



HAL
open science

Mathematical Programs with Vanishing Constraints: Constraint Qualifications, their Applications and a New Regularization Method

Jean-Pierre Dussault, Mounir Haddou, Tangi Migot

► **To cite this version:**

Jean-Pierre Dussault, Mounir Haddou, Tangi Migot. Mathematical Programs with Vanishing Constraints: Constraint Qualifications, their Applications and a New Regularization Method. *Optimization*, 2019, 68 (2-3), pp.509-538. 10.1080/02331934.2018.1542531 . hal-01701461v2

HAL Id: hal-01701461

<https://hal.science/hal-01701461v2>

Submitted on 16 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mathematical Programs with Vanishing Constraints: Constraint Qualifications, their Applications and a New Regularization Method

Dussault, J.-P.*

Haddou, M.†

Migot, T.‡

July 2018

Preprint version of the article published in *Optimization*
doi:10.1080/02331934.2018.1542531

Abstract

We propose a new family of relaxation schemes for mathematical programs with vanishing constraints that extend the relaxation of Hoheisel, Kanzow & Schwartz from 2012. We discuss the properties of the sequence of relaxed non-linear programs as well as stationary properties of limiting points. Our relaxation schemes have the desired property of converging to an M-stationary point. A main advantage of this new method is to converge to an S-stationary point satisfying MPVC-LICQ for a large class of problem. We also study MPVC constraint qualification connected to this study and prove convergence of the method under the new MPVC-CRSC. Additionally, we obtain the new MPVC-wGCQ and prove that it is the weakest MPVC constraint qualification.

Keywords: non-linear programming; MPCC; MPEC; MPVC; relaxation methods

1 Introduction

We consider the Mathematical Program with Vanishing Constraints defined as

$$\begin{aligned} \min_{x \in \mathbb{R}^n} f(x) \\ \text{s.t. } h(x) = 0, g(x) \leq 0, H(x) \geq 0, \\ G_i(x)H_i(x) \leq 0, i = 1, \dots, q, \end{aligned} \tag{MPVC}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}, g : \mathbb{R}^n \rightarrow \mathbb{R}^p, h : \mathbb{R}^n \rightarrow \mathbb{R}^m, G, H : \mathbb{R}^n \rightarrow \mathbb{R}^q$ are continuously differentiable functions. This problem has been recently introduced in [2], motivated by several real-world applications, mainly for topology design problems in mechanical structures as described in [2, 14], but also for robots motion planning [24, 25]. In general, the vanishing constraint can be interpreted as a logic constraint of the form $0 < H_i(x) \implies G_i(x) \leq 0$, which can appear in many challenging applications.

The MPVC can be reformulated as the well-studied mathematical program with complementarity constraints. However, this reformulation may induce some difficulties regarding to the constraint qualifications as pointed in early studies of the problem in [2] or in the thesis [14].

This observation has motivated enhanced studies to derive stationary conditions and numerical methods to find a local minimum for the MPVC. Indeed, classical optimality conditions in non-linear programming cannot be applied here due to the degeneracy of the vanishing constraints. It has been shown in [2] that most

*Département d'Informatique, faculté des Sciences, Université de Sherbrooke, Canada. This research was partially supported by NSERC grant.

†Univ Rennes, INSA Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France.

‡Univ Rennes, INSA Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France. E-mail: tangi.migot@insa-rennes.fr. This research was partially supported by a French grant from "l'Ecole des Docteurs de l'UBL" and "le Conseil Régional de Bretagne".

of the classical constraint qualifications used in non-linear programming does not hold here in a generic way. Thus, enhanced stationary conditions and their corresponding constraint qualifications have been defined in the literature in [2, 10, 15, 16, 17, 20, 21, 23].

From a geometrical point of view, these difficulties can be explained by the fact that the feasible domain is in general non-convex (possibly non-connected see Example 8.1) and has an empty relative interior.

In view of the constraint qualifications issues pledge to the MPVC the relaxation methods provide an intuitive answer. The vanishing constraint, $G_i(x)H_i(x) \leq 0$, is relaxed using a parameter so that the new feasible domain is not thin anymore. It is assumed here that the classical constraints $g(x) \leq 0$ and $h(x) = 0$ are not more difficult to handle than the vanishing constraint. Finally, the relaxation parameter is reduced to converging to the feasible set of (MPVC) in a similar way as an homotopy technique. This method is still one of the most popular in the literature [1, 3, 18, 19, 23].

We mention here that other methods have also been proposed: a Newton with active-set strategy in [22], and a SQP method in [24] that considers sequences of quadratic programs with vanishing constraints. However, both methods require stronger assumptions than the ones used for regularization method in general and in particular in this paper.

Due to the non-convexity of the constraints the natural goal is to compute a local minimum of the problem. Algorithms for non-convex problems usually guarantee the convergence to necessary optimality conditions holding at a local minimum whenever a constraint qualification hold. In the context of MPVC, the genuine optimality conditions are the M-stationary conditions, see [17], as long as the problem satisfies some enhanced MPVC constraint qualification. So, we expect numerical methods designed for MPVC to guarantee convergence to M-stationary points.

The main aim of this paper focus on a new relaxation method to solve the MPVC and in particular on a method with guarantees of convergence to the desired M-stationary conditions. A first step in this path has been proposed in [19] and we continue the discussion by introducing the new butterfly relaxation method.

In the special case, where MPVC-LICQ holds at a local minimum, then S-stationary conditions, that are stronger than M-stationary conditions, are expected to hold at a local minimum. A main advantage of the butterfly relaxation method compared to others is to offer convergence to S-stationary point under this condition for a large class of problem.

One of the main advantages of the relaxation methods is that they require weak constraint qualification to converge. In this paper, we also focus on the assumptions to ensure convergence of the relaxation scheme. This analyze leads us to introduce a new constraint qualification so-called MPVC-CRSC and a new constraint qualification connected with the M-stationary conditions, called MPVC-wGCQ.

