Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 9, No. 2, pp. 201–224, April 2013

LOCAL MODELS IN EQUILIBRIUM OPTIMIZATION

D. Dorsch*, H.Th. Jongen and V. Shikhman

Abstract: We study equilibrium optimization from a structural point of view. For that, we consider equilibrium optimization problems up to smooth local coordinate transformations at their solutions. The latter equivalence relation induces classes of equilibrium optimization problems. We focus on stable classes corresponding to a dense set of data functions. We prove that these classes are unique and call them "basic classes". Their representatives in the simplest form are called local models. For particular realizations of equilibrium optimization basic classes and their local models are elaborated. The latter include bilevel optimization, general semi-infinite programming and Nash optimization. In bilevel optimization we discuss new phenomena arising from the appearance of multiple global minimizers as well as their possible bifurcation.

Key words: *classification problem, basic classes, local models, bilevel optimization, general semi-infinite programming, generalized Nash equilibrium problem, parametric optimization, singularities*

Mathematics Subject Classification: *90C31, 90C33, 90C34*

1 Introduction

Equilibrium problems arise quite naturally in many applications, such as finance, economics, physics, chemistry, biology, engineering and computer science. As relevant applications we mention the traffic flow in networks, contract design agreements, the control of frictional contacts in robotics, nonlinear obstacle problem from mathematical physics, design centering problem, Chebyshev and reverse Chebyshev approximation, robustness of mechanical structures, truss optimization, options pricing, Stackelberg games and others. For literature we refer to [3, 10, 21, 23, 27]. From the modeling perspective, it is crucial to appropriately describe the **coupling behavior** of different agents (or systems) in equilibrium problems. For that, one assumes that each agent steers his particular group of variables. Then, the coupling behavior leads to **parametric dependencies** in the sense that the variables of one agent play a role of parameters for other agents' subproblems. The latter subproblems are usually optimization problems. They arise due the fact that the agents' behavior is **goal-oriented** w.r.t. their particular group of variables. Altogether, the agents are facing individual parametric optimization problems. Mathematically, the qualitative nature of the coupling behavior of agents can be analyzed by **optimal value functions** and **solution set mappings** of these problems. Actually, the relationship between the agents' parametric optimization problems determines the particular subclass of equilibrium optimization. In this article we consider the following subclasses of equilibrium optimization:

*Corresponding author

- Bilevel Optimization Problem (BOP),
- General Semi-Infinite Programming (GSIP), and
- Nash Optimization Problem (NOP).

We study equilibrium optimization from a structural point of view. Our goal is to indicate so-called *local models* in bilevel optimization, GSIP and Nash optimization. In what follows, we define precisely what we mean by the term "local models". To this aim, we consider a general optimization problem

$$P(f,F): \quad \min_{x} f(x) \quad \text{s.t.} \quad x \in M[F], \tag{1.1}$$

where $f \in C^\infty(\mathbb{R}^n)$ is a real-valued objective function, $F \in C^\infty(\mathbb{R}^n, \mathbb{R}^l)$ is a vector-valued function, and $M[F] \subseteq \mathbb{R}^n$ is a feasible set defined by F in some structured way. For simplicity reasons we assume that the functions f, F are smooth. We emphasize that bilevel optimization, GSIP and Nash optimization can be written as $P(f,F)$ with appropriate data functions f, F (cf. formulas (2.3), (3.1), (4.1) later on). Now, let x be a local minimizer of $P(f,F)$. We consider the optimization problem $P(f,F)$ locally at x. For that, we define an equivalence relation on the triples (f,F,x).

Definition 1.1 (Equivalence relation)**.** Two triples (f,F,x) and (g,G,y) are called equivalent w.r.t. (1.1), written

$$(f,F,x) \sim (g,G,y),$$

if there exist open $\mathbb{R}^n$-neighborhoods U_x of x, V_y of y, a C^∞-diffeomorphism $\Psi : U_x \longrightarrow V_y$ and $c \in \mathbb{R}$ such that

(i) $\Psi(x) = y$,

(ii) $f \circ \Psi^{-1}(\cdot) = g(\cdot) + c$ on V_y, and

(iii) $\Psi(M[F] \cap U_x) = M[G] \cap V_y$.

It is easy to see that $\sim$ is an equivalence relation. Note that the latter can be interpreted as a local equivariant morphism between (f,F,x) and (g,G,y). We denote the equivalence classes w.r.t. $\sim$ as $[(f,F,x)]$.

For a subset of data functions $\mathcal{H} \subset C^\infty(\mathbb{R}^n) \times C^\infty(\mathbb{R}^n, \mathbb{R}^l)$ we put

$$T_{\mathcal{H}} := \{[(f,F,x)] \,|\, (f,F) \in \mathcal{H},\, x \text{ is a local minimizer of } P(f,F)\},$$

as the set of equivalence classes $[(f,F,x)]$ corresponding to $\mathcal{H}$.

Further, we define special classification sets of data functions for (1.1). For that, we endow the space $C^\infty(\mathbb{R}^n) \times C^\infty(\mathbb{R}^n, \mathbb{R}^l)$ with the strong (or Whitney) C^∞-topology, denoted by C^∞_s (see [13, 15] and below).

Definition 1.2 (Classification set)**.** $\mathcal{H} \subset C^\infty(\mathbb{R}^n) \times C^\infty(\mathbb{R}^n, \mathbb{R}^l)$ is called a classification set of data functions for (1.1) if

(A) $\mathcal{H}$ is C^∞_s-dense in the strong (or Whitney) topology, and

(B) $\mathcal{H}$ is stable w.r.t. (1.1), i.e. for each $(f, F) \in \mathcal{H}$ and each local minimizer x of $P(f, F)$ there exist neighborhoods $\mathcal{O}_{(f,F)}$ of (f, F) and $\mathcal{O}_x$ of x, such that for all $(\tilde{f}, \tilde{F}) \in \mathcal{O}_{(f,F)}$ it holds

$$(f, F, x) \sim (\tilde{f}, \tilde{F}, \tilde{x})$$

with the unique local minimizer $\tilde{x} \in \mathcal{O}_x$ of $P(\tilde{f}, \tilde{F})$.

It turns out that every classification set $\mathcal{H}$ defines the same set $T_{\mathcal{H}}$.

Theorem 1.3 (Uniqueness of equivalence classes)**.** *Let $\mathcal{H}_1, \mathcal{H}_2 \subset C^\infty(\mathbb{R}^n) \times C^\infty(\mathbb{R}^n, \mathbb{R}^l)$ be classification sets of data functions for (1.1). Then, $T_{\mathcal{H}_1} = T_{\mathcal{H}_2}$.*

Proof. Let $[(f_1, F_1, x_1)] \in T_{\mathcal{H}_1}$. We take neighborhoods $\mathcal{O}_{(f_1,F_1)}$ of (f_1, F_1) and $\mathcal{O}_{x_1}$ of x_1 from (B) in Definition 1.2 applied to $\mathcal{H}_1$. Due to (A) in Definition 1.2 applied to $\mathcal{H}_2$, there exist $(f_2, F_2) \in \mathcal{H}_2 \cap \mathcal{O}_{(f_1,F_1)}$. Then, again using (B) for $\mathcal{H}_1$, we get the unique local minimizer $x_2 \in \mathcal{O}_{x_1}$ of $P(f_2, F_2)$, such that

$$(f_1, F_1, x_1) \sim (f_2, F_2, x_2).$$

Hence, $T_{\mathcal{H}_1} \subset T_{\mathcal{H}_2}$. Analogously, $T_{\mathcal{H}_2} \subset T_{\mathcal{H}_2}$ and we are done. □

For a classification set $\mathcal{H}$ we call the elements of $T_{\mathcal{H}}$ the basic classes of the optimization problem (1.1). Note that the basic classes are unique due to Theorem 1.3. Their representatives, having the simplest form, are called local models of (1.1). Altogether, the following questions are to be addressed for a particular realization of (1.1):

(i) Does there exist a classification set?

(ii) How can we describe the basic classes in initial coordinates?

(iii) What are the local models after a change of coordinates?

We refer to the questions (i)-(iii) as a **classification problem**. Regarding the analysis of equilibrium problems we will see that the solution of the classification problem leads especially to explicit local descriptions of feasible sets. In fact, after diffeomorphism the implicit information which come from the solution sets and optimal value functions of parametric problems are described directly. In this article, we focus on classification problems for bilevel optimization, GSIP and Nash optimization. For bilevel optimization (x = upper level, y = lower level), we describe basic classes and derive local models in the following cases:

- $\dim(y) = 1$ without constraints (see Classification Theorem 2.2), and
- $\dim(x) = 1$ (see Classification Theorem 2.3).

For GSIP, the classification problem is presented in Classification Theorem 3.4. It turns out that for bilevel optimization and GSIP the number of basic classes is finite. As it will be shown, the feasible sets of local models in the considered bilevel optimization problems and GSIP are given by

$$\{0\}^{q_1} \times \mathbb{H}^{q_2} \times \mathbb{L}^{q_3} \times \mathbb{R}^{q_4}, \qquad \text{with } q_1, q_2, q_3, q_4 \in \mathbb{N}, \tag{1.2}$$

where

$$\mathbb{H} := \{x \in \mathbb{R} \mid x \geq 0\}, \quad \mathbb{L} := \left\{(a, b) \in \mathbb{R}^2 \mid a \geq 0,\, b \geq 0,\, a \cdot b = 0\right\}.$$

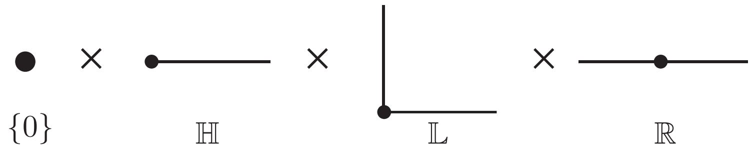

Figure 1: Typical feasible set

This means that the local models represent nonlinear programming problems and mathematical programs with complementarity constraints (e.g., [25]). Hence, by considering local models, we reduce the parametric dependencies given in the formulations of the considered equilibrium problems to explicitly given constraints. By a further diffeomorphism we can find a simple description of the objective function f. If, for example, a feasible set is given by $\{0\} \times \mathbb{H} \times \mathbb{L} \times \mathbb{R} \subset \mathbb{R}^5$ (see Figure 1) and zero is a nondegenerate local minimizer, then (after a further diffeomorphism leaving the feasible set invariant) the objective function of this local model has the form

$$f(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2 + (x_3 + x_4) + x_5^2. \tag{1.3}$$

Objective functions for other local models with feasible sets as in (1.2) can be described analogously. We point out that the above representation of f can be obtained via the so-called generalized Morse Lemma (see [15, 19]).

