A semi-infinite approach to design centering

Oliver Stein

RWTH Aachen University, Aachen, Germany stein@mathc.rwth-aachen.de

Summary. We consider design centering problems in their reformulation as general semi-infinite optimization problems. The main goal of the article is to show that the Reduction Ansatz of semi-infinite programming generically holds at each solution of the reformulated design centering problem. This is of fundamental importance for theory and numerical methods which base on the intrinsic bilevel structure of the problem.

For the genericity considerations we prove a new first order necessary optimality condition in design centering. Since in the course of our analysis also a certain standard semi-infinite programming problem turns out to be related to design centering, the connections to this problem are studied, too.

Key words: Optimality conditions, Reduction Ansatz, Jet transversality, Genericity.

1 Introduction

Design Centering. A design centering problem considers a container set $C \subset \mathbb{R}^m$ and a parametrized body $B(x) \subset \mathbb{R}^m$ with parameter vector $x \in \mathbb{R}^n$. The task is to inscribe $B(x)$ into C such that some functional f, e.g. the volume of $B(x)$, is maximized:

$$DC: \quad \max_{x \in \mathbb{R}^n} \; f(x) \quad \text{subject to} \quad B(x) \subset C\,.$$

In Figure 1 $B(x)$ is a disk in $\mathbb{R}^2$, parametrized by its midpoint and its radius. The parameter vector $x \in \mathbb{R}^3$ is chosen such that $B(x)$ has maximal area in the nonconvex container set C.

A straightforward extension of the model is to inscribe finitely many nonoverlapping bodies into C such that some total measure is maximized. Figure 2 shows the numerical solution of such a multi-body design centering

S. Dempe and V. Kalashnikov (eds.), *Optimization with Multivalued Mappings*, pp. 209-228

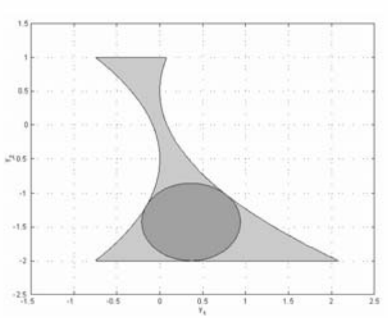

Fig. 1. A disk with maximal area in a nonconvex container

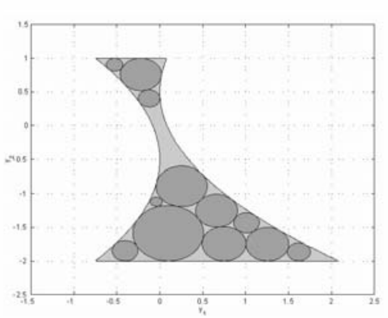

Fig. 2. Twelve disks with maximal total area in a nonconvex container

problem with the same container set as in Figure 1 and twelve nonoverlapping disks.

Single-body design centering problems with special sets $B(x)$ and C have been studied extensively, see e.g. [5] for the complexity of inscribing a convex body into a convex container, [12] for maximization of a production yield under uncertain quality parameters, and [18] for the problem of cutting a diamond with prescribed form and maximal volume from a raw diamond. The cutting stock problem ([2]) is an example of multi-body design centering.

To give an example of a design centering problem with a rather intricate container set, consider the so-called maneuverability problem of a robot from [4]:

Example 1. A robot may be viewed as a structure of connected links, where some geometrical parameters $\theta_1, ..., \theta_R$, such as lengths of the links or angles in the joints, can be controlled by drive motors (cf. Figure 3 which is taken from [8]).

The equations of motion for a robot have the form

$$F \;=\; A(\theta)\cdot\ddot{\theta} \;+\; H(\theta,\dot{\theta}) \;,$$

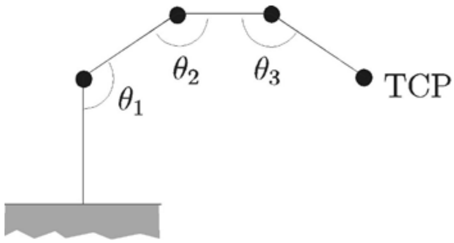

Fig. 3. A robot with connected links and a tool center point

where $F \in \mathrm{IR}^R$ denotes the vector of forces (torques), $A(\theta)$ is the inertia matrix, and $H(\theta, \dot{\theta})$ is the vector of friction, gravity, centrifugal and Coriolis forces. Given vectors $F^-, F^+ \in \mathrm{IR}^R$ of lower and upper bounds of F as well as an operating region $\Omega \subset \mathrm{IR}^R \times \mathrm{IR}^R$, the set

$$C \;=\; \{\, \ddot{\theta} \in \mathrm{IR}^R |\; F^- \leq A(\theta)\ddot{\theta} + H(\theta, \dot{\theta}) \leq F^+ \;\text{ for all }\; (\theta, \dot{\theta}) \in \Omega \,\}$$

describes the accelerations which can be realized in every point $(\theta, \dot{\theta}) \in \Omega$. Since the size of C is a measure for the usefulness of a given robot for certain tasks, an approximation for the volume of C is sought in [4]: Find a simple body B which is parametrized by a vector x such that $B(x)$ is as large as possible and contained in C. In this way we arrive at a design centering problem DC.

The aim of this article is to use techniques from general semi-infinite programming to treat a broad class of design centering problems theoretically as well as numerically. In fact, Example 1 gave rise to one of the first formulations of a general semi-infinite optimization problem in [8].

Semi-infinite Programming. The connection of design centering to semi-infinite programming is straightforward: let C be described by the inequality constraint $c(y) \leq 0$. Then the inclusion

$$B(x) \;\subset\; C \;=\; \{\, y \in \mathrm{IR}^m |\; c(y) \leq 0 \,\}$$

is trivially equivalent to the semi-infinite constraint

$$c(y) \;\leq\; 0 \quad \forall\, y \in B(x)\;.$$

Thus the design centering problem DC is equivalent to the general semi-infinite problem

$$GSIP_{DC}: \qquad \max_x f(x) \quad \text{subject to} \quad c(y) \;\leq\; 0 \quad \forall\, y \in B(x)\;.$$

Problems of this type are called semi-infinite as they involve a finite-dimensional decision variable x and possibly infinitely many inequality constraints

$$g(x,y) \le 0 \quad \forall\, y \in B(x)\ ,$$

where in design centering the function $g(x,y) := c(y)$ does not depend on x.

On the other hand, in a so-called standard semi-infinite optimization problem there is no $x-$dependence in the set $B(x)$, i.e. the semi-infinite index set $B(x) \equiv B$ is fixed. Standard semi-infinite optimization problems have been studied systematically since the early 1960s. For an extensive survey on standard semi-infinite programming see [7].

As it turned out more recently in [16], general semi-infinite programming is intrinsically more complicated than standard semi-infinite programming, so that some basic theoretical and numerical strategies cannot be transferred from the standard to the general case. In particular, the feasible set M of *GSIP* may be nonclosed and exhibit a disjunctive structure even for defining functions in general position. An introduction to general semi-infinite programming is given in [21].

Bilevel Programming. The key to the theoretical treatment of general semi-infinite programming and to the conceptually new solution method from [23] lies in the bilevel structure of semi-infinite programming. In the following we briefly sketch the main ideas of this approach.

Consider the general semi-infinite program

$$GSIP: \qquad \max_x f(x) \quad \text{subject to} \quad g(x,y) \le 0 \quad \forall\, y \in B(x)\ ,$$

where for all $x \in \mathbb{R}^n$ we have

$$B(x) = \{\, y \in \mathbb{R}^m \,|\, w(x,y) \le 0 \,\}\ .$$

Let the defining functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g, w : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be at least once continuously differentiable, and let $\nabla_x g$ denote the column vector of partial derivatives of g with respect to x, etc. Then the set-valued mapping $B : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is closed. Let B also be locally bounded, i.e. for all $\bar{x} \in \mathbb{R}^n$ there exists a neighborhood U of $\bar{x}$ and a bounded set $Y \subset \mathbb{R}^m$ with $B(x) \subset Y$ for all $x \in U$. Note that then $B(x)$ is compact for each $x \in \mathbb{R}^n$. We also assume that $B(x)$ is nonempty for all $x \in \mathbb{R}^n$.