One difficulty of this class of method can be to compute the sequence of stationary points of the regularized sub-problems. We present in Section 7 some preliminary result on this question.

The article is organized in 9 sections. We now present a hierarchical overview of these sections.

1. **Introduction.**
2. **Non-linear Programming.** Introduce some preliminary notions from non-linear programming.
3. **MPVC: Definitions, Stationarity and Constraint Qualifications.** Define the various stationary conditions and some constraint qualifications tailored for MPVC.
4. **Algorithmic Applications of MPVC Constraint Qualifications.** Split into two parts, this section discusses the weakest MPVC constraint qualification guaranteeing M-stationary conditions to hold at a local minimum, and, presents a stronger condition, so-called MPVC-CRSC, that can be used to guarantee convergence of some sequence to an M-stationary point.
5. **The New Butterfly Relaxation Method for MPVC.** Introduce the new butterfly relaxation method.
6. **Convergence of the Butterfly Relaxation.** Prove convergence of the method to M-stationary points under the new MPVC-CRSC condition and to S-stationary points for a special class of problem.

7. **Existence of Lagrange Multipliers of the Relaxed Sub-Problems.** Discuss the existence of Lagrange multipliers of the relaxed sub-problems.
8. **Numerical Examples.** Provide numerical illustrations of this new method.
9. **Conclusion.** Present some conclusions and outline perspectives.
- A. **MPVC-CRSC and Error Bound for MPVC.** Show additional error bound and exact penalty results connected with the new constraint qualification MPVC-CRSC.

2 Non-linear Programming

The MPVC is obviously a non-linear program. Even so we cannot directly apply non-linear programming (NLP) techniques, we will use some of the tools from NLP to analyze this problem.

Consider the following non-linear program (without vanishing constraints)

$$\min_{x \in \mathbb{R}^n} f(x) \text{ s.t. } g(x) \leq 0, h(x) = 0. \quad (1)$$

In a usual way, we denote $\mathcal{I}_g(x^*) := \{i \mid g_i(x^*) = 0\}$ and \mathcal{F} the feasible set of (1). A classical tool in the analysis of this problem is the computation of stationary points of (1). We call x^* a stationary point of the non-linear program (1), if there exists $\lambda := (\lambda^g, \lambda^h) \in \mathbb{R}_+^p \times \mathbb{R}^m$ such that $-\nabla f(x^*) = \nabla g(x^*)^T \lambda^g + \nabla h(x^*)^T \lambda^h$ and $\lambda_i^g = 0$ for all $i \notin \mathcal{I}_g(x^*)$.

It is well-known that any local minimum that satisfies the Guignard CQ (GCQ) is a KKT-point. GCQ is said to hold at a point x^* feasible for (1) if the following tangent cone

$$\mathcal{T}_{\mathcal{F}}(x^*) = \{d \in \mathbb{R}^n \mid \exists t_k \geq 0 \text{ and } \mathcal{F} \ni x^k \rightarrow x^* \text{ s.t. } t_k(x^k - x^*) \rightarrow d\},$$

and linearized cone

$$\mathcal{L}(x^*) = \{d \in \mathbb{R}^n \mid \nabla g_i(x)^T d \leq 0 \ (i \in \mathcal{I}_g(x^*)), \nabla h_i(x)^T d = 0 \ (i = 1, \dots, m)\},$$

have the same polar cones. In other words, GCQ holds at x^* if

$$\mathcal{T}_{\mathcal{F}}(x^*)^\circ = \mathcal{L}(x^*)^\circ,$$

where the polar cone of a cone K is given as $K^\circ := \{d \in \mathbb{R}^n \mid d^T y \leq 0, \forall y \in K\}$. It is also well-known since Gould and Tolle in [12], that GCQ is the weakest constraint qualification ensuring that the KKT conditions are necessary conditions independently of the objective function. Up to this point, we may point out that we say that $x^* \in \mathcal{F}$ satisfies Abadie CQ (ACQ) if $\mathcal{T}(x^*) = \mathcal{L}(x^*)$. This condition is stronger than GCQ.

In general, it is not convenient to verify GCQ or ACQ and these conditions may be too weak to induce algorithmically useful properties. Thus, a whole family of stronger constraint qualifications have been studied in the literature. We restrict ourselves to define only the ones that will be useful in the sequel.

Definition 2.1. Let $x^* \in \mathcal{F}$.

- (a) *Linear Independence CQ (LICQ) holds at x^* if the family of gradients $\{\nabla g_i(x^*) \ (i \in \mathcal{I}_g(x^*)), \nabla h_i(x^*) \ (i = 1, \dots, m)\}$ is linearly independent.*
- (b) *Constant Positive-Linear Dependence constraint qualification (CPLD) holds at x^* if, for any subsets $I_1 \subset \mathcal{I}_g(x^*)$ and $I_2 \subset \{1, \dots, m\}$ such that the gradients $\{\nabla g_i(x^*) \ (i \in I_1)\} \cup \{\nabla h_i(x^*) \ (i \in I_2)\}$ are positively linearly dependent, there exists $\delta > 0$ such that they remain linearly dependent for every $x \in \mathcal{B}_\delta(x^*)$.*
- (c) *Constant Rank in the Subspace of Components (CRSC) holds at x^* if there exists $\delta > 0$ such that the family of gradients $\{\nabla g_i(x) \ (i \in J_-), \nabla h_i(x^*) \ (i = 1, \dots, m)\}$ has the same rank for every $x \in \mathcal{B}_\delta(x^*)$, where $J_- := \{i \in \mathcal{I}_g(x^*) \mid -\nabla g_i(x^*) \in \mathcal{L}(x^*)^\circ\}$.*

(d) The Cone-Continuity Property (CCP) holds at x^* if the set-valued mapping $\mathbb{R}^n \ni x \rightrightarrows K(x)$ defined by

$$K(x) := \left\{ \sum_{i \in \mathcal{I}_g(x^*)} \lambda_i \nabla g_i(x) + \sum_{i=1}^m \mu_i \nabla h_i(x) \mid \lambda \in \mathbb{R}_+^p, \mu \in \mathbb{R}^m \right\}$$

is outer semicontinuous.