Convention 1.4. Throughout this article we omit the dependence of f on the first q_1 variables belonging to $\{0\}^{q_1}$ in (1.2). Thus the function in (1.3) would become

$$f = f(0, x_2, x_3, x_4, x_5) = x_2 + (x_3 + x_4) + x_5^2.$$

In the context of the optimization problem this represents the essential information.

For bilevel optimization, we also discuss the classification problem in the following cases:

- $\dim(y) = 1$ with a one-side constraint (see Section 2.2),
- $\dim(y) > 1$ without constraints (see Section 2.3).
- $\dim(x) = \dim(y)$ with linear independent constraints (see Section 2.5).

It turns out that the obstacle here comes from **globality** issues. In fact, the appearance of multiple global minimizers as well as their possible **bifurcation** still need to be understood for a complete classification analysis. For details we refer to Sections 2.2, 2.1, 2.5.

In Nash optimization we treat the classification problem for two its subclasses: NEP and GNEP_0 (see Theorem 4.2). Moreover, we indicate how basic classes and local models can be obtained for general Nash optimization (see Examples 4.3, 4.4).

The article is organized as follows. Section 2 is devoted to bilevel optimization. In Section 3 we address the classification problem for GSIP. Finally, Section 4 deals with Nash optimization.

Our notation is standard. The n-dimensional Euclidean space is denoted by $\mathbb{R}^n$. Given an arbitrary set $K \subseteq \mathbb{R}^n$ we denote its topological closure by $\overline{K}$. For $x \in \mathbb{R}^n$ we write

$\dim(x)$ to denote the number of components of x, i.e. $\dim(x) = n$. Given a differentiable function $F : \mathbb{R}^n \to \mathbb{R}^l$, DF denotes its Jacobian matrix. Given a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, Df denotes the row vector of partial derivatives of first order and $D^T f$ stands for the transposed vector. For a matrix A we write $A \succeq 0$, $A \succ 0$ if A is positive semi-definite or positive definite, respectively.

Let $C^\infty(\mathbb{R}^n)$ denote the space of smooth real-valued functions. Let $C^\infty(\mathbb{R}^n)$ be endowed with the strong (or Whitney) C^∞-topology, denoted by C^∞_s (see [13, 15]). Roughly speaking, the C^∞_s-topology is generated by allowing perturbations of the functions and their derivatives which are controlled by means of continuous positive functions. The product space of continuously differentiable functions will be topologized with the corresponding product topology. The space of C^∞-functions endowed with the strong C^∞_s-topology constitutes a Baire space. A set is called C^∞_s-generic if it contains a countable intersection of C^∞_s-open, -dense subsets. A generic subset of a Baire space is dense as well.

2 Bilevel Optimization

We consider bilevel optimization problems as hierarchical problems of two decision makers, the so-called leader and follower. The follower selects his decision knowing the choice of the leader, whereas the latter has to anticipate the follower's response in his decision. Bilevel optimization problems have been studied in the monographs [1] and [3]. We model the bilevel optimization problem in the so-called optimistic formulation. To this aim, assume that the follower solves the parametric optimization problem (lower level problem L)

$$L(x): \quad \min_y g(x,y) \quad \text{s.t.} \quad h_j(x,y) \geq 0,\, j \in J \tag{2.1}$$

and that the leader's optimization problem (upper level problem U) is the following:

$$U: \quad \min_{(x,y)} f(x,y) \quad \text{s.t.} \quad y \in \text{Argmin } L(x). \tag{2.2}$$

Above we have $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ and the real-valued mappings $f, g, h_j, j \in J$ belong to $C^\infty(\mathbb{R}^n \times \mathbb{R}^m)$. Argmin $L(x)$ denotes the solution set of the optimization problem $L(x)$.

Note that U can be written as follows

$$U: \quad \min_{(x,y)} f(x,y) \quad \text{s.t.} \quad (x,y) \in M[g, h_j, j \in J], \tag{2.3}$$

where

$$M = M[g, h_j, j \in J] := \{(x,y) \,|\, y \in \text{Argmin } L(x)\} \tag{2.4}$$

is the bilevel feasible set given by data functions $g, h_j, j \in J$.

In order to avoid asymptotic effects, we assume the following technical assumption (cf. [4]) for all considered bilevel optimization problems:

Assumption 2.1. The function g belongs to the set $\mathcal{O} \subset C^\infty(\mathbb{R}^n \times \mathbb{R}^m)$ with

$$\mathcal{O} := \left\{ g \in C^\infty(\mathbb{R}^n \times \mathbb{R}^m) \;\middle|\; \begin{array}{c} \text{for all } (\bar{x}, c) \in \mathbb{R}^n \times \mathbb{R} \text{ the set} \\ \{(x,y) \mid \|x - \bar{x}\| \leq 1,\, g(x,y) \leq c\} \text{ is compact} \end{array} \right\}.$$

Let $\mathcal{G}$ be the subset of all data functions where g belongs to $\mathcal{O}$. In the following we consider classification sets $\mathcal{H} \subset \mathcal{G}$. Accordingly, the density property (A) in Definition 1.2 is meant w.r.t. $\mathcal{G}$. Since $\mathcal{O}$ is a C^∞-open, Theorem 1.3 remains true, i.e. the basic classes are unique.

Throughout the whole article we use the notion of a nondegenerate (local) minimizer for nonlinear programming problems (NLPs), see [15] for details. The latter refers to

ND1: linear independence constraint qualification (LICQ),
ND2: strict complementarity slackness, and
ND3: the restriction of the Lagrange function's Hessian to the tangent space to the feasible set is positive definite.

2.1 Unconstrained Lower Level ($\dim(y) = 1$)

In this section we consider the case where $J = \emptyset$ in (2.1) and where $\dim(y) = 1$. This case was already treated in [4]. There it has been shown, that, generically, at a solution $(\bar{x}, \bar{y})$ of the bilevel problem, all lower level solutions, i.e. points from the set

$$\text{Argmin } L(\bar{x}) = \{\bar{y}^{(1)}, \dots, \bar{y}^{(p)}\}, \quad \text{where } \bar{y} = \bar{y}^{(p)},$$

are nondegenerate (global) minimizers. This means that locally for $x \approx \bar{x}$ the local minimizers emerging from $\bar{y}^{(i)}$ can be described by a smooth function $y^{(i)} = y^{(i)}(x)$. Moreover, it has been shown that in the latter situation the bilevel problem can locally be reduced to a nondegenerate NLP in the variable x. The following Classification Theorem 2.2 is a result of [4, Theorem 3].

Classification Theorem 2.2 ($\dim(y) = 1$, y unconstrained)**.**
Let the lower level problem $L(\cdot)$ in (2.1) be unconstrained ($J = \emptyset$) and let $\dim(y) = 1$. Then, the bilevel optimization problem (2.3) has the basic classes K_p, $1 \leq p \leq n+1$, with:

K_p :

- *Argmin $L(\bar{x}) = \{\bar{y}^{(1)}, \dots, \bar{y}^{(p)}\}$, where $\bar{y} = \bar{y}^{(p)}$.*
- *Each $\bar{y}^{(i)}$ is a nondegenerate (global) minimizer of $L(\bar{x})$, i.e. there are smooth functions $y^{(i)} = y^{(i)}(x)$ with $y^{(i)}(\bar{x}) = \bar{y}^{(i)}$, such that $y^{(i)}(x)$ is the unique local minimizer near $\bar{y}^{(i)}$ (for x close to $\bar{x}$).*
- *The point $\bar{x}$ is a nondegenerate local minimizer of the following NLP:*

$$\min_x \bar{f}(x) \quad s.t. \quad g_j(x) \geq 0, \; 1 \leq j \leq p-1,$$

where $g_j(x) := g(x, y^{(j)}(x)) - g(x, y^{(p)}(x))$, $1 \leq j \leq p-1$, and $\bar{f}(x) := f(x, y^{(p)}(x))$.

For $1 \leq p \leq n+1$ the class K_p is represented[†] *by the following local model:*

$$\min_{z \in \mathbb{R}^{n+m}} \left(\sum_{i=m+1}^{m+(p-1)} z_i + \sum_{i=m+(p-1)+1}^{n+m} z_i^2 \right) \quad s.t. \quad z \in \{0\}^m \times \mathbb{H}^{p-1} \times \mathbb{R}^{n-(p-1)}$$

2.2 Constrained One-Dimensional Lower Level ($\dim(y) = 1$)

We consider the case, where the lower level problem $L(\cdot)$ from (2.1) is one-dimensional, i.e. $\dim(y) = 1$, and exactly one constraint is given:

$$L(x): \quad \min_{y \in \mathbb{R}} g(x, y) \quad \text{s.t.} \quad y \geq 0 \tag{2.5}$$

Problem (2.5) is the simplest form of dealing with lower level constraints. It turns out that in this case a **kink** might occur as a new situation in the lower level. As this kink originated

[†]According to Convention 1.4.

as a preimage of the typical set $\mathbb{L}$, the new local models turn out to be MPCCs in general. Thus in the following we use the notion of a nondegenerate minimizer for mathematical programs with complementarity constraints (MPCCs). The latter refers to linear independence constraint qualification, strict complementarity, positivity of biactive Lagrange multipliers and second order sufficiency condition (see [9, 19, 24, 26] for details).

We present the following Classification Conjecture. Here, we omitted the explicit description of the basic classes in initial data, since the combinatorial number of possibilities is too high.

Classification Conjecture 1 ($\dim(y) = 1$, $y \geq 0$)**.**
Let the lower level problem given as in (2.5). Then, the basic classes of the bilevel optimization problem are represented[‡] by the following local models:

$$\min_{z \in \mathbb{R}^{n+1}} \left(\sum_{i=2-\delta}^{p-\delta} z_i + \sum_{i=p-\delta+1}^{p} (z_i + z_{i+1}) + \sum_{i=p+\delta+1}^{n+1} z_i^2 \right) \quad \text{s.t.} \quad z \in \{0\}^{1-\delta} \times \mathbb{H}^{p-1} \times \mathbb{L}^{\delta} \times \mathbb{R}^{n-(p-1)-\delta},$$

where $\delta \in \{0, 1\}$, and $1 \leq p \leq n + 1 - \delta$.

Classification Conjecture holds if Argmin $L(\bar{x}) = \{\bar{y}\}$ is a singleton, i.e. $p = 1$. In fact, without loss of generality, we may assume that $\bar{x} = 0$, $\bar{y} = 0$ and, by a change of coordinates in y only, $g(0, y) = y^k$, $k \geq 1$. Analogously to [4], it is sufficient to focus on the universal unfolding of the (constrained) singularity y^k, $y \geq 0$:

$$g(x, y) = y^k + x_{k-1} y^{k-1} + x_{k-2} y^{k-2} + \ldots + x_1 y, \; y \geq 0, \tag{2.6}$$

where $k \geq 1$ and $x = (x_1, \ldots, x_{k-1}, x_k, \ldots, x_n)$.