Under these assumptions it is easy to see that the semi-infinite constraint in *GSIP* is equivalent to

$$\varphi(x) := \max_{y \in B(x)} g(x,y) \le 0\ ,$$

which means that the feasible set M of *GSIP* is the lower level set of some optimal value function. In fact, φ is the optimal value function of the so-called lower level problem

$$Q(x): \quad \max_{y \in \mathbb{R}^m} g(x,y) \quad \text{subject to} \quad w(x,y) \le 0\ .$$

In contrast to the upper level problem which consists in maximizing f over M, in the lower level problem x plays the role of an $n-$dimensional parameter, and y is the decision variable. The main computational problem in semi-infinite programming is that the lower level problem has to be solved to global optimality, even if only a stationary point of the upper level problem is sought.

Since under the assumptions of closedness and local boundedness of the set-valued mapping B and the continuity of g the optimal value function φ is at least upper semi-continuous, points $x \in \mathbb{R}^n$ with $\varphi(x) < 0$ belong to the topological interior of M. For investigations of the local structure of M or of local optimality conditions we are only interested in points from the boundary ∂M of M, so that it suffices to consider the zeros of φ, i.e. points $x \in \mathbb{R}^n$ for which $Q(x)$ has vanishing maximal value. We denote the corresponding globally maximal points of $Q(x)$ by

$$B_0(x) \;=\; \{\; y \in B(x) |\; g(x, y) = 0 \;\} \;.$$

The Reduction Ansatz. When studying semi-infinite problems, it is of crucial importance to control the elements of $B_0(x)$ for varying x. This can be achieved, for example, by means of the implicit function theorem. For $\bar{x} \in M$ a local maximizer $\bar{y}$ of $Q(\bar{x})$ is called nondegenerate in the sense of Jongen/Jonker/Twilt ([14]), if the linear independence constraint qualification (LICQ), strict complementary slackness (SCS) and the second order sufficiency condition $D_y^2 \Lambda(\bar{x}, \bar{y}, \bar{\gamma})|_{T_{\bar{y}} B(\bar{x})} \prec 0$ are satisfied. Here $\Lambda(x, y, \gamma) = g(x, y) - \gamma\, w(x, y)$ denotes the lower level Lagrangian, $T_{\bar{y}} B(\bar{x})$ is the tangent space to $B(\bar{x})$ at $\bar{y}$, and $A \prec 0$ stands for the negative definiteness of a matrix A. The *Reduction Ansatz* is said to hold at $\bar{x} \in M$ if all global maximizers of $Q(\bar{x})$ are nondegenerate. Since nondegenerate maximizers are isolated, and $B(\bar{x})$ is a compact set, the set $B_0(\bar{x})$ can only contain finitely many points. By a result from [3] the local variation of these points with x can be described by the implicit function theorem.

The Reduction Ansatz was originally formulated for standard semi-infinite problems in [6] and [24] under weaker regularity assumptions. It was transferred to general semi-infinite problems in [9]. For *standard* semi-infinite problems the Reduction Ansatz is a natural assumption in the sense that for problems with defining functions in general position it holds at each local maximizer ([19, 25]). For *GSIP* this result can be transferred to local maximizers $\bar{x}$ with $|B_0(\bar{x})| \geq n$ ([20]). Moreover, in [22] it is shown that it holds in the "completely linear" case, i.e. when the defining functions f, g and w of *GSIP* are affine linear on their respective domains. For *GSIP* without these special structures, until now it is not known whether the Reduction Ansatz generically holds at all local maximizers. Note that even if this general result was true, it would not necessarily mean that the Reduction Ansatz holds generically at local maximizers of $GSIP_{DC}$. In fact, only such specially structured perturbations of the defining functions of $GSIP_{DC}$ are allowed which leave the function c independent of x.

Under the Reduction Ansatz it was not only shown that M can locally be described by finitely many smooth inequality constraints ([9]), but it also serves as a regularity condition for the convergence proof of the numerical solution method from [23]. For completeness, we briefly sketch the main idea of this bilevel method.

A numerical method for *GSIP*. To make the global solution of the lower level problem computationally tractable, we assume that $Q(x)$ is a regular convex problem for all $x \in \mathbb{R}^n$, i.e. the functions $-g(x,\cdot)$ and $w(x,\cdot)$ are convex in y, and $B(x)$ possesses a Slater point. It is well-known that then the global solutions of the problem $Q(x)$ are exactly its Karush-Kuhn-Tucker points: y solves $Q(x)$ if and only if there exists some $\gamma \in \mathbb{R}$ such that

$$\begin{aligned} \nabla_y \Lambda(x,y,\gamma) &= 0 \\ \gamma \cdot w(x,y) &= 0 \\ \gamma,\ -w(x,y) &\geq 0\ . \end{aligned}$$

For this reason it makes sense to replace the problem *GSIP*, in which only optimal *values* of the lower problem enter, by a problem which also uses lower level optimal *points*. In fact, we first consider the Stackelberg game

$$SG: \qquad \max_{x,y} f(x) \quad \text{subject to} \quad g(x,y) \leq 0, \quad y \ \text{ solves } \ Q(x)\ .$$

Note that the decision variable of SG resides in the higher-dimensional space $\mathbb{R}^n \times \mathbb{R}^m$, i.e. *GSIP* is lifted. In [22] it is shown that under our assumptions the orthogonal projection of the feasible set of SG to $\mathbb{R}^n$ coincides with the feasible set of *GSIP*, so that the $x-$component of any solution of SG is a solution of *GSIP*.

In a second step we replace the restriction that y solves $Q(x)$ in SG equivalently by the corresponding Karush-Kuhn-Tucker condition:

$$\begin{aligned} MPCC: \qquad \max_{x,y,\gamma} f(x) \quad \text{subject to} \quad g(x,y) &\leq 0 \\ \nabla_y \Lambda(x,y,\gamma) &= 0 \\ \gamma \cdot w(x,y) &= 0 \\ \gamma,\ -w(x,y) &\geq 0\ . \end{aligned}$$

The resulting mathematical program with complementarity constraints lifts the problem again to a higher-dimensional space, but now *MPCC* solution techniques may be applied. One possibility is to reformulate the complementarity conditions in *MPCC* by means of an NCP function Φ like the Fischer-Burmeister function $\Phi(a,b) = a+b-||(a,b)||_2$, and then to regularize the necessarily nonsmooth or degenerate NCP function by a one-dimensional parameter $\tau > 0$, e.g. to $\Phi_\tau(a,b) = a+b-||(a,b,\tau)||_2$. An obvious idea for a numerical method is to solve the finite and regular optimization problems

$$\begin{aligned} P_\tau: \qquad \max_{x,y,\gamma} f(x) \quad \text{subject to} \quad g(x,y) &\leq 0 \\ \nabla_y \Lambda(x,y,\gamma) &= 0 \\ \Phi_\tau(\,\gamma, -w(x,y)\,) &= 0 \end{aligned}$$

for $\tau \searrow 0$. For details and for a convergence proof of this method see [21].

As mentioned before, this convergence proof relies on the Reduction Ansatz in the solution point. Although for general semi-infinite problems it is not clear yet whether the Reduction Ansatz holds generically in each local solution, in numerical tests convergence can usually be observed. The numerical examples in Figures 1 and 2 were actually generated by this algorithm, applied to the general semi-infinite reformulation $GSIP_{DC}$ of DC.

The present article will show that for the specially structured problems $GSIP_{DC}$ which stem from a reformulation of DC, the Reduction Ansatz in each local maximizer is generic. In Section 2 we derive a first order necessary optimality condition for DC which will be the basis of the genericity considerations in Section 3. Section 4 presents some connections to a *standard* semi-infinite problem that can be associated with DC, before Section 5 closes the article with some final remarks.

2 First order optimality conditions

Let us consider the slightly more general design centering problem

$$DC: \quad \max_{x \in \mathbb{R}^n} \; f(x) \quad \text{subject to} \quad B(x) \subset C$$

with

$$C \;=\; \{\, y \in \mathbb{R}^m |\; c_j(y) \leq 0,\; j \in J \,\}$$

and

$$B(x) \;=\; \{\, y \in \mathbb{R}^m |\; v_\ell(y) \leq 0,\; \ell \in L,\; w(x,y) \leq 0 \,\}$$

with finite index sets J and L, and with at least once continuously differentiable defining functions f, c_j , $j \in J$, v_ℓ , $\ell \in L$, and w. We assume that C and

$$Y \;=\; \{\, y \in \mathbb{R}^m |\; v_\ell(y) \leq 0,\; \ell \in L \,\}$$

are nonempty and compact sets. In applications the set Y can often be chosen to contain C so that the compactness of C follows from the compactness of Y. Moreover, the local boundedness of the set-valued mapping B is a trivial consequence of the boundedness of Y.