Constant rank of the subspace component, CRSC, was introduced recently in [4]. This latter definition considers an unusual set denoted J_- , that can be viewed as the set of indices of the gradients of the active constraints whose Lagrange multiplier, if they exist, may be non-zero.

It is to be noted that in the definition of CCP, $K(x)$ depends on x^* , since it considers only active constraints at x^* . Clearly, $K(x^*)$ is a closed convex cone and coincides with the polar of the linearized cone $\mathcal{L}(x^*)^\circ$. Moreover, $K(x)$ is always inner semicontinuous due to the continuity of the gradients and the definition of $K(x)$. For this reason, outer semicontinuity is sufficient to get the continuity of $K(x)$ at x^* . Finally, it has been shown in [5] that CCP is strictly stronger than ACQ and weaker than CRSC, see Figure 1.

$$\text{CRSC} \implies \text{CCP} \implies \text{ACQ}$$

Figure 1: Relations between the constraint qualifications.

3 MPVC: Definitions, Stationarity and Constraint Qualifications

This section defines the stationary conditions and some constraint qualifications that have been used in the literature to study the MPVC. Up to this point, we notice that similar but fundamentally different notions are given in the literature for non-linear programs and mathematical programs with complementarity constraints (MPCC).

Beforehand, let us define some notations. Given $x^* \in \mathcal{Z} := \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0, H(x) \geq 0, G_i(x)H_i(x) \leq 0 \ (i = 1, \dots, q)\}$, we denote

$$\begin{aligned} \mathcal{I}_g &:= \{i \mid g_i(x^*) = 0\}, \quad \mathcal{I}^0 := \{i \mid H_i(x^*) = 0\}, \quad \mathcal{I}^+ := \{i \mid H_i(x^*) > 0\}, \\ \mathcal{I}^{-+} &:= \{i \mid G_i(x^*) < 0, H_i(x^*) > 0\}, \quad \mathcal{I}^{0+} := \{i \mid G_i(x^*) = 0, H_i(x^*) > 0\}, \\ \mathcal{I}^{+0} &:= \{i \mid G_i(x^*) > 0, H_i(x^*) = 0\}, \quad \mathcal{I}^{00} := \{i \mid G_i(x^*) = 0, H_i(x^*) = 0\}, \\ \mathcal{I}^{-0} &:= \{i \mid G_i(x^*) < 0, H_i(x^*) = 0\}. \end{aligned}$$

Note that the dependence of x^* is omitted here, since these sets are always used at this point. Let the enhanced Lagrangian of (MPVC) be defined as $\mathcal{L}_{MPVC}(x^*, \lambda) := f(x^*) + g(x^*)^T \lambda^g + h(x^*)^T \lambda^h + G(x^*)^T \lambda^G - H(x^*)^T \lambda^H$ with $\lambda := (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$.

3.1 Enhanced Stationarity Conditions

We now define the alphabet of stationary conditions for (MPVC). They differ in the sign of λ_i^G and λ_i^H for indices $i \in \mathcal{I}^{00}$.

Definition 3.1. A point $x^* \in \mathcal{Z}$, is called

(a) *W-stationary* [23], if there exists multipliers $\lambda \in \mathbb{R}_+^p \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^q$ such that

$$\nabla_x \mathcal{L}_{MPVC}(x^*, \lambda) = 0,$$

and

$$\begin{aligned} \lambda_i^g &= 0 \ \forall i \notin \mathcal{I}_g, \\ \lambda_i^G &= 0 \ \forall i \in \mathcal{I}^{-+} \cup \mathcal{I}^{-0} \cup \mathcal{I}^{+0}, \quad \lambda_i^G \geq 0 \ \forall i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}, \\ \lambda_i^H &= 0 \ \forall i \in \mathcal{I}^+, \quad \lambda_i^H \geq 0 \ \forall i \in \mathcal{I}^{-0}. \end{aligned}$$

- (b) T-stationary [10], if x^* is W-stationary and $\lambda_i^G \lambda_i^H \leq 0$ for all $i \in \mathcal{I}^{00}$.
- (c) M-stationary [16], if x^* is W-stationary and $\lambda_i^G \lambda_i^H = 0$ for all $i \in \mathcal{I}^{00}$.
- (d) S-stationary, if x^* is W-stationary and $\lambda_i^H \geq 0, \lambda_i^G = 0$ for all $i \in \mathcal{I}^{00}$.

The relations between these conditions follow in a straightforward way from the definition: S-stationary \implies M-stationary \implies T-stationary \implies W-stationary. These conditions are in general weaker than the classical KKT condition, which is equivalent here to the S-stationary condition. The T-stationary condition defined in [10] plays here the role of the C-stationary condition in the context of MPCCs.

We conclude this section by giving a new interpretation of M-stationary based on non-linear programming.

Proposition 3.1. x^* is an M-stationary point of (MPVC) if and only if there exists a partition $\mathcal{I}^{00} = A \cup B$ with $A \cap B = \emptyset$ such that x^* is a stationary point of the following non-linear program, denoted $MNLP_{A,B}(x^*)$,

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f(x) \\ & \text{s.t. } g(x) \leq 0, h(x) = 0, \\ & \quad G_i(x) \leq 0 \ \forall i \in \mathcal{I}^{0+} \cup A, H_i(x) = 0 \ \forall i \in \mathcal{I}^{+0} \cup B, \\ & \quad H_i(x) \geq 0 \ \forall i \in \mathcal{I}^{-0}. \end{aligned} \tag{MNLP_{A,B}(x^*)}$$

Based on this observation, we can derive constraint qualifications tailored for MPVC by applying classical constraint qualifications from non-linear programming to $(MNLP_{A,B}(x^*))$. This will be discussed with more details in the following section. This philosophy to derive constraint qualifications based on classical constraint qualifications applied on a specific non-linear program has been used in [18] to derive stronger conditions than the ones presented here.