The situation $k = 1$ corresponds to the case $\delta = 0$ in the assertion of Classification Conjecture. Indeed, $\bar{y}$ is a nondegenerate minimizer of $L(\bar{x})$, hence, M is locally diffeomorphic to $\{0\} \times \mathbb{R}^n$. Next, $k = 2$ corresponds to the case $\delta = 1$. Here, $\bar{y}$ is a degenerate minimizer for $L(\bar{x})$ of the so-called Type 2 (see below), i.e. the Lagrange multiplier vanishes. M is locally diffeomorphic to $\mathbb{L} \times \mathbb{R}^{n-1}$. Now, we turn our attention to $k \geq 3$ and show that this situation can be generically avoided. The key idea is indicated in [18] and [4]. In order to avoid certain higher order singularities in the description of the feasible set M, we have to focus on a neighborhood of (local) solutions of the bilevel problem. Suppose that the feasible set M contains a smooth curve, say C, through the point $(\bar{x}, \bar{y}) \in M$. Let the point $(\bar{x}, \bar{y})$ be a local solution of the bilevel problem U, i.e. $(\bar{x}, \bar{y})$ is a local minimizer for the objective function f on the set M. Then, $(\bar{x}, \bar{y})$ is also a local minimizer for $f_{|C}$, the restriction of f to the curve C. If, in addition, $(\bar{x}, \bar{y})$ is a nondegenerate local minimizer for $f_{|C}$, then we may shift this local minimizer along C by means of a linear perturbation of f. After that perturbation with resulting $\widetilde{f}$, the point $(\bar{x}, \bar{y})$ is not any more a local minimizer for $\widetilde{f}_{|C}$ and, hence, it is not any more a local minimizer for $\widetilde{f}_{|M}$. Now, if the singularities in M outside of the point $(\bar{x}, \bar{y})$ are of lower order, then in this way we are able to move away from the higher order singularity. The key point however is to find a smooth curve through a given point of the feasible set M.

In order to do so, we put

$$x_1 = x_2 = \ldots = x_{k-3} = 0.$$

[‡]According to Convention 1.4.

So, we are left with the reduced lower level problem function

$$\widetilde{L}(x_{k-2}, x_{k-1}) : \quad \min_{y} \widetilde{g}(x_{k-2}, x_{k-1}, y) \text{ s. t. } y \geq 0$$

with reduced feasible set $\widetilde{M}$, where

$$\widetilde{g}(x_{k-2}, x_{k-1}, y) = y^k + x_{k-1} y^{k-1} + x_{k-2} y^{k-2}.$$

Firstly, let $x_{k-1} < 0$ and $x_{k-2} > 0$ and consider the curve defined by the equation

$$x_{k-2} - \frac{1}{4} x_{k-1}^2 = 0.$$

One calculates, that for points on this curve, the lower level $\widetilde{L}$ has two different global minimizers on the set $y \geq 0$ (with $\widetilde{g}$-value zero), one of them being $y = 0$. Secondly, we note that the set $\{(x_{k-1}, x_{k-2}, 0) \,|\, x_{k-1} \geq 0, \, x_{k-2} \geq 0\}$ belongs to $\widetilde{M}$. Altogether, we obtain that the curve C defined by the equations

$$x_1 = x_2 = \ldots = x_{k-3} = y = 0, \; x_{k-2} - \frac{1}{4} x_{k-1}^2 = 0,$$

belongs to M, and we are done.

The proof of Classification Conjecture in case of multiple set Argmin $L(\bar{x})$ is still open. In fact, in order to apply the above "curve"-construction, we need to guarantee that minima along this curve remain **global** ones for the lower level objective function. However, the latter property could be affected by curves starting at other points from Argmin $L(\bar{x})$. Here, a careful multi-jet analysis as in [4] has to be performed. This is a topic of current research.

2.3 Unconstrained Lower Lever ($\dim(y) > 1$)

In [4] it has been conjectured that Classification Theorem 2.2 also holds true for arbitrary dimensions of y. However, the techniques which have been used in [4] can not directly be generalized. In the latter paper the proof was based on explicit unfoldings of singularities with finite codimension. For $\dim(y) > 1$ this is not possible anymore.

We present a new idea, how to prove the general result for the case where global optimality in the lower level is replaced by local optimality. This is of course a relaxation since the feasible set is enlarged. But since the arguments we use hold true for arbitrary dimensions of y, it might be possible to treat the globality aspect by some additional arguments as indicated below. Here, the problem of possible **bifurcations of global minimizers** is the main challenge.

First, we present the arguments for the case of local minimizers. We replace M in (2.4) by

$$M_{loc} := \{(x, y) \,|\, y \text{ is a local minimizer of } L(x)\}.$$

For the sake of simplicity we restrict our considerations to $\dim(y) = 2$, since the higher dimensions can be treated analogously. Further, we assume that $(\bar{x}, \bar{y}) \in M_{loc}$ minimizes f on M_{loc}. The optimality of $\bar{y}$ for the lower level problem $L(\bar{x})$ implies the following necessary optimality conditions:

$$D_y g(\bar{x}, \bar{y}) = 0 \in \mathbb{R}^m \quad \text{and} \quad D^2_{yy} g(\bar{x}, \bar{y}) \succeq 0.$$

Note that if $D^2_{yy} g(\bar{x}, \bar{y})$ is positive definite, then $\bar{y}$ is a nondegenerate local minimizer of $L(\bar{x})$. Our goal is to show that the cases where $D^2_{yy} g(\bar{x}, \bar{y})$ is not positive definite do not occur generically. For that, the space of symmetric $(2,2)$-matrices (where the Hessian $D^2_{yy}g$ lies in) can naturally be identified with $\mathbb{R}^3$ by using the data $(\partial^2_{y_1y_1}g, \partial^2_{y_1y_2}g, \partial^2_{y_2y_2}g)$. This way the cone of positive semi-definite matrices $\mathcal{C}$ is diffeomorphic to the so-called "ice cream cone" C^*, i.e.

$$\mathcal{C} = \phi(C^*), \quad C^* := \left\{(a,b,c) \in \mathbb{R}^3 \mid \sqrt{a^2+b^2} \le c,\, c \ge 0\right\},$$

where ϕ is a smooth diffeomorphism. We can partition $\mathcal{C}$ into three submanifolds of $\mathbb{R}^3$ as follows:

$$\mathcal{C} = \mathcal{C}_0 \cup \mathcal{C}_1 \cup \mathcal{C}_3, \quad \mathcal{C}_0 := \phi(\mathrm{int}(C^*)), \quad \mathcal{C}_1 := \phi(\partial C^* \setminus \{0\}), \quad \mathcal{C}_3 := \phi(\{0\}).$$

Note that the codimension of $\mathcal{C}_i \subset \mathbb{R}^3$ is

$$\mathrm{cod}(\mathcal{C}_i) = i,\ i \in \{0,1,3\}.$$

The fact that the Hessian $D^2_{yy}g$ belongs to $\mathcal{C}_i$ for a $i \in \{0,1,3\}$ corresponds to three different cases:

Case A: both eigenvalues of $D^2_{yy}g$ are positive (regular case),
Case B: exactly one eigenvalue of $D^2_{yy}g$ vanishes,
Case C: both eigenvalues of $D^2_{yy}g$ vanish.

In Case B the other eigenvalue is positive due to the positive semi-definiteness of $D^2_{yy}g$. After a change of coordinates (in y only) it holds:

$$g(0,y) = y_1^2 + \eta(y_2), \quad \frac{d\eta}{dy_2}(0) = \frac{d^2\eta}{dy_2^2}(0) = 0.$$

Thus, the y_1-dependence can be neglected here and Case B can be treated by using arguments from Classification Theorem 2.2. Hence, Case B does not occur in a generic situation.

In the following we will show how Case C can be excluded. For that, we apply the Jet Transversality Theorem (cf. [15]). The latter says that generically the jet mapping

$$J(x,y) := \left(f, g, Df, Dg, D^2f, D^2g, D^3g\right)_{|(x,y)} \in \mathbb{R}^\sigma$$

meets a manifold $\mathcal{M} \subseteq \mathbb{R}^\sigma$ transversally at $(\bar{x}, \bar{y})$. In particular this means

$$\mathrm{cod}(\mathcal{M}) \le \#\text{variables} = n+2. \tag{2.7}$$

By showing that Case C corresponds to an intersection of $J(\cdot)$ with a manifold $\mathcal{M}$ of $\mathrm{cod}(\mathcal{M}) > n+2$, the former can be generically excluded (cf. [15] for details). To do so, we may generically assume that the (reduced) jet mapping

$$j : (x,y) \mapsto \left(\partial_{y_1}g, \partial_{y_2}g, \partial^2_{y_1y_1}g, \partial^2_{y_1y_2}g, \partial^2_{y_2y_2}g\right)_{|(y_1,y_2)}$$

meets the manifold $\{0\}^2 \times \mathcal{C}_3 \subset \mathbb{R}^5$ transversally, i.e. the matrix $Dj(\bar{x},\bar{y})$ has full rank. This implies that there exist smooth mappings $z_i = z_i(x,y)$, $6 \le i \le n+2$, such that

$$\Psi(x,y) := \left(\partial_{y_1}g, \partial_{y_2}g, \partial^2_{y_1y_1}g, \partial^2_{y_1y_2}g, \partial^2_{y_2y_2}g, z_6, z_7, \ldots, z_{n+2}\right)_{|(x,y)},$$

is a local diffeomorphism with $\Psi(\bar{x},\bar{y}) = 0 \in \mathbb{R}^{n+2}$. We set $z = \Psi(x,y)$ as new local coordinates. Since the conditions $D_y g(x,y) = 0$ and $D^2_{yy} g(x,y) \succ 0$ are locally sufficient for y to be a local minimizer of $L(x)$ we obtain for the feasible set M_{loc}:

$$\Psi^{-1}\left(\{0\}^2 \times \mathcal{C}_0 \times \mathbb{R}^{n-3}\right) \subseteq M_{loc}. \tag{2.8}$$

(Note that $D^2_{yy} g \in \mathcal{C}_0$ means that $D^2_{yy} g$ is positive definite.) The latter inclusion implies

$$\overline{\Psi^{-1}\left(\{0\}^2 \times \mathcal{C}_0 \times \mathbb{R}^{n-3}\right)} = \Psi^{-1}\left(\{0\}^2 \times \mathcal{C} \times \mathbb{R}^{n-3}\right) \subseteq \overline{M_{loc}} \tag{2.9}$$

Since the optimality of f at $(\bar{x},\bar{y}) \in M_{loc}$ is equivalent to the optimality of f at $(\bar{x},\bar{y}) \in \overline{M_{loc}}$, we get that $(\bar{x},\bar{y})$ is a minimizer of f on $\overline{M_{loc}}$. In new coordinates this means, in particular, that

$$\frac{\partial (f \circ \Psi^{-1})}{\partial z_i}(0) = 0, \quad \text{for } 6 \le i \le n+2. \tag{2.10}$$

Thus $Df(\bar{x},\bar{y})$ lies in a linear space with codimension $n-3$. The situation of Case C now corresponds to an intersection of jet mapping $j(\cdot)$ with a manifold $\widetilde{\mathcal{M}}$ with

$$\text{cod}(\widetilde{\mathcal{M}}) = \text{cod}(\underbrace{\{0\}^2 \times \mathcal{C}_3}_{\text{Case C}}) + \underbrace{(n-3)}_{\text{optimality of } f} = 5 + (n-3) = n+2. \tag{2.11}$$

Again using the optimality of $\bar{y}$ for $L(\bar{x})$ and applying a Taylor approximation, one easily verifies that $D^3_{yyy}(\bar{x},\bar{y})$ must be equal to the 0-tensor. This implies that all third-order partial derivatives (w.r.t. y) of g vanish at $(\bar{x},\bar{y})$. The latter fact corresponds to an intersection of the jet mapping $J(\cdot)$ with a manifold of additional codimension 4. In fact, together with (2.11) we find that Case C corresponds to an intersection with a manifold $\mathcal{M}$ of $\text{cod}(\mathcal{M}) = (n+2) + 4 = n+6$. Thus, we obtain a contradiction to (2.7). We conclude that for local optimality in the lower level problem the result in Classification Theorem 2.2 can be generalized for $\dim(y) > 1$ as outlined above.

