

Lämmel, S.; Shikhman, V.

Article — Published Version

On nondegenerate M-stationary points for sparsity constrained nonlinear optimization

Journal of Global Optimization

Provided in Cooperation with:

Springer Nature

Suggested Citation: Lämmel, S.; Shikhman, V. (2021) : On nondegenerate M-stationary points for sparsity constrained nonlinear optimization, Journal of Global Optimization, ISSN 1573-2916, Springer US, New York, NY, Vol. 82, Iss. 2, pp. 219-242, <https://doi.org/10.1007/s10898-021-01070-7>

This Version is available at:

<https://hdl.handle.net/10419/287145>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



On nondegenerate M-stationary points for sparsity constrained nonlinear optimization

S. Lämmel¹ · V. Shikhman¹

Received: 5 June 2020 / Accepted: 5 August 2021 / Published online: 23 August 2021
© The Author(s) 2021

Abstract

We study sparsity constrained nonlinear optimization (SCNO) from a topological point of view. Special focus will be on M-stationary points from Burdakov et al. (SIAM J Optim 26:397–425, 2016), also introduced as N^C -stationary points in Pan et al. (J Oper Res Soc China 3:421–439, 2015). We introduce nondegenerate M-stationary points and define their M-index. We show that all M-stationary points are generically nondegenerate. In particular, the sparsity constraint is active at all local minimizers of a generic SCNO. Some relations to other stationarity concepts, such as S-stationarity, basic feasibility, and CW-minimality, are discussed in detail. By doing so, the issues of instability and degeneracy of points due to different stationarity concepts are highlighted. The concept of M-stationarity allows to adequately describe the global structure of SCNO along the lines of Morse theory. For that, we study topological changes of lower level sets while passing an M-stationary point. As novelty for SCNO, multiple cells of dimension equal to the M-index are needed to be attached. This intriguing fact is in strong contrast with other optimization problems considered before, where just one cell suffices. As a consequence, we derive a Morse relation for SCNO, which relates the numbers of local minimizers and M-stationary points of M-index equal to one. The appearance of such saddle points cannot be thus neglected from the perspective of global optimization. Due to the multiplicity phenomenon in cell-attachment, a saddle point may lead to more than two different local minimizers. We conclude that the relatively involved structure of saddle points is the source of well-known difficulty if solving SCNO to global optimality.

Keywords Sparsity constraint · M-stationarity, M-index · Nondegeneracy · Genericity · Morse theory · Saddle points

✉ V. Shikhman
vladimir.shikhman@mathematik.tu-chemnitz.de

S. Lämmel
sebastian.laemmel@mathematik.tu-chemnitz.de

¹ Department of Mathematics, Chemnitz University of Technology, Reichenhainer Str. 41, 09126 Chemnitz, Germany

1 Introduction

We consider the sparsity constrained nonlinear optimization:

$$\text{SCNO: } \min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad \|x\|_0 \leq s,$$

where the so-called ℓ_0 "norm" counts non-zero entries of x :

$$\|x\|_0 = |\{i \in \{1, \dots, n\} \mid x_i \neq 0\}|,$$

the objective function $f \in C^2(\mathbb{R}^n, \mathbb{R})$ is twice continuously differentiable, and $s \in \{0, 1, \dots, n-1\}$ is an integer. The difficulty of solving SCNO comes from the combinatorial nature of the sparsity constraint $\|x\|_0 \leq s$. The requirement of sparsity is however motivated by various applications, such as compressed sensing, model selection, image processing etc. We refer e.g. to [7,25], and [22] for further details on the relevant applications.

In the seminal paper [1], necessary optimality conditions for SCNO have been stated. Namely, the notions of basic feasibility (BF-vector), L -stationarity and CW-minimality have been introduced and studied there. Note that the formulation of L -stationarity mimics the techniques from convex optimization by using the orthogonal projection on the SCNO feasible set. The notion of CW-minimum incorporates the coordinate-wise optimality along the axes. Based on both stationarity concepts, algorithms that find points satisfying these conditions have been developed. Those are the iterative hard thresholding method, as well as the greedy and partial sparse-simplex methods. In a series of subsequent papers [2,3] elaborated the algorithmic approach for SCNO which is based on L -stationarity and CW-minimality.

Another line of research started with [5], where additionally smooth equality and inequality constraints have been incorporated into SCNO. For that, the authors coin the new term of mathematical programs with cardinality constraints (MPCC). The key idea in [5] is to provide a mixed-integer formulation whose standard relaxation still has the same solutions as MPCC. For the relaxation the notion of S-stationary points is proposed. S-stationarity corresponds to the standard Karush–Kuhn–Tucker condition for the relaxed program. The techniques applied follow mainly those for mathematical programs with complementarity constraints. In particular, an appropriate regularization method for solving MPCC is suggested. The latter is proved to converge towards so-called M-stationary points. M-stationarity corresponds to the standard Karush–Kuhn–Tucker condition of the tightened program, where zero entries of an MPCC feasible point remain locally vanishing. Further research in this direction is presented in a series of subsequent papers [4,6].

Finally, we would like to mention stationarity concepts for SCNO based on the normal cones of the sparsity constrained feasible set. In [20], the Bouligand and Clarke normal cones of the SCNO feasible set are used to derive N^B - and N^C -stationarity, respectively. Corresponding second-order necessary and sufficient optimality conditions are stated there. These findings were generalized for MPCC in [19]. In [16], the Fréchet and Mordukhovich normal cones of the SCNO feasible set are used to derive \widehat{N} - and N -stationarity, respectively. These notions were generalized for the intersection of the sparsity constrained feasible set with a polyhedral set. In [17], a penalty decomposition method essentially based on the notion of N -stationarity is proposed for solving MPCC under the Robinson's constraint qualification.

The goal of this paper is the study of SCNO from a topological point of view. The topological approach to optimization has been pioneered by [12,13] for nonlinear programming problems, and successfully developed for mathematical programs with complementarity constraints, mathematical problems with vanishing constraints, general semi-infinite pro-

gramming, bilevel optimization, semi-definite programming, disjunctive programming etc., see e. g. [23] and references therein. The main idea of the topological approach is to identify stationary points which roughly speaking induce the global structure of the underlying optimization problem. The stationary points include minimizers, but also all kinds of saddle points—just in analogy to the unconstrained case. It turns out that for SCNO the concept of M-stationarity from [5]—coinciding with N^C -stationarity from [20]—is the adequate stationarity concept at least from the topological perspective. We outline our main findings and results:

1. We introduce nondegenerate M-stationary points along with their associated M-indices. The latter subsume as usual the quadratic part—the number of negative eigenvalues of the objective’s Hessian restricted to non-vanishing variables. As novelty, the sparsity constraint provides an addition to the M-index, namely, the difference between the bound and the current number of non-zero variables at a nondegenerate M-stationary point. We prove that all M-stationary points are generically nondegenerate. In particular, it follows that all local minimizers of SCNO are nondegenerate with vanishing M-index, hence, the sparsity constraint is active. Note that M-stationary points with non-vanishing M-index correspond to saddle points. The local structure of SCNO around a nondegenerate M-stationary point is fully described just by its M-index, at least up to a differentiable change of coordinates.
2. We thoroughly discuss the relation of M-stationarity to S-stationarity, basic feasibility, and CW-minimality for SCNO. It turns out that nondegenerate M-stationary points may cause degeneracies of S-stationary points viewed as Karush–Kuhn–Tucker-points for the relaxed problem. Moreover, even under the cardinality constrained second-order sufficient optimality condition from [4] assumed to hold at an S-stationary point, the corresponding M-stationary point does not need to be a nondegenerate local minimizer for SCNO. As for CW-minima, we show that they are not stable with respect to data perturbations in SCNO. After an arbitrarily small C^2 -perturbation of f a locally unique CW-minimum may bifurcate into multiple CW-minima. More importantly, this bifurcation unavoidably causes the emergence of M-stationary points, being different from the CW-minima. Despite of this instability phenomenon, if a BF-vector and, hence, CW-minimum, happens to be nondegenerate as an M-stationary point, then the sparsity constraint is necessarily active.
3. We use the concept of M-stationarity in order to describe the global structure of SCNO. To this aim the study of topological properties of its lower level sets is undertaken. As in the standard Morse theory, see e. g. [10,18], we focus on the topological changes of the lower level sets as their levels vary. Appropriate versions of deformation and cell-attachment theorems are shown to hold for SCNO. Whereas the deformation is standard, the cell-attachment reveals an essentially new phenomenon not observed in nonsmooth optimization before. In SCNO, *multiple* cells of the same dimension need to be attached, see Theorem 5. To determine the number of these attached cells turns out to constitute a challenging combinatorial problem from algebraic topology, see Lemma 1.
4. As a consequence of proposed Morse theory, we derive a Morse relation for SCNO, which relates the numbers of local minimizers and M-stationary points of M-index equal to one. The appearance of such saddle points cannot be thus neglected from the perspective of global optimization. As novelty for SCNO, a saddle point may lead to more than two different local minimizers. This is in strong contrast with other nonsmooth optimization problems studied before, see e. g. [23], where a saddle point leads to at most two of

them. We conclude that the relatively involved structure of saddle points is the source of well-known difficulty if solving SCNO to global optimality.

The paper is organized as follows. In Sect. 2 we discuss the notion of M-stationarity for SCNO. Section 3 is devoted to the relation of M-stationarity to other stationarity concepts from the literature. In Sect. 4 the global structure of SCNO is described within the scope of Morse theory.

Our notation is standard. The cardinality of a finite set S is denoted by $|S|$. The n -dimensional Euclidean space is denoted by \mathbb{R}^n with the coordinate vectors $e_i, i = 1, \dots, n$. For $J \subset \{1, \dots, n\}$ we denote by $\text{conv}(e_j, j \in J)$ and $\text{span}(e_j, j \in J)$ the convex and linear combination of the coordinate vectors $e_j, j \in J$, respectively. Given a twice continuously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, ∇f denotes its gradient, and $D^2 f$ stands for its Hessian.

