bar i}\xi_{t,n+i}^{\bar i}+\sum_{i\in\mathcal H^{\le}(x^t,y^t)}2\eta_{\bar i}^{\le,t}\xi_{t,i}^{\bar i}\xi_{t,n+i}^{\bar i}$$
$$=\xi_t^{\bar i^T}D^2L^{\mathcal R}\left(x^t,y^t\right)\xi_t^{\bar i}+2\sum_{i\in\mathcal H(x^t,y^t)}\left(\eta_{\bar i}^{\le,t}-\eta_{\bar i}^{\ge,t}\right)\xi_{t,i}^{\bar i}\xi_{t,n+i}^{\bar i}$$
$$\overset{c)}{\le}\xi_t^{\bar i^T}D^2L^{\mathcal R}\left(x^t,y^t\right)\xi_t^{\bar i}+2\left(\eta_{\bar i}^{\le,t}-\eta_{\bar i}^{\ge,t}\right)\xi_{t,\bar i}^{\bar i}\xi_{t,n+\bar i}^{\bar i}.$$

Thus, due to d) and e), $\xi_t^{\bar i^T}D^2L^{\mathcal S}\left(x^t,y^t\right)\xi_t^{\bar i}$ has to be negative for t small enough. As we have seen in Step 2c, it holds $\mathcal T'\subset\mathcal T_{(x^t,y^t)}^{\mathcal S}$. Then, due a) and b), we have

$$QI_{t,\mathcal T_{(x^t,y^t)}^{\mathcal S}}^{\mathcal S}\ge QI_{t,\mathcal T'}^{\mathcal S}+|a_{00}(\bar x,\bar y)|.$$

By Step 2a, we have $QI_{t,\mathcal T_{(\bar x,\bar y)}^{\mathcal R}}^{\mathcal S}=\overline{QI}_{\mathcal T_{(\bar x,\bar y)}^{\mathcal R}}^{\mathcal R}$, and by Step 2b, $QI_{t,\mathcal T_{(\bar x,\bar y)}^{\mathcal R}}^{\mathcal S}=QI_{t,\mathcal T'}^{\mathcal S}$. Overall, we obtain:

$$m=QI_{t,\mathcal T_{(x^t,y^t)}^{\mathcal S}}^{\mathcal S}\ge\overline{QI}_{\mathcal T_{(\bar x,\bar y)}^{\mathcal R}}^{\mathcal R}+|a_{00}(\bar x,\bar y)|=TI.$$

$\square$

Proof of Theorem 5 (i) We show the existence of nondegenerate Karush–Kuhn–Tucker points of $\mathcal{S}$ in a neighborhood of $(\bar{x}, \bar{y})$.

Step 1. First, we show that for all $i \in a_{10}(\bar{x}, \bar{y})$ it holds $\bar{\sigma}_{2,i} \neq 0$. Assume contrarily that $\bar{\sigma}_{2,\tilde{i}} = 0$ for some $\tilde{i} \in a_{10}(\bar{x}, \bar{y})$. We then have due to T-stationarity, cf. (13):

$$c_{\tilde{i}} = \bar{\sigma}_{2,\tilde{i}} + \bar{\mu}_3 = \bar{\mu}_3.$$

Moreover, we have in view of Lemma 1c) an index $\tilde{\tilde{i}} \in a_{01}(\bar{x}, \bar{y}) \setminus \mathcal{E}(\bar{y})$. Thus it holds, cf. (13), $c_{\tilde{\tilde{i}}} = \bar{\mu}_3$. Due to the assumption on c, we have $\tilde{i} = \tilde{\tilde{i}}$, a contradiction. Hence, we may write:

$$
\begin{aligned}
a_{10}(\bar{x}, \bar{y}) &= \left\{ i \in a_{10}(\bar{x}, \bar{y}) \,\middle|\, \bar{\sigma}_{2,i} < 0 \right\} \cup \left\{ i \in a_{10}(\bar{x}, \bar{y}) \,\middle|\, \bar{\sigma}_{2,i} > 0 \right\} \\
&= a_{10}^{<}(\bar{x}, \bar{y}) \cup a_{10}^{>}(\bar{x}, \bar{y}).
\end{aligned}
$$

Due to NDT6 and NDT3, we may split the other index sets as

$$
\begin{aligned}
a_{01}(\bar{x}, \bar{y}) &= \left\{ i \in a_{01}(\bar{x}, \bar{y}) \,\middle|\, \bar{\sigma}_{1,i} < 0 \right\} \cup \left\{ i \in a_{01}(\bar{x}, \bar{y}) \,\middle|\, \bar{\sigma}_{1,i} > 0 \right\} \\
&= a_{01}^{<}(\bar{x}, \bar{y}) \cup a_{01}^{>}(\bar{x}, \bar{y}), \\
a_{00}(\bar{x}, \bar{y}) &= \left\{ i \in a_{00}(\bar{x}, \bar{y}) \,\middle|\, \bar{\varrho}_{1,i} < 0 \right\} \cup \left\{ i \in a(\bar{x}, \bar{y}) \,\middle|\, \bar{\varrho}_{1,i} > 0 \right\} \\
&= a_{00}^{<}(\bar{x}, \bar{y}) \cup a_{00}^{>}(\bar{x}, \bar{y}).
\end{aligned}
$$