The adjustment of the arguments above for M itself (rather than M_{loc}) remains challenging. A submanifold of an appropriate dimension has to be identified within the set M (instead of M_{loc}) in order to obtain additional codimensions from the optimality of f on this submanifold (as in (2.10)). Instead of analyzing M explicitly, we propose to identify directions in the parameter space where bifurcations can be excluded. This means that only the subset of M_{loc} is considered where local optimality and global optimality coincide. For the latter it becomes crucial to analyze how the mapping

$$z \mapsto (\Pi_x \circ \Psi^{-1})(z), \quad \Pi_x : (x,y) \mapsto x$$

acts on the set S, where S represents the sufficient optimality conditions in the jet space (as the set $\{0\}^2 \times \mathcal{C}_0 \times \mathbb{R}^{n-3}$ in (2.8)). The latter issue is a topic of current research.

2.4 One-Dimensional Upper Level ($\dim(x) = 1$)

In the case where $\dim(x) = 1$ the lower level is one-parametric. Now, the classification can be done by using the so-called five types of parametric optimization given in [14]. In the following the five Types will shortly be introduced. For that, we consider parametric optimization problems for $x \in \mathbb{R}^1$

$$\text{NLP}(x): \quad \min_y f(x,y) \quad \text{s.t.} \quad g_j(x,y) \ge 0,\ j \in J.$$

Here, $J = \{1, \ldots, |J|\}$ is a finite index set and all appearing functions f, g_j, $j \in J$, are real-valued and from $C^\infty(\mathbb{R} \times \mathbb{R}^m)$.

A classification of *generalized critical points* was given in [14]. We recall that y is called a generalized critical point for NLP(x), if the vectors

$$D_y f(x,y),\ D_y g_j(x,y),\ j \in J_0(x,y)$$

are linearly dependent, i.e. there exist real numbers λ_0, λ_j, $j \in J_0(x,y)$ (Lagrange multipliers) - not all vanishing - such that

$$\lambda_0 D_y f(x,y) = \sum_{j \in J_0(x,y)} \lambda_j D_y g_j(x,y).$$

It was shown in [14] that for and open and dense subset of problem data, every generalized critical point $\bar{y}$ for NLP($\bar{x}$) is one of the *five types*. The Five Types originate from degeneracies of the generalized critical points occurring in one-parametric problems. A degeneracy describes the failure of at least one of the nondegeneracy properties ND1-ND3 from above. The following table introduces the five types of a generalized critical point $\bar{y}$ for NLP($\bar{x}$):

	ND1	ND2	ND3	description
Type 1	$\surd$	$\surd$	$\surd$	nondegenerate generalized critical point
Type 2	$\surd$	fails	$\surd$	exactly one active λ_j vanishes
Type 3	$\surd$	$\surd$	fails	exactly one eigenvalue (in ND3) vanishes
Type 4	fails	$\surd$	$\surd$	rank of active $D_y g_j$ is $\|J_0\| - 1$
Type 5	fails	$\surd$	$\surd$	$\|J_0\| = m + 1$

Note, that the description above only indicates the failure of the properties ND1-ND3 (for fixed x). For a precise definition some additional regularity conditions on the derivatives w.r.t. x and y have to be imposed (see [14]). In what follows, we are only interested in those generalized critical points $\bar{y}$ for NLP($\bar{x}$), which are global minimizers of NLP($\bar{x}$). This means that Type 3, which is not a global minimizer, will not be considered. The remaining generalized critical points are subdivided into Types 1, 2, 4, 5.1, and 5.2, where Types 5.1 and 5.2 are subtypes of Type 5. We recall that in Type 5 exactly $m+1$ constraints in the lower level problem are active, where m is the dimension of the y-variable. This leads to two possible scenarios: The feasible set ends at the critical value $\bar{x}$ (Type 5.1) or the local exchange of active constraints leads to a kink (Type 5.2). For details we refer to [18]. The corresponding typical sets of local minimizers are depicted in Figure 2.

Note that outside $\bar{x}$ the sets of local minimizers are graphs of smooth functions (cf. [18]). For those sets we write locally around $(\bar{x}, \bar{y})$

$$\begin{array}{rc} \text{Type 1:} & \{(x, y(x)) \,|\, x \approx \bar{x}\} \\ \text{Type 2, Type 5.2:} & \left\{ (x, y(x)) \,\middle|\, y(x) = \left\{ \begin{array}{ll} y^{(1)}(x), & x \le \bar{x}, \\ y^{(2)}(x), & x \ge \bar{x}. \end{array} \right. \right\} \\ \text{Type 4, Type 5.1:} & \{(x, y(x)) \,|\, x \ge \bar{x}\} \end{array}$$

where $y(x)$, $y^{(1)}(x)$, $y^{(2)}(x)$ are unique local minimizers for NLP(x) in a neighborhood of $\bar{x}$ with $y(\bar{x}) = y^{(1)}(\bar{x}) = y^{(2)}(\bar{x}) = \bar{y}$ corresponding to the particular type. The description of the lower level solutions sets leads to the following classification theorem:

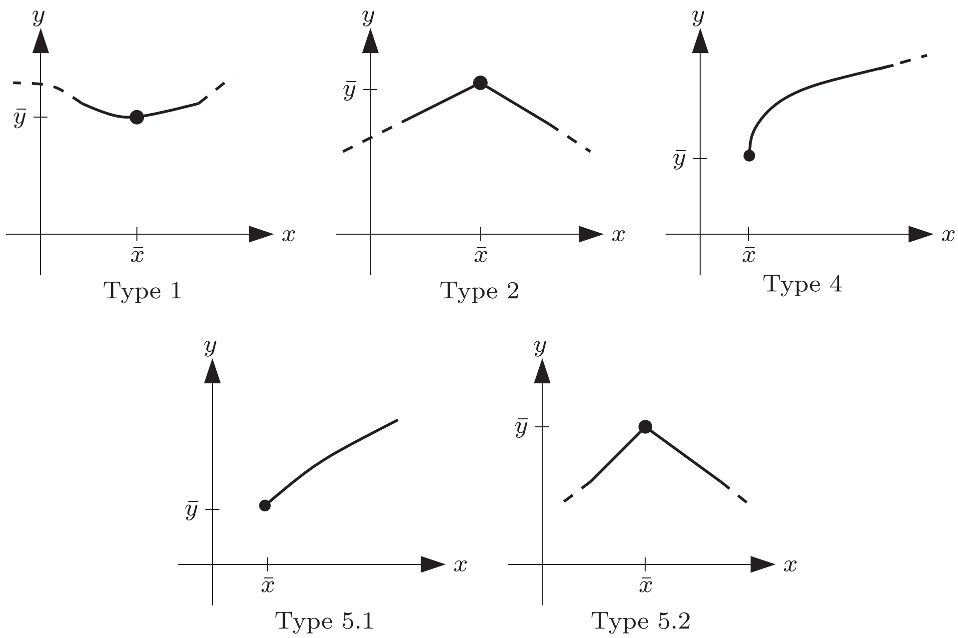

Figure 2: Types of local minimizers

Classification Theorem 2.3 (dim$(x) = 1$, y constrained)**.**
Let $\dim(x) = 1$*. Then, the bilevel optimization problem has the following basic classes:*

$$K_1 : \begin{cases} \text{The point } (\bar{x}, \bar{y}) \text{ is a nondegenerate minimizer of the (implicitly given) NLP} \\ \qquad \min\limits_{(x,y)} f(x,y) \quad s.t. \quad (x,y) \in \{(x, y(x)) \mid x \approx \bar{x}\}, \\ \text{where } y = y(x) \text{ is a smooth curve.} \end{cases}$$

$$K_{2/5.2} : \begin{cases} \text{The point } (\bar{x}, \bar{y}) \text{ is a nondegenerate minimizer of the (implicitly given) MPCC} \\ \qquad \min\limits_{(x,y)} f(x,y) \quad s.t. \quad (x,y) \in \{(x, y(x)) \mid x \approx \bar{x}\}, \\ \text{where } y(x) = \begin{cases} y^{(1)}(x) & x \le \bar{x} \\ y^{(2)}(x) & x \ge \bar{x} \end{cases} \text{ is a piecewise smooth curve.} \end{cases}$$

$$K_{4/5.1} : \begin{cases} \text{The point } (\bar{x}, \bar{y}) \text{ is a nondegenerate minimizer of the (implicitly given) NLP} \\ \qquad \min\limits_{(x,y)} f(x,y) \quad s.t. \quad (x,y) \in \{(x, y(x)) \mid x \approx \bar{x},\ x \ge \bar{x}\}, \\ \text{where } y = y(x) \text{ is a smooth curve.} \end{cases}$$

These basic classes are represented[§] *by the following local models:*

$$\min_{z\in\mathbb{R}^{1+m}} \left(\sum_{i=m-\delta_1+1}^{m-\delta_1+\delta_2} z_i + \sum_{i=m-\delta_1+\delta_2+1}^{m+\delta_2} (z_i + z_{i+1}) + \sum_{i=m+\delta_1+\delta_2}^{1+m} z_i^2 \right)$$
$$s.t. \quad z \in \{0\}^{m-\delta_1} \times \mathbb{H}^{\delta_2} \times \mathbb{L}^{\delta_1} \times \mathbb{R}^{1-\delta_1-\delta_2},$$

where $(\delta_1, \delta_2) \in \{(0,0),(1,0),(0,1)\}$.