2 M-stationarity

For $0 \leq k \leq n$ we use the notation

$$\mathbb{R}^{n,k} = \{x \in \mathbb{R}^n \mid \|x\|_0 \leq k\}.$$

Using the latter, the feasible set of SCNO can be written as

$$\mathbb{R}^{n,s} = \{x \in \mathbb{R}^n \mid \|x\|_0 \leq s\}.$$

For a feasible point $x \in \mathbb{R}^{n,s}$ we define the following complementary index sets:

$$I_0(x) = \{i \in \{1, \dots, n\} \mid x_i = 0\}, \quad I_1(x) = \{i \in \{1, \dots, n\} \mid x_i \neq 0\}.$$

Without loss of generality, we assume throughout the whole paper that at the particular point of interest $\bar{x} \in \mathbb{R}^{n,s}$ with $\|\bar{x}\|_0 = k$ it holds:

$$I_0(\bar{x}) = \{1, \dots, n - k\}, \quad I_1(\bar{x}) = \{n - k + 1, \dots, n\}.$$

Using this convention, the following local description of SCNO feasible set can be deduced. Let $\bar{x} \in \mathbb{R}^{n,s}$ be a feasible point for SCNO with $\|\bar{x}\|_0 = k$. Then, there exist neighborhoods $U_{\bar{x}}$ and V_0 of \bar{x} and 0, respectively, such that under the linear coordinate transformation $\Phi(x) = x - \bar{x}$ we have locally:

$$\Phi(\mathbb{R}^{n,s} \cap U_{\bar{x}}) = (\mathbb{R}^{n-k,s-k} \times \mathbb{R}^k) \cap V_0, \quad \Phi(\bar{x}) = 0. \tag{1}$$

Definition 1 (*M-stationarity*, [5]) A feasible point $\bar{x} \in \mathbb{R}^{n,s}$ is called M-stationary for SCNO if

$$\frac{\partial f}{\partial x_i}(\bar{x}) = 0 \text{ for all } i \in I_1(\bar{x}).$$

Obviously, a local minimizer of SCNO is an M-stationary point.

Definition 2 (*Nondegenerate M-stationarity*) An M-stationary point $\bar{x} \in \mathbb{R}^{n,s}$ with $\|\bar{x}\|_0 = k$ is called nondegenerate if the following conditions hold:

- ND1: if $k < s$ then $\frac{\partial f}{\partial x_i}(\bar{x}) \neq 0$ for all $i \in I_0(\bar{x})$,
- ND2: the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})\right)_{i,j \in I_1(\bar{x})}$ is nonsingular.

Otherwise, we call \bar{x} degenerate.

Definition 3 (*M-Index*) Let $\bar{x} \in \mathbb{R}^{n,s}$ be a nondegenerate M-stationary point with $\|\bar{x}\|_0 = k$. The number of negative eigenvalues of the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})\right)_{i,j \in I_1(\bar{x})}$ is called its quadratic index (*QI*). The number $s - k + QI$ is called the M-index of \bar{x} .

Theorem 1 (Morse-Lemma for SCNO) Suppose that \bar{x} is a nondegenerate M-stationary point for SCNO with $\|\bar{x}\|_0 = k$ and quadratic index *QI*. Then, there exist neighborhoods $U_{\bar{x}}$ and V_0 of \bar{x} and 0, respectively, and a local C^1 -coordinate system $\Psi : U_{\bar{x}} \rightarrow V_0$ of \mathbb{R}^n around \bar{x} such that:

$$f \circ \Psi^{-1}(y) = f(\bar{x}) + \sum_{i=1}^{n-k} y_i + \sum_{j=n-k+1}^n \pm y_j^2, \tag{2}$$

where $y \in \mathbb{R}^{n-k,s-k} \times \mathbb{R}^k$. Moreover, there are exactly *QI* negative squares in (2).

Proof Without loss of generality, we may assume $f(\bar{x}) = 0$. By using Φ from (1), we put $\bar{f} := f \circ \Phi^{-1}$ on the set $(\mathbb{R}^{n-k,s-k} \times \mathbb{R}^k) \cap V_0$. At the origin we have:

- (i) if $k < s$ then $\frac{\partial \bar{f}}{\partial y_i} \neq 0$ for all $i = 1, \dots, n - k$,
- (ii) $\frac{\partial \bar{f}}{\partial y_i} = 0$ for all $i = n - k + 1, \dots, n$,
- (iii) the matrix $\left(\frac{\partial^2 \bar{f}}{\partial y_i \partial y_j}\right)_{i,j=n-k+1,\dots,n}$ is nonsingular.

We denote \bar{f} by f again. Under the following coordinate transformations the set $\mathbb{R}^{n-k,s-k} \times \mathbb{R}^k$ will be equivariantly transformed in itself. We put $y = (Y_{n-k}, Y^k)$, where $Y_{n-k} = (y_1, \dots, y_{n-k})$ and $Y^k = (y_{n-k+1}, \dots, y_n)$. It holds:

$$f(Y_{n-k}, Y^k) = \int_0^1 \frac{d}{dt} f(tY_{n-k}, Y^k) dt + f(0, Y^k) = \sum_{i=1}^{n-k} y_i d_i(y) + f(0, Y^k),$$

where

$$d_i(y) = \int_0^1 \frac{\partial f}{\partial y_i}(tY_{n-k}, Y^k) dt, \quad i = 1, \dots, n - k.$$

Note that $d_i \in C^1, i = 1, \dots, n - k$. Due to (ii)-(iii), we may apply the standard Morse lemma on the C^2 -function $f(0, Y^k)$ without affecting the coordinates Y_{n-k} , see e. g. [13]. The corresponding coordinate transformation is of class C^1 . Denoting the transformed functions again by f and d_i , we obtain

$$f(y) = \sum_{i=1}^{n-k} y_i d_i(y) + \sum_{j=n-k+1}^n \pm y_j^2.$$

In case $k = s$, we need to consider f locally around the origin on the set

$$\mathbb{R}^{n-k,s-k} \times \mathbb{R}^k = \mathbb{R}^{n-k,0} \times \mathbb{R}^k = \{0\}^{n-k} \times \mathbb{R}^k.$$

Hence, $y_i = 0$ for $i = 1, \dots, n - k$, and we immediately obtain the representation (2).

In case $k < s$, (i) provides that $d_i(0) = \frac{\partial f}{\partial y_i}(0) \neq 0, i = 1, \dots, n - k$. Hence, we may take

$$y_i d_i(y), i = 1, \dots, n - k, \quad y_j, j = n - k + 1, \dots, n$$

as new local C^1 -coordinates by a straightforward application of the inverse function theorem. Denoting the transformed function again by f , we obtain (2). Here, the coordinate transformation Ψ is understood as the composition of all previous ones. \square

Proposition 1 (Nondegenerate minimizers) *Let \bar{x} be a nondegenerate M-stationary point for SCNO. Then, \bar{x} is a local minimizer for SCNO if and only if its M-index vanishes.*

Proof Let \bar{x} be a nondegenerate M-stationary point for SCNO. The application of Morse Lemma from Theorem 1 says that there exist neighborhoods $U_{\bar{x}}$ and V_0 of \bar{x} and 0, respectively, and a local C^1 -coordinate system $\Psi : U_{\bar{x}} \rightarrow V_0$ of \mathbb{R}^n around \bar{x} such that:

$$f \circ \Psi^{-1}(y) = f(\bar{x}) + \sum_{i=1}^{n-k} y_i + \sum_{j=n-k+1}^n \pm y_j^2, \tag{3}$$

where $y \in \mathbb{R}^{n-k, s-k} \times \mathbb{R}^k$. Therefore, \bar{x} is a local minimizer for SCNO if and only if 0 is a local minimizer of $f \circ \Psi^{-1}$ on the set $(\mathbb{R}^{n-k, s-k} \times \mathbb{R}^k) \cap V_0$. If the M-index of \bar{x} vanishes, we have $k = s$ and $QI = 0$, and (3) reads as

$$f \circ \Psi^{-1}(y) = f(\bar{x}) + \sum_{j=n-s+1}^n y_j^2, \tag{4}$$

where $y \in \{0\}^{n-s} \times \mathbb{R}^s$. Thus, 0 is a local minimizer for (4). Vice versa, if 0 is a local minimizer for (3), then obviously $k = s$ and $QI = 0$, hence, the M-index of \bar{x} vanishes. \square

Let $C^2(\mathbb{R}^n, \mathbb{R})$ be endowed with the strong (or Whitney) C^2 -topology, denoted by C_s^k (see e. g. [11]). The C_s^k -topology is generated by allowing perturbations of the functions, their gradients and Hessians, which are controlled by means of continuous positive functions. We say that a set is C_s^2 -generic if it contains a countable intersection of C_s^2 -open and -dense subsets. Since $C^2(\mathbb{R}^n, \mathbb{R})$ endowed with the C_s^2 -topology is a Baire space, generic sets are in particular dense.

Theorem 2 (Genericity for SCNO) *Let $\mathcal{F} \subset C^2(\mathbb{R}^n, \mathbb{R})$ denote the subset of objective functions in SCNO for which each M-stationary point is nondegenerate. Then, \mathcal{F} is C_s^2 -open and -dense.*

Proof Let us fix a number of non-zero entries $k \in \{0, \dots, s\}$, an index set of k non-zero entries $D \subset \{1, \dots, n\}$, i.e. $|D| = k$, an index subset of zero entries $E \subset \{1, \dots, n\} \setminus D$, and a rank $r \in \{0, \dots, k\}$. For this choice we consider the set $\Gamma_{k,D,E,r}$ of x such that the following conditions are satisfied:

- (m1) $x_i \neq 0$ for all $i \in D$, and $x_i = 0$ for all $i \in \{1, \dots, n\} \setminus D$,
- (m2) $\frac{\partial f}{\partial x_i}(x) = 0$ for all $i \in D$,
- (m3) if $k < s$ then $\frac{\partial f}{\partial x_i}(x) = 0$ for all $i \in E$,
- (m4) the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i,j \in D}$ has rank r .