3.2 Constraint Qualifications for MPVC

Let us introduce the MPVC-weak Guignard CQ (MPVC-wGCQ) that will be useful in stating optimality conditions for (MPVC).

Definition 3.2. Let $x^* \in \mathcal{Z}$. MPVC-wGCQ holds at x^* if for any partition $\mathcal{I}^{00} = A \cup B$ with $A \cap B = \emptyset$, GCQ holds at x^* for $(MNLP_{A,B}(x^*))$.

In the sequel, we show that this is a weaker constraint qualification than MPVC-GCQ. As a consequence of Proposition 3.1, we have the following equivalent characterization of M-stationary points.

Corollary 3.1. Let $x^* \in \mathcal{Z}$ and let the cone \mathcal{P}_M be defined as

$$\mathcal{P}_M := \bigcup_{A \in \mathcal{P}(\mathcal{I}^{00}), B = \mathcal{I}^{00} \setminus A} \mathcal{L}_{MNLP_{A,B}}(x^*)^\circ.$$

Then, it holds true that

$$-\nabla f(x^*) \in \mathcal{P}_M \iff x^* \text{ is M-stationary,}$$

In the previous statement $\mathcal{P}(\mathcal{I}^{00})$ denotes the set of all subsets of \mathcal{I}^{00} .

Proof. By straightforward computation, we can give an explicit form of the polar cone of $\mathcal{L}_{MNLP_{A,B}}$ for all

A, B . Then, computing the union of these cones yields to

$$\begin{aligned} \bigcup_{A \in \mathcal{P}(\mathcal{I}^{00}), B = \mathcal{I}^{00} \setminus A} \mathcal{L}_{MNLPA,B}(x^*)^\circ &= \{d \in \mathbb{R}^n \mid \exists(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \text{ s.t. } d = \mathcal{L}_{MPVC}(x^*, \lambda), \\ &\quad \text{with } \lambda_i^g \geq 0 \ (i \in \mathcal{I}_g), \\ &\quad \lambda_i^H = 0 \ (i \in \mathcal{I}^+), \\ &\quad \lambda_i^H \geq 0 \ (i \in \mathcal{I}^{-0}), \\ &\quad \lambda_i^G = 0 \ (i \in \mathcal{I}^{-+} \cup \mathcal{I}^{-0} \cup \mathcal{I}^{+0}), \\ &\quad \lambda_i^G \geq 0 \ (i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}), \\ &\quad \lambda_i^G \lambda_i^H = 0 \ (i \in \mathcal{I}^{00})\}. \end{aligned} \quad (2)$$

Thus, the result follows by combining this computation with Proposition 3.1. \square

MPVC-wGCQ is different from the MPVC-GCQ used for instance in [16] or in [17]. First, let us define the MPVC-linearized cone

$$\begin{aligned} \mathcal{L}_{MPVC}(x^*) &:= \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0 \ (i \in \mathcal{I}_g), \nabla h_j(x^*)^T d = 0 \ (j = 1, \dots, m), \\ &\quad \nabla H_i(x^*)^T d = 0 \ (i \in \mathcal{I}^{+0}), \nabla H_i(x^*)^T d \geq 0 \ (i \in \mathcal{I}^{00} \cup \mathcal{I}^{-0}), \\ &\quad \nabla G_i(x^*)^T d \leq 0 \ (i \in \mathcal{I}^{0+}), \\ &\quad (\nabla H_i(x^*)^T d)(\nabla G_i(x^*)^T d) \leq 0 \ (i \in \mathcal{I}^{00})\}. \end{aligned} \quad (3)$$

Note that $\mathcal{L}_{MPVC}(x^*)$ is, in general, a non-convex cone and then, differs from $\mathcal{L}(x^*)$ applied to (MPVC). However, due to Corollary 2.5 of [17] it holds

$$\mathcal{T}_{\mathcal{Z}}(x^*) \subseteq \mathcal{L}_{MPVC}(x^*) \subseteq \mathcal{L}(x^*). \quad (4)$$

Definition 3.3. Let $x^* \in \mathcal{Z}$. MPVC-GCQ holds at x^* if $\mathcal{T}_{\mathcal{Z}}(x^*)^\circ = \mathcal{L}_{MPVC}(x^*)^\circ$.

These definitions have been introduced earlier in [17]. The following lemma gives a relation between $\mathcal{L}_{MPVC}(x^*)^\circ$ and the set \mathcal{P}_M used in Corollary 3.1.

Lemma 3.1. Let $x^* \in \mathcal{Z}$. The following inclusion holds true

$$\mathcal{L}_{MPVC}(x^*)^\circ \subseteq \mathcal{P}_M.$$

Proof. The proof of Corollary 3.1 already states an explicit formula for the right-hand side of the inclusion in equation (2). Let us now compute explicitly $\mathcal{L}_{MPVC}(x^*)^\circ$.

First, it is to be observed that the cone \mathcal{L}_{MPVC} can be rewritten as a union of polyhedral cones in the following way.

$$\mathcal{L}_{MPVC}(x^*) = \bigcup_{I \subset \mathcal{I}^{00}} \mathcal{L}_{MPVC}(x^*, I),$$

where $\mathcal{L}_{MPVC}(x^*, I)$ stands for

$$\begin{aligned} \mathcal{L}_{MPVC}(x^*, I) &:= \{d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0 \ (i \in \mathcal{I}_g), \\ &\quad \nabla h_i(x^*)^T d = 0 \ (i = 1, \dots, m), \\ &\quad \nabla H_i(x^*)^T d = 0 \ (i \in \mathcal{I}^{+0}), \\ &\quad \nabla H_i(x^*)^T d \geq 0 \ (i \in I \cup \mathcal{I}^{-0}), \\ &\quad \nabla G_i(x^*)^T d \leq 0 \ (i \in \mathcal{I}^{0+}), \\ &\quad \nabla G_i(x^*)^T d \leq 0 \ (i \in I), \\ &\quad \nabla H_i(x^*)^T d = 0, \nabla G_i(x^*)^T d \geq 0 \ (i \in \mathcal{I}^{00} \setminus I)\}. \end{aligned}$$