The general semi-infinite reformulation of DC now becomes a problem with finitely many semi-infinite constraints,

$$GSIP_{DC}: \qquad \max_x \; f(x) \quad \text{subject to} \quad c_j(y) \;\leq\; 0 \quad \forall\, y \in B(x)\,,\; j \in J\,,$$

and finitely many lower level problems $Q^j(x)$ with optimal value functions $\varphi_j(x)$ and optimal points $B_0^j(x)$, $j \in J$. For $\bar{x} \in M$ we denote by

$$J_0(\bar{x}) \;=\; \{\, j \in J |\; \varphi_j(\bar{x}) = 0 \,\}$$

the set of active semi-infinite constraints. From the upper semi-continuity of the functions φ_j, $j \in J$, it is clear that at each feasible boundary point $\bar{x} \in M \cap \partial M$ the set $\bigcup_{j \in J_0(\bar{x})} B_0^j(\bar{x})$ is nonempty. For the problem $GSIP_{DC}$ we can show that an even smaller set is nonempty. In fact, with

$$B_{00}^j(\bar{x}) \;=\; \{\; y \in B_0^j(\bar{x}) |\; w(\bar{x}, y) = 0 \;\}$$

the following result holds.

Lemma 1. *The set* $\bigcup_{j \in J_0(\bar{x})} B_{00}^j(\bar{x})$ *is nonempty for each feasible boundary point* $\bar{x} \in M \cap \partial M$.

Proof. For $\bar{x} \in \partial M$ there exists a sequence $x^\nu \to \bar{x}$ with $x^\nu \notin M$ for all $\nu \in \mathbb{N}$. By definition of M, for all $\nu \in \mathbb{N}$ there exists some $y^\nu \in B(x^\nu)$ and some $j_\nu \in J$ with $c_{j_\nu}(y^\nu) > 0$.

As J is a finite set, the sequence $(j_\nu)_{\nu \in \mathbb{N}}$ contains some index $j_0 \in J$ infinitely many times. Taking the corresponding subsequence if necessary, we may assume $j_\nu \equiv j_0$ without loss of generality.

Moreover, as B is locally bounded at $\bar{x}$, the sequence $(y^\nu)_{\nu \in \mathbb{N}}$ is bounded and, thus, without loss of generality convergent to some $\bar{y} \in \mathbb{R}^m$. From the closedness of the set-valued mapping B and $x^\nu \to \bar{x}$ we also obtain $\bar{y} \in B(\bar{x})$. The feasibility of $\bar{x}$ means that for all $j \in J$ and all $y \in B(\bar{x})$ we have $c_j(y) \le 0$, so that we arrive at

$$0 \;\le\; \lim_{\nu \to \infty} c_{j_0}(y^\nu) \;=\; c_{j_0}(\bar{y}) \;\le\; 0\,.$$

This implies $\bar{y} \in B_0^{j_0}(\bar{x})$ as well as $j_0 \in J_0(\bar{x})$.

Next, assume that for some $\nu \in \mathbb{N}$ it is $w(\bar{x}, y^\nu) \le 0$. Since we have $y^\nu \in Y$, it follows $y^\nu \in B(\bar{x})$. From $\bar{x} \in M$ we conclude that $c_{j_0}(y^\nu) \le 0$, in contradiction to the construction of y^ν. Consequently we have

$$\text{for all } \nu \in \mathbb{N} : \quad 0 \;<\; w(\bar{x}, y^\nu)\;. \tag{1}$$

Together with $y^\nu \in B(x^\nu)$ for all $\nu \in \mathbb{N}$ it follows

$$0 \;\le\; \lim_{\nu \to \infty} w(\bar{x}, y^\nu) \;=\; w(\bar{x}, \bar{y}) \;=\; \lim_{\nu \to \infty} w(x^\nu, y^\nu) \;\le\; 0$$

and thus $\bar{y} \in B_{00}^{j_0}(\bar{x})$. □

A usual starting point for genericity considerations is a first order optimality condition which holds without any regularity assumptions. For general semi-infinite problems

$$GSIP: \qquad \max_x f(x) \quad \text{subject to} \quad g_j(x, y) \;\le\; 0 \quad \forall\, y \in B(x)\,,\; j \in J\,,$$

such a condition is given in [16]. To formulate this condition, we denote by

$$\Lambda_j(x,y,\alpha,\beta,\gamma) \;=\; \alpha\, g_j(x,y) - \beta^\top v(y) - \gamma\, w(x,y)\,,\; j \in J,$$

the Fritz-John type lower level Lagrangians, and for $\bar{x} \in M$, $j \in J_0(\bar{x})$ and $\bar{y} \in B_0^j(\bar{x})$ by

$$\begin{aligned} FJ^j(\bar{x},\bar{y}) \;=& \{(\alpha,\beta,\gamma) \in \mathbb{R} \times \mathbb{R}^{|L|} \times \mathbb{R}|\; (\alpha,\beta,\gamma) \geq 0,\; ||(\alpha,\beta,\gamma)||_1 = 1, \\ & \nabla_y \Lambda_j(\bar{x},\bar{y},\alpha,\beta,\gamma) = 0,\; \Lambda_j(\bar{x},\bar{y},\alpha,\beta,\gamma) = 0\;\} \end{aligned}$$

the corresponding sets of Fritz-John multipliers.

Theorem 1 ([16]). *Let $\bar{x} \in M \cap \partial M$ be a local maximizer of GSIP. Then there exist $p_j \in \mathbb{N}$, $\bar{y}^{j,k} \in B_0^j(\bar{x})$, $(\alpha_{j,k},\beta_{j,k},\gamma_{j,k}) \in FJ^j(\bar{x},\bar{y}^{j,k})$, and nontrivial multipliers $\kappa \geq 0$, $\lambda_{j,k} \geq 0$, $1 \leq k \leq p_j$, $j \in J_0(\bar{x})$, such that $\sum_{j \in J_0(\bar{x})} p_j \leq n+1$ and*

$$\kappa \nabla f(\bar{x}) \;-\; \sum_{j \in J_0(\bar{x})} \sum_{k=1}^{p_j} \lambda_{j,k}\, \nabla_x \Lambda_j(\bar{x},\bar{y}^{j,k},\alpha_{j,k},\beta_{j,k},\gamma_{j,k}) \;=\; 0\;.$$

This condition simplifies significantly for the problem $GSIP_{DC}$. In fact, in the lower level Lagrangians

$$\Lambda_j(x,y,\alpha,\beta,\gamma) \;=\; \alpha\, c_j(y) - \beta^\top v(y) - \gamma\, w(x,y)\,,\; j \in J\,,$$

only the function w depends on x, so that we obtain

$$\nabla_x \Lambda_j(x,y,\alpha,\beta,\gamma) \;=\; -\gamma\, \nabla_x w(x,y)\;.$$

The following result is thus immediate.

Corollary 1. *Let $\bar{x} \in M \cap \partial M$ be a local maximizer of DC. Then there exist $p_j \in \mathbb{N}$, $\bar{y}^{j,k} \in B_0^j(\bar{x})$, $(\alpha_{j,k},\beta_{j,k},\gamma_{j,k}) \in FJ^j(\bar{x},\bar{y}^{j,k})$, and nontrivial multipliers $\kappa \geq 0$, $\lambda_{j,k} \geq 0$, $1 \leq k \leq p_j$, $j \in J_0(\bar{x})$, such that $\sum_{j \in J_0(\bar{x})} p_j \leq n+1$ and*

$$\kappa \nabla f(\bar{x}) \;+\; \sum_{j \in J_0(\bar{x})} \sum_{k=1}^{p_j} \lambda_{j,k}\, \gamma_{j,k}\, \nabla_x w(\bar{x},\bar{y}^{j,k}) = \; 0\;. \tag{2}$$

A major disadvantage of condition (2) is that it does not guarantee the linear dependence of the vectors $\nabla f(\bar{x})$, $\nabla_x w(\bar{x},\bar{y}^{j,k})$, $1 \leq k \leq p_j$, $j \in J_0(\bar{x})$. In fact, it is easy to construct situations in which $\kappa = 0$ and $\gamma^{j,k} = 0$, $1 \leq k \leq p_j$, $j \in J_0(\bar{x})$. Since the linear dependence of these vectors is crucial for genericity investigations, next we will give a stronger optimality condition.