Note that the basic class $K_{2/5.2}$ is not only induced by the Types 2 and 5.2, but also the case where Argmin $L(\bar{x}) = \{\bar{y}^{(1)}, \bar{y}^{(2)}\}$ is included here; cf. [5], and the nondegenerate case in Classification Theorem 2.2.

2.5 Copositive Lower Level

In this section we assume that the linear independence constraint qualification at the lower level is fulfilled, that the dimensions of the variables x and y coincide (i.e. $n = m$), that $J_0(\bar{x}, \bar{y}) = m$ and that $\bar{x} = \bar{y} = 0$. Taking the constraints h_j in (2.1) as new coordinates, we may assume that the lower level feasible set is the nonnegative orthant. In this setting, the Lagrange multipliers of the lower level function g at the origin just become the partial derivatives with respect to the coordinates y_j, $j = 1, \ldots, m$. Now we suppose that all these partial derivatives vanish (generalization of Type 2, cf. Section 2.4). This means, that the Hessian $D^2_{yy}g(0,0)$ comes into play and we assume that it is nonsingular. A stable condition for the origin to be a (local) minimizer for $L(0)$ is that the positive cone of the Hessian $D^2_{yy}g(0,0)$ contains the nonnegative orthant with deleted origin. The latter fact refers to the so-called **copositivity** of the matrix $D^2_{yy}g(0,0)$ (cf. [22] and also [2, 7]). There are several combinatorial possibilities, depending on the number of negative eigenvalues of $D^2_{yy}g(0,0)$. In the next Examples 2.4,2.5, we restrict ourselves to two dimensions, i.e. $n = m = 2$. Note that the feasible sets of the corresponding local models now **can not** be written as

$$\{0\}^{q_1} \times \mathbb{H}^{q_2} \times \mathbb{L}^{q_3} \times \mathbb{R}^{q_4}, \quad \text{with } q_1, q_2, q_3, q_4 \in \mathbb{N}$$

anymore. Instead, **multidimensional complementarity constraints** appear.

Example 2.4 (cf. [18])**.** Let $\dim(x) = \dim(y) = 2$ and the bilevel optimization problem be given as follows:

$$f(x_1, x_2, y_1, y_2) = (-x_1 + 2y_1) + (-x_2 + 2y_2)$$
$$L(x_1, x_2): \quad \min_y g(x_1, x_2, y_1, y_2) := (y_1 - x_1)^2 + (y_1 - x_1)\cdot(y_2 - x_2) + (y_2 - x_2)^2$$
$$\text{s.t.} \quad y_1 \geq 0,\ y_2 \geq 0.$$

In this example the Hessian $D^2_{yy}g(0,0)$ has two distinct positive eigenvalues. In particular, $D^2_{yy}g(0,0)$ is positive definite. In order to obtain the feasible set M, we have to consider critical points of $L(x_1, x_2)$ for the following four cases I-IV. These cases result from the natural stratification of the nonnegative orthant in y-space:

$$I: y_1 > 0,\ y_2 > 0 \quad II: y_1 = 0,\ y_2 > 0$$
$$III: y_1 > 0,\ y_2 = 0 \quad IV: y_1 = 0,\ y_2 = 0.$$

It turns out that the feasible set M is piecewise smooth two-dimensional manifold. Moreover, it can be parametrized via the x-variable by means of a subdivision of the x-space into four

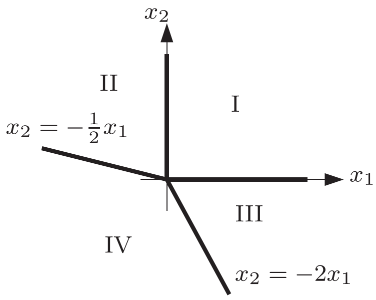

Figure 3: Illustration of Example 2.4

regions according to the above cases I-IV, see Figure 3. On the regions I-IV the corresponding global minimizer $(y_1(\cdot), y_2(\cdot))$ is given by:

$$(y_1(x), y_2(x)) = \begin{cases} (x_1, x_2), & \text{if } (x_1, x_2) \in I, \\ (0, \frac{x_1}{2} + x_2), & \text{if } (x_1, x_2) \in II, \\ (\frac{x_2}{2} + x_1, 0), & \text{if } (x_1, x_2) \in III, \\ (0, 0), & \text{if } (x_1, x_2) \in IV. \end{cases} \tag{2.12}$$

In particular, we obtain $M = \{(x, y(x)) \,|\, y(x) \text{ as in } (2.12)\}$. A few calculations show that the origin $(0, 0)$ solves the corresponding bilevel problem U.

The corresponding local model is

$$\min_z \sum_{i=1}^{4} z_i \quad \text{s.t.} \quad \begin{cases} (z_1 + z_2)(z_2 + z_3)(z_3 + z_4)(z_4 + z_1) = 0, \\ z_1 \geq 0,\ z_2 \geq 0,\ z_3 \geq 0,\ z_4 \geq 0. \end{cases}$$

Example 2.5 (cf. [5, 18])**.** Let $\dim(x) = \dim(y) = 2$ and the bilevel optimization problem be given as follows:

$$f(x_1, x_2, y_1, y_2) = -3x_1 + x_2 + 4y_1 + 5y_2$$

$$L(x_1, x_2): \quad \min_y g(x_1, x_2, y_1, y_2) := (y_1 - x_1)^2 + 4(y_1 - x_1) \cdot y_2 + 3(y_2 + \frac{1}{3}x_2)^2$$

$$\text{s.t.} \quad y_1 \geq 0,\ y_2 \geq 0.$$

It is easy to see that $(y_1, y_2) = (0, 0)$ is the global minimizer for $L(0, 0)$. In order to obtain the feasible set M, we have to consider critical points of $L(x_1, x_2)$ for the four cases I-IV as in Example 2.4. We subdivide the parameter space (x_1, x_2) into regions on which the global minimizer $(y_1(x), y_2(x))$ for L(x) is a smooth function. Here, we obtain three regions II-IV, see Figure 4. Note that the region corresponding to the case I is empty. In addition, for the parameters (x_1, x_2) lying on the half-line

$$G:\ x_1 = (2 + \sqrt{3})x_2,\ x_1 \geq 0$$

the problem $L(x_1, x_2)$ exhibits two different global minimizers. It is due to the fact that $(y_1, y_2) = (0, 0)$ is a saddlepoint of the objective function $g(0, y_1, y_2)$. Moreover, $(y_1, y_2) = (0, 0)$ is not strongly stable for $L(0, 0)$.

§According to Convention 1.4.

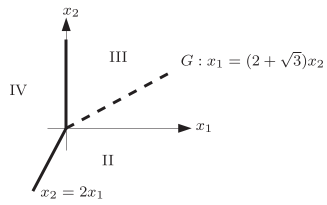

Figure 4: Illustration for Example 2.5

On the regions II-IV and on G the corresponding global minimizers $(y_1(\cdot),\ y_2(\cdot))$ are given by:

$$(y_1(x), y_2(x)) = \begin{cases} (0, \frac{2}{3}x_1 - \frac{1}{3}x_2), & \text{if } (x_1, x_2) \in II, \\ (x_1, 0), & \text{if } (x_1, x_2) \in III, \\ (0, 0), & \text{if } (x_1, x_2) \in IV, \\ \big\{(0, \frac{2}{3}x_1 - \frac{1}{3}x_2), (x_1, 0)\big\} & \text{if } (x_1, x_2) \in G. \end{cases} \tag{2.13}$$

Here, $M = \{(x, y(x)) \,|\, y(x) \text{ as in (2.13)}\}$. We point out that along the half-line G the bifurcation of lower level solutions occurs. The origin (0,0) solves the corresponding bilevel problem U.

The corresponding local model is

$$\min_z \sum_{i=1}^{4} z_i \quad \text{s.t.} \quad \begin{cases} (z_1 + z_2)(z_3 + z_4)(z_4 + z_1) = 0, \\ z_1 \geq 0,\, z_2 \geq 0,\, z_3 \geq 0,\, z_4 \geq 0. \end{cases}$$

3 General Semi-Infinite Programming

General semi-infinite programming problems (GSIPs) have the form

$$\text{GSIP:} \quad \min_x f(x) \quad \text{s.t.} \quad x \in M[g_0, \ldots, g_s] \tag{3.1}$$

with

$$M[g_0, \ldots, g_s] := \{x \in \mathbb{R}^n \mid g_0(x, y) \geq 0 \text{ for all } y \in Y(x)\}$$

and

$$Y(x) := \{y \in \mathbb{R}^m \mid g_k(x, y) \leq 0,\, 1 \leq k \leq s\}.$$

Here, $f \in C^\infty(\mathbb{R}^n)$ and $g_k \in C^\infty(\mathbb{R}^n \times \mathbb{R}^m)$, $0 \leq k \leq s$, are real-valued functions. In the case of a constant mapping $Y(\cdot) = Y$, we refer to semi-infinite programming problems (SIPs).

Note that testing feasibility for x means that $\psi(x) := \inf_{y \in Y(x)} g_0(x, y) \geq 0$. The function $\psi(x)$ is the optimal value of the parametric optimization problem

$$\min_y g_0(x, y) \quad \text{s.t.} \quad g_k(x, y) \leq 0,\, 1 \leq k \leq s.$$

From this perspective, one might think of GSIP as a game played by an agent against the nature. Indeed, if the agent changes the x-variables, the nature affects his feasible set by excluding or including some constraints. This can also be seen as a back coupling behavior.

Throughout this chapter we assume that the following standard assumption is fulfilled.

Assumption 3.1. The set-valued mapping $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is locally bounded.

Recall that the set-valued mapping $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called locally bounded if for each $\bar{x} \in \mathbb{R}^n$ there exists a neighborhood U of $\bar{x}$ such that $\bigcup_{x\in U} Y(x)$ is bounded in $\mathbb{R}^m$.

It is well-known that the feasible set M does not need to be closed (cf. [27]). Therefore, one considers the topological closure $\overline{M}$ of M instead. In [12, Theorem 2.2] an explicit description of $\overline{M}$ is provided. In fact, under Assumption 3.1 and the so-called Symmetric Mangasarian-Fromovitz Constraint Qualification (Sym-MFCQ) (see Definition 3.3 below) the closure of the feasible set is given by

$$\overline{M} = M^{max},$$

where

$$M^{max} = \{x \in \mathbb{R}^n \mid \sigma(x,y) \geq 0 \text{ for all } y \in \mathbb{R}^m\} \text{ and}$$
$$\sigma(x,y) := \max_{0\leq k\leq s} g_k(x,y).$$

Note that the above description of $\overline{M}$ is symmetric in the functions g_k, $0 \leq k \leq s$. This means that the function g_0 does not play any special role. It is shown in [12] that Sym-MFCQ on M^{max} is generic and stable under C^1-perturbations of the data functions. Thus, let the following assumption be fulfilled throughout this article.