Now, let us compute the polar of \mathcal{L}_{MPVC} at x^* using [6, Proposition 1.1.16],

$$\mathcal{L}_{MPVC}(x^*)^\circ = \bigcap_{I \subset \mathcal{I}^{00}} \mathcal{L}_{MPVC}(x^*, I)^\circ.$$

We can compute the polar cone of $\mathcal{L}_{MPVC}(x^*, I)$ given by

$$\begin{aligned} \mathcal{L}_{MPVC}(x^*, I)^\circ := \{d \in \mathbb{R}^n \mid & \exists(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}_+^p \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^q \\ & \text{with } \lambda_i^G \geq 0 \ \forall i \in I \cup \mathcal{I}^{0+}, \lambda_i^G \leq 0 \\ & \forall i \in \mathcal{I}^{00} \setminus I, \lambda_i^H \geq 0 \ \forall i \in I \cup \mathcal{I}^{-0}, \\ d = & \sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x^*) \\ & + \sum_{i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}} \lambda_i^G \nabla G_i(x^*) \\ & - \sum_{i \in \mathcal{I}^{+0} \cup \mathcal{I}^{00} \cup \mathcal{I}^{-0}} \lambda_i^H \nabla H_i(x^*)\}. \end{aligned}$$

Now, computing the intersection in the equation above leads to

$$\begin{aligned} \mathcal{L}_{MPVC}(x^*)^\circ := \{d \in \mathbb{R}^n \mid & \exists(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}_+^p \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^q \\ & \text{with } \lambda_i^G \geq 0 \ \forall i \in \mathcal{I}^{0+}, \lambda_i^G = 0 \ \forall i \in \mathcal{I}^{00}, \\ & \lambda_i^H \geq 0 \ \forall i \in \mathcal{I}^{00} \cup \mathcal{I}^{-0}, \\ d = & \sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x^*) \\ & + \sum_{i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}} \lambda_i^G \nabla G_i(x^*) \\ & - \sum_{i \in \mathcal{I}^{+0} \cup \mathcal{I}^{00} \cup \mathcal{I}^{-0}} \lambda_i^H \nabla H_i(x^*)\}. \end{aligned}$$

From the two explicit formulas of \mathcal{L}_{MPVC}° and \mathcal{P}_M , it is now clear that $\mathcal{L}_{MPVC}(x^*)^\circ \subseteq \mathcal{P}_M$. \square

A consequence of this result is that MPVC-GCQ implies MPVC-wGCQ.

Definition 3.4. A feasible point x^* of (MPVC) satisfies

(a) *MPVC-Linear Independence CQ (MPVC-LICQ) if the gradients*

$$\{\nabla g_i(x^*) \ (i \in \mathcal{I}_g), \nabla h_i(x^*) \ (i = 1, \dots, m), \nabla G_i(x^*) \ (i \in I_{00} \cup I_{+0}), \nabla H_i(x^*) \ (i \in \mathcal{I}^{00})\}$$

are linearly independent.

(b) *MPVC-Constant Positive-Linear Dependence constraint qualification (MPVC-CPLD) if, for all subsets $I_1 \subseteq \mathcal{I}_g$, $I_2 \subseteq I_{0-}$, $I_3 \subseteq I_{+0} \cup I_{00}$, $I_4 \subseteq \{1, \dots, p\}$, $I_5 \subseteq I_{0+} \cup I_{00}$, the following implication holds true: if the gradients*

$$\begin{aligned} & \{\nabla g_i(x^*) \ (i \in \mathcal{I}_g), -\nabla H_i(x^*) \ (i \in I_2), \nabla G_i(x^*) \ (i \in I_3)\} \cup \\ & \{\{\nabla h_i(x^*) \ (i \in I_4), \nabla H_i(x^*) \ (i \in I_5)\}\} \end{aligned}$$

are positively linearly dependent, then they remain linearly dependent for all x in some neighborhood of x^ .*

The MPVC-LICQ, introduced in [2], is a strong condition that guarantee that a local minimum is an S-stationary point [17]. Under this condition, it has been proved in [2] that the MPVC Lagrange multipliers are unique. The MPVC-CPLD introduced in [18] is very popular in the context of relaxation methods to prove their convergence, see [18, 19]. In the sequel, we introduce a weaker condition, denoted as MPVC-Constant Rank in the Subspace of Component (MPVC-CRSC), to prove similar results of convergence.

Definition 3.5. *Let $x^* \in \mathcal{Z}$. MPVC-CRSC holds at x^* if for any partition $\mathcal{I}^{00} = A \cup B$ with $A \cap B = \emptyset$, CRSC holds at x^* for $(MNLP_{A,B}(x^*))$.*

This definition is an extension of the one given in [4] for non-linear programming and [11] for the MPCC. It follows in a straightforward way from their definitions that MPVC-CRSC is stronger than MPVC-wGCQ.

In a similar way as in Definition 3.5, applying the CCP at $(MNLP_{A,B}(x^*))$ for any partition $\mathcal{I}^{00} = A \cup B$ yields a new MPVC constraint qualification, denoted MPVC-CCP.