Note that (m1) refers to feasibility, (m2) to M-stationarity, and (m3)-(m4) describe possible violations of ND1-ND2, respectively.

Now, it suffices to show that all $\Gamma_{k,D,E,r}$ are generically empty whenever E is nonempty or the rank r is less than k . By setting $I_1(x) = D$ and $I_0(x) = \{1, \dots, n\} \setminus D$, this would mean, respectively, that at least one of the derivatives $\frac{\partial f}{\partial x_i}(x)$ vanishes for $i \in E \subset I_0(x)$ in

ND1 if $k < s$, or the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i,j \in I_1(x)}$ is singular in ND2. In fact, the available

degrees of freedom of the variables involved in each $\Gamma_{k,D,E,r}$ are n . The loss of freedom caused by (m1) is $n - k$, and the loss of freedom caused by (m2) is k . Hence, the total loss of freedom is n . We conclude that a further nondegeneracy would exceed the total available degrees of freedom n . By virtue of the jet transversality theorem from [13], generically the sets $\Gamma_{k,D,E,r}$ must be empty.

For the openness result, we argue in a standard way. Locally, M-stationarity can be written via stable equations. Then, the implicit function theorem for Banach spaces can be applied to follow M-stationary points with respect to (local) C^2 -perturbations of defining functions. Finally, a standard globalization procedure exploiting the specific properties of the strong C_s^2 -topology can be used to construct a (global) C_s^2 -neighborhood of problem data for which the nondegeneracy property is stable. □

Theorem 3 (Genericity for minimizers) *Generically, all minimizers of SCNO are nondegenerate with the vanishing M-index.*

Proof Note that every local minimizer of SCNO has to be M-stationary. Nondegenerate M-stationary points are generic by Theorem 2. Hence, generically, local minimizers are nondegenerate. Due to Proposition 1, they have vanishing M-index. □

By recalling Definition 3 of M-index, we deduce the following important Corollary 1 on the structure of minimizers for SCNO.

Corollary 1 (Sparsity constraint at minimizers) *At each generic local minimizer $\bar{x} \in \mathbb{R}^{n,s}$ of SCNO the sparsity constraint is active, i.e. $\|\bar{x}\|_0 = s$.*

3 Relation to other stationarity concepts

We relate M-stationarity to other well-known stationarity concepts for SCNO from the literature. First, we focus on S-stationarity introduced in [5]. Then, the notions of basic feasibility and CW-minimality from [1] will be discussed.

3.1 S-stationarity

In [5] the following observation has been made: \bar{x} solves SCNO if and only if there exists \bar{y} such that (\bar{x}, \bar{y}) solves the following mixed-integer program:

$$\min_{x,y} f(x) \quad \text{s.t.} \quad \sum_{i=1}^n y_i \geq n - s, \quad y_i \in \{0, 1\}, \quad x_i y_i = 0, \quad i = 1, \dots, n. \quad (5)$$

Using the standard relaxation of the binary constraints $y_i \in \{0, 1\}$, the authors arrive at the following continuous optimization problem:

$$\min_{x,y} f(x) \quad \text{s. t.} \quad \sum_{i=1}^n y_i \geq n - s, \quad y_i \in [0, 1], \quad x_i y_i = 0, \quad i = 1, \dots, n. \quad (6)$$

As pointed out in [5], SCNO and the optimization problem (6) are closely related: \bar{x} solves SCNO if and only if there exists a vector \bar{y} such that (\bar{x}, \bar{y}) solves (6). Additionally, the concept of S-stationarity is proposed for (6). For its formulation the following index sets are needed:

$$\begin{aligned} I_{\pm 0}(\bar{x}, \bar{y}) &= \{i \in \{1, \dots, n\} \mid \bar{x}_i \neq 0, \bar{y}_i = 0\}, \\ I_{00}(\bar{x}, \bar{y}) &= \{i \in \{1, \dots, n\} \mid \bar{x}_i = 0, \bar{y}_i = 0\}. \end{aligned}$$

Definition 4 (*S-stationarity*, [5]) A feasible point (\bar{x}, \bar{y}) of (6) is called S-stationary if there exist real multipliers $\gamma_1, \dots, \gamma_n$, such that

$$\nabla f(\bar{x}) + \sum_i^n \gamma_i e_i = 0, \quad \gamma_i = 0 \text{ for all } i \in I_{\pm 0}(\bar{x}, \bar{y}), \quad (7)$$

and, additionally, it holds:

$$\gamma_i = 0 \text{ for all } i \in I_{00}(\bar{x}, \bar{y}).$$

Remark 1 (M-stationarity) We point out that initially [5] defined the concept of M-stationarity for the relaxed optimization problem (6). Namely, a feasible point (\bar{x}, \bar{y}) of (6) is called M-stationary if just (7) is valid. Due to the feasibility of (\bar{x}, \bar{y}) , we have $\bar{y}_i = 0$ if $\bar{x}_i \neq 0$ for all $i = 1, \dots, n$. Hence, it holds:

$$I_{\pm 0}(\bar{x}, \bar{y}) = I_1(\bar{x}),$$

and M-stationarity is independent from the auxiliary variable \bar{y} . Thus, already in [4] it is sometimes said that a feasible point \bar{x} of SCNO is M-stationary itself. We use M-stationarity exactly in this sense, cf. Definition 1. □

In order to relate M- and S-stationarity, we introduce the canonical choice of the auxiliary variables \bar{y} for a feasible point \bar{x} of SCNO:

$$\bar{y}_i = \begin{cases} 0, & \text{if } i \in I_1(\bar{x}), \\ 1, & \text{if } i \in I_0(\bar{x}). \end{cases} \quad (8)$$

The auxiliary variables \bar{y} can be seen as counters of the zero elements of \bar{x} . Note that (\bar{x}, \bar{y}) becomes feasible for (6).

Proposition 2 (M- and S-stationarity) *If (\bar{x}, \bar{y}) is S-stationary for (6) then \bar{x} is M-stationary for SCNO. Vice versa, for any M-stationary point \bar{x} the canonical choice (8) of auxiliary variables \bar{y} provides an S-stationary point (\bar{x}, \bar{y}) for (6).*

Proof Let (\bar{x}, \bar{y}) be S-stationary for (6). After a moment of reflection we see that $I_{\pm 0}(\bar{x}, \bar{y}) = I_1(\bar{x})$ is the support of \bar{x} , and (7) reads as the M-stationarity of \bar{x} :

$$\nabla_i f(\bar{x}) = 0 \text{ for all } i \in I_1(\bar{x}).$$

Vice versa, let \bar{x} be an M-stationary point for SCNO with the canonical choice (8) of \bar{y} . Then, (\bar{x}, \bar{y}) is feasible for (6), since

$$\sum_{i=1}^n \bar{y}_i = |I_0(\bar{x})| = n - |I_1(\bar{x})| \geq n - s.$$

The last inequality is due to $\|\bar{x}\|_0 \leq s$ or, equivalently, $|I_1(\bar{x})| \leq s$. Moreover, by the choice of \bar{y} we have $I_{\pm 0}(\bar{x}, \bar{y}) = I_1(\bar{x})$ and $I_{00}(\bar{x}, \bar{y}) = \emptyset$. Thus, due to the M-stationarity of \bar{x} , (7) is fulfilled, and (\bar{x}, \bar{y}) is S-stationary. \square

The importance of S-stationary points is due to the following Proposition 3.

Proposition 3 (S-stationarity and KKT-points, [5]) *A feasible point (\bar{x}, \bar{y}) satisfies the Karush–Kuhn–Tucker condition if and only if it is S-stationary for (6).*

Despite this appealing relation, nondegenerate M-stationary points of SCNO may cause degeneracies of the corresponding S-stationary points. This means that they become degenerate Karush–Kuhn–Tucker-points for (6), i. e. the linear independent constraint qualification is not fulfilled, strict complementarity is violated, or the second derivative of the corresponding Lagrange function restricted to the tangential space becomes singular. The appearance of these degeneracies is mainly due to the fact that the objective function in (6) does not depend on y -variables. We illustrate this phenomenon by means of the following Example 1.

Example 1 (S-stationarity and degeneracies) We consider the following SCNO with $n = 2$ and $s = 1$:

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

It is easy to see that the feasible point $\bar{x} = (0, 0)$ is M-stationary with $\|\bar{x}\|_0 = k = 0$. Moreover, it is nondegenerate with quadratic index $QI = 0$. For its M-index we have

$$s - k + QI = 1 - 0 + 0 = 1,$$

meaning that \bar{x} is a saddle point which connects two minimizers $(1, 0)$ and $(0, 1)$. Further, by the canonical choice (8) of auxiliary y -variables, we obtain the corresponding S-stationary point $(\bar{x}, \bar{y}) = (0, 0, 1, 1)$. Due to Proposition 3, (\bar{x}, \bar{y}) is also a Karush–Kuhn–Tucker-point for the optimization problem (6):

$$\min_{x, y} (x_1 - 1)^2 + (x_2 - 1)^2 \quad \text{s. t.} \quad y_1 + y_2 \geq 1, \quad y_1, y_2 \in [0, 1], \quad x_1 y_1 = 0, \quad x_2 y_2 = 0.$$

The gradients of the active constraints at (\bar{x}, \bar{y}) are linearly independent:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ 0 \\ \bar{x}_1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \bar{y}_2 \\ 0 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Hence, the linear independent constraint qualification holds at (\bar{x}, \bar{y}) . Let us determine the unique Lagrange multipliers from the Karush–Kuhn–Tucker condition:

$$\begin{pmatrix} 2(\bar{x}_1 - 1) \\ 2(\bar{x}_2 - 1) \\ 0 \\ 0 \end{pmatrix} = \mu_1 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mu_1, \mu_2 \leq 0.$$