It is not surprising that this strengthening is possible if one compares the situation to that of standard semi-infinite programming: also there only one of the lower level defining functions depends on x, namely $g_j(x,y)$. The corresponding first order optimality condition deduced from Theorem 1 involves multiplier products $\lambda_{j,k}\,\alpha_{j,k}$ as coefficients of the vectors $\nabla_x g_j(\bar{x},\bar{y}^{j,k})$, whereas from John's original condition for standard semi-infinite programs ([13]) it is clear that a single coefficient $\mu_{j,k}$ would suffice.

Theorem 2. *Let $\bar{x} \in M \cap \partial M$ be a local maximizer of DC. Then there exist $p_j \in I\!N$, $\bar{y}^{j,k} \in B^j_{00}(\bar{x})$, and nontrivial multipliers $\kappa \geq 0$, $\mu_{j,k} \geq 0$, $1 \leq k \leq p_j$, $j \in J_0(\bar{x})$, such that $\sum_{j \in J_0(\bar{x})} p_j \leq n+1$ and*

$$\kappa \nabla f(\bar{x}) \ + \sum_{j \in J_0(\bar{x})} \sum_{k=1}^{p_j} \mu_{j,k} \, \nabla_x w(\bar{x}, \bar{y}^{j,k}) = \ 0 \ . \tag{3}$$

The proof of Theorem 2 needs some preparation. Recall that the outer tangent cone (contingent cone) $\Gamma^\star(\bar{x}, M)$ to a set $M \subset \mathbb{R}^n$ at $\bar{x} \in \mathbb{R}^n$ is defined by $\bar{d} \in \Gamma^\star(\bar{x}, M)$ if and only if there exist sequences $(t^\nu)_{\nu \in \mathbb{N}}$ and $(d^\nu)_{\nu \in \mathbb{N}}$ such that

$$t^\nu \searrow 0, \ d^\nu \ \to \ \bar{d} \quad \text{and} \quad \bar{x} + t^\nu d^\nu \in M \ \text{ for all } \ \nu \in \mathbb{N} \, .$$

Moreover, we define the inner tangent cone $\Gamma(\bar{x}, M)$ to M at $\bar{x} \in \mathbb{R}^n$ as: $\bar{d} \in \Gamma(\bar{x}, M)$ if and only if there exist some $\bar{t} > 0$ and a neighborhood D of $\bar{d}$ such that

$$\bar{x} + t \, d \in M \ \text{ for all } \ t \in (0, \bar{t}), \ d \in D \ .$$

It is well-known ([17]) that $\Gamma(\bar{x}, M) \subset \Gamma^\star(\bar{x}, M)$ and that $\Gamma(\bar{x}, M)^c = \Gamma^\star(\bar{x}, M^c)$, where A^c denotes the set complement of a set $A \subset \mathbb{R}^n$. Furthermore, the following primal first order necessary optimality condition holds.

Lemma 2 ([17]). *Let $\bar{x}$ be a local maximizer of f over M. Then there exists no contingent direction of first order ascent in $\bar{x}$:*

$$\{ \ d \in I\!R^n | \ \langle \nabla f(\bar{x}), d \rangle > 0 \ \} \ \cap \ \Gamma^\star(\bar{x}, M) \ = \ \emptyset \ .$$

Lemma 3. *For $\bar{x} \in M$ each solution $d^0 \in I\!R^n$ of the system*

$$\langle \, \nabla_x w(\bar{x}, y), d \, \rangle \ > \ 0 \quad \textit{for all } y \in B^j_{00}(\bar{x}), \ j \in J_0(\bar{x}) \tag{4}$$

is an element of $\Gamma(\bar{x}, M)$.

Proof. Let d^0 be a solution of (4) and assume that $d^0 \in \Gamma(\bar{x}, M)^c$. Then we have $d^0 \in \Gamma^\star(\bar{x}, M^c)$, so that there exist sequences $(t^\nu)_{\nu \in \mathbb{N}}$ and $(d^\nu)_{\nu \in \mathbb{N}}$ such that $t^\nu \searrow 0$, $d^\nu \to d^0$ and $x^\nu := \bar{x} + t^\nu d^\nu \in M^c$ for all $\nu \in \mathbb{N}$.

Exactly like in the proof of Lemma 1 we can now construct some $j_0 \in J_0(\bar{x})$ and a sequence $y^\nu \in B(x^\nu)$ with $y^\nu \to \bar{y} \in B^{j_0}_{00}(\bar{x})$. For all $\nu \in \mathbb{N}$ the mean value theorem guarantees the existence of some $\theta^\nu \in [0, 1]$ with

$$0 \ \geq \ w(\bar{x} + t^\nu d^\nu, y^\nu) \ = \ w(\bar{x}, y^\nu) + t^\nu \langle \, \nabla_x w(\bar{x} + \theta^\nu t^\nu d^\nu, y^\nu), d^\nu \, \rangle \ .$$

From (1) and $t^\nu > 0$ we conclude $0 > \langle \, \nabla_x w(\bar{x} + \theta^\nu t^\nu d^\nu, y^\nu), d^\nu \, \rangle$ for all $\nu \in \mathbb{N}$ which implies $0 \geq \langle \, \nabla_x w(\bar{x}, \bar{y}), d^0 \, \rangle$. Hence we have constructed some $j_0 \in J_0(\bar{x})$ and $\bar{y} \in B^{j_0}_{00}(\bar{x})$ with $\langle \, \nabla_x w(\bar{x}, \bar{y}), d^0 \, \rangle \leq 0$, in contradiction to the assumption. □

A combination of Lemma 2, the inclusion $\Gamma(\bar{x},M) \subset \Gamma^{\star}(\bar{x},M)$, and Lemma 3 yields that at a local maximizer $\bar{x}$ of DC the system

$$\langle \nabla f(\bar{x}), d \rangle > 0, \quad \langle \nabla_x w(\bar{x},y), d \rangle \ > \ 0 \quad \text{for all } y \in B^j_{00}(\bar{x}),\ j \in J_0(\bar{x})$$

is not soluble in d. By a theorem of the alternative this result is equivalent to the assertion of Theorem 2. In the following $\text{conv}(S)$ denotes the convex hull of a set $S \subset \mathbb{R}^n$, i.e. the set of all finite convex combinations of elements from S.

Lemma 4 (Lemma of Gordan, [1, 10]). *Let $S \subset \mathbb{R}^n$ be nonempty and compact. Then the inequality system*

$$s^\top d \ > \ 0 \ \ \textit{for all} \ \ s \in S$$

is inconsistent for $d \in \mathbb{R}^n$ if and only if $0 \in \text{conv}(S)$.

Recall that in the case $0 \in \text{conv}(S)$ it is possible to express the origin as the convex combination of at most $n+1$ elements from S, due to Carathéodory's theorem.

Since the set $\bigcup_{j \in J_0(\bar{x})} B^j_{00}(\bar{x})$ is compact as the finite union of closed subsets of the compact set $B(\bar{x})$, Lemma 4 implies Theorem 2. Note that if the latter union of sets was empty, we would simply obtain the condition $\nabla f(\bar{x}) = 0$ from unconstrained optimization. However, in view of Lemma 1 under the assumption $\bar{x} \in M \cap \partial M$ of Theorem 2 this is not possible.

3 Genericity of the Reduction Ansatz

Multi-jet transversality. In the following we give a short introduction to transversality theory, as far as we need it for our analysis. For details, see [11, 15]. Two smooth manifolds V, W in $\mathbb{R}^N$ are said to intersect transversally (notation: $V \pitchfork W$) if at each intersection point $u \in V \cap W$ the tangent spaces $T_u V,\ T_u W$ together span the embedding space:

$$T_u V \ + \ T_u W \ = \ \mathbb{R}^N \ . \tag{5}$$

The number $N - \dim V$ is called the codimension of V in $\mathbb{R}^N$, shortly $\text{codim}\, V$, and we have

$$\text{codim}\, V \ \leq \ \dim W \tag{6}$$

whenever $V \pitchfork W$ and $V \cap W \neq \emptyset$. For our purpose, the manifold W is induced by the 1-jet extension of a function $F \in C^\infty(\mathbb{R}^N, \mathbb{R}^M)$, i.e. by the mapping

$$j^1 F : \ \mathbb{R}^N \longrightarrow J(N,M,1),\ z \longmapsto (z, F(z), F_z(z))$$

where $J(N,M,1) \ = \ \mathbb{R}^{N+M+N \cdot M}$ and the partial derivatives are listed according to some order convention ([15]). Choosing W as the graph of $j^1 F$

(notation: $W = j^1F(\mathbb{R}^N)$) it is easily shown that W is a smooth manifold of dimension N in $J(N, M, 1)$. Given another smooth manifold V in $J(N, M, 1)$, we define the set

$$\pitchfork^1 V \;=\; \{F \in C^\infty(\mathbb{R}^N, \mathbb{R}^M) |\; j^1F(\mathbb{R}^N) \pitchfork V\} \;.$$

Our analysis bases on the following theorem which is originally due to R. Thom. For proofs see [11, 15].