Definition 3.6. *We say that a feasible point x^* satisfies the MPVC-CCP if the set-valued mapping $\mathbb{R}^n \ni x \rightrightarrows K_{MPVC}(x)$ defined by*

$$K_{MPVC}(x) := \left\{ \sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) + \sum_{i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}} \lambda_i^G \nabla G_i(x) - \sum_{i \in \mathcal{I}^{+0} \cup \mathcal{I}^{00} \cup \mathcal{I}^{-0}} \lambda_i^H \nabla H_i(x) \mid \right. \\ \left. \lambda_i^g \geq 0 (i \in \mathcal{I}_g) \text{ and } , \lambda_i^H \geq 0 (i \in \mathcal{I}^{-0}), \lambda_i^G \geq 0 (i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}), \right. \\ \left. \lambda_i^G \lambda_i^H = 0 (i \in \mathcal{I}^{00}) \right\}$$

is outer semicontinuous at x^* , that is

$$\limsup_{x \rightarrow x^*} K_{MPVC}(x) \subset K_{MPVC}(x^*).$$

In this context, the outer limit is taken in the sense of Painlevé-Kuratowski corresponding to the Definition 5.4 given in [28]. By construction, it holds that MPVC-CRSC implies MPVC-CCP, which itself implies MPVC-wGCQ.

As we already remarked the new conditions defined here (MPVC-wGCQ, MPVC-CCP and MPVC-CRSC) are weak conditions compared to existing ones in the literature, except for MPVC-GCQ and MPVC-ACQ. In a same way, they are weaker than other conditions that have been omitted here, MPVC-CRCQ defined in [18], MPVC-MFCQ defined in [2] and MPVC-GMFCQ first introduced in [20]. Besides in the following section, we show that these new weak conditions have a theoretical and algorithmic impacts.

4 Algorithmic Applications of MPVC Constraint Qualifications

In this section, we prove algorithmic applications of the new MPVC constraint qualifications defined in the previous section. First, we show that MPVC-wGCQ is the weakest MPVC constraint qualification in the sense that it is the weakest condition independent of the objective function such that M-stationary conditions are necessary optimality conditions. Then, in Section 4.2 we show convergence of a specific sequence converging to an M-stationary point under MPCC-CRSC or MPCC-CCP.

4.1 MPVC-wGCQ as the Weakest MPVC Constraint Qualification

The following shows that MPVC-wGCQ is an MPVC-CQ.

Theorem 4.1. *Let $x^* \in \mathcal{Z}$ be a local minimum of (MPVC) that satisfies MPVC-wGCQ. Then, x^* is an M-stationary point.*

Proof. By Equation 4 and Lemma 3.1, MPVC-wGCQ yields to $\mathcal{T}_{\mathcal{Z}}(x^*)^\circ \subset \mathcal{P}_M$.

Thus, under MPVC-wGCQ, it holds that it holds that

$$\begin{aligned} -\nabla f(x^*) \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ &\implies -\nabla f(x^*) \in \cup \mathcal{P}_M, \\ &\iff x^* \text{ M-stationary,} \end{aligned}$$

where the last equivalence comes from Corollary 3.1. \square

A similar result has been proved under MPVC-GCQ in [17]. In particular, this result gives that in our context the natural goal of any numerical method designed for (MPVC) is to compute an M-stationary point. Indeed, it was already pointed out in [17] that S-stationary conditions are unlikely to hold in general at a local minimum even for problems with g, h, G, H affines.

The fact that this condition is the weakest constraint qualification for (MPVC) is actually a consequence of the similar result in non-linear programming stated in [12].

Theorem 4.2. *MPVC-wGCQ is the weakest MPVC constraint qualification. In other words, for all function f that admits a local constraint minimum at x^* then*

$$x^* \text{ is an M-stationary point} \iff \text{MPVC-wGCQ holds at } x^*.$$

Before moving to the proof, let us introduce a theorem from [12]. The original result is stated for local maximum, but there is no loss of generality to write it for a local minimum.

Theorem 4.3. *For every, $y \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ$ there exists, an objective function f , which is differentiable at x^* , which has a local constrained minimum at x^* and for which $-\nabla f(x^*) = y$.*

We can now go on the proof of Theorem 4.2.

Proof of Theorem 4.2. The " \iff " part is given by Theorem 4.1.

So, let us consider the " \implies " part. Assume that for all function f that admits a local constrained minimum at x^* , x^* is an M-stationary point. We prove that for any $y \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ$, then $y \in \mathcal{P}_M(x^*)$ and so MPVC-wGCQ holds at x^* .

By Theorem 4.3, for any $y \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ$ there exists a function f such that $y = -\nabla f(x^*)$. Now, since x^* is an M-stationary point, Corollary 3.1 gives that $-\nabla f(x^*) \in \mathcal{P}_M(x^*)$, which proves that MPVC-wGCQ holds at x^* . \square

4.2 Convergence of a Primal-Dual Sequence

During the process of an iterative algorithm, we are interested in the study of accumulation points of sequences computed by the relaxation method. It is common to compute sequences that satisfy the following assumptions.

Assumption 4.1. *Let $\{x^k\}$ and $0 \neq \{\lambda^k\} \in \mathbb{R}_+^p \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^q$ be such that $x^k \rightarrow x^*$ and*

$$\nabla f(x^k) + \sum_{i=1}^p \lambda_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i(x^k) + \sum_{i=1}^q \lambda_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^q \lambda_i^{H,k} \nabla H_i(x^k) \rightarrow 0, \quad (\text{i})$$

$$\lim_{k \rightarrow \infty} \frac{\lambda_i^g}{\|\lambda^k\|_\infty} = 0 \quad \forall i \notin \mathcal{I}_g, \quad \lim_{k \rightarrow \infty} \frac{\lambda_i^H}{\|\lambda^k\|_\infty} \geq 0 \quad \forall i \in \mathcal{I}^{-0}, \quad (\text{ii})$$

$$\lim_{k \rightarrow \infty} \frac{\lambda_i^{G,k}}{\|\lambda^k\|_\infty} = 0 \quad \forall i \in \mathcal{I}^{-+} \cup \mathcal{I}^{-0} \cup \mathcal{I}^{+0} \quad \text{and} \quad \lim_{k \rightarrow \infty} \frac{\lambda_i^{H,k}}{\|\lambda^k\|_\infty} = 0 \quad \forall i \in \mathcal{I}^{+}, \quad (\text{iii})$$

$$\lim_{k \rightarrow \infty} \frac{\lambda_i^{G,k} \lambda_i^{H,k}}{\|\lambda^k\|_\infty^2} = 0 \quad \forall i \in \mathcal{I}^{00}, \quad (\text{iv})$$

$$\lim_{k \rightarrow \infty} \frac{\lambda_i^{G,k}}{\|\lambda^k\|_\infty} \geq 0 \quad \forall i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}, \quad (\text{v})$$

the family of gradients of non-vanishing multipliers in (i) are linearly independent. (vi)