Step 2. We consider the auxiliary system of equations $F(t, x, y, \lambda, \mu, \sigma, \varrho) = 0$ given by (16)–(22), which mimics stationarity and feasibility. For stationarity we use:

$$-L(t, x, y, \lambda, \mu, \sigma, \varrho) = 0, \tag{16}$$

where

$$
\begin{aligned}
L = &\begin{pmatrix} \nabla f(x) \\ c \end{pmatrix} - \sum_{p \in P} \lambda_p \begin{pmatrix} \nabla h_p(x) \\ 0 \end{pmatrix} - \sum_{q \in Q_0(\bar{x})} \mu_{1,q} \begin{pmatrix} \nabla g_q(x) \\ 0 \end{pmatrix} + \sum_{i \in \mathcal{E}(\bar{y})} \mu_{2,i} \begin{pmatrix} 0 \\ e_i \end{pmatrix} \\
&- \mu_3 \begin{pmatrix} 0 \\ e \end{pmatrix} - \sum_{i \in a_{01}(\bar{x}, \bar{y})} \frac{\sigma_{1,i}}{\bar{y}_i} \begin{pmatrix} y_i e_i \\ x_i e_i \end{pmatrix} - \sum_{i \in a_{10}^{<}(\bar{x}, \bar{y})} \frac{\sigma_{2,i}}{\bar{x}_i} \begin{pmatrix} y_i e_i \\ x_i e_i \end{pmatrix} \\
&- \sum_{i \in a_{10}^{>}(\bar{x}, \bar{y})} \sigma_{2,i} \begin{pmatrix} 0 \\ e_i \end{pmatrix} - \sum_{i \in a_{00}(\bar{x}, \bar{y})} \left(\varrho_{1,i} \begin{pmatrix} e_i \\ 0 \end{pmatrix} + \varrho_{2,i} \begin{pmatrix} 0 \\ e_i \end{pmatrix} \right).
\end{aligned}
$$

For feasibility we use:

$$h_p(x) = 0, \; p \in P, \quad g_q(x) = 0, \; q \in Q_0(\bar{x}), \tag{17}$$

$$1 + \varepsilon - y_i = 0, \ i \in \mathcal{E}(\bar{y}), \quad \sum_{i=1}^{n} y_i - (n - s) = 0, \tag{18}$$

$$\frac{1}{\bar{y}_i}(x_i y_i - t) = 0, \ i \in a_{01}^{<}(\bar{x}, \bar{y}), \quad \frac{1}{\bar{y}_i}(x_i y_i + t) = 0, \ i \in a_{01}^{>}(\bar{x}, \bar{y}), \tag{19}$$

$$\frac{1}{|\bar{x}_i|}(\mathrm{sgn}\,(\bar{x}_i) x_i y_i - t) = 0, \ i \in a_{10}^{<}(\bar{x}, \bar{y}), \quad y_i = 0, \ i \in a_{10}^{>}(\bar{x}, \bar{y}), \tag{20}$$

$$x_i + \frac{\varrho_{2,i} \sqrt{t}}{\sqrt{\varrho_{1,i} \varrho_{2,i}}} = 0, \ i \in a_{00}^{<}(\bar{x}, \bar{y}), \quad x_i - \frac{\varrho_{2,i} \sqrt{t}}{\sqrt{-\varrho_{1,i} \varrho_{2,i}}} = 0, \ i \in a_{00}^{>}(\bar{x}, \bar{y}), \tag{21}$$

$$y_i + \frac{\varrho_{1,i} \sqrt{t}}{\sqrt{\varrho_{1,i} \varrho_{2,i}}} = 0, \ i \in a_{00}^{<}(\bar{x}, \bar{y}), \quad y_i - \frac{\varrho_{1,i} \sqrt{t}}{\sqrt{-\varrho_{1,i} \varrho_{2,i}}} = 0, \ i \in a_{00}^{>}(\bar{x}, \bar{y}). \tag{22}$$

In view of feasibility and T-stationarity of $(\bar{x}, \bar{y})$ for $\mathcal{R}$, the vector $(0, \bar{x}, \bar{y}, \bar{\lambda}, \bar{\mu}, \bar{\sigma}, \bar{\varrho})$ solves (16)–(22).

Step 3. We consider the Jacobian matrix $DF(t, x, y, \lambda, \mu, \sigma, \rho) = \begin{bmatrix} A & B \\ B^T & D \end{bmatrix}$, where

$$A = -D^2 f(x) + \sum_{p \in P} \lambda_p D^2 h_p(x) + \sum_{q \in Q_0(\bar{x})} \mu_{1,q} D^2 g_q(x)$$

$$+ \sum_{i \in a_{01}(\bar{x}, \bar{y})} \frac{1}{\bar{y}_i} E(i) + \sum_{i \in a_{10}^{<}(\bar{x}, \bar{y})} \frac{1}{\bar{x}_i} E(i),$$

the columns of B are given by the vectors:

$$\begin{pmatrix} \nabla h_p(x) \\ 0 \end{pmatrix}, \ p \in P, \quad \begin{pmatrix} \nabla g_q(x) \\ 0 \end{pmatrix}, \ q \in Q_0(\bar{x}), \quad \begin{pmatrix} 0 \\ -e_i \end{pmatrix}, \ i \in \mathcal{E}(\bar{y}), \quad \begin{pmatrix} 0 \\ e \end{pmatrix},$$

$$\begin{pmatrix} \frac{y_i}{\bar{y}_i} e_i \\ \frac{x_i}{\bar{y}_i} e_i \end{pmatrix}, \ i \in a_{01}(\bar{x}, \bar{y}), \quad \begin{pmatrix} \frac{y_i}{\bar{x}_i} e_i \\ \frac{x_i}{\bar{x}_i} e_i \end{pmatrix}, \ i \in a_{10}^{<}(\bar{x}, \bar{y}), \quad \begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in a_{10}^{>}(\bar{x}, \bar{y}),$$

$$\begin{pmatrix} e_i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_i \end{pmatrix}, \ i \in a_{00}(\bar{x}, \bar{y}),$$