We get $\mu_1 = \mu_2 = 0$ and $\lambda_1 = \lambda_2 = -2$. Hence, the strict complementarity is violated at (\bar{x}, \bar{y}) . Finally, the tangential space on the feasible set vanishes at (\bar{x}, \bar{y}) . Hence, the second derivative of the corresponding Lagrange function restricted to the tangential space is trivially nonsingular. Overall, we claim that (\bar{x}, \bar{y}) is a degenerate Karush–Kuhn–Tucker-point for (6) due to the lack of strict complementarity. It remains to note that the degeneracy of S-stationary points (\bar{x}, y) prevails if other choices of auxiliary y -variables are made. \square

An attempt to define a tailored notion of nondegeneracy for S-stationary points of (6) has been recently undertaken in [4]. Let us briefly recall their main idea. For that, the so-called CC-linearization cone $\mathcal{L}^{CC}(\bar{x}, \bar{y})$ at a feasible point (\bar{x}, \bar{y}) of (6) is used, cf. [6]. Namely,

$$(d_x, d_y) \in \mathcal{L}^{CC}(\bar{x}, \bar{y}) \subset \mathbb{R}^n \times \mathbb{R}^n$$

satisfies by definition the following conditions:

$$\left\{ \begin{array}{l} \sum_{i=1}^n (d_y)_i \geq 0 \text{ if } \sum_{i=1}^n \bar{y}_i = n - s, \\ (d_y)_i = 0 \text{ for all } i \in I_{\pm 0}(\bar{x}, \bar{y}), \\ (d_y)_i \geq 0 \text{ for all } i \in I_{00}(\bar{x}, \bar{y}), \\ (d_y)_i \leq 0 \text{ for all } i \in I_{01}(\bar{x}, \bar{y}), \\ (d_x)_i = 0 \text{ for all } i \in I_{\pm 0}(\bar{x}, \bar{y}) \cup I_{0+}(\bar{x}, \bar{y}), \\ (d_x)_i (d_y)_i = 0 \text{ for all } i \in I_{00}(\bar{x}, \bar{y}). \end{array} \right. \tag{9}$$

Here, the new index sets are

$$\begin{aligned} I_{01}(\bar{x}, \bar{y}) &= \{i \in \{1, \dots, n\} \mid \bar{x}_i = 0, \bar{y}_i = 1\}, \\ I_{0+}(\bar{x}, \bar{y}) &= \{i \in \{1, \dots, n\} \mid \bar{x}_i = 0, \bar{y}_i \in (0, 1)\}. \end{aligned}$$

Definition 5 (CC-SOSC, [4]) Let (\bar{x}, \bar{y}) be an S-stationary point for (6). If for all directions $(d_x, d_y) \in \mathcal{L}^{CC}(\bar{x}, \bar{y})$ with $d_x \neq 0$, we have

$$d_x^T \cdot D^2 f(\bar{x}) \cdot d_x > 0,$$

then the cardinality constrained second-order sufficient optimality condition (CC-SOSC) is said to hold at (\bar{x}, \bar{y}) .

The role of CC-SOSC can be seen from the following Proposition 4.

Proposition 4 (Sufficient optimality condition, [4]) Let (\bar{x}, \bar{y}) be an S-stationary point for (6) satisfying CC-SOSC. Then, (\bar{x}, \bar{y}) is a strict local minimizer of (6) with respect to x , i. e.

$$f(\bar{x}) < f(x)$$

for all feasible points (x, y) of (6) taken sufficiently close to (\bar{x}, \bar{y}) , and fulfilling $x \neq \bar{x}$.

We relate the concepts of nondegeneracy for M-stationary points and of CC-SOSC for S-stationary points. Next, Proposition 5 mainly follows from Corollary 3.2 a) in [4]. We prove it here for the sake of completeness.

Proposition 5 (Nondegeneracy and CC-SOSC) Let \bar{x} be an M-stationary point for SCNO with $\|\bar{x}\|_0 = s$. Assume that CC-SOSC holds at the S-stationary point (\bar{x}, \bar{y}) for (6) with the canonical choice (8) of auxiliary variables \bar{y} . Then, \bar{x} is a nondegenerate local minimizer for SCNO.

Proof By Proposition 4, (\bar{x}, \bar{y}) is a local minimizer of (6) with respect to x . For all $x \in \mathbb{R}^{n,s}$ sufficiently close to \bar{x} we have because of $\|\bar{x}\|_0 = s$ that (x, \bar{y}) is feasible for (6). Thus, \bar{x} is a local minimizer for SCNO. Due to the canonical choice (8) of auxiliary variables \bar{y} , the index sets from the definition of the CC-linearization cone $\mathcal{L}^{CC}(\bar{x}, \bar{y})$ are

$$I_{\pm 0}(\bar{x}, \bar{y}) = I_1(\bar{x}), \quad I_{00}(\bar{x}, \bar{y}) = I_{0+}(\bar{x}, \bar{y}) = \emptyset, \quad I_{01}(\bar{x}, \bar{y}) = I_0(\bar{x}).$$

Due to $\|\bar{x}\|_0 = s$, we additionally have $\sum_{i=1}^n \bar{y}_i = n - s$. Recalling (9), $(d_x, d_y) \in \mathcal{L}^{CC}(\bar{x}, \bar{y})$ if and only if

$$\begin{cases} \sum_{i=1}^n (d_y)_i \geq 0, \\ (d_y)_i = 0 \text{ for all } i \in I_1(\bar{x}), \\ (d_y)_i \leq 0 \text{ for all } i \in I_0(\bar{x}), \\ (d_x)_i = 0 \text{ for all } i \in I_0(\bar{x}). \end{cases}$$

Hence, it holds:

$$\mathcal{L}^{CC}(\bar{x}, \bar{y}) = \{(d_x, 0) \mid (d_x)_i = 0 \text{ for all } i \in I_0(\bar{x})\},$$

so that CC-SOSC says that the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})\right)_{i,j \in I_1(\bar{x})}$ is positive definite. Hence, the minimizer \bar{x} is nondegenerate. □

If the sparsity constraint is not active for an M-stationary point \bar{x} of SCNO, i.e. $\|\bar{x}\|_0 < s$, the implication in Proposition 5 does not hold in general anymore. Namely, \bar{x} does not need to be a local minimizer for SCNO, even if CC-SOSC holds at the corresponding S-stationary point (\bar{x}, \bar{y}) with the canonical choice (8) of auxiliary variables \bar{y} . This is illustrated by means of the following Example 2.

Example 2 (Sparsity constraint and CC-SOSC) We consider the following SCNO with $n = 2$ and $s = 1$:

$$\min_{x_1, x_2} x_1 + x_2 \quad \text{s.t.} \quad \|(x_1, x_2)\|_0 \leq 1.$$

It is easy to see that the feasible point $\bar{x} = (0, 0)$ is M-stationary. Note that the sparsity constraint is not active for \bar{x} , since $k = \|\bar{x}\|_0 = 0 < 1 = s$. By the canonical choice (8) of auxiliary y -variables, we obtain the corresponding S-stationary point $(\bar{x}, \bar{y}) = (0, 0, 1, 1)$. Analogously to the proof of Proposition 4 and by recalling (9), $(d_x, d_y) \in \mathcal{L}^{CC}(\bar{x}, \bar{y})$ if and only if

$$\begin{cases} (d_y)_i = 0 \text{ for all } i \in I_1(\bar{x}), \\ (d_y)_i \leq 0 \text{ for all } i \in I_0(\bar{x}), \\ (d_x)_i = 0 \text{ for all } i \in I_0(\bar{x}). \end{cases}$$

Note that here $I_1(\bar{x}) = \emptyset$ and $I_0(\bar{x}) = \{1, 2\}$. Hence, the CC-linearization cone is

$$\mathcal{L}^{CC}(\bar{x}, \bar{y}) = \{(0, d_y) \mid (d_y)_1, (d_y)_2 \leq 0\}.$$

Overall, CC-SOSC trivially holds at (\bar{x}, \bar{y}) , and as follows from Proposition 4, it is a strict local minimizer of (6) with respect to x . Nevertheless, \bar{x} is not a local minimizer. Actually,

it is a nondegenerate M-stationary point with the quadratic index $QI = 0$. For its M-index we have

$$s - k + QI = 1 - 0 + 0 = 1.$$

We conclude that \bar{x} is rather a saddle point for SCNO. □

3.2 Basic feasibility and CW-minimality

We proceed by discussing stationarity concepts from [1]. Inspired by linear programming terminology, they first introduce the notion of a basic feasible vector for SCNO.

Definition 6 (*Basic feasibility*, [1]) A vector $\bar{x} \in \mathbb{R}^{n,s}$ with $\|\bar{x}\|_0 = k$ is called basic feasible (BF) for SCNO if the following conditions are fulfilled:

BF1: in case $k < s$, it holds:

$$\frac{\partial f}{\partial x_i}(\bar{x}) = 0 \text{ for all } i = 1, \dots, n,$$

BF2: in case $k = s$, it holds:

$$\frac{\partial f}{\partial x_i}(\bar{x}) = 0 \text{ for all } i \in I_1(\bar{x}).$$

Attention has been also paid to the notion of coordinate-wise minimum for SCNO.

Definition 7 (*CW-minimality*, [1]) A vector $\bar{x} \in \mathbb{R}^{n,s}$ with $\|\bar{x}\|_0 = k$ is called coordinate-wise (CW) minimum for SCNO if the following conditions are fulfilled:

– CW1: in case $k < s$, it holds:

$$f(\bar{x}) = \min_{t \in \mathbb{R}} f(\bar{x} + te_i) \text{ for all } i = 1, \dots, n,$$

– CW2: in case $k = s$, it holds:

$$f(\bar{x}) \leq \min_{t \in \mathbb{R}} f(\bar{x} - \bar{x}_i e_i + te_j) \text{ for all } i \in I_1(\bar{x}) \text{ and } j = 1, \dots, n.$$

Basic feasibility and CW-minimality can be viewed as necessary optimality condition for SCNO.

Proposition 6 (BF-vector and CW-minimum, [1]) *Every global minimizer for SCNO is a CW-minimum, and every CW-minimum for SCNO is a BF-vector.*

It is claimed in [1] that the basic feasibility condition is quite weak, namely, there are many BF-points that are not optimal for SCNO. The notion of CW-minimum provides a much stricter necessary optimality condition. Based on the latter, a greedy sparse-simplex method for the numerical treatment of SCNO is proposed by [1]. Let us now examine the relation between M-stationarity, basic feasibility, and CW-minimality.