Theorem 3 (Jet transversality). *With respect to the C^∞_s-topology, the set $\pitchfork^1 V$ is generic in $C^\infty(\mathbb{R}^N, \mathbb{R}^M)$.*

Here, C^∞_s denotes the Whitney topology ([11, 15]). In particular, $\pitchfork^1 V$ is C^∞_s-dense in $C^\infty(\mathbb{R}^N, \mathbb{R}^M)$ and hence, C^d_s-dense in $C^d(\mathbb{R}^N, \mathbb{R}^M)$ for any $d \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ ([11]).

Since jet transversality gives information about certain properties of the functions under investigation only at every *single* point we apply the concept of multi-jet transversality instead ([15]). Thereby we are able to study properties that have to be satisfied at all global maximizers of the lower level problem at the same time. Let D be a positive integer and define

$$\mathbb{R}^N_D \;=\; \left\{(z^1, \ldots, z^D) \in \textstyle\prod_{k=1}^D \mathbb{R}^N |\; z^i \neq z^j \text{ for } 1 \leq i < j \leq D\right\}$$

as well as the multi-jet space

$$\begin{aligned} &J_D(N, M, 1) \;= \\ &\quad \left\{(z^1, u^1, \ldots, z^D, u^D) \in \textstyle\prod_{k=1}^D J(N, M, 1) |\; (z^1, \ldots, z^D) \in \mathbb{R}^N_D\right\} \;. \end{aligned}$$

The multi-jet extension $j^1_D F : \; \mathbb{R}^N_D \longrightarrow J_D(N, M, 1)$ is the mapping

$$j^1_D F : \; (z^1, \ldots, z^D) \;\longmapsto\; \left(j^1F(z^1), \ldots, j^1F(z^D)\right) \;,$$

and for a smooth manifold V in $J_D(N, M, 1)$ we define the set

$$\pitchfork^1_D V \;=\; \{F \in C^\infty(\mathbb{R}^N, \mathbb{R}^M) |\; j^1_D F(\mathbb{R}^N_D) \pitchfork V\} \;.$$

Theorem 4 (Multi-jet transversality). *With respect to the C^∞_s-topology, the set $\pitchfork^1_D V$ is generic in $C^\infty(\mathbb{R}^N, \mathbb{R}^M)$.*

Rank conditions. For $M, N \in \mathbb{N}$ and $R \leq \min(M, N)$ let us define the set of matrices of rank R,

$$\mathbb{R}^{M \times N}_R \;=\; \left\{A \in \mathbb{R}^{M \times N} \,\middle|\, \operatorname{rank}(A) = R\right\} \;.$$

Moreover, for $M, N \in \mathbb{N}$, $R \leq \min(M, N)$, $\mathcal{I} \subset \{1, ..., M\}$ and

$$\max(R + |\mathcal{I}| - M, 0) \;\leq\; S \;\leq\; \min(R, |\mathcal{I}|)$$

we let

$$\mathbb{R}^{M\times N}_{R,\mathcal{I},S} \;=\; \left\{ A \in \mathbb{R}^{M\times N}_{R} \,\middle|\, A^{(\mathcal{I})} \in \mathbb{R}^{(M-|\mathcal{I}|)\times N}_{R-S} \right\} \,,$$

where the matrix $A^{(\mathcal{I})}$ results from A by deletion of the rows with indices in $\mathcal{I}$. Observe that the above restrictions on S follow from the trivial relations $0 \le R - S \le M - |\mathcal{I}|$ and $R - |\mathcal{I}| \le R - S \le R$.

These definitions are intimately related to the Reduction Ansatz in the lower level problem. In fact, for $\bar{x} \in M$ and some $j \in J_0(\bar{x})$ let $\bar{y}$ be a maximizer of $Q^j(\bar{x})$. From the first order necessary optimality condition of Fritz John we know that then the gradient $\nabla c_j(\bar{y})$ and the gradients of the active inequality constraints are linearly dependent. To identify these constraints conveniently we put $L = \{1, ..., s\}$ with $s \in \mathbb{N}$, $v_{s+1}(x, y) := w(x, y)$, $\Lambda = L \cup \{s + 1\}$, $\Lambda_0(\bar{x}, \bar{y}) = \{\ell \in \Lambda |\; v_\ell(\bar{x}, \bar{y}) = 0\}$, and $s_0 = |\Lambda_0(\bar{x}, \bar{y})|$. Let $D_y v_{\Lambda_0}(\bar{x}, \bar{y})$ denote the matrix with rows $D_y v_\ell(\bar{x}, \bar{y}) := \nabla^\top_y v_\ell(\bar{x}, \bar{y})$, $\ell \in \Lambda_0(\bar{x}, \bar{y})$. We obtain

$$\begin{pmatrix} D_y c_j(\bar{x}, \bar{y}) \\ D_y v_{\Lambda_0}(\bar{x}, \bar{y}) \end{pmatrix} \;\in\; \mathbb{R}^{(1+s_0)\times m}_{\rho_j}$$

with $\rho_j \le s_0$. With this notation, LICQ is equivalent to

$$\begin{pmatrix} D_y c_j(\bar{x}, \bar{y}) \\ D_y v_{\Lambda_0}(\bar{x}, \bar{y}) \end{pmatrix} \;\in\; \mathbb{R}^{(1+s_0)\times m}_{s_0\,,\{0\},0} \;,$$

if we identify the first row of the matrix with the index $\ell = 0$. Moreover, SCS implies

$$\begin{pmatrix} D_y c_j(\bar{x}, \bar{y}) \\ D_y v_{\Lambda_0}(\bar{x}, \bar{y}) \end{pmatrix} \;\in\; \mathbb{R}^{(1+s_0)\times m}_{s_0\,,\{\ell\},0} \;,$$

for all $\ell \in \Lambda_0(\bar{x}, \bar{y})$.

For a matrix $A \in \mathbb{R}^{M\times N}$ with rows $A^1, ..., A^M$ we define the function

$$\mathrm{vec}:\; \mathbb{R}^{M\times N} \longrightarrow \mathbb{R}^{M\cdot N},\; A \longmapsto (A^1, ..., A^M)\;.$$

Lemma 5 ([15, 20]).

(i) The set vec $\left(\mathbb{R}^{M\times N}_R\right)$ *is a smooth manifold of codimension* $(M - R)\cdot(N - R)$ *in* $\mathbb{R}^{M\cdot N}$.

(ii) The set vec $\left(\mathbb{R}^{M\times N}_{R,\mathcal{I},S}\right)$ *is a smooth manifold of codimension* $(M - R)\cdot(N - R) + S\cdot(M - R + S - |\mathcal{I}|)$ *in* $\mathbb{R}^{M\cdot N}$.

A codimension formula. Let $J = \{1, ..., p\}$ as well as $p_0 = |J_0(\bar{x})|$. By Lemma 1, for $\bar{x} \in M \cap \partial M$ the set $\bigcup_{j\in J_0(\bar{x})} B^j_{00}(\bar{x})$ is nonempty. We consider the case in which it contains at least r different elements, say $\bar{y}^{j,k} \in B^j_{00}(\bar{x})$, $1 \le k \le p_j$, $j \in J_0(\bar{x})$, with $\sum_{j=1}^{p_0} p_j = r$.