This condition may correspond to some kind of sequential optimality conditions. Note that assumption (vi) is not restrictive. As a corollary of Caratheodory's Theorem (see Exercise B.1.7 in [8]), we can build a sequence of multipliers that satisfies conditions (i) to (v), such that the gradients corresponding to non-vanishing multipliers in equation (i) are linearly independent for all $k \in \mathbb{N}$. This may change the multipliers, but previously positive ones will stay at least non-negative and vanishing multipliers will remain zero.

Theorem 4.4. *Given two sequences $\{x^k\}, \{\lambda^k\}$ that satisfy Assumption 4.1. Suppose that $x^k \rightarrow x^* \in \mathcal{Z}$, and MPCC-CRSC holds at x^* . Then, the sequence $\{\lambda^k\}$ is bounded.*

Proof. Let $\{w^k\}$ be a sequence defined such that

$$w^k := \sum_{i \in \mathcal{I}_g} \lambda_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i(x^k) + \sum_{i \in \mathcal{I}^0 \cup \mathcal{I}^{00}} \lambda_i^{G,k} \nabla G_i(x^k) - \sum_{i \in \mathcal{I}^{-0} \cup \mathcal{I}^{00}} \lambda_i^{H,k} \nabla H_i(x^k). \quad (5)$$

We prove by contradiction that the sequence $\{\lambda^k\}$ is bounded. If λ^k were not bounded, there would exist a subsequence such that

$$\frac{\lambda^k}{\|\lambda^k\|_\infty} \rightarrow \bar{\lambda} \neq 0.$$

Here we consider a subsequence K , where the family of linearly independent gradients of non-vanishing multipliers is the same for all $k \in K$. Note that this can be done with no loss of generality, since there is a finite number of such subsequences and altogether they form a partition of the sequence.

Note here that conditions (i) to (v) give that $\lim_{k \rightarrow \infty} w^k = \lim_{k \rightarrow \infty} -\nabla f(x^k) / \|\lambda^k\|_\infty = 0$.

Dividing by $\|\lambda^k\|_\infty$ and passing to the limit in (5),

$$w^* = \sum_{i \in \mathcal{I}_g} \bar{\lambda}_i^g \nabla g_i(x^*) + \sum_{i=1}^m \bar{\lambda}_i^h \nabla h_i(x^*) + \sum_{i \in \mathcal{I}^0 \cup \mathcal{I}^{00}} \bar{\lambda}_i^G \nabla G_i(x^*) - \sum_{i \in \mathcal{I}^{-0} \cup \mathcal{I}^{00}} \bar{\lambda}_i^H \nabla H_i(x^*),$$

with $\bar{\lambda}_i^G = 0$ for $i \in \mathcal{I}^{-+} \cup \mathcal{I}^{-0} \cup \mathcal{I}^{+0}$, $\bar{\lambda}_i^H = 0$ for $i \in \mathcal{I}^{+}$ by assumptions (ii) to (v).

It follows that the gradients with non-zero multipliers involved in the previous equation are linearly dependent.

MPVC-CRSC guarantees that these gradients remain linearly dependent in a whole neighborhood. This, however, is a contradiction to the linear independence of these gradients in x^k since $w^k \neq 0$ for k sufficiently large. Here, we used that for all k sufficiently large $\text{supp}(\bar{\lambda}) \subseteq \text{supp}(\lambda^k)$. Consequently, we get a contradiction and thus the sequence $\{\lambda^k\}$ is bounded. \square

A direct consequence of this theorem is that the sequence herein converge to an M-stationary point. Indeed, by the previous theorem any limit point of sequences $\{x^k, \lambda^k\}$ verifying Assumption 4.1 satisfy the M-stationary conditions from Definition 3.1.

Corollary 4.1. *Let x^* be in \mathcal{Z} such that MPVC-CRSC holds at x^* . Given two sequences $\{x^k\}, \{\lambda^k\}$ such that $x^k \rightarrow x^*$ that satisfy Assumption 4.1. Then, x^* is an M-stationary point.*

By considering the weaker condition MPVC-CCP, one can make a similar study.

Theorem 4.5. *Let $x^* \in \mathcal{Z}$ such that MPVC-CCP holds at x^* . Given two sequences $\{\delta^k\}, \{x^k\}$ such that $\delta^k \rightarrow 0, x^k \rightarrow x^*$ and for all $k \in \mathbb{N}$*

$$-\nabla f(x^k) \in K_{MPVC}(x^k) + \mathcal{B}_{\delta^k}(x^k).$$

Then, x^ is an M-stationary point of (MPVC).*

Proof. By continuity of the gradients involved in (MPVC) and MPVC-CCP assumption, it holds that

$$\limsup_{x^k \rightarrow x^*} K_{MPVC}(x^k) = K_{MPVC}(x^*).$$

Therefore, it follows that $-\nabla f(x^*) \in K_{MPVC}(x^*)$. So, x^* is an M-stationary point of (MPVC). \square

The gap between Theorem 4.4 and Theorem 4.5 is that the former guarantee boundedness of the dual sequence, while the latter does not. This observation will be of importance while studying the convergence of the butterfly relaxation in Section 6.

5 The New Butterfly Relaxation Method for MPVC

We now propose an adaptation of the butterfly relaxation method introduced for MPCC in [11] to solve MPVC. As pointed out earlier, Theorem 4.1 gives that our goal is to compute an M-stationary point.