Proposition 7 (M-stationarity, BF-vector, and CW-minimum) *Every BF-vector for SCNO is an M-stationary point, in particular, so is every CW-minimum.*

Proof Let \bar{x} be a BF-vector for SCNO with $\|\bar{x}\|_0 = k$. If $k < s$, then BF1 implies M-stationarity of \bar{x} . If $k = s$, then BF2 coincides with the latter property. Since every CW-minimum for SCNO is a BF-vector according to Proposition 6, the assertion follows. □

Proposition 7 says that M-stationarity is an even weaker condition than basic feasibility and CW-minimality. Why should we care about M-stationarity then? Is it not enough to rather focus on the stricter necessary optimality condition of CW-minimality as in [1]? It turns out that CW-minima need not to be stable with respect to data perturbations. Namely, after an arbitrarily small C^2 -perturbation of f a locally unique CW-minimum may bifurcate into multiple CW-minima. More importantly, this bifurcation unavoidably causes the emergence of M-stationary points, being different from CW-minima. Next Example 3 illustrates this instability phenomenon.

Example 3 (CW-minimum and instability) We consider the following SCNO with $n = 2$ and $s = 1$:

$$\min_{x_1, x_2} x_1^2 + x_2^2 \quad \text{s. t.} \quad \|(x_1, x_2)\|_0 \leq 1. \tag{10}$$

Obviously, $\bar{x} = (0, 0)$ is the unique minimizer of (10). Due to Proposition 6, it is also a CW-minimum, as well as a BF-vector. Further, let us perturb (10) by using an arbitrarily small $\varepsilon > 0$ as follows:

$$\min_{x_1, x_2} (x_1 - \varepsilon)^2 + (x_2 - \varepsilon)^2 \quad \text{s. t.} \quad \|(x_1, x_2)\|_0 \leq 1. \tag{11}$$

It is easy to see that the perturbed problem (11) has now two solutions $\bar{x}^1 = (\varepsilon, 0)$ and $\bar{x}^2 = (0, \varepsilon)$. Both are CW-minima, and, hence, BF-points. Here, we observe a bifurcation of the CW-minimum \bar{x} of the original problem (10) into two CW-minima \bar{x}^1 and \bar{x}^2 of the perturbed problem (11). Let us explain this bifurcation in terms of M-stationarity. The bifurcation is caused by the degeneracy of \bar{x} viewed as an M-stationary point of the original problem (10). Note that ND1 is violated at the M-stationary point \bar{x} of the original problem (10). More interestingly, although \bar{x} is neither a CW-minimum nor a BF-vector of (11) anymore, it becomes a new M-stationary point for the perturbed problem. In fact, due to $\|\bar{x}\|_0 = k = 0$ and the validity of ND1, \bar{x} is a nondegenerate M-stationary point of (11) with the quadratic index $QI = 0$. For its M-index we have

$$s - k + QI = 1 - 0 + 0 = 1,$$

meaning that \bar{x} is a saddle point which connects two nondegenerate minimizers \bar{x}^1 and \bar{x}^2 of (11). Overall, we conclude that the degenerate CW-minimum \bar{x} of the original problem (10) is not stable. Moreover, it bifurcates into two nondegenerate CW-minima \bar{x}^1 and \bar{x}^2 , as well as leads to one nondegenerate saddle point \bar{x} of the perturbed problem (10). □

Example 3 suggests to consider nondegenerate BF-vectors or nondegenerate CW-minima for SCNO, in order to guarantee their stability with respect to sufficiently small data perturbations. Then, however, the sparsity constraint turns out to be active. This means that BF1 in Definition 6 and CW1 in Definition 7 become redundant.

Proposition 8 (BF-vector, CW-minimum and nondegeneracy) *Let \bar{x} be a BF-vector for SCNO with $\|\bar{x}\|_0 = k$. If it is nondegenerate as an M-stationary point for SCNO, then $k = s$. The same applies for CW-minima.*

Proof Assume that $k < s$, then ND1 contradicts BF1, whenever $I_0(\bar{x}) \neq \emptyset$. Otherwise, we have $k = n$, and, hence, $n < s$, a contradiction. It remains to note that every CW-minimum for SCNO is a BF-vector due to Proposition 6. □

3.3 Normal cone stationarity

In [20], the Bouligand and Clarke normal cones of the SCNO feasible set are used to derive corresponding stationarity concepts. Let $N_{\mathbb{R}^{n,s}}^B(\bar{x})$ stand for the Bouligand and $N_{\mathbb{R}^{n,s}}^C(\bar{x})$ for the Clarke normal cone of $\mathbb{R}^{n,s}$ at \bar{x} , see e. g. [21] for details.

Definition 8 (*N^B - and N^C -stationarity*, [20]) A feasible point $\bar{x} \in \mathbb{R}^{n,s}$ is called N^B - and N^C -stationary for SCNO if it respectively holds:

$$0 \in \nabla f(\bar{x}) + N_{\mathbb{R}^{n,s}}^B(\bar{x}) \quad \text{and} \quad 0 \in \nabla f(\bar{x}) + N_{\mathbb{R}^{n,s}}^C(\bar{x}).$$

We relate the normal cone stationarity to the previously discussed concepts in the context of SCNO. As consequence, N^B - and N^C -stationarity can be viewed as necessary optimality conditions for SCNO. Note that the equivalence of basic feasibility and N^B -stationarity for SCNO has been already mentioned in [20].

Proposition 9 *The notions of basic feasibility and N^B -stationarity for SCNO coincide, so do the notions of M - and N^C -stationarity.*

Proof Theorems 2.1 and 2.2 from [20] provide explicit formulas for the Bouligand and Clarke normal cones of $\mathbb{R}^{n,s}$ at a SCNO feasible point \bar{x} with $\|\bar{x}\|_0 = k$:

$$N_{\mathbb{R}^{n,s}}^B(\bar{x}) = \begin{cases} \{0\}, & \text{if } k < s, \\ \text{span}\{e_i \mid i \in I_0(\bar{x})\}, & \text{if } k = s \end{cases} \quad \text{and} \quad N_{\mathbb{R}^{n,s}}^C(\bar{x}) = \text{span}\{e_i \mid i \in I_0(\bar{x})\}.$$

Thus, the conclusion immediately follows. □

Let us now comment on the second-order sufficient condition introduced in [20] for N^C -stationary points. For that, we denote by $T_{\mathbb{R}^{n,s}}^C$ the Clarke tangential cone of $\mathbb{R}^{n,s}$ at \bar{x} , see e. g. [21] for details.

Proposition 10 (Second-order sufficient optimality, [20]) *Let \bar{x} be an N^C -stationary point for SCNO. Assume the second-order sufficient condition (SOSC) to hold at \bar{x} :*

$$d^T \cdot D^2 f(\bar{x}) \cdot d > 0 \text{ for all } d \in T_{\mathbb{R}^{n,s}}^C(\bar{x}) \text{ with } d \neq 0. \tag{12}$$

Then, \bar{x} is a strict local minimizer of f restricted to the set $\mathbb{R}^{n,s} \cap \text{span}\{e_i \mid i \in I_1(\bar{x})\}$.

It turns out that CC-SOSC from [4] and SOSC from [20] are equivalent.

Proposition 11 (CC-SOSC and SOSC) *SOSC holds at an M -stationary point \bar{x} for SCNO if and only if CC-SOSC holds at the S -stationary point (\bar{x}, \bar{y}) for (6) with the canonical choice (8) of auxiliary variables \bar{y} .*

Proof Due to the canonical choice (8) of auxiliary variables \bar{y} , the index sets from the definition of the CC-linearization cone $\mathcal{L}^{CC}(\bar{x}, \bar{y})$ are

$$I_{\pm 0}(\bar{x}, \bar{y}) = I_1(\bar{x}), \quad I_{00}(\bar{x}, \bar{y}) = I_{0+}(\bar{x}, \bar{y}) = \emptyset, \quad I_{01}(\bar{x}, \bar{y}) = I_0(\bar{x}).$$

Case 1: $\|\bar{x}\|_0 = s$. Then, we additionally have $\sum_{i=1}^n \bar{y}_i = n - s$. Recalling (9), $(d_x, d_y) \in \mathcal{L}^{CC}(\bar{x}, \bar{y})$ if and only if

$$\begin{cases} \sum_{i=1}^n (d_y)_i \geq 0, \\ (d_y)_i = 0 \text{ for all } i \in I_1(\bar{x}), \\ (d_y)_i \leq 0 \text{ for all } i \in I_0(\bar{x}), \\ (d_x)_i = 0 \text{ for all } i \in I_0(\bar{x}). \end{cases}$$

Hence, it holds:

$$\mathcal{L}^{CC}(\bar{x}, \bar{y}) = \{(d_x, 0) \mid (d_x)_i = 0 \text{ for all } i \in I_0(\bar{x})\}.$$

Case 2: $\|\bar{x}\|_0 < s$. Then, we additionally have $\sum_{i=1}^n \bar{y}_i > n - s$. Recalling (9), $(d_x, d_y) \in \mathcal{L}^{CC}(\bar{x}, \bar{y})$ if and only if

$$\begin{cases} (d_y)_i = 0 \text{ for all } i \in I_1(\bar{x}), \\ (d_y)_i \leq 0 \text{ for all } i \in I_0(\bar{x}), \\ (d_x)_i = 0 \text{ for all } i \in I_0(\bar{x}). \end{cases}$$

Hence, it holds:

$$\mathcal{L}^{CC}(\bar{x}, \bar{y}) = \{(d_x, d_y) \mid (d_x)_i = 0 \text{ and } (d_y)_i \leq 0 \text{ for all } i \in I_0(\bar{x}), (d_y)_i = 0 \text{ for all } i \in I_1(\bar{x})\}.$$

In both case, CC-SOSC is to say that the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})\right)_{i,j \in I_1(\bar{x})}$ is positive definite. This is exactly what SOSC requires. In fact, Theorem 2.2 from [20] gives the explicit representation of the Clarke tangential cone of $\mathbb{R}^{n,s}$ at \bar{x} :

$$T_{\mathbb{R}^{n,s}}^C(\bar{x}) = \text{span}\{e_i \mid i \in I_1(\bar{x})\}.$$

Now, the assertion follows due to Proposition 2. □

We can easily relate the concepts of nondegeneracy and SOSC.