As $\bar{y}^{j,k}$ is a maximizer of $Q^j(\bar{x})$ we find a unique number $\rho_{j,k} \le s^{j,k}_0 := |\Lambda_0(\bar{x}, \bar{y}^{j,k})|$ such that

$$\begin{pmatrix} D_y c_j(\bar{x}, \bar{y}^{j,k}) \\ D_y v_{\Lambda_0}(\bar{x}, \bar{y}^{j,k}) \end{pmatrix} \in \mathbb{R}^{(1+s_0^{j,k})\times m}_{\rho_{j,k}} ,$$

and we define the rank defect $d_{j,k} = s_0^{j,k} - \rho_{j,k}$. Moreover, we have

$$\begin{pmatrix} D_y c_j(\bar{x}, \bar{y}^{j,k}) \\ D_y v_{\Lambda_0}(\bar{x}, \bar{y}^{j,k}) \end{pmatrix} \in \mathbb{R}^{(1+s_0^{j,k})\times m}_{\rho_{j,k}\, ,\, D_{j,k}\, ,\, \sigma_{j,k}}$$

for several choices of $D_{j,k}$ and $\sigma_{j,k}$, where we can always choose $D_{j,k} = \emptyset$ and $\sigma_{j,k} = 0$.

Furthermore, if $\bar{x}$ is a local maximizer of DC, Theorem 2 guarantees that for some choice $\bar{y}^{j,k} \in B_{00}^j(\bar{x})$, $1 \le k \le p_j$, $j \in J_0(\bar{x})$ with $\sum_{j=1}^{p_0} p_j = r \le n+1$ we also have

$$\begin{pmatrix} Df(\bar{x}) \\ D_x w(\bar{x}, \bar{y}^{j,k})_{1\le k\le p_j\, ,\, 1\le j\le p_0} \end{pmatrix} \in \mathbb{R}^{(1+r)\times n}_{\rho_0\, ,\, D_0\, ,\, \sigma_0}$$

with $\rho_0 \le r$. We denote the corresponding rank defect by $d_0 = r - \rho_0$. Our subsequent analysis bases on the following relation:

$$\begin{aligned} 0 \ \ge\ & d_0 + d_0(n - r + d_0) + \sigma_0(1 + d_0 + \sigma_0 - |D_0|) \\ +\sum_{j=1}^{p_0}\sum_{k=1}^{p_j} & \Big[\ d_{j,k} + d_{j,k}(m - s_0^{j,k} + d_{j,k}) + \sigma_{j,k}(1 + d_{j,k} + \sigma_{j,k} - |D_{j,k}|)\ \Big] \ . \end{aligned} \tag{7}$$

Put $\mathbb{Q}^d = C^d(\mathbb{R}^n, \mathbb{R}) \times \mathbb{C}^d(c) \times \mathbb{C}^d(v) \times C^d(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R})$, where $\mathbb{C}^d(c)$ and $\mathbb{C}^d(v)$ are defined to be the set of vector functions $c \in C^d(\mathbb{R}^m, \mathbb{R}^p)$ and $v \in C^d(\mathbb{R}^m, \mathbb{R}^s)$ such that C and Y are nonempty and compact, respectively. Define

$$\begin{aligned} \mathcal{F}^d = \ \{\ & (f, c, v, w) \in \mathbb{Q}^d |\ \text{any choice of } r \text{ elements} \\ & \text{from } \textstyle\bigcup_{j\in J_0(\bar{x})} B_{00}^j(\bar{x}) \text{ corresponding to a point} \\ & \bar{x} \in M \cap \partial M \text{ satisfies relation (7) } \}\ . \end{aligned}$$

Theorem 5. *$\mathcal{F}^\infty$ is C_s^∞-dense in $\mathbb{Q}^\infty$.*

Proof. For $r \in \mathbb{N}$ and $K := \{1, ..., r\}$ consider the reduced multi-jet

$$\begin{aligned} j_r^1(f, c, v, w)(x^1, y^1, ..., x^r, y^r) \ = \ & (x^k, y^k, Df^k, c_1^k, ..., c_p^k, Dc_1^k, ..., Dc_p^k, \\ & v_1^k, ..., v_s^k, Dv_1^k, ..., Dv_s^k, w^k, D_x w^k, D_y w^k,\ k \in K\ ) \end{aligned}$$

with $(x^1, y^1, ..., x^r, y^r) \in \mathbb{R}_r^{n+m}$ and $Df^k = Df(x^k)$, etc. In the following we call $K_j\ ,\ j \in \tilde{J}_0$, a partition of K if $\bigcup_{j\in\tilde{J}_0} K_j = K$ and if the sets $K_j\ ,\ j \in \tilde{J}_0$, are pairwise distinct. For

$$\left.\begin{array}{l} r \in \mathbb{N} \\ \tilde{J}_0 \subset J \\ K_j\,,\ j \in \tilde{J}_0, \quad \text{a partition of} \ \ K = \{1, ..., r\} \\ 0 \le \rho_0 \le \min(1+r, n) \\ D_0 \subset \{0, ..., r\} \\ \max(\rho_0 + |D_0| - 1 - r, 0) \le \sigma_0 \le \min(\rho_0, |D_0|) \\ \tilde{\Lambda}_0^{j,k} \subset \Lambda \\ 0 \le \rho_{j,k} \le \min(1 + s_0^{j,k}, m) \\ D_{j,k} \subset \{0, ..., s_0^{j,k}\} \\ \max(\rho_{j,k} + |D_{j,k}| - 1 - s_0^{j,k}, 0) \le \sigma_{j,k} \le \min(\rho_{j,k}, |D_{j,k}|) \\ k \in K_j,\ j \in \tilde{J}_0 \end{array}\right\} \qquad (8)$$

we define the C^∞-manifold $\mathcal{N}_{r,(K_j,j\in\tilde{J}_0),\rho_0,D_0,\sigma_0,\ (\tilde{\Lambda}_0^{j,k},\rho_{j,k},D_{j,k},\sigma_{j,k},\ k\in K_j,j\in\tilde{J}_0)}$ to be the set of points

$$(\tilde{x}^k, \tilde{y}^k, \tilde{F}^k, \tilde{c}_1^k, ..., \tilde{c}_p^k, \tilde{C}_1^k, ..., \tilde{C}_p^k, \tilde{v}_1^k, ..., \tilde{v}_s^k, \tilde{V}_1^k, ..., \tilde{V}_s^k, \tilde{w}^k, \tilde{X}^k, \tilde{Y}^k,\ k \in K\ )$$

with the following properties:

- dimensions:

$$\begin{aligned} (\tilde{x}^1, \tilde{y}^1, ..., \tilde{x}^r, \tilde{y}^r) &\in \mathbb{R}_r^{n+m}, \\ \tilde{c}_j^k,\ j \in J,\ \tilde{v}_\ell^k,\ \ell \in L,\ \tilde{w}^k &\in \mathbb{R},\ k \in K \\ \tilde{F}^k,\ \tilde{X}^k &\in \mathbb{R}^n,\ k \in K \\ \tilde{C}_j^k,\ j \in J,\ \tilde{V}_\ell^k,\ \ell \in L,\ \tilde{Y}^k &\in \mathbb{R}^m,\ k \in K \end{aligned}$$

- conditions on the independent variables:

$$\tilde{x}^1 \ = \ ... \ = \ \tilde{x}^r$$

- conditions on the functional values:

$$\tilde{c}_j^k = 0,\ k \in K_j\,,\ j \in \tilde{J}_0\,, \quad \tilde{v}_\ell^k = 0, \quad \ell \in \tilde{\Lambda}_0^{j,k},\ k \in K_j\,,\ j \in \tilde{J}_0$$

- conditions on the gradients:

$$\begin{pmatrix} \tilde{F}^1 \\ (\tilde{X}^k)_{k\in K_j\,,\,j\in\tilde{J}_0} \end{pmatrix} \in \mathbb{R}^{(1+r)\times n}_{\rho_0\,,\,D_0\,,\,\sigma_0}\ ,$$

$$\begin{pmatrix} \tilde{C}_j^k \\ \tilde{V}^k_{\tilde{\Lambda}_0^{j,k}} \end{pmatrix} \in \mathbb{R}^{(1+s_0^{j,k})\times m}_{\rho_{j,k}\,,\,D_{j,k}\,,\,\sigma_{j,k}}\ , \quad k \in K_j\,,\ j \in \tilde{J}_0\ .$$