and D consists of $|P| + |Q_0(\bar{x})| + |\mathcal{E}(\bar{y})| + 1 + |a_{01}(\bar{x}, \bar{y})| + |a_{10}(\bar{x}, \bar{y})|$ vanishing rows. The remaining rows of D are given by the vectors:

$$\begin{pmatrix} 0 \\ -\frac{\sqrt{-\varrho_{2,i} t}}{2\sqrt{-\varrho_{1,i}^3}} e_i \\ \frac{\sqrt{t}}{2\sqrt{\varrho_{1,i} \varrho_{2,i}}} e_i \end{pmatrix}, \ i \in a_{10}^{<}(\bar{x}, \bar{y}), \quad \begin{pmatrix} 0 \\ \frac{\sqrt{-\varrho_{2,i} t}}{2\sqrt{\varrho_{1,i}^3}} e_i \\ -\frac{\sqrt{t}}{2\sqrt{-\varrho_{1,i} \varrho_{2,i}}} e_i \end{pmatrix}, \ i \in a_{10}^{>}(\bar{x}, \bar{y}),$$

$$\begin{pmatrix} 0 \\ \dfrac{\sqrt{t}}{2\sqrt{\varrho_{1,i}\varrho_{2,i}}} e_i \\ -\dfrac{\sqrt{-\varrho_{1,i}t}}{2\sqrt{-\varrho_{2,i}^3}} e_i \end{pmatrix}, \; i \in a_{10}^{\leqq}(\bar{x}, \bar{y}), \qquad \begin{pmatrix} 0 \\ -\dfrac{\sqrt{t}}{2\sqrt{-\varrho_{1,i}\varrho_{2,i}}} e_i \\ -\dfrac{\sqrt{\varrho_{1,i}t}}{2\sqrt{-\varrho_{2,i}^3}} e_i \end{pmatrix}, \; i \in a_{10}^{>}(\bar{x}, \bar{y}).$$

Additionally we have $D = 0$ at $(0, \bar{x}, \bar{y}, \bar{\lambda}, \bar{\mu}, \bar{\sigma}, \bar{\varrho})$. Hence, we can apply Theorem 2.3.2 from [13], which says that $DF(0, \bar{x}, \bar{y}, \bar{\lambda}, \bar{\mu}, \bar{\sigma}, \bar{\rho}) = \begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix}$ is nonsingular if and only if $\xi^T A\xi \neq 0$ for all $\xi \in B^\perp$, the orthogonal complement of the subspace spanned by the columns of B. In view of $B^\perp = T_{(\bar{x},\bar{y})}^{\mathcal{R}}$, we check:

$$\xi^T A\xi = \xi^T \left(-D^2 L^{\mathcal{R}}(\bar{x}, \bar{y}) + \sum_{i \in a_{01}(\bar{x},\bar{y})} \frac{1}{\bar{y}_i} E(i) + \sum_{i \in a_{10}^{\leqq}(\bar{x},\bar{y})} \frac{1}{\bar{x}_i} E(i) \right) \xi$$

$$= -\underbrace{\xi^T D^2 L^{\mathcal{R}}(\bar{x}, \bar{y}) \xi}_{\neq 0 \text{ due to NDT4}} + \sum_{i \in a_{01}(\bar{x},\bar{y})} \frac{1}{\bar{y}_i} 2\xi_i \underbrace{\xi_{n+i}}_{=0} + \sum_{i \in a_{10}^{\leqq}(\bar{x},\bar{y})} \frac{1}{\bar{x}_i} 2\xi_i \underbrace{\xi_{n+i}}_{=0}.$$

Hence, by means of the implicit function theorem we obtain for any $t > 0$ sufficiently small a solution $(t, x^t, y^t, \lambda^t, \mu^t, \sigma^t, \varrho^t)$ of the system of equations (16)–(22).

Step 4. By choosing t even smaller, if necessary, we can ensure due to continuity reasons as well as NDT2, NDT3, and NDT6 that the following holds:

(i) $g_q(x^t) > 0, q \in Q \backslash Q_0(\bar{x})$ and $\mu_{1,q}^t > 0, q \in Q_0(\bar{x})$,

(ii) $\text{sgn}\left(\sigma_{1,i}^t \right) = \text{sgn}\left(\bar{\sigma}_{1,i} \right), i \in a_{01}(\bar{x}, \bar{y})$,

(iii) $\text{sgn}\left(\sigma_{2,i}^t \right) = \text{sgn}\left(\bar{\sigma}_{2,i} \right), \; \text{sgn}\left(x_i^t \right) = \text{sgn}(\bar{x}_i), i \in a_{10}(\bar{x}, \bar{y})$,

(iv) $\text{sgn}\left(\varrho_{1,i}^t \right) = \text{sgn}\left(\bar{\varrho}_{1,i} \right), \varrho_{2,i}^t < 0, i \in a_{00}(\bar{x}, \bar{y})$,

(v) $y_i^t \geq 0, i \in \{1, \ldots, n\}$ and $y_i^t < 1 + \varepsilon, i \in \{1, \ldots, n\} \backslash \mathcal{E}(\bar{y})$.

From here it is straightforward to see that (x^t, y^t) is feasible for $\mathcal{S}$ and we have:

(i) $Q_0(x^t) = Q_0(\bar{x}), \mathcal{E}(y^t) = \mathcal{E}(\bar{y}), \mathcal{N}(y^t) = a_{10}^{>}(\bar{x}, \bar{y}), \mathcal{O}(x^t, y^t) = \emptyset$,

(ii) $\mathcal{H}^{\geqq}(x^t, y^t) = a_{01}^{>}(\bar{x}, \bar{y}) \cup \left\{ i \in a_{10}^{\leqq}(\bar{x}, \bar{y}) \mid x_i^t < 0 \right\} \cup a_{00}^{>}(\bar{x}, \bar{y})$,

(iii) $\mathcal{H}^{\leqq}(x^t, y^t) = a_{01}^{<}(\bar{x}, \bar{y}) \cup \left\{ i \in a_{10}^{<}(\bar{x}, \bar{y}) \mid x_i^t > 0 \right\} \cup a_{00}^{<}(\bar{x}, \bar{y})$.