Proposition 12 (Nondegeneracy and SOSC) *Let \bar{x} be an M-stationary point for SCNO with $\|\bar{x}\|_0 = s$. Assume that SOSC holds at \bar{x} . Then, \bar{x} is a nondegenerate local minimizer for SCNO.*

Proof Due to Proposition 10, \bar{x} is a local minimizer of f on the set $\mathbb{R}^{n,s} \cap \text{span}\{e_i \mid i \in I_1(\bar{x})\}$. Since $\|\bar{x}\|_0 = s$, we have for all $x \in \mathbb{R}^{n,s}$ sufficiently close to \bar{x} that $I_1(x) = I_1(\bar{x})$. Hence, \bar{x} is actually a local minimizer for SCNO. Moreover, it is nondegenerate, because SOSC means that the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})\right)_{i,j \in I_1(\bar{x})}$ is positive definite. □

If the sparsity constraint is not active for an M-stationary point \bar{x} of SCNO, i.e. $\|\bar{x}\|_0 < s$, the implication in Proposition 12 does not hold in general anymore. Namely, \bar{x} does not need to be a local minimizer for SCNO, even in presence of SOSC. We note that this observation has been already made in Example 2.12 from [20]. Let us reconsider this example by using the notion of nondegeneracy.

Example 4 (Sparsity constraint and SOSC, [20]) We consider the following SCNO with $n = 3$ and $s = 2$:

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

It is easy to see that the feasible point $\bar{x} = (0, 0, 1)$ is M-stationary. Note that the sparsity constraint is not active for \bar{x} , since $k = \|\bar{x}\|_0 = 1 < 2 = s$. Here, $I_1(\bar{x}) = \{3\}$ and, hence,

$$T_{\mathbb{R}^{3,2}}^C(\bar{x}) = \text{span}\{e_i \mid i \in I_1(\bar{x})\} = \text{span}\{e_3\}.$$

Overall, SOSC holds at \bar{x} , but \bar{x} is not a local minimizer. This is due to $f(0, \varepsilon, 1) < f(\bar{x})$ for all $\varepsilon \in (0, 1]$, and $(0, \varepsilon, 1)^T \in \mathbb{R}^{3,2}$. Actually, \bar{x} is a nondegenerate M-stationary point with the quadratic index $QI = 0$. For its M-index we have

$$s - k + QI = 2 - 1 + 0 = 1.$$

We conclude that \bar{x} is rather a saddle point for SCNO. □

In [16], the Fréchet and Mordukhovich normal cones of the SCNO feasible set are used to derive corresponding stationarity concepts. Let $\widehat{N}_{\mathbb{R}^{n,s}}(\bar{x})$ stand for the Fréchet and $N_{\mathbb{R}^{n,s}}(\bar{x})$ for the Mordukhovich normal cone of $\mathbb{R}^{n,s}$ at \bar{x} , see e.g. [21] for details.

Definition 9 (*\widehat{N} - and N -stationarity*, [16]) A feasible point $\bar{x} \in \mathbb{R}^{n,s}$ is called \widehat{N} - and N -stationary for SCNO if it respectively holds:

$$0 \in \nabla f(\bar{x}) + \widehat{N}_{\mathbb{R}^{n,s}}(\bar{x}) \quad \text{and} \quad 0 \in \nabla f(\bar{x}) + N_{\mathbb{R}^{n,s}}(\bar{x}).$$

Note that \widehat{N} - and N -stationarity can be viewed as necessary optimality conditions for SCNO, see [16]. The relation of \widehat{N} - and N -stationarity to the previously discussed concepts in the context of SCNO has been essentially elaborated in [16].

Proposition 13 *The notions of basic feasibility and \widehat{N} -stationarity coincide. Every N -stationary point for SCNO is also M-stationary.*

Proof Theorem 3.1 from [16] provides explicit formulas for the Fréchet and Mordukhovich normal cones of $\mathbb{R}^{n,s}$ at a SCNO feasible point \bar{x} with $\|\bar{x}\|_0 = k$:

$$\widehat{N}_{\mathbb{R}^{n,s}}(\bar{x}) = \begin{cases} \{0\}, & \text{if } k < s, \\ \text{span}\{e_i \mid i \in I_0(\bar{x})\}, & \text{if } k = s \end{cases}$$

and

$$N_{\mathbb{R}^{n,s}}(\bar{x}) = \begin{cases} \bigcup_{J \in \mathcal{J}(\bar{x})} \text{span}\{e_j \mid j \notin J\}, & \text{if } k < s, \\ \text{span}\{e_i \mid i \in I_0(\bar{x})\}, & \text{if } k = s, \end{cases}$$

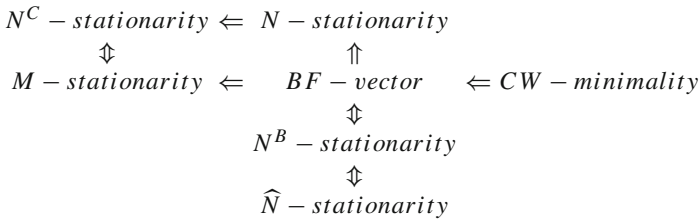
where

$$\mathcal{J}(\bar{x}) = \{J \in \mathcal{J} \mid \bar{x} \in S_J\}, \quad \mathcal{J} = \{J \subset \{1, \dots, n\} \mid |J| = s\}, \quad S_J = \text{span}\{e_j \mid j \in J\}.$$

The equivalence of basic feasibility and \widehat{N} -stationarity follows immediately. For the second assertion, we assume that \bar{x} is N -stationary and let $i \in I_1(\bar{x})$ be arbitrarily, but fixed. Then, $i \in J$ for every $J \in \mathcal{J}(\bar{x})$, and hence, $\frac{\partial f}{\partial x_i}(\bar{x}) = 0$. Thus, \bar{x} is M-stationary. □

Remark 2 (*N -stationarity and instability*) Proposition 13 says that M-stationarity is a weaker condition than N -stationarity. However, it turns out that N -stationary points need not to be stable with respect to data perturbations. Namely, after an arbitrarily small C^2 -perturbation of f a locally unique N -stationary point may bifurcate into multiple N -stationary points. More importantly, this bifurcation unavoidably causes the emergence of M-stationary points, not being N -stationary. This is in full analogy with CW-minima and BF-vectors. The same Example 3 illustrates the instability phenomenon for N -stationary points as well. It is worth to mention that there bifurcation happens even though the N -stationary point under consideration fulfils SOSC. □

In order to better understand the relations of M-stationarity to other stationarity concepts discussed in Sect. 3, we provide the following diagram:



4 Global results

Let us study the topological properties of lower level sets

$$M^a = \{x \in \mathbb{R}^{n,s} \mid f(x) \leq a\},$$

where $a \in \mathbb{R}$ is varying. For that, we define intermediate sets for $a < b$:

$$M_a^b = \{x \in \mathbb{R}^{n,s} \mid a \leq f(x) \leq b\}.$$

For the topological concepts used below we refer to [24].

Let us start with Assumption 1 which is usual within the scope of Morse theory, cf. [10]. It prevents from considering asymptotic effects at infinity.

Assumption 1 The restriction of the objective function $f|_{\mathbb{R}^{n,s}}$ on the SCNO feasible set is proper, i. e. $f^{-1}(K) \cap \mathbb{R}^{n,s}$ is compact for any compact set $K \subset \mathbb{R}$.

Theorem 4 (Deformation for SCNO) *Let Assumption 1 be fulfilled and M_a^b contain no M-stationary points for SCNO. Then, M^a is homeomorphic to M^b .*

Proof We apply Proposition 3.2 from Part I in [10]. The latter provides the deformation for general Whitney stratified sets with respect to critical points of proper maps. Note that the SCNO feasible set admits a Whitney stratification:

$$\mathbb{R}^{n,s} = \bigcup_{\substack{I \subset \{1, \dots, n\} \\ |I| \leq s}} \bigcup_{J \subset I} Z_{I,J},$$

where

$$Z_{I,J} = \{x \in \mathbb{R}^n \mid x_{I^c} = 0, x_J > 0, x_{I \setminus J} < 0\}.$$

The notion of criticality used in [10] can be stated for SCNO as follows. A point $\bar{x} \in \mathbb{R}^{n,s}$ is called critical for $f|_{\mathbb{R}^{n,s}}$ if it holds:

$$\nabla f(\bar{x})|_{T_{\bar{x}}Z} = 0,$$

where Z is the stratum of $\mathbb{R}^{n,s}$ which contains \bar{x} , and $T_{\bar{x}}Z$ is the tangent space of Z at \bar{x} . By identifying $I = I_1(\bar{x})$ and, hence, $I^c = I_0(\bar{x})$, we see that the concepts of criticality and M-stationarity coincide. This concludes the assertion. □

Let us now turn our attention to the topological changes of lower level sets when passing an M-stationary level. Traditionally, they are described by means of the so-called cell-attachment. We first consider a special case of cell-attachment. For that, let N^ϵ denote the lower level set of a special linear function on $\mathbb{R}^{p,q}$, i.e.

$$N^\epsilon = \left\{ x \in \mathbb{R}^{p,q} \mid \sum_{i=1}^p x_i \leq \epsilon \right\},$$

where $\epsilon \in \mathbb{R}$, and the integers $q < p$ are nonnegative.