With the help of Lemma 5(ii) we can calculate the codimension of this manifold:

$$
\begin{aligned}
&\operatorname{codim} \mathcal{N}_{r,(K_j, j\in \tilde{J}_0), \rho_0, D_0, \sigma_0,\ (\tilde{\Lambda}_0^{j,k}, \rho_{j,k}, D_{j,k}, \sigma_{j,k},\ k\in K_j, j\in \tilde{J}_0)} = \\
&= (r-1)n + r + \sum_{j\in \tilde{J}_0} \sum_{k\in K_j} s_0^{j,k} \\
&\quad + (1+r-\rho_0)(n-\rho_0) + \sigma_0(1+r-\rho_0+\sigma_0-|D_0|) \\
&\quad + \sum_{j\in \tilde{J}_0} \sum_{k\in K_j} \Big[(1+s_0^{j,k}-\rho_{j,k})(m-\rho_{j,k}) \\
&\quad + \sigma_{j,k}(1+s_0^{j,k}-\rho_{j,k}+\sigma_{j,k}-|D_{j,k}|) \Big] \ .
\end{aligned} \tag{9}
$$

Define the set

$$
\mathcal{F}^\star = \bigcap_{r=1}^{\infty} \bigcap_{(K_j \cdots \tilde{J}_0)} \pitchfork_r^1 \mathcal{N}_{r,(K_j, j\in \tilde{J}_0), \rho_0, D_0, \sigma_0,\ (\tilde{\Lambda}_0^{j,k}, \rho_{j,k}, D_{j,k}, \sigma_{j,k},\ k\in K_j, j\in \tilde{J}_0)}
$$

where the inner intersection ranges over all possible choices of K_1, etc., according to (8). $\mathcal{F}^\star$ is C_s^∞-dense in $\mathbb{Q}^\infty$ by Theorem 4. It remains to be shown that $\mathcal{F}^\star \subset \mathcal{F}^\infty$. Choose a function vector $(f,c,v,w) \in \mathcal{F}^\star$ as well as a local maximizer $\bar{x}$ of DC. By Lemma 1 the set $\bigcup_{j\in J_0(\bar{x})} B_{00}^j(\bar{x})$ is non-empty. From each nonempty $B_{00}^j(\bar{x})$ choose some (pairwise distinct) $\bar{y}^{j,k}$, $k \in K_j$, and put $K_j = \emptyset$ if $B_{00}^j(\bar{x}) = \emptyset$. Denote the total number of chosen elements by r and put $K = \{1,...,r\}$. Then $K_j\,, j \in J_0(\bar{x})$, forms a partition of K, $(\bar{x},\bar{y}^1,...,\bar{x},\bar{y}^r) \in \mathbb{R}_r^{n+m}$, and $j_r^1(f,c,v,w)(\bar{x},\bar{y}^1,...,\bar{x},\bar{y}^r)$ is contained in some set $\mathcal{N}_{r,(\cdots\, \tilde{J}_0)}$. As the intersection of $j_r^1(f,c,v,w)(\mathbb{R}_r^{n+m})$ with $\mathcal{N}_{r,(\cdots\, \tilde{J}_0)}$ is transverse, (6) yields $r\,(n+m) \ge \operatorname{codim} \mathcal{N}_{r,(\cdots\, \tilde{J}_0)}$. Inserting (9) now yields (7) after a short calculation. □

Note that the statement of Theorem 5 is equivalent to saying that $\mathcal{F}^\infty$ is C_s^d-dense in $\mathbb{Q}^\infty$ for each $d \in \mathbb{N}_0$. Since the set $C^\infty(\mathbb{R}^N,\mathbb{R})$ is also C_s^d-dense in $C^d(\mathbb{R}^N,\mathbb{R})$ ([11]), it is no restriction to consider the space of smooth defining functions $\mathbb{Q}^\infty$ instead of the space $\mathbb{Q}^d$, $d \ge 2$.

Corollary 2. *For $(f,c,v,w) \in \mathcal{F}^\star$ let $\bar{x} \in M \cap \partial M$ be a local maximizer of DC. Then the set $\bigcup_{j\in J_0(\bar{x})} B_{00}^j(\bar{x})$ contains at most n elements $\bar{y}^1,...,\bar{y}^r$, and for each $1 \le k \le r$ LICQ and SCS hold at $\bar{y}^k$ in the corresponding lower level problem.*

Proof. One can easily conclude from the relations in (8) that each factor in the right hand side of (7) is nonnegative. Consequently, all summands have to vanish. In particular we find $d_0 = d_{j,k} = 0$ for all $1 \le k \le p_j\,,\ j \in J_0(\bar{x})$. This implies $0 \le n - \rho_0 = n - r + d_0 = n - r$ which is the first part of the assertion.

A second consequence is $\sigma_{j,k}(1+\sigma_{j,k}-|D_{j,k}|)=0$ for all $1\leq k\leq p_j$, $j\in J_0(\bar{x})$. Hence, $|D_{j,k}|=1$ implies $\sigma_{j,k}=0$. This means that LICQ and SCS hold at each $\bar{y}^{j,k}$ in $Q^j(\bar{x})$. □

With a tedious evaluation of the tangent space condition (5) it is also possible to show that for $(f,c,v,w)\in\mathcal{F}^\star$ and a local maximizer $\bar{x}\in M\cap\partial M$ of DC at each $\bar{y}\in\bigcup_{j\in J_0(\bar{x})}B^j_{00}(\bar{x})$ the second order sufficiency condition holds. Altogether this means that for $(f,c,v,w)\in\mathcal{F}^\star$ the Reduction Ansatz is valid at each local maximizer of DC.

4 An associated standard semi-infinite problem

The first order necessary optimality condition in Theorem 2 has the typical structure of an optimality condition for some *standard* semi-infinite program. In fact, we can construct a certain standard semi-infinite problem which is strongly related to DC.

For the following arguments we put $C_j^{\leq}=\{y\in\mathbb{R}^m|\ c_j(y)\leq 0\}$, $C_j^{<}=\{y\in\mathbb{R}^m|\ c_j(y)<0\}$, etc. for $j\in J$ as well as $W^{\leq}(x)=\{y\in\mathbb{R}^m|\ w(x,y)\leq 0\}$ etc. The main idea is to rewrite the inclusion constraint $B(x)\subset C$ of DC in an equivalent form like $C^c\subset B(x)^c$.

Slightly modified this idea proceeds as follows. By definition we have $B(x)\subset C$ if and only $Y\cap W^{\leq}(x)\subset\bigcap_{j\in J}C_j^{\leq}$. The latter is equivalent to $Y\cap W^{\leq}(x)\cap\bigcup_{j\in J}C_j^{>}=\emptyset$ and, thus, to $\bigcup_{j\in J}\left(Y\cap C_j^{>}\right)\subset W^{>}(x)$.

This means that an equivalent formulation of the constraint $B(x)\subset C$ is given by

$$w(x,y)\ >\ 0\quad\text{for all}\ \ y\in Y\cap C_j^{>}\,,\ j\in J\,.$$

Due to the strict inequalities these are not semi-infinite constraints in the usual sense. We can, however, formulate an *associated standard semi-infinite problem* for DC:

$$SIP_{DC}:\quad\max_x f(x)\quad\text{subject to}\quad w(x,y)\ \geq\ 0\quad\forall\,y\in Y\cap C_j^{\geq}\,,\ j\in J\,.$$

Note that the index sets $Y\cap C_j^{\geq}$, $j\in J$, of the finitely many semi-infinite constraints are compact, and certainly nonempty if $C\subset Y$. Recall that we defined the optimal value functions

$$\varphi_j(x)\ =\ \max_{y\in Y\cap W^{\leq}(x)}\ c_j(y)\,,\ j\in J\,,$$

and the active index set $J_0(x)=\{j\in J|\ \varphi_j(x)=0\}$ for the problem $GSIP_{DC}$. For the problem SIP_{DC} we put analogously

$$\psi_j(x)\ =\ \min_{y\in Y\cap C_j^{\geq}}\ w(x,y)\,,\ j\in J\,,$$

$J_0^{SIP}(x) = \{j \in J|\ \psi_j(x) = 0\}$, and $Q^j_{SIP}(x)$, $j \in J$, for the corresponding lower level problems. For $j \in J_0^{SIP}(x)$ the optimal points of $Q^j_{SIP}(x)$ form the set $\{y \in Y \cap C_j^{\geq}|\ w(x,y) = 0\} = Y \cap C_j^{\geq} \cap W^{=}(x)$. Fritz John's first order optimality condition for standard semi-infinite problems thus yields the following result.