Thus, it holds:

$$\begin{pmatrix} \nabla f(x^t) \\ c \end{pmatrix} = \sum_{p \in P} \lambda_p^t \begin{pmatrix} \nabla h_p(x^t) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(\bar{x})} \mu_{1,q}^t \begin{pmatrix} \nabla g_q(x^t) \\ 0 \end{pmatrix} - \sum_{i \in \mathcal{E}(\bar{y})} \mu_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}$$

$$+ \mu_3^t \begin{pmatrix} 0 \\ e \end{pmatrix} + \sum_{i \in a_{01}(\bar{x},\bar{y})} \sigma_{1,i}^t \frac{1}{\bar{y}_i} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{i \in a_{10}^{\leqq}(\bar{x},\bar{y})} \sigma_{2,i}^t \frac{1}{\bar{x}_i} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}$$

$$+ \sum_{i \in a_{10}^{>}(\bar{x},\bar{y})} \sigma_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix} + \sum_{i \in a_{00}(\bar{x},\bar{y})} \left(\varrho_{1,i}^t \begin{pmatrix} e_i \\ 0 \end{pmatrix} + \varrho_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix} \right)$$

$$= \sum_{p \in P} \lambda_p^t \begin{pmatrix} \nabla h_p\left(x^t\right) \\ 0 \end{pmatrix} + \sum_{q \in Q_0(x^t)} \mu_{1,q}^t \begin{pmatrix} \nabla g_q\left(x^t\right) \\ 0 \end{pmatrix} - \sum_{i \in \mathcal{E}(y^t)} \mu_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}$$

$$+ \mu_3^t \begin{pmatrix} 0 \\ e \end{pmatrix} + \sum_{\substack{i \in \mathcal{H}^{\geq}(x^t, y^t), \\ i \in a_{01}(\bar{x}, \bar{y})}} \sigma_{1,i}^t \frac{1}{\bar{y}_i} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{\substack{i \in \mathcal{H}^{\geq}(x^t, y^t), \\ i \in a_{10}(\bar{x}, \bar{y})}} \sigma_{2,i}^t \frac{1}{\bar{x}_i} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}$$

$$+ \sum_{\substack{i \in \mathcal{H}^{\geq}(x^t, y^t), \\ i \in a_{00}(\bar{x}, \bar{y})}} \frac{\sqrt{-\varrho_{1,i}^t \varrho_{2,i}^t}}{\sqrt{t}} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{\substack{i \in \mathcal{H}^{\leq}(x^t, y^t), \\ i \in a_{01}(\bar{x}, \bar{y})}} \sigma_{1,i}^t \frac{1}{\bar{y}_i} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}$$

$$+ \sum_{\substack{i \in \mathcal{H}^{\leq}(x^t, y^t), \\ i \in a_{10}(\bar{x}, \bar{y})}} \sigma_{2,i}^t \frac{1}{\bar{x}_i} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix} + \sum_{\substack{i \in \mathcal{H}^{\leq}(x^t, y^t), \\ i \in a_{00}(\bar{x}, \bar{y})}} \frac{-\sqrt{\varrho_{1,i}^t \varrho_{2,i}^t}}{\sqrt{t}} \begin{pmatrix} y_i^t e_i \\ x_i^t e_i \end{pmatrix}$$

$$+ \sum_{i \in \mathcal{N}(y^t)} \sigma_{2,i}^t \begin{pmatrix} 0 \\ e_i \end{pmatrix}.$$

We rename the multipliers as

$$\eta_i^{\geq, t} = \begin{cases} \sigma_{1,i}^t \dfrac{1}{\bar{y}_i}, & \text{for } i \in \mathcal{H}^{\geq}\left(x^t, y^t\right) \cap a_{01}\left(\bar{x}, \bar{y}\right), \\ \sigma_{2,i}^t \dfrac{1}{\bar{x}_i}, & \text{for } i \in \mathcal{H}^{\geq}\left(x^t, y^t\right) \cap a_{10}\left(\bar{x}, \bar{y}\right), \\ \dfrac{\sqrt{-\varrho_{1,i}^t \varrho_{2,i}^t}}{\sqrt{t}}, & \text{for } i \in \mathcal{H}^{\geq}\left(x^t, y^t\right) \cap a_{00}\left(\bar{x}, \bar{y}\right), \\ 0, & \text{else,} \end{cases}$$

$$\eta_i^{\leq, t} = \begin{cases} -\sigma_{1,i}^t \dfrac{1}{\bar{y}_i}, & \text{for } i \in \mathcal{H}^{\leq}\left(x^t, y^t\right) \cap a_{01}\left(\bar{x}, \bar{y}\right), \\ -\sigma_{2,i}^t \dfrac{1}{\bar{x}_i}, & \text{for } i \in \mathcal{H}^{\leq}\left(x^t, y^t\right) \cap a_{10}\left(\bar{x}, \bar{y}\right), \\ \dfrac{\sqrt{\varrho_{1,i}^t \varrho_{2,i}^t}}{\sqrt{t}}, & \text{for } i \in \mathcal{H}^{\leq}\left(x^t, y^t\right) \cap a_{00}\left(\bar{x}, \bar{y}\right), \\ 0, & \text{else,} \end{cases}$$

$$\nu_i^t = \begin{cases} \sigma_{2,i}^t, & \text{for } i \in \mathcal{N}\left(y^t\right) \cap a_{10}\left(\bar{x}, \bar{y}\right), \\ 0, & \text{else.} \end{cases}$$

Hence, $\left(x^t, y^t\right)$ fulfills (9). Also, it is straightforward to check that (10) and (11) are fulfilled. Thus, $\left(x^t, y^t\right)$ is a Karush–Kuhn–Tucker point of $\mathcal{S}$.