Lemma 1 (Normal Morse data) *For any $\epsilon > 0$ the set N^ϵ is homotopy-equivalent to $N^{-\epsilon}$ with $\binom{p-1}{q}$ cells of dimension q attached. The latter cells are the q -dimensional simplices from the collection*

$$\{ \text{conv}(e_j, j \in J) \mid J \subset \{1, \dots, p\}, 1 \in J, |J| = q + 1 \}.$$

Proof Let N_ϵ denote the upper level set of a special linear function on $\mathbb{R}^{p,q}$, i.e.

$$N_\epsilon = \left\{ x \in \mathbb{R}^{p,q} \mid \sum_{i=1}^p x_i \geq \epsilon \right\}.$$

In terms of upper level sets Lemma 1 can be obviously reformulated as follows: For any $\epsilon > 0$ the set $N_{-\epsilon}$ is homotopy-equivalent to N_ϵ with $\binom{p-1}{q}$ cells of dimension q attached. Let us show the latter assertion.

First, we note that the sets N_0 and $N_{-\epsilon}$ are contractible. The contraction is performed via the mapping

$$(x, t) \mapsto (1 - t) \cdot x, \quad t \in [0, 1].$$

For the lower level set N^ϵ we have the representation

$$N^\epsilon = \bigcup_{\substack{J \subset \{1, \dots, p\} \\ |J| = q}} N^{\epsilon, J},$$

where

$$N^{\epsilon, J} = \left\{ x \in \mathbb{R}^{p,q} \mid x_{J^c} = 0, \sum_{i \in J} x_i \geq \epsilon \right\}.$$

Note that $N^{\epsilon, J}$ is homotopy-equivalent to the set N^J , where

$$N^J = \left\{ x \in \mathbb{R}^{p,q} \mid x_{J^c} = 0, \sum_{i \in J} x_i = 1 \right\}$$

is the $(|J| - 1)$ -dimensional simplex $\text{conv}(e_j, j \in J)$ of \mathbb{R}^p . In fact, the map

$$(x, t) \mapsto t \cdot \frac{x}{\sum_{i=1}^p x_i} + (1 - t) \cdot x, \quad t \in [0, 1]$$

can be used for all N^J . Altogether, N_ϵ is homotopy-equivalent to

$$\bigcup_{\substack{J \subset \{1, \dots, p\} \\ |J|=q}} \text{conv}(e_j, j \in J). \tag{13}$$

Note that the set in (13) is the $(q - 1)$ -skeleton of the $(p - 1)$ -dimensional simplex of \mathbb{R}^p . The $(q - 1)$ -skeleton of the $(p - 1)$ -dimensional simplex is the union of its simplices up to dimension $q - 1$, see e.g. [9].

Within the $(q - 1)$ -skeleton (13), we close all q -dimensional holes by attaching q -dimensional cells from the collection of simplices

$$\{ \text{conv}(e_j, j \in J) \mid J \subset \{1, \dots, p\}, |J| = q + 1 \}.$$

The attachment should result in a contractible set, as it is actually N_0 . We note that the union of the subdivision

$$\{ \text{conv}(e_j, j \in J) \mid J \subset \{1, \dots, p\}, 1 \in J, |J| = q + 1 \} \tag{14}$$

is also contractible, namely, to e_1 . To see this, we may use the map

$$(x, t) \mapsto t \cdot e_1 + (1 - t) \cdot x, \quad t \in [0, 1].$$

Furthermore, none of the relative interiors of the simplices in (14) can be deleted. In fact, deleting gives rise to the boundary of a q -dimensional simplex and the latter is not contractible.

On the other hand, for any $J^* \subset \{1, \dots, p\} \setminus \{1\}$ with $|J^*| = q + 1$ the union

$$\text{conv}(e_j, j \in J^*) \cup \bigcup_{\substack{J^{**} \subset J^* \\ |J^{**}|=q}} \text{conv}(e_j, j \in J^{**} \cup \{1\}) \tag{15}$$

forms the boundary of the $(q + 1)$ -dimensional simplex $\text{conv}(e_j, j \in J^* \cup \{1\})$. Hence, the set in (15) is not contractible. Altogether, precisely the q -dimensional cells in (14) can be attached to the $(q - 1)$ -skeleton (13) in order to obtain a contractible set. Its number obviously equals $\binom{p-1}{q}$. This completes the proof. \square

Theorem 5 (Cell-Attachment for SCNO) *Let Assumption 1 be fulfilled and M_a^b contain exactly one M -stationary point \bar{x} for SCNO with $\|\bar{x}\|_0 = k$ and the M -index equal to $s - k + QI$. If $a < f(\bar{x}) < b$, then M^b is homotopy-equivalent to M^a with $\binom{n-k-1}{s-k}$ cells of dimension $s - k + QI$ attached, namely:*

$$\bigcup_{\substack{J \subset \{1, \dots, n-k\} \\ 1 \in J, |J|=s-k+1}} \text{conv}(e_j, j \in J) \times [0, 1]^{QI}.$$

Proof Theorem 4 allows deformations up to an arbitrarily small neighborhood of the M -stationary point \bar{x} . In such a neighborhood, we may assume without loss of generality that $\bar{x} = 0$ and f has the following form as from Theorem 1:

$$f(x) = f(\bar{x}) + \sum_{i=1}^{n-k} x_i + \sum_{j=n-k+1}^n \pm x_j^2, \tag{16}$$

where $x \in \mathbb{R}^{n-k, s-k} \times \mathbb{R}^k$, and the number of negative squares in (16) equals QI . In terms of [10] the set $\mathbb{R}^{n-k, s-k} \times \mathbb{R}^k$ can be interpreted as the product of the tangential part \mathbb{R}^k

and the normal part $\mathbb{R}^{n-k,s-k}$. The cell-attachment along the tangential part is standard. Analogously to the unconstrained case, one QI -dimensional cell has to be attached on \mathbb{R}^k . The cell-attachment along the normal part is more involved. Due to Lemma 1, we need to attach $\binom{n-k-1}{s-k}$ cells on $\mathbb{R}^{n-k,s-k}$, each of dimension $s - k$. Finally, we apply Theorem 3.7 from Part I in [10], which says that the local Morse data is the product of tangential and normal Morse data. Hence, the dimensions of the attached cells add together. Here, we have then to attach $\binom{n-k-1}{s-k}$ cells on $\mathbb{R}^{n-k,s-k} \times \mathbb{R}^k$, each of dimension $s - k + QI$. \square

Let us put Theorem 5 into the context of Morse theory as developed in the literature for other nonsmooth optimization problems. The new issue for SCNO is the multiplicity of attached cells.

Remark 3 (Multiplicity of attached cells) We recall that for nonlinear programming problems (NLP) the dimension of the cell to be attached while passing a critical point equals to its quadratic index, see e. g. [13]. The situation changes if we consider mathematical programs with complementarity constraints (MPCC). Here, the dimension of attached cells equals to the so-called C-index of C-stationary points, see [14]. In addition to quadratic, the C-index also has a bi-active part. The latter counts negative pairs of Lagrange multipliers corresponding to the bi-active complementarity constraints. The cell-attachment for mathematical programs with vanishing constraints (MPVC) is similar, see [8]. The dimension of attached cells equals here to the so-called T-index of T-stationary points. The T-index consists again of quadratic and bi-active parts. We emphasize that the cell-attachment for SCNO considerably differs from the described cases of NLP, MPCC, and MPVC. The main difference is that *multiple* cells are involved into the cell-attachment procedure for SCNO. The multiplicity of attached cells is a novel and striking phenomenon in nonsmooth optimization not observed in the literature before. From the technical point of view, this makes the cell-attachment result for SCNO to appear rather challenging. Note that the determination of the number of attached cells becomes an involved combinatorial problem from algebraic topology, see Lemma 1.

Let us present a global interpretation of our results for SCNO. For that, we need to state another assumption. Following Assumption 2 is standard in the context of SCNO, cf. [1], and gives a necessary condition for its solvability.

Assumption 2 The restriction of the objective function $f|_{\mathbb{R}^{n,s}}$ on the SCNO feasible set is lower bounded.

Now, we consider M-stationary points \bar{x} for SCNO with $\|\bar{x}\|_0 = k$ and the M-index equal to one, thus, fulfilling $s - k + QI = 1$. These so-called saddle points can be of two types:

(I) with active sparsity constraint and quadratic index equal to one, i. e.

$$k = s, \quad QI = 1,$$

(II) with exactly $s - 1$ non-zero entries and vanishing quadratic index, i. e.

$$k = s - 1, \quad QI = 0.$$

Theorem 6 (Morse relation for SCNO) *Let Assumptions 1 and 2 be fulfilled, and all M-stationary points of SCNO be nondegenerate. Additionally, we assume that there exists a connected lower level set which contains all M-stationary points. Then, it holds:*

$$r_I + (n - s)r_{II} \geq r - 1, \tag{17}$$

where r is the number of local minimizers of SCNO, r_I and r_{II} are the numbers of M-stationary points with M-index equal to one, which correspond to the types (I) and (II), respectively.

Proof We assume without loss of generality that the objective function f has pairwise different values at all M-stationarity points of SCNO. If it is not the case, we may enforce this property by sufficiently small perturbations of the objective function. Due to the openness part in Theorem 2, all M-stationarity points of such a perturbed SCNO remain nondegenerate. Moreover, the formula (17) is still valid since it does not depend on the functional values of f .