Assumption 3.2. Sym-MFCQ is satisfied at all points of M^{max}.

Note that $\bar{x} \in M$ is a local minimizer of the continuous function f on M if and only if it is a local minimizer of f on $\overline{M}$. Hence, in the following we consider the relaxed problem

$$\overline{\text{GSIP}}: \quad \min_x f(x) \quad \text{s.t.} \quad x \in \overline{M}.$$

Let $\bar{x} \in \overline{M}$. We set

$$M(\bar{x}) := \{y \in \mathbb{R}^m \mid \sigma(\bar{x},y) = 0\}.$$

Note that $M(\bar{x})$ consists of the global minimizers of $\sigma(\bar{x},\cdot)$ with vanishing optimal value. We consider the well-known epigraph reformulation: $\bar{y}$ is a global minimizer of $\sigma(\bar{x},\cdot)$ with vanishing optimal value if and only if $(\bar{y},0)$ is a global minimizer of

$$Q(\bar{x}): \quad \min_{(y,z)\in\mathbb{R}^m\times\mathbb{R}} z \quad \text{s.t.} \quad z - g_k(\bar{x},y) \geq 0,\ 0 \leq k \leq s.$$

From the first-order optimality condition for $(\bar{y},0)$ we obtain that the corresponding polytope of Lagrange multipliers $\Delta(\bar{x},\bar{y})$ is nonempty:

$$\Delta(\bar{x},\bar{y}) := \left\{ (\gamma_k)_{k\in K_0(\bar{x},\bar{y})} \in \mathbb{R}^{|K_0(\bar{x},\bar{y})|} \;\middle|\; \begin{array}{l} \sum\limits_{k\in K_0(\bar{x},\bar{y})} \gamma_k D_y g_k(\bar{x},\bar{y}) = 0, \\ \sum\limits_{k\in K_0(\bar{x},\bar{y})} \gamma_k = 1,\ \gamma_k \geq 0,\ k \in K_0(\bar{x},\bar{y}) \end{array} \right\}.$$

$K_0(\bar{x},\bar{y}) := \{k \in \{0,\dots,s\} \mid g_k(\bar{x},\bar{y}) = 0\}$ is the active index set for $(\bar{y},0)$.

We set for $\bar{x} \in \overline{M}$ and $\bar{y} \in M(\bar{x})$

$$V(\bar{x},\bar{y}) := \left\{ \sum_{k\in K_0(\bar{x},\bar{y})} \gamma_k D_x g_k(\bar{x},\bar{y}) \;\middle|\; (\gamma_k)_{k\in K_0(\bar{x},\bar{y})} \in \Delta(\bar{x},\bar{y}) \right\}.$$

Moreover, we put: $V(\bar{x}) := \bigcup_{\bar{y}\in M(\bar{x})} V(\bar{x},\bar{y})$.

We mention the definition of Sym-MFCQ from Assumption 3.2.

Definition 3.3 (Sym-MFCQ, cf. [12]). Let $\bar{x} \in M^{max}$. The Symmetric Mangasarian-Fromovitz Constraint Qualification (Sym-MFCQ) is said to hold at $\bar{x}$ if there exists a vector $\xi \in \mathbb{R}^n$ such that for all $v \in V(\bar{x})$ it holds:

$$v \cdot \xi > 0.$$

Now, we are ready to specify the basic classes and local models for $\overline{\text{GSIP}}$. The Classification Theorem 3.4 follows mainly from [11] and [16].

Classification Theorem 3.4 ($\overline{\text{GSIP}}$). *Let Assumptions 3.1 and 3.2 be fulfilled. Then, $\overline{GSIP}$ has the following basic classes D_p, $1 \le p \le n$:*

$$D_p : \begin{cases} M(\bar{x}) = \{\bar{y}^{(1)}, \dots, \bar{y}^{(p)}\}, \\ (\bar{y}^{(j)}, 0),\ 1 \le j \le p, \text{ are nondegenerate minimizers of } Q(\bar{x}), \\ \text{The point } \bar{x} \text{ is a nondegenerate local minimizer of the following NLP:} \\ \qquad \min\limits_{x} f(x) \quad s.t. \quad z_j(x) \ge 0,\ 1 \le j \le p, \\ \text{where } (y^{(j)}(x), z_j(x)) \text{ are unique local minimizers of } Q(x) \text{ in a neighbor-} \\ \text{hood of } \bar{x} \text{ with } (y^{(j)}(\bar{x}), z_j(\bar{x})) = (\bar{y}^{(j)}, 0),\ 1 \le j \le p. \end{cases}$$

Moreover, for a given $1 \le p \le n$ the class D_p is represented by the following local model:

$$\min_{u} \quad \sum_{i=1}^{p} u_i + \sum_{i=p+1}^{n} u_i^2 \quad s.t. \quad u \in \mathbb{H}^p \times \mathbb{R}^{n-p}.$$

Remark 3.5 (Basic classes for KKT points in $\overline{\text{GSIP}}$). In [17] the critical point theory for GSIP has been established. There, Karush-Kuhn-Tucker points for $\overline{\text{GSIP}}$ are considered. The point $\bar{x} \in \overline{M}$ is called a Karush-Kuhn-Tucker (KKT) point if there exist $\bar{y}^{(1)}, \dots, \bar{y}^{(l)} \in M(\bar{x})$, $\bar{v}^{(i)} \in V(\bar{x}, \bar{y}^{(i)})$ and $\bar{\mu}_i \ge 0$, $1 \le i \le l$ such that

$$Df(\bar{x}) = \sum_{i=1}^{l} \bar{\mu}_i \bar{v}^{(i)}.$$

We point out that the basic classes for KKT points can be easily deduced from [16]. The corresponding local models fit into the context of the so-called disjunctive optimization (see [20]).

4 Nash Optimization

In Nash optimization there are N agents competitive to each other. Their behavior is governed by the so-called generalized Nash equilibrium problem (e.g., [8]). The latter is given as follows

$$\text{GNEP:} \quad \begin{array}{rl} \text{Find a vector} & x = (x^1, \dots, x^N) \in \mathbb{R}^n \\ \text{such that} & x^\nu \text{ solves } P_\nu(x^{-\nu}) \text{ for each } \nu \in \mathcal{N}. \end{array}$$

Such a solution x of GNEP is called a (generalized) Nash equilibrium. Here, each player ν from a finite player set $\mathcal{N} = \{1, \dots, N\}$, $N \in \mathbb{N}$, tries to solve the following parametric optimization problem:

$$\text{P}_\nu(x^{-\nu}): \quad \min_{x^\nu} f^\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad x^\nu \in M^\nu(x^{-\nu}),$$

where the feasible set $M^\nu(x^{-\nu})$ is defined as

$$M^\nu(x^{-\nu}) := \left\{x^\nu \in \mathbb{R}^{n_\nu} \mid g_j^\nu(x^\nu, x^{-\nu}) \geq 0,\, j \in J^\nu,\, G_j(x^\nu, x^{-\nu}) \geq 0,\, j \in \mathcal{J}\right\}.$$

For a given player $\nu \in \mathcal{N}$, the finite set $J^\nu = \{1, \ldots, |J^\nu|\}$ indexes player ν's *individual constraints*, and $\mathcal{J} = \{1, \ldots, |\mathcal{J}|\}$ is the index set for the *common constraints*, shared by all players. All data functions f^ν, $\nu \in \mathcal{N}$, g_j^ν, $j \in J^\nu$, $\nu \in \mathcal{N}$, and G_j, $j \in \mathcal{J}$, are real-valued and belong to $C^\infty(\mathbb{R}^n)$, with $n := \sum_{n \in \mathcal{N}} n_\nu$. As in standard notation, the symbol $x^{-\nu}$ denotes the vector formed by all players' variables except those of player ν. Occasionally, we write $(x^\nu, x^{-\nu})$ for $x \in \mathbb{R}^n$ to emphasize the ν-th players variables within x. We emphasize that in the description of the feasible set M^ν we explicitly distinguish between the player ν's *individual constraints* $g_j^\nu \geq 0$, $j \in J^\nu$, and the *common constraints* $G_j \geq 0$, $j \in \mathcal{J}$, shared by all players. The latter issue is motivated by various applications (see e.g. [8]).

Finally, we state the Nash optimization problem as follows

$$\text{N:} \quad \min_x f(x) \quad \text{s.t.} \quad x \in NE[f^\nu, g_j^\nu, j \in J^\nu, \nu \in \mathcal{N}, G_j, j \in \mathcal{J}], \tag{4.1}$$

where $NE[f^\nu, g_j^\nu, j \in J^\nu, \nu \in \mathcal{N}, G_j, j \in \mathcal{J}]$ is the set of Nash equilibria w.r.t GNEP given by data functions $f^\nu, g_j^\nu, j \in J^\nu, \nu \in \mathcal{N}, G_j, j \in \mathcal{J}$ and f is a real-valued objective function belonging to $C^\infty(\mathbb{R}^n)$. We refer to [21] for a generalization of Nash optimization to a hierarchical case.

We identify two subclasses of Nash optimization w.r.t. NEP and GNEP_0, where NEP $\subset$ $\text{GNEP}_0 \subset$ GNEP. Here, GNEP_0 is the subclass of GNEP, where only *individual constraints* occur in the definition of the players' feasible sets. This means that there are no *common constraints*, which have to be fulfilled by all players. NEP consists of those problems from GNEP_0, where the constraint functions of different players' subproblems only depend on the player's own decision variables. I.e., the choice of the opposing players has no influence on the single player's feasible set. Thus, the coupling in NEP is only due to the dependence of the players' objective functions on the other players' decisions.

The subclasses NEP (the classical Nash equilibrium problem) and GNEP_0 are defined as follows:

GNEP_0:	As GNEP, but $\mathcal{J} = \emptyset$ (i.e. no *common constraints*),
NEP:	As GNEP_0, but, additionally it holds: $g_j^\nu(x^\nu, x^{-\nu}) = g_j^\nu(x^\nu)$ for all $x \in \mathbb{R}^n$ and all $\nu \in \mathcal{N}$ (no coupling of the feasible sets).

Nash optimization w.r.t. NEP and GNEP_0 are defined as in (4.1).