Step 5. Moreover, ND1 is satisfied as well in view of Theorem 2. Similar to (10) and (11), ND2 holds at $\left(x^t, y^t\right)$. It remains to show ND3, i.e.

$$\xi_k^T D^2 L^{\mathcal{S}}\left(x^t, y^t\right) \xi_k \neq 0 \quad \text{for } k = 1, \ldots, 2n - \alpha_t^{\mathcal{S}},$$

where $\xi_1, \ldots \xi_{2n-\alpha_t^{\mathcal{S}}}$ form a basis of $\mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$, cf. Step 1a from the proof of Theorem 3. Note, that by construction $\alpha_t^{\mathcal{S}}$ is constant for t sufficiently small. Thus, we refer to it as $\alpha^{\mathcal{S}}$. Next, we construct for $t > 0$ sufficiently small such a basis as follows. First, we choose eigenvectors $\bar{\xi}_1, \ldots, \bar{\xi}_{2n-\alpha^{\mathcal{R}}}$ of $D^2 L^{\mathcal{R}} (\bar{x}, \bar{y})$ forming a basis of $\mathcal{T}_{(\bar{x}, \bar{y})}^{\mathcal{R}}$, cf. Step 1b from the proof of Theorem 4. With similar arguments as in Step 2b of the proof of Theorem 4 and by using the implicit function theorem, we find $\xi_j^t \in \mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$, $j = 1, \ldots, 2n - \alpha^{\mathcal{R}}$, still linearly independent. The remaining $2n - \alpha^{\mathcal{S}} - (2n - \alpha^{\mathcal{R}}) = |a_{00}(\bar{x}, \bar{y})|$ vectors are chosen as follows. Namely, for $\bar{i} \in a_{00}(\bar{x}, \bar{y})$ we consider

$$
\mathcal{T}_t^{\bar{i}} = \left\{ \xi \in \mathbb{R}^{2n} \left| \begin{array}{l} (Dh_p(x^t), 0)\,\xi = 0,\, p \in P,\, (Dg_q(x^t), 0)\,\xi = 0,\, q \in Q_0(\bar{x}), \\ (0, e)\,\xi = 0,\, (0, e_i)\,\xi = 0,\, i \in \mathcal{E}(\bar{y}), \\ (e_i, 0)\,\xi = 0,\, i \in a_{00}(\bar{x}, \bar{y})\backslash\{\bar{i}\},\, (0, e_i)\,\xi = 0,\, i \in a_{00}(\bar{x}, \bar{y})\backslash\{\bar{i}\}, \\ \left(e_{\bar{i}}, \frac{\varrho_{2,\bar{i}}^t}{\varrho_{1,\bar{i}}^t} e_{\bar{i}}\right)\xi = 0,\, (y_i^t e_i, x_i^t e_i)\,\xi = 0,\, i \in a_{01}(\bar{x}, \bar{y}) \cup a_{10}(\bar{x}, \bar{y}) \end{array} \right. \right\}.
$$

As in Step 4 of the proof of Theorem 4, we can find $\bar{\xi}_{\bar{i}} \in \mathcal{T}_0^{\bar{i}}\backslash\mathcal{T}_{(\bar{x}, \bar{y})}^{\mathcal{R}}$. Especially, $(e_{\bar{i}}, 0)\,\bar{\xi}_{\bar{i}} \neq 0$. We note that $\lim_{t\to 0} \frac{x_{\bar{i}}^t}{y_{\bar{i}}^t} = \frac{\bar{\varrho}_{2,\bar{i}}}{\bar{\varrho}_{1,\bar{i}}}$. Using this and the implicit function theorem, we find for t sufficiently small a vector $\xi_{\bar{i}}^t \in \mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$. It is then straightforward to check, that ξ_j^t, $j \in \{1, \ldots, 2n - \alpha^{\mathcal{R}}\}$ and $\xi_{\bar{i}}^t$, $\bar{i} \in a_{00}(\bar{x}, \bar{y})$, indeed form a basis of $\mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$.

We continue by considering the following limits with respect to the subsets of $\mathcal{H}(x^t, y^t)$ for any sequence of vectors $\xi^t \in \mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$ from the constructed base, cf. the definition of η-multipliers:

$$
\lim_{t\to 0} \sum_{i \in a_{01}^{\leq}(\bar{x}, \bar{y})} \eta_i^{\leq, t} \xi_i^t \xi_{n+i}^t = \lim_{t\to 0} \sum_{i \in a_{01}^{\leq}(\bar{x}, \bar{y})} \sigma_{1,i}^t \frac{1}{\bar{y}_i} \frac{x_i^t}{y_i^t} \left(\xi_{n+i}^t\right)^2 = 0,
$$

$$
\lim_{t\to 0} \sum_{i \in a_{01}^{>}(\bar{x}, \bar{y})} \eta_i^{\geq, t} \xi_i^t \xi_{n+i}^t = \lim_{t\to 0} \sum_{i \in a_{01}^{>}(\bar{x}, \bar{y})} -\sigma_{1,i}^t \frac{1}{\bar{y}_i} \frac{x_i^t}{y_i^t} \left(\xi_{n+i}^t\right)^2 = 0,
$$

$$
\lim_{t\to 0} \sum_{i \in a_{10}^{\leq}(\bar{x}, \bar{y}), x_i^t > 0} \eta_i^{\leq, t} \xi_i^t \xi_{n+i}^t = \lim_{t\to 0} \sum_{i \in a_{10}^{\leq}(\bar{x}, \bar{y}), x_i^t > 0} \sigma_{2,i}^t \frac{1}{\bar{x}_i} \frac{y_i^t}{x_i^t} \left(\xi_i^t\right)^2 = 0,
$$

$$
\lim_{t\to 0} \sum_{i \in a_{10}^{\leq}(\bar{x}, \bar{y}), x_i^t < 0} \eta_i^{\geq, t} \xi_i^t \xi_{n+i}^t = \lim_{t\to 0} \sum_{i \in a_{10}^{\leq}(\bar{x}, \bar{y}), x_i^t < 0} -\sigma_{2,i}^t \frac{1}{\bar{x}_i} \frac{y_i^t}{x_i^t} \left(\xi_i^t\right)^2 = 0.
$$