Further, let q_a denote the number of connected components of the lower level set M^a . We focus on how q_a changes as $a \in \mathbb{R}$ increases. Due to Theorem 4, q_a can change only if passing through a value corresponding to an M-stationary point \bar{x} , i.e. $a = f(\bar{x})$. In fact, Theorem 4 allows homeomorphic deformations of lower level sets up to an arbitrarily small neighborhood of the M-stationary point \bar{x} . Then, we have to estimate the difference between q_a and $q_{a-\varepsilon}$, where $\varepsilon > 0$ is arbitrarily, but sufficiently small, and $a = f(\bar{x})$. This is done by a *local argument*. For that, let the M-index of \bar{x} be $s - k + QI$ with $\|\bar{x}\|_0 = k$. We use Theorem 5 which says that M^a is homotopy-equivalent to $M^{a-\varepsilon}$ with a cell-attachment of

$$\bigcup_{\substack{J \subset \{1, \dots, n-k\} \\ 1 \in J, |J|=s-k+1}} \text{conv}(e_j, j \in J) \times [0, 1]^{QI}. \tag{18}$$

Let us distinguish the following cases:

- 1) \bar{x} is a local minimizer with vanishing M-index, i.e. $k = s$ and $QI = 0$. Then, by (18) we attach to $M^{a-\varepsilon}$ the cell $\text{conv}(e_1)$ of dimension zero. Consequently, a new connected component is created, and it holds:

$$q_a = q_{a-\varepsilon} + 1.$$

- 2) \bar{x} is of type (I) with M-index equal to one, i.e. $k = s$ and $QI = 1$. Then, by (18) we attach to $M^{a-\varepsilon}$ the cell $\text{conv}(e_1) \times [0, 1]$ of dimension one. Consequently, at most one connected component disappears, and it holds:

$$q_{a-\varepsilon} - 1 \leq q_a \leq q_{a-\varepsilon}.$$

This case is well known from nonlinear programming, see e.g. [13].

- 3) \bar{x} is of type (II) with M-index equal to one, i.e. $k = s - 1$ and $QI = 0$. Then, by (18) we attach to $M^{a-\varepsilon}$ as many as $n - s$ cells of dimension one, namely:

$$\bigcup_{j=2, \dots, n-s+1} \text{conv}(e_1, e_j).$$

Consequently, at most $n - s$ connected components disappear, and it holds:

$$q_{a-\varepsilon} - (n - s) \leq q_a \leq q_{a-\varepsilon}.$$

For illustration we refer to Fig. 1. Case 3) is new and characteristic for SCNO.

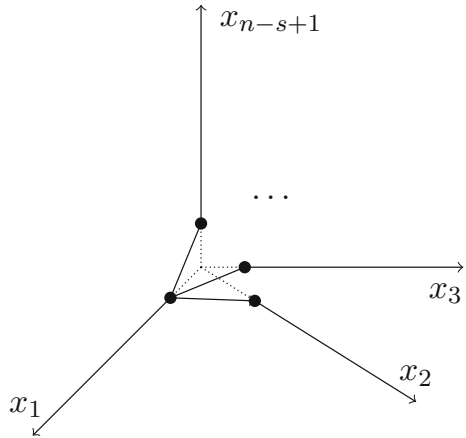
- 4) \bar{x} is M-stationary with M-index greater than one, i.e. $s - k + QI > 1$. The boundary of the cell-attachment in (18) is

$$\bigcup_{\substack{J \subset \{1, \dots, n-k\} \\ 1 \in J, |J|=s-k+1}} \left(\partial \text{conv}(e_j, j \in J) \times [0, 1]^{QI} \right) \cup \left(\text{conv}(e_j, j \in J) \times \{0, 1\}^{QI} \right).$$

The latter set is connected if $s - k + QI > 1$. Consequently, the number of connected components of M^a remains unchanged, and it holds:

$$q_a = q_{a-\varepsilon}.$$

Fig. 1 Cell-attachment for type (II)



Now, we proceed with the *global argument*. Assumption 2 implies that there exists $c \in \mathbb{R}$ such that M^c is empty, thus, $q_c = 0$. Additionally, there exists $d \in \mathbb{R}$ such that M^d is connected and contains all M-stationary points, thus, $q_d = 1$. Due to Assumption 1, M_c^d is compact, moreover, it contains all M-stationary points. Since nondegenerate M-stationary points are in particular isolated, we conclude that there must be finitely many of them. Let us now increase the level a from c to d and describe how the number q_a of connected components of the lower level sets M^a changes. It follows from the local argument that r new connected components are created, where r is the number of local minimizers for SCNO. Let q_I and q_{II} denote the actual number of disappearing connected components if passing the levels corresponding to M-stationary points of types (I) and (II), respectively. The local argument provides that at most r_I and $(n - s)r_{II}$ connected components might disappear while doing so, i. e.

$$q_I \leq r_I, \quad q_{II} \leq (n - s)r_{II}.$$

Altogether, we have:

$$r - r_I - (n - s)r_{II} \leq r - q_I - q_{II} = q_d - q_c.$$

By recalling that $q_d = 1$ and $q_c = 0$, we get Morse relation (17). □

We illustrate Theorem 6 by discussing the same SCNO as in Example 1.

Example 5 (Saddle point) We consider the following SCNO with $n = 2$ and $s = 1$:

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

As we have seen in Example 1, both M-stationary points $(1, 0)$ and $(0, 1)$ are nondegenerate minimizers. Thus, we have $r = 2$. Morse relation (17) from Theorem 6 provides:

$$r_I + r_{II} \geq 1.$$

Hence, there should exist an additional M-stationary point with M-index one. In fact, $(0, 0)$ is this nondegenerate M-stationary point of type (II), cf. Example 1. Note that, due to $r_I = 0$ and $r_{II} = 1$, Morse relation (17) holds with equality here. □

Let us briefly comment on the applicability of deformation and cell-attachment results in Theorems 4 and 5, respectively, for the least squares loss function.

Remark 4 (Least squares loss function) We take the least squares as the objective function in SCNO, i. e.

$$f(x) = \|Ax - b\|_2^2,$$

where $A \in \mathbb{R}^{m \times n}$ can be viewed as a sensing matrix and $b \in \mathbb{R}^m$ as a measurement vector. Then, SCNO corresponds to the problem of sparse recovery from compressed sensing, see e. g. [1]. Here, it is convenient to assume that the bound on the number of non-zero entries of the signal does not exceed the number of measurements, i. e. $s \leq m$. Let us examine whether Assumption 1 is fulfilled for the least squares loss function. It turns out that the so-called s -regularity of A is sufficient for the latter. Recall from [1] that a matrix $A \in \mathbb{R}^{m \times n}$ is called s -regular if for every index set $I \subset \{1, \dots, n\}$ with $|I| = s$ it holds:

$$\text{rank}(A_I) = s,$$

where A_I denotes the submatrix of A with the columns corresponding to the set I , and $\text{rank}(A_I)$ stands for its rank. In presence of s -regularity of A , it is shown in [15] that the lower level sets

$$M^a = \{x \in \mathbb{R}^{n,s} \mid \|Ax - b\|_2^2 \leq a\}$$

are bounded for all $a \in \mathbb{R}$. Hence, the restriction of the least squares loss function on $\mathbb{R}^{n,s}$ is in this case proper, i. e. Assumption 1 is satisfied. Note that Assumption 2 trivially holds for the least squares loss function, since it is nonnegative. Finally, we refer to [15] for the detailed exposition of the topological approach as applied to sparse recovery. \square

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Acknowledgements The authors would like to thank Hubertus Th. Jongen for fruitful discussions, as well as the anonymous referee for the precise and constructive remarks.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Beck, A., Eldar, Y.C.: Sparsity constrained nonlinear optimization: optimality conditions and algorithms. *SIAM J. Optim.* **24**, 1480–1509 (2013)
2. Beck, A., Hallak, N.: On the minimization over sparse symmetric sets: projections, optimality conditions, and algorithms. *Math. Oper. Res.* **41**, 196–223 (2016)
3. Beck, A., Hallak, N.: Proximal mapping for symmetric penalty and sparsity. *SIAM J. Optim.* **28**, 496–527 (2018)
4. 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)

5. Burdakov, O., 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. Červinka, M., Kanzow, C., Schwartz, A.: Constraint qualifications and optimality conditions for optimization problems with cardinality constraints. *Math. Progr.* **160**, 353–377 (2016)
7. Donoho, D.L.: Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006)
8. Dorsch, D., Shikhman, V., Stein, O.: Mathematical programs with vanishing constraints: critical point theory. *J. Glob. Optim.* **52**, 591–605 (2012)
9. Goerss, P.G., Jardine, J.F.: *Simplicial Homotopy Theory*. Birkhäuser, Basel (2009)
10. Goresky, M., MacPherson, R.: *Stratified Morse Theory*. Springer, New York (1988)
11. Hirsch, M.W.: *Differential Topology*. Springer, Berlin-Heidelberg-New York (1976)
12. Jongen, H.T.: On non-convex optimization. Dissertation, University of Twente, The Netherlands (1977)
13. Jongen, H.T., Jonker, P., Twilt, F.: *Nonlinear Optimization in Finite Dimensions*. Kluwer Academic Publishers, Dordrecht (2000)
14. Jongen, H.T., Shikhman, V., Ruckmann, J.-J.: MPCC: critical point theory. *SIAM J. Optim.* **20**, 473–484 (2009)
15. Lämmel, S., Shikhman, V.: Critical point theory for sparse recovery. [arXiv:2002.10913](https://arxiv.org/abs/2002.10913) (2020)
16. Li, X., Song, W.: The first-order necessary conditions for sparsity constrained optimization. *J. Oper. Res. Soc. China* **3**, 521–535 (2015)
17. Lu, Z., Zhang, Y.: Sparse approximation via penalty decomposition methods. *SIAM J. Optim.* **23**, 2448–2478 (2013)
18. Milnor, J.: *Morse Theory*. Princeton University Press, Princeton (1963)
19. Pan, L., Xiu, N., Fan, J.: Optimality conditions for sparse nonlinear programming. *Sci. China Math.* **5**, 1–18 (2017)
20. Pan, L., Xiu, N., Zhou, S.: On solutions of sparsity constrained optimization. *J. Oper. Res. Soc. China* **3**, 421–439 (2015)
21. Rockafellar, R., Wets, R.: *Variational Analysis*. Springer, Berlin (1998)
22. Shechtman, Y., Eldar, Y.C., Szameit, A., Segev, M.: Sparsity-based sub-wavelength imaging with partially spatially incoherent light via quadratic compressed sensing. *Opt. Express* **19**, 14807–14822 (2011)
23. Shikhman, V.: *Topological Aspects of Nonsmooth Optimization*. Springer, New York (2012)
24. Spanier, E.H.: *Algebraic Topology*. McGraw-Hill Book Company, New York (1966)
25. Tibshirani, R.: Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* **58**, 267–288 (1996)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.