Now, suppose that $\mathcal{J} = \emptyset$ and $\bar{x}$ is a Nash equilibrium of GNEP_0, moreover, for all $\nu \in \mathcal{N}$ the point $\bar{x}^\nu$ is a nondegenerate minimizer for $P_\nu(\bar{x}^{-\nu})$. Then, the system of necessary optimality conditions for $\bar{x}$ is given by concatenating the players' KKT conditions. For simplicity, we assume $N = 2$ and denote $(\bar{x}^1, \bar{x}^2)$ as $(\bar{x}, \bar{y})$. Thus, we obtain

$$\left\{\begin{array}{rcll} D_x f^1(\bar{x}, \bar{y}) - \displaystyle\sum_{j \in J_0^1(\bar{x}, \bar{y})} \bar{\lambda}_j^1 \cdot D_x g_j^1(\bar{x}, \bar{y}) & = & 0, & \\ D_y f^2(\bar{x}, \bar{y}) - \displaystyle\sum_{j \in J_0^2(\bar{x}, \bar{y})} \bar{\lambda}_j^2 \cdot D_y g_j^2(\bar{x}, \bar{y}) & = & 0, & \\ g_j^1(\bar{x}, \bar{y}) & = & 0 & \text{for all } j \in J_0^1(\bar{x}, \bar{y}), \\ g_j^2(\bar{x}, \bar{y}) & = & 0 & \text{for all } j \in J_0^2(\bar{x}, \bar{y}) \end{array}\right.$$

with multipliers $\bar{\lambda}_j^1 > 0$, $j \in J_0^1(\bar{x},\bar{y})$ and $\bar{\lambda}_j^2 > 0$, $j \in J_0^2(\bar{x},\bar{y})$.

Furthermore, we define the mapping

$$\mathcal{F}(x,y,\lambda^1,\lambda^2) := \begin{pmatrix} D_x f^1(x,y) - \sum\limits_{j\in J_0^1(\bar{x},\bar{y})} \lambda_j^1 \cdot D_x g_j^1(x,y) \\ D_y f^2(x,y) - \sum\limits_{j\in J_0^2(\bar{x},\bar{y})} \lambda_j^2 \cdot D_y g_j^2(x,y) \\ g_j^1(x,y),\ j \in J_0^1(\bar{x},\bar{y}), \\ g_j^2(x,y),\ j \in J_0^2(\bar{x},\bar{y}) \end{pmatrix},$$

where

$$\lambda^1 := \left(\lambda_j^1, j \in J_0^1(\bar{x},\bar{y})\right) \text{ and } \lambda^2 := \left(\lambda_j^2, j \in J_0^2(\bar{x},\bar{y})\right).$$

Note that $\mathcal{F}(\bar{x},\bar{y},\bar{\lambda}^1,\bar{\lambda}^2) = 0$. Moreover, every Nash equilibrium, sufficiently close to $(\bar{x},\bar{y})$, solves $\mathcal{F} = 0$ together with the corresponding unique multipliers. However, the Jacobian $D\mathcal{F}$ might be singular at $(\bar{x},\bar{y},\bar{\lambda}^1,\bar{\lambda}^2)$, and - even more - the Nash equilibrium $(\bar{x},\bar{y})$ need not to be isolated (in $\mathbb{R}^n$). This occurs despite of the fact that $\bar{x}$ is an isolated minimizer (in $\mathbb{R}^{n_1}$) for $P_1(\bar{y})$ and $\bar{y}$ is an isolated minimizer (in $\mathbb{R}^{n_2}$) for $P_2(\bar{x})$ (cf. [5]). This fact motivates the introduction of *jointly nondegenerate Nash equilibria* for GNEP_0s. For the case $N > 2$ the system $\mathcal{F} = 0$ can be defined in an analogous way.

Definition 4.1 (Jointly nondegenerate Nash equilibrium)**.** A Nash equilibrium $\bar{x}$ of GNEP_0 is called jointly nondegenerate, if for all $\nu \in \mathcal{N}$ the point $\bar{x}^\nu$ is a nondegenerate minimizer for $P_\nu(\bar{x}^{-\nu})$ and, additionally, it holds:

ND3* The matrix $D\mathcal{F}$ is nonsingular at $(\bar{x},\bar{\lambda})$,
where $\bar{\lambda}$ is the corresponding vector of unique multipliers.

Note that jointly nondegenerate Nash equilibria are isolated (in $\mathbb{R}^n$). We mention that the property ND3* does not imply ND3 (see Section 2.1), moreover, ND3* and ND3 are independent from each other (cf. [5]).

Classification Theorem 4.2 presents the solution of the classification problem for GNEP_0. For its proof we refer to [5].

Classification Theorem 4.2 (GNEP_0)**.** *$GNEP_0$ has the unique basic class*

$$C: \begin{cases} \bar{x} \textit{ is a jointly nondegenerate Nash equilibrium of } GNEP_0, \\ \textit{locally around } \bar{x} \textit{ the set of Nash equilibria is given by} \\ \qquad \bigcap\limits_{\nu=1}^{N} \left\{ \left(\xi\left(x^{-\nu}\right), x^{-\nu}\right) \mid x^{-\nu} \textit{ close to } \bar{x}^{-\nu} \right\} = \{\bar{x}\}, \\ \textit{where } \xi\left(x^{-\nu}\right) \textit{ is the unique local minimizers for } P_\nu(x^{-\nu}) \textit{ in a neighbor-} \\ \textit{hood of } \bar{x}^\nu \textit{ with } \xi\left(\bar{x}^{-\nu}\right) = \bar{x}^\nu \textit{ for all } \nu \in \mathcal{N}. \end{cases}$$

Moreover, the corresponding local model is trivial.

The same result as in Theorem 4.2 holds for NEP (cf. [5]).

Now, we turn our attention to Nash optimization with $\mathcal{J} \neq \emptyset$. The presence of common constraints makes the local structure of the set of Nash equilibria become rather involved. To see this, let $N = 2$, $\mathcal{J} \neq \emptyset$. Let $(\bar{x},\bar{y})$ be a Nash equilibrium. Moreover, we assume that

$\bar{x}$ is a nondegenerate minimizer for $P_1(\bar{y})$ and $\bar{y}$ is a nondegenerate minimizer for $P_2(\bar{x})$. We denote the vectors of corresponding Lagrange multipliers as $(\bar{\lambda}^1, \bar{\mu}^1)$ for $\bar{x}$ and $(\bar{\lambda}^2, \bar{\mu}^2)$ for $\bar{y}$. Hence, for $(\bar{x}, \bar{y}, \bar{\lambda}^1, \bar{\mu}^1, \bar{\lambda}^2, \bar{\mu}^2)$ the first order necessary optimality conditions are given by the following KKT system

$$\left\{ \begin{array}{l} D_x f^1(x,y) - \sum\limits_{j \in J_0^1(\bar{x},\bar{y})} \lambda_j^1 \cdot D_x g_j^1(x,y) - \sum\limits_{j \in \mathcal{J}_0(\bar{x},\bar{y})} \mu_j^1 \cdot D_x G_j(x,y) = 0, \\ D_y f^2(x,y) - \sum\limits_{j \in J_0^2(\bar{x},\bar{y})} \lambda_j^2 \cdot D_y g_j^2(x,y) - \sum\limits_{j \in \mathcal{J}_0(\bar{x},\bar{y})} \mu_j^2 \cdot D_y G_j(x,y) = 0, \\ \qquad g_j^1(x,y) = 0, \ j \in J_0^1(\bar{x},\bar{y}), \\ \qquad g_j^2(x,y) = 0, \ j \in J_0^2(\bar{x},\bar{y}), \\ \qquad G_j(x,y) = 0, \ j \in \mathcal{J}_0(\bar{x},\bar{y}). \end{array} \right. \tag{4.2}$$

Note that $(\bar{x}, \bar{y}, \bar{\lambda}^1, \bar{\mu}^1, \bar{\lambda}^2, \bar{\mu}^2)$ solves (4.2). Moreover, every Nash equilibrium, sufficiently close to $(\bar{x}, \bar{y})$, is a solution of (4.2) together with the corresponding unique multipliers. However, locally around $(\bar{x}, \bar{y})$ the set of Nash equilibria need not to be a singleton. Indeed, a calculation of the number of variables and equations in (4.2) yields:

$$\begin{array}{ll} \text{variables:} & \underbrace{n_1 + |J_0^1(\bar{x},\bar{y})| + |\mathcal{J}_0(\bar{x},\bar{y})|}_{\text{player 1}} + \underbrace{n_2 + |J_0^2(\bar{x},\bar{y})| + |\mathcal{J}_0(\bar{x},\bar{y})|}_{\text{player 2}}, \\ \text{equations:} & n_1 + n_2 + |J_0^1(\bar{x},\bar{y})| + |J_0^2(\bar{x},\bar{y})| + |\mathcal{J}_0(\bar{x},\bar{y})|. \end{array}$$

The KKT system (4.2) is underdetermined, namely, $|\mathcal{J}_0(\bar{x},\bar{y})|$ degrees of freedom are available in (4.2). This explains that locally around $(\bar{x}, \bar{y})$ the set of Nash equilibria constitutes a manifold of dimension $|\mathcal{J}_0(\bar{x},\bar{y})|$ (see Example 4.3). We recall that it happens despite of the fact that $\bar{x}$ (resp., $\bar{y}$) is a nondegenerate minimizer for $P_1(\bar{y})$ (resp., $P_2(\bar{x})$). For $N > 2$, the degrees of freedom add up to $(N-1) \cdot |\mathcal{J}_0(\bar{x},\bar{y})|$. These available degrees of freedom not only cause the local non-uniqueness of the Nash equilibria set as described above. But, in addition, possible violations of ND1-ND3 for players' subproblems correspond to $(N-1) \cdot |\mathcal{J}_0(\bar{x},\bar{y})|$ degrees of freedom (see Examples 4.3 and 4.4).

We present *stable* examples of GNEPs illustrating the phenomena described above. Note that these examples provide some basic classes of the Nash optimization problem along with their local models.