If for some $i \in a_{00}^{<}(\bar{x}, \bar{y})$ or $i \in a_{00}^{>}(\bar{x}, \bar{y})$ it holds $\lim_{t \to 0} (e_i, 0) \, \xi^t \neq 0$, we observe:

$$\lim_{t \to 0} \sum_{i \in a_{00}^{<}(\bar{x}, \bar{y})} \eta_i^{\leq, t} \xi_i^t \xi_{n+i}^t = \lim_{t \to 0} \sum_{i \in a_{00}^{<}(\bar{x}, \bar{y})} -\frac{\sqrt{\varrho_{1,i}^t \varrho_{2,i}^t}}{\sqrt{t}} \frac{\varrho_{1,i}^t}{\varrho_{2,i}^t} \left(\xi_i^t\right)^2 = -\infty,$$

$$\lim_{t \to 0} \sum_{i \in a_{00}^{>}(\bar{x}, \bar{y})} \eta_i^{\geq, t} \xi_i^t \xi_{n+i}^t = \lim_{t \to 0} \sum_{i \in a_{00}^{>}(\bar{x}, \bar{y})} -\frac{\sqrt{-\varrho_{1,i}^t \varrho_{2,i}^t}}{\sqrt{t}} \frac{\varrho_{1,i}^t}{\varrho_{2,i}^t} \left(\xi_i^t\right)^2 = \infty.$$

Finally, we calculate as in (15):

$$\left(\xi^t\right)^T D^2 L^{\mathcal{S}} \left(x^t, y^t\right) \xi^t = \left(\xi^t\right)^T D^2 L^{\mathcal{R}} \left(x^t, y^t\right) \xi^t - \sum_{i \in \mathcal{H}^{\geq}(x^t, y^t)} 2\eta_i^{\geq, t} \xi_i^t \xi_{n+i}^t$$
$$+ \sum_{i \in \mathcal{H}^{\leq}(x^t, y^t)} 2\eta_i^{\leq, t} \xi_i^t \xi_{n+i}^t.$$

Altogether, we obtain for any basis vector ξ_j^t, $j \in \{1, \ldots, 2n - \alpha^{\mathcal{R}}\}$ of $\mathcal{T}_{(x^t, y^t)}^{\mathcal{S}}$:

$$\lim_{t \to 0} \left(\xi_j^t\right)^T D^2 L^{\mathcal{S}} \left(x^t, y^t\right) \xi_j^t = \bar{a}_j \left\|\bar{\xi}_j\right\|^2, \tag{23}$$

where $\bar{a}_j$ is a nonzero eigenvalue of $D^2 L^{\mathcal{R}} (\bar{x}, \bar{y})$ due to the choice of $\bar{\xi}_j$ and NDT4. Let us now focus on the basis vectors $\xi_{\bar{i}}^t$, $\bar{i} \in a_{00}(\bar{x}, \bar{y})$. We then have:

$$\lim_{t \to 0} \left(\xi_{\bar{i}}^t\right)^T D^2 L^{\mathcal{S}} \left(x^t, y^t\right) \xi_{\bar{i}}^t = -\infty. \tag{24}$$

This is due to the following reasoning. First, $\left(\xi_{\bar{i}}^t\right)^T D^2 L^{\mathcal{R}} \left(x^t, y^t\right) \xi_{\bar{i}}^t$ is bounded due to the construction of $\xi_{\bar{i}}^t$, and, moreover, $\lim_{t \to 0} (e_{\bar{i}}, 0) \, \xi_{\bar{i}}^t \neq 0$. We conclude that ND3 is fulfilled. Additionally, the T-index of $(\bar{x}, \bar{y})$ is equal to the sum of its quadratic index and its biactive index, i.e. $m = \overline{QI}^{\mathcal{R}} + |a_{00}(\bar{x}, \bar{y})|$. In view of (23) and (24), the quadratic index $QI_t^{\mathcal{S}}$ of (x^t, y^t) is then exactly m for t sufficiently small.
(ii) Next, we elaborate on the uniqueness of Karush–Kuhn–Tucker points (x^t, y^t) constructed above.
Step 6. As preliminary considerations we show Steps 6.1–6.3.
Step 6.1. For all sufficiently small t we show

$$a_{10}^{>}(\bar{x}, \bar{y}) = \mathcal{N}\left(y^t\right), \quad a_{10}^{<}(\bar{x}, \bar{y}) = a_{10}(\bar{x}, \bar{y}) \backslash \mathcal{N}\left(y^t\right).$$

First we note, that due to Remark 1 we have $\lim_{t\to 0}\mu_3^t = \bar{\mu}_3$. Suppose $\bar{i} \in a_{10}^{>}(\bar{x}, \bar{y})$. Due to (13), we have $c_{\bar{i}} = \bar{\sigma}_{2,\bar{i}} + \bar{\mu}_3 > \bar{\mu}_3$. Assume $\bar{i} \notin \mathcal{N}(y^t)$. Since (x^t, y^t) is a Karush–Kuhn–Tucker point (12), $c_{\bar{i}} \leq \mu_3^t$. Taking the limit yields a contradiction. Thus, $a_{10}^{>}(\bar{x}, \bar{y}) \subset \mathcal{N}(y^t)$. The reverse inclusion follows from Lemma 3d). For the other assertion, we just recall $a_{10}(\bar{x}, \bar{y}) = a_{10}^{<}(\bar{x}, \bar{y}) \cup a_{10}^{>}(\bar{x}, \bar{y})$, cf. Step 1.