Proposition 4.1 *Let $\bar{x} \in \partial M_{SIP}$ be a local maximizer of SIP_{DC}. Then there exist $p_j \in \mathbb{N}$, $\bar{y}^{j,k} \in Y \cap C_j^{\geq} \cap W^{=}(x)$, and nontrivial multipliers $\kappa \geq 0$, $\mu_{j,k} \geq 0$, $1 \leq k \leq p_j$, $j \in J_0^{SIP}(\bar{x})$, such that $\sum_{j \in J_0^{SIP}(\bar{x})} p_j \leq n+1$ and*

$$\kappa \nabla f(\bar{x}) \ + \sum_{j \in J_0^{SIP}(\bar{x})} \sum_{k=1}^{p_j} \mu_{j,k} \nabla_x w(\bar{x}, \bar{y}^{j,k}) = \ 0\ .$$

The resemblance of this result with Theorem 2 is obvious. We emphasize that we relaxed strict to nonstrict inequalities while deriving the problem SIP_{DC} from DC, so that an identical result for both problems cannot be expected. More precisely, the feasible sets

$$M \ = \ \{\, x \in \mathbb{R}^n|\ \varphi_j(x) \leq 0\,,\ j \in J\,\} \ = \ \bigcap_{j \in J} \Phi_j^{\leq}$$

and

$$M_{SIP} \ = \ \{\, x \in \mathbb{R}^n|\ \psi_j(x) \geq 0\,,\ j \in J\,\} \ = \ \bigcap_{j \in J} \Psi_j^{\geq}$$

do not necessarily coincide. Their relation is clarified by the next results.

Lemma 6.
(i) For all $j \in J$ we have $\Phi_j^{<} = \Psi_j^{>}$.
(ii) For all $j \in J$ and $x \in \Phi_j^{=}$ we have $x \in \Psi_j^{=}$ if and only if $w(x,\cdot)$ is active in all global solutions of $Q^j(x)$.
(iii) For all $j \in J$ and $x \in \Psi_j^{=}$ we have $x \in \Phi_j^{=}$ if and only if c_j is active in all global solutions of $Q^j_{SIP}(x)$.

Proof. For all $j \in J$ we have $x \in \Phi_j^{<}$ if and only if $Y \cap W^{\leq}(x) \subset C_j^{<}$, and we have $x \in \Psi_j^{>}$ if and only if $Y \cap C_j^{\geq} \subset W^{>}(x)$. Since both characterizations are equivalent to $Y \cap C_j^{\geq} \cap W^{\leq}(x) = \emptyset$, the assertion of part (i) follows.

From part (i) it is clear that for each $j \in J$ the set $\Phi_j^{=}$ is necessarily contained in $\Psi_j^{\leq}$. We have $x \in \Psi_j^{<}$ if and only if $Y \cap C_j^{\geq} \cap W^{<}(x) \neq \emptyset$. On the other hand, for $x \in \Phi_j^{=}$ the set $Y \cap C_j^{\geq} \cap W^{\leq}(x)$ is the set of global solutions of $Q^j(x)$. This shows the assertion of part (ii). The proof of part (iii) is analogous. □

Theorem 6.
(i) Let $x \in M$ and for each $j \in J_0(x)$ let $w(x,\cdot)$ be active in all global solutions of $Q^j(x)$. Then we have $x \in M_{SIP}$.

(ii) *Let $x \in M_{SIP}$ and for each $j \in J_0^{SIP}(x)$ let c_j be active in all global solutions of $Q^j_{SIP}(x)$. Then we have $x \in M$.*

Proof. Lemma 6. □

Note that under the assumption of Theorem 6(ii) the global solution set $Y \cap C_j^{\geq} \cap W^{=}(x)$ can be replaced by $Y \cap C_j^{=} \cap W^{=}(x) = B^j_{00}(x)$, so that the difference between Theorem 2 and Proposition 4.1 disappears.

5 Final remarks

A main technical assumption for the genericity proof in Section 3 is that only one of the smooth constraints in the description of $B(x)$ actually depends on x. There are, of course, design centering problems which cannot be formulated this way. These problems appear to be as difficult as the general semi-infinite optimization problem without any additional structure, so that genericity results for this case can be expected as soon as the generic validity of the Reduction Ansatz at all solutions of *GSIP* has been shown.

Under the Reduction Ansatz, locally around a local solution $\bar{x}$ the problem $GSIP_{DC}$ can be rewritten as a smooth problem with finitely many constraints. We point out that our genericity proof from Section 3 also shows that for $(f, c, v, w) \in \mathcal{F}^\star$ a local maximizer $\bar{x} \in M \cap \partial M$ of *DC* is nondegenerate for this locally reduced problem.

The results of the present article for single-body design centering problems can be transferred to the multi-body case with some additional technical effort. This and efficient numerical methods for multi-body design centering will be subject of future research.

References

1. E.W. Cheney, *Introduction to Approximation Theory,* McGraw-Hill, New York, 1966.
2. V. Chvatal, *Linear Programming,* Freeman, New York, 1983.
3. A.V. Fiacco, G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques,* Wiley, New York, 1968.
4. T.J. Graettinger, B.H. Krogh, *The acceleration radius: a global performance measure for robotic manipulators,* IEEE Journal of Robotics and Automation, Vol. 4 (1988), 60-69.
5. P. Gritzmann, V. Klee, *On the complexity of some basic problems in computational convexity. I. Containment problems,* Discrete Mathematics, Vol. 136 (1994), 129-174.
6. R. Hettich, H.Th. Jongen, *Semi-infinite programming: conditions of optimality and applications,* in: J. Stoer (ed): Optimization Techniques, Part 2, Lecture Notes in Control and Information Sciences, Vol. 7, Springer, Berlin, 1978, 1-11.

7. R. Hettich, K.O. Kortanek, *Semi-infinite programming: theory, methods, and applications,* SIAM Review, Vol. 35 (1993), 380-429.
8. R. Hettich, G. Still, *Semi-infinite programming models in robotics,* in: J. Guddat, H.Th. Jongen, B. Kummer, F. Nožička (eds): Parametric Optimization and Related Topics II, Akademie Verlag, Berlin, 1991, 112-118.
9. R. Hettich, G. Still, *Second order optimality conditions for generalized semi-infinite programming problems,* Optimization, Vol. 34 (1995), 195-211.
10. R. Hettich, P. Zencke, *Numerische Methoden der Approximation und semi-infiniten Optimierung,* Teubner, Stuttgart, 1982.
11. M.W. Hirsch, *Differential Topology,* Springer, New York, 1976.
12. R. Horst, H. Tuy, *Global Optimization,* Springer, Berlin, 1990.
13. F. John, *Extremum problems with inequalities as subsidiary conditions,* in: Studies and Essays, R. Courant Anniversary Volume, Interscience, New York, 1948, 187-204.
14. H.Th. Jongen, P. Jonker, F. Twilt, *Critical sets in parametric optimization,* Mathematical Programming, Vol. 34 (1986), 333-353.
15. H.Th. Jongen, P. Jonker, F. Twilt, *Nonlinear Optimization in Finite Dimensions,* Kluwer, Dordrecht, 2000.
16. H.Th. Jongen, J.-J. Rückmann, O. Stein, *Generalized semi-infinite optimization: a first order optimality condition and examples,* Mathematical Programming, Vol. 83 (1998), 145-158.
17. P.-J. Laurent, *Approximation et Optimisation,* Hermann, Paris, 1972.
18. V.H. Nguyen, J.J. Strodiot, *Computing a global optimal solution to a design centering problem,* Mathematical Programming, Vol. 53 (1992), 111-123.
19. O. Stein, *On parametric semi-infinite optimization,* Thesis, Shaker, Aachen, 1997.
20. O. Stein, *On level sets of marginal functions,* Optimization, Vol. 48 (2000), 43-67.
21. O. Stein, *Bi-level Strategies in Semi-infinite Programming,* Kluwer, Boston, 2003.
22. O. Stein, G. Still, *On generalized semi-infinite optimization and bilevel optimization,* European Journal of Operational Research, Vol. 142 (2002), 444-462.
23. O. Stein, G. Still, *Solving semi-infinite optimization problems with interior point techniques,* SIAM Journal on Control and Optimization, Vol. 42 (2003), 769-788.
24. W. Wetterling, *Definitheitsbedingungen für relative Extrema bei Optimierungs- und Approximationsaufgaben,* Numerische Mathematik, Vol. 15 (1970), 122-136.
25. G. Zwier, *Structural Analysis in Semi-Infinite Programming,* Thesis, University of Twente, 1987.