Example 4.3 (cf. [5, 6])**.** Let $N = 2$, $J^1 = J^2 = \emptyset$, $\mathcal{J} = \{1, 2\}$ and Nash optimization problem be given by the following data functions

$$f(x,y) = -y,$$

$$f^1(x,y) = -x, \quad f^2(x,y) = -y,$$

$$G_1(x,y) = 1 - x - y, \quad G_2(x,y) = x - y.$$

The global minimizers for both players are depicted in Figure 5. The set of Nash equilibria is the intersection of both sets, i.e. the half-line starting at $\left(\frac{1}{2}, \frac{1}{2}\right)$ as its boundary point. Note that $x = \frac{1}{2}$ is of Type 5.1 for $P_1(\frac{1}{2})$, and $y = \frac{1}{2}$ is of Type 5.2 for $P_2(\frac{1}{2})$. In particular, Mangasarian-Fromovitz constraint qualification is violated at $x = \frac{1}{2}$ for player 1, and the linear independence constraint qualification is violated at $y = \frac{1}{2}$ for player 2. These two degeneracies correspond to the number $(N-1)|\mathcal{J}_0(\frac{1}{2}, \frac{1}{2})| = 2$. The Nash equilibrium $\left(\frac{1}{2}, \frac{1}{2}\right)$ is the solution of the Nash optimization problem. The corresponding local model is

$$\min_z z \quad \text{s.t.} \quad z \in \mathbb{H}.$$

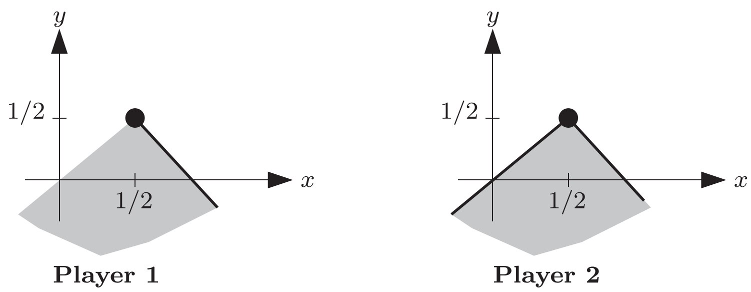

Figure 5: Global minimizers from Examples 4.3

Now, we consider the Nash equilibrium $(1, 0)$. Note that $x = 1$ is of Type 1 for $P_1(0)$, and $y = 0$ is also of Type 1 for $P_2(1)$. Hence, we have nondegenerate minimizers for both players' subproblems. Further, locally around $(1, 0)$ the set of Nash equilibria constitutes a one-dimensional manifold. This correspond to the crucial number $(N-1)|\mathcal{J}_0(1,0)| = 1$. The Nash equilibrium $(1, 0)$ is the solution of the Nash optimization problem with the new objective function

$$f(x, y) = (x-1)^2 + y^2.$$

The corresponding local model is

$$\min_z z^2 \quad \text{s.t.} \quad z \in \mathbb{R}.$$

Example 4.4 (cf. [5, 6])**.** Let $N = 2$, $J^1 = J^2 = \emptyset$, $\mathcal{J} = \{1, 2\}$ Nash optimization problem be given by the following data functions

$$f(x, y, t) = t,$$

$$f^1((x, y), t) = x, \quad f^2((x, y), t) = t,$$

$$G_1((x, y), t) = 1 - (x-t)^2 - (y-(1-2t))^2, \quad G_2((x, y), t) = 1 - x^2 - (y+1)^2.$$

The feasible sets $M^1(t)$ of player 1's subproblem $P_1(t)$ are depicted in Figure 6. The set of Nash equilibria is a half-parabola NE starting at $((0, 0), 0)$ as its boundary point (see Figure 7). Here, $(0, 0)$ is of Type 4.1 for $P_1(0)$. The point 0 can be seen as a point of Type 5.2 for $P_2(0, 0)$, since $\dim(t) + 1$ constraints are active there. These two degeneracies correspond to the number $(N-1)|\mathcal{J}_0((0, 0), 0)| = 2$. The Nash equilibrium $((0, 0), 0)$ is the solution of the Nash optimization problem. The corresponding local model is

$$\min_z z \quad \text{s.t.} \quad z \in \mathbb{H}.$$

Now, we consider an arbitrary Nash equilibrium $((\bar{x}, \bar{y}), \bar{t}) \in NE$ with $((\bar{x}, \bar{y}), \bar{t}) \neq ((0, 0), 0)$. Note that $(\bar{x}, \bar{y})$ is of Type 1 for $P_1(\bar{t})$ (i.e., a nondegenerate minimizer), and $\bar{t}$ is of Type 5.2 for $P_2(\bar{x}, \bar{y})$. Locally around $((\bar{x}, \bar{y}), \bar{t}))$ the set of Nash equilibria is one-dimensional. Altogether, one degeneracy of Type 5.2 and one dimension of the Nash equilibria set correspond to the number $(N-1)|\mathcal{J}_0((\bar{x}, \bar{y}), \bar{t})| = 2$. The Nash equilibrium $((\bar{x}, \bar{y}), \bar{t})$ is the solution of the Nash optimization problem with the new objective function

$$f(x, y, t) = (x-\bar{x})^2 + (y-\bar{y})^2 + (t-\bar{t})^2.$$

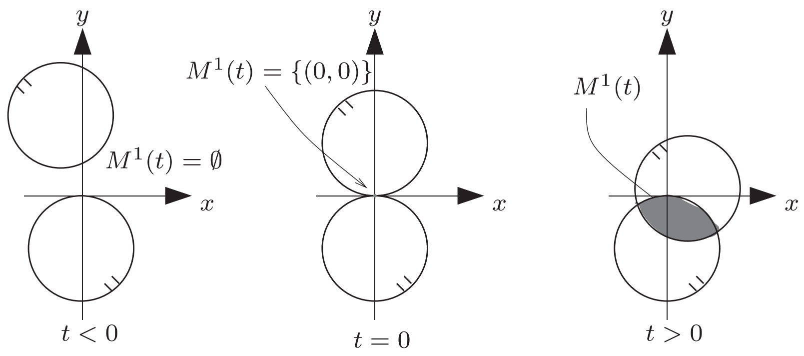

Figure 6: Feasible sets $M^1(t)$ from Example 4.4

The corresponding local model is

$$\min_z z^2 \quad \text{s.t.} \quad z \in \mathbb{R}.$$

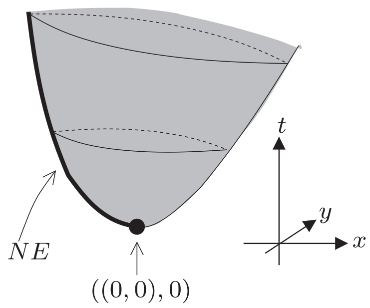

Figure 7: Set of Nash equilibria from Example 4.4

Note that local models in Examples 4.3 and 4.4 coincide despite of the fact that minimizers of the players' subproblems are of different types. We point out that the complete solution of the classification problem for Nash optimization is a topic of current research.

Acknowledgments

The authors would like to thank Harald Günzel for pointing out some gaps in the first version of the paper. Also, the authors would like to thank both anonymous referees for their valuable comments.

References

[1] J. F. Bard, *Practical Bilevel Optimization: Algorithms and Applications*, Kluwer Academic Publishers, Dordrecht, 1998.

[2] A. Berman and N. Shaked-Monderer, *Completely Positive Matrices,* World Scientific, Singapore, 2003.

[3] S. Dempe, *Foundations of Bilevel Programming*, Kluwer Academic Publishers, Dordrecht, 2002.

[4] S. Dempe, H. Günzel and H.Th. Jongen, On reducibility in bilevel problems, *SIAM Journal on Optimization* 20 (2009) 718–727.

[5] D. Dorsch, H. Th. Jongen and V. Shikhman, On intrinsic complexity of Nash equilibrium problems and bilevel optimization, *Journal of Optimization Theory and Applications*, 2012, DOI: 10.1007/s10957-012-0210-7.

[6] D. Dorsch, H.Th. Jongen and V. Shikhman, On structure and computation of generalized Nash equilibria, *SIAM Journal on Optimization* 23 (2013) 452474.

[7] M. Dür, Copositive Programming - a Survey, in *Recent Advances in Optimization and its Applications in Engineering*, M. Diehl, F. Glineur, E. Jarlebring, W. Michiels (eds.), Springer, Berlin, 2010, pp. 3–20.

[8] F. Facchinei and C. Kanzow, Generalized Nash equilibrium problems, *Annals of Operations Research* 175 (2010) 177–211.

[9] F. Facchinei, H. Jiang and L. Qi, A smoothing method for mathematical programs with equilibrium constraints, *Math. Program., Ser. A* 85 (1999 107–134.

[10] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems I, II,* Springer, Berlin, 2003.

[11] H. Günzel, H.Th. Jongen and O. Stein, Generalized semi-infinite programming: the symmetric reduction ansatz, *Optimization Letters* 2 (2008) 415–424.

[12] F. Guerra-Vázquez, H.Th. Jongen and V. Shikhman, General semi-infinite programming: symmetric Mangasarian Fromovitz constraint qualification and the closure of the feasible set, *SIAM Journal on Optimization* 20 (2010) 2487–2503.

[13] M. W. Hirsch, *Differential Topology*, Springer, Berlin-Heidelberg-New York, 1976.

[14] H.Th. Jongen, P. Jonker, F. Twilt, Critical sets in parametric optimization, *Mathematical Programming* 34 (1986) 333–353.

[15] H.Th. Jongen, P. Jonker and F. Twilt, *Nonlinear Optimization in Finite Dimensions*, Kluwer Academic Publishers, Dordrecht, 2000.

[16] H.Th. Jongen and V. Shikhman, Generalized semi-infinite programming: the nonsmooth symmetric reduction ansatz, *SIAM Journal on Optimization* 21 (2011) 193–211.

[17] H.Th. Jongen and V. Shikhman, General semi-infinite programming: critical point theory, *Optimization* 60 (2011) 859–873.

[18] H.Th. Jongen and V. Shikhman, Bilevel optimization: on the structure of the feasible set, *Mathematical Programming* 136 (2012) 65–89.

[19] H.Th. Jongen, J.-J. Rückmann and V. Shikhman, MPCC: critical point theory, *SIAM Journal on Optimization* 20 (2009) 473–484.

[20] H.Th. Jongen, J.-J. Rückmann and O. Stein, Disjunctive optimization: critical point theory, *Journal of Optimization Theory and Applications* 93 (1997) 321–336.

[21] Z.-Q. Luo, J.-S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[22] T.S. Motzkin, Copositive quadratic forms, National Bureau of Standards Report, No. 1818, 1952, pp. 11–12.

[23] J. V. Outrata, M. Kočvara and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints,* Kluwer Academic Publishers, Dordrecht, 1998.

[24] D. Ralph and O. Stein, The C-index: a new stability concept for quadratic programs with complementarity constraints, *Mathematics of Operations Research* 36 (2011) 504–526.

[25] H. Scheel and S. Scholtes, Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity, *Mathematics of Operations Research* 25 (2000) 1–22.

[26] S. Scholtes and M. Stöhr, How stringent is the linear independence assumption for mathematical programs with complementarity constraints?, *Math. Oper. Res.* 26 (2001) 851–863.

[27] O. Stein, *Bi-level Strategies in Semi-infinite Programming*, Kluwer, Boston, 2003.

Manuscript received 26 April 2012
revised 6 August 2012
accepted for publication 29 October 2012

D. Dorsch
Department of Mathematics – C, RWTH Aachen University
D-52056 Aachen, Germany
E-mail address: dorsch@mathc.rwth-aachen.de

H.Th. Jongen
RWTH Aachen University, D-52056 Aachen, Germany
E-mail address: jongen@rwth-aachen.de

V. Shikhman
Center for Operations Research and Econometrics (CORE)
Catholic University of Louvain (UCL)
1348 Louvain-la-Neuve, Belgium
E-mail address: vladimir.shikhman@uclouvain.be