Step 6.2. For all sufficiently small t we show

$$a_{00}^{>}(\bar{x}, \bar{y}) = a_{00}(\bar{x}, \bar{y}) \cap \mathcal{H}^{\geq}(x^t, y^t), \quad a_{00}^{<}(\bar{x}, \bar{y}) = a_{00}(\bar{x}, \bar{y}) \cap \mathcal{H}^{\leq}(x^t, y^t).$$

Lemma 3c) provides $a_{00}(\bar{x}, \bar{y}) \subset \mathcal{H}(x^t, y^t)$. Suppose $\bar{i} \in a_{00}(\bar{x}, \bar{y}) \cap \mathcal{H}^{\geq}(x^t, y^t)$. Then, the $\bar{i}$-th row of (9) provides

$$\frac{\partial f}{\partial x_{\bar{i}}}(x^t) = \sum_{p \in P} \lambda_p^t \frac{\partial h_p}{\partial x_{\bar{i}}}(x^t) + \sum_{q \in Q_0(x^t)} \mu_{1,q}^t \frac{\partial g_q}{\partial x_{\bar{i}}}(x^t) + \mu_3^t + \eta_{\bar{i}}^{\geq,t} y_{\bar{i}}^t. \quad (25)$$

However, we also have due to T-stationarity of $(\bar{x}, \bar{y})$ that the $\bar{i}$-th row of (5) reads as

$$\frac{\partial f}{\partial x_{\bar{i}}}(\bar{x}) = \sum_{p \in P} \bar{\lambda}_p \frac{\partial h_p}{\partial x_{\bar{i}}}(\bar{x}) + \sum_{q \in Q_0(\bar{x})} \bar{\mu}_{1,q} \frac{\partial g_q}{\partial x_{\bar{i}}}(\bar{x}) + \bar{\mu}_3 + \bar{\varrho}_{1,\bar{i}}. \quad (26)$$

The application of Lemma 3a) yields $Q_0(\bar{x}) = Q_0(x^t)$ for t sufficiently small. Remark 1a) states $\lim_{t\to 0} \lambda^t = \bar{\lambda}$, $\lim_{t\to 0} \mu^t = \bar{\mu}$. By taking $t \to 0$ in (25) and comparing to (26), we obtain

$$\lim_{t\to 0} \eta_{\bar{i}}^{\geq,t} y_{\bar{i}}^t = \varrho_{1,\bar{i}}.$$

In view of NDT3, we know that $\varrho_{1,\bar{i}} \neq 0$. Thus, we conclude, due to $\eta_{\bar{i}}^{\geq,t} y_{\bar{i}}^t \geq 0$, that $\varrho_{1,\bar{i}} > 0$ and, thus, by definition $\bar{i} \in a_{00}^{>}(\bar{x}, \bar{y})$. Analogously, we have $i \in a_{00}^{<}(\bar{x}, \bar{y})$ for all $i \in a_{00}(\bar{x}, \bar{y}) \cap \mathcal{H}^{\leq}(x^t, y^t)$. Both assertions then follow.

Step 6.3. For all sufficiently small t we show

$$a_{01}^{<}(\bar{x}, \bar{y}) = a_{01}(\bar{x}, \bar{y}) \cap \mathcal{H}^{\leq}(x^t, y^t), \quad a_{01}^{>}(\bar{x}, \bar{y}) = a_{01}(\bar{x}, \bar{y}) \cap \mathcal{H}^{\geq}(x^t, y^t).$$

Suppose $\bar{i} \in a_{01}(\bar{x}, \bar{y}) \cap \mathcal{H}^{\geq}(x^t, y^t)$. The $\bar{i}$-th row of (9) provides again

$$\frac{\partial f}{\partial x_{\bar{i}}}(x^t) = \sum_{p \in P} \lambda_p^t \frac{\partial h_p}{\partial x_{\bar{i}}}(x^t) + \sum_{q \in Q_0(x^t)} \mu_{1,q}^t \frac{\partial g_q}{\partial x_{\bar{i}}}(x^t) + \mu_3^t + \eta_{\bar{i}}^{\geq,t} y_{\bar{i}}^t. \quad (27)$$

The $\bar{i}$-th row of (5) provides

$$\frac{\partial f}{\partial x_{\bar{i}}}(\bar{x}) = \sum_{p \in P} \bar{\lambda}_p \frac{\partial h_p}{\partial x_{\bar{i}}}(\bar{x}) + \sum_{q \in Q_0(\bar{x})} \bar{\mu}_{1,q} \frac{\partial g_q}{\partial x_{\bar{i}}}(\bar{x}) + \bar{\mu}_3 + \bar{\sigma}_{1,\bar{i}}. \tag{28}$$

As above, we apply Lemma 3a) and Remark 1a). By taking $t \to 0$ in (27) and comparing to (28), we obtain

$$\lim_{t \to 0} \eta_{\bar{i}}^{\geq,t} y_{\bar{i}}^t = \sigma_{1,\bar{i}}.$$

We conclude due to $\eta_{\bar{i}}^{\geq,t} y_{\bar{i}}^t \geq 0$ and NDT6, i.e. $\sigma_{1,\bar{i}} \neq 0$, that it holds $\sigma_{1,\bar{i}} > 0$ and, thus, by definition $\bar{i} \in a_{01}^>(\bar{x}, \bar{y})$. Analogously, we have $i \in a_{01}^<(\bar{x}, \bar{y})$ for all $i \in a_{01}(\bar{x}, \bar{y}) \cap \mathcal{H}^\leq(x^t, y^t)$. We conclude the proof of Step 6.2 by noting that $a_{01}(\bar{x}, \bar{y}) \backslash \mathcal{H}(x^t, y^t) = \emptyset$. Assume instead there exists $\bar{i} \in a_{01}(\bar{x}, \bar{y}) \backslash \mathcal{H}(x^t, y^t)$, then we derive as before $\sigma_{1,\bar{i}} = 0$, a contradiction to NDT6.

Step 7. We recall that any Karush–Kuhn–Tucker point $(\tilde{x}^t, \tilde{y}^t)$ of $\mathcal{S}$ with multipliers $(\tilde{\lambda}^t, \tilde{\mu}^t, \tilde{\eta}^t, \tilde{\nu}^t)$ has to fulfill (9)–(11). We define the multipliers

$$\tilde{\sigma}_{1,i}^t = \left(\tilde{\eta}_i^\geq - \tilde{\eta}_i^\leq\right) \tilde{y}_i^t, i \in a_{01}(\bar{x}, \bar{y}) \cap \mathcal{H}\left(\tilde{x}^t, \tilde{y}^t\right),$$
$$\tilde{\sigma}_{2,i}^t = \left(\tilde{\eta}_i^\geq - \tilde{\eta}_i^\leq\right) \tilde{x}_i^t, i \in a_{10}(\bar{x}, \bar{y}) \backslash \mathcal{N}(\tilde{y}^t),$$
$$\tilde{\sigma}_{2,i}^t = \nu_i, i \in \mathcal{N}(\tilde{y}^t),$$
$$\tilde{\varrho}_{1,i}^t = \left(\tilde{\eta}_i^\geq - \tilde{\eta}_i^\leq\right) \tilde{y}_i^t, i \in a_{00}(\bar{x}, \bar{y}),$$
$$\tilde{\varrho}_{2,i}^t = \left(\tilde{\eta}_i^\geq - \tilde{\eta}_i^\leq\right) \tilde{x}_i^t, i \in a_{00}(\bar{x}, \bar{y}).$$

Let us consider the point $\left(t, \tilde{x}^t, \tilde{y}^t, \tilde{\lambda}^t, \tilde{\mu}^t, \tilde{\sigma}^t, \tilde{\varrho}^t\right)$. In view of Lemma 3 and Lemma 2, we obtain for t sufficiently small

$$Q_0(\bar{x}) = Q_0(\tilde{x}^t), \quad \mathcal{E}(\bar{y}) = \mathcal{E}(\tilde{y}^t), \quad \sum_{i=1}^n \tilde{y}_i^t = n - s.$$

Due to Step 6, an immediate calculation shows then that $\left(t, \tilde{x}^t, \tilde{y}^t, \tilde{\lambda}^t, \tilde{\mu}^t, \tilde{\sigma}^t, \tilde{\varrho}^t\right)$ fulfills equations (16)–(22) for t sufficiently small. However, the inverse function theorem was used in Step 3 to find the solution of this system of equations in the neighborhood of $(\bar{x}, \bar{y})$. Hence, (x^t, y^t) must be unique. $\qquad\square$

References

1. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM J. Optim. **11**, 918–936 (2001)

2. Izmailov, A.F., Solodov, M.V.: Mathematical programs with vanishing constraints: Optimality conditions, sensitivity, and a relaxation method. J. Optim. Theory Appl. **142**, 501–532 (2009)
3. Kanzow, C., Mehlitz, P., Steck, D.: Relaxation schemes for mathematical programmes with switching constraints. Optim. Methods Softw. **36**, 1223–1258 (2021)
4. Lämmel, S., Shikhman, V.: Optimality conditions for mathematical programs with orthogonality type constraints. Set-Valued Variat. Anal. **31**(9) (2023)
5. Burdakov, O.P., Kanzow, C., Schwartz, A.: Mathematical programs with cardinality constraints: reformulation by complementarity-type conditions and a regularization method. SIAM J. Optim. **26**, 397–425 (2016)
6. Bucher, M., Schwartz, A.: Second-order optimality conditions and improved convergence results for regularization methods for cardinality-constrained optimization problems. J. Optim. Theory Appl. **178**, 383–410 (2018)
7. Branda, M., Bucher, M., Červinka, M., Schwartz, A.: Convergence of a Scholtes-type regularization method for cardinality-constrained optimization problems with an application in sparse robust portfolio optimization. Comput. Optim. Appl. **70**, 503–530 (2018)
8. Jongen, H.T., Jonker, P., Twilt, F.: Nonlinear Optimization in Finite Dimensions. Kluwer Academic Publishers, Dordrecht (2000)
9. Lämmel, S., Shikhman, V.: Global aspects of the continuous reformulation for cardinality-constrained optimization problems. Optimization (2023). https://doi.org/10.1080/02331934.2023.2249014. (**Online first**)
10. Lämmel, S., Shikhman, V.: Cardinality-constrained optimization problems in general position and beyond. Pure Appl. Funct. Anal. **8**(4), 1107–1133 (2023)
11. Červinka, M., Kanzow, C., Schwartz, A.: Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. Math. Program. **160**, 353–377 (2016)
12. Günzel, H.: The structured jet transversality theorem. Optimization **57**, 159–164 (2008)
13. Jongen, H.T., Meer, K., Triesch, E.: Optimization Theory. Kluwer Academic Publishers, Dordrecht (2004)