Let $t := (t_1, t_2) \in \mathbb{R}_+^2$. We consider the following relaxed butterfly problem, denoted as (B_t) , parametrized by t

$$\begin{aligned} \min_{x \in \mathbb{R}^n} f(x) \\ \text{s.t. } x \in \mathcal{X}_t^B := \{x \mid h(x) = 0, g(x) \leq 0, H(x) \geq -t_2, \Phi^B(G(x), H(x); t) \leq 0.\} \end{aligned} \quad (B_t)$$

where for all $i = 1, \dots, q$

$$\Phi_i^B(G(x), H(x); t) := G_i(x)F_i(x; t),$$

and

$$F_i(x; t) := H_i(x) - t_2 - t_2\theta_{t_1}(G_i(x) - t_2).$$

Here, the functions $\theta : \mathbb{R} \rightarrow]-\infty, 1]$ are continuously differentiable non-decreasing concave functions with $\theta(0) = 0$, $\theta_{t_1}(x) := \theta(x/t_1)$ for all $t_1 > 0$ and $\lim_{t_1 \rightarrow 0} \theta_{t_1}(x) = 1 \forall x \in \mathbb{R}_{++}$ completed in a smooth way for negative values by considering $\theta_{t_1}(z < 0) = z\theta'(0)/t_1$. Examples of such functions are $\theta_{t_1}^1(x) = \frac{x}{x+t_1}$ and $\theta_{t_1}^2(x) = 1 - \exp^{-\frac{x}{t_1}}$. These functions have already been used in the context of complementarity constraints in [11, 13, 26]. This new relaxation handles two parameters t_1 and t_2 chosen such that

$$t_2 = o(t_1). \quad (6)$$

A figure illustrating the feasible set of the (MPVC) and the relaxed feasible set is given in Figure 1. This method is an extension of the method introduced in [19]. A typical choice to satisfy (6) is to choose $t_2 = t_1^\alpha$ with $0 < \alpha < 1$. We could be even more general by considering a third parameter for the constraints $H_i(x) \geq -\bar{t}$, but we decided not to follow this for simplicity.

The sets of indices used for the analysis of this relaxation are defined in the following way

$$\begin{aligned} \mathcal{I}_H(x; t) &:= \{i \mid H_i(x) = -t_2\}, \\ \mathcal{I}_{GH}(x; t) &:= \{i \mid \Phi_i^B(G(x), H(x); t) = 0\}, \\ \mathcal{I}_{GH}^{+0}(x; t) &:= \{i \in \mathcal{I}_{GH}(x; t) \mid G(x) > 0, F_i(x; t) = 0\}, \\ \mathcal{I}_{GH}^{0+}(x; t) &:= \{i \in \mathcal{I}_{GH}(x; t) \mid G(x) = 0, F_i(x; t) > 0\}, \\ \mathcal{I}_{GH}^{00}(x; t) &:= \{i \in \mathcal{I}_{GH}(x; t) \mid G(x) = 0, F_i(x; t) = 0\}. \end{aligned}$$

The following lemma sum up some of the key features of the relaxation.

Lemma 5.1. *The following properties on \mathcal{X}_t^B hold:*

1. $\mathcal{X}_0^B = \mathcal{Z}$;
2. $\cap_{t \geq 0} \mathcal{X}_t^B = \mathcal{Z}$.

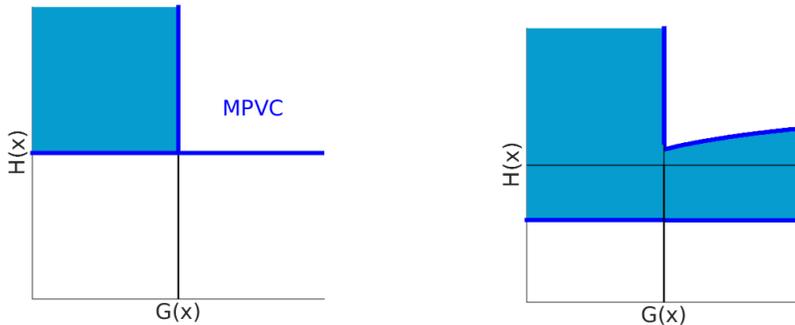


Table 1: Feasible set of the MPVC and the butterfly relaxation for $\theta_{t_1}(z) = \frac{z}{z+t_1}$ with $t_2 = t_1^{3/2}$.

If the feasible set of the (MPVC) is non-empty, then the feasible set of the relaxed sub-problems are also non-empty for all $t \geq 0$. If for some parameter $t \geq 0$ the set \mathcal{X}_t^B is empty, then it implies that \mathcal{Z} is empty. If a local minimum of (B_t) already belongs to \mathcal{Z} , then it is a local minimum of the (MPVC).

Finally, the following lemma from [11] gives more information on the asymptotic convergence of the θ 's functions.

Lemma 5.2. *Given two positive sequences $\{t_{1k}\}$ and $\{t_{2k}\}$ converging to zero. Then, for any $z \in \mathbb{R}_+$ it holds*

$$\lim_{k \rightarrow \infty} t_{2k} \theta_{t_{1k}}(z) = 0,$$

and

$$\lim_{k \rightarrow \infty} t_{2k} \theta'_{t_{1k}}(z) \leq \lim_{k \rightarrow \infty} t_{2k} \theta'_{t_{1k}}(0).$$

Since we assume that $t_{2k} = o(t_{1k})$ in (6), it implies that $t_{2k} \theta'_{t_{1k}}(z_k) \rightarrow 0$.

A generic relaxation algorithm to compute a stationary point of (MPVC) is Algorithm 1.

<p>Data: x^0 an initial point, t_0 initial parameters, $\sigma_t \in (0, 1)$ parameters update;</p> <p>1 Set $k := 0, t_k := t_0$;</p> <p>2 repeat</p> <p>3 $t_{k+1} := \sigma_t t_k$;</p> <p>4 $x^{k+1} :=$ stationary point of (B_t);</p> <p>5 $k := k + 1$;</p> <p>6 until x^{k+1} is a "stationary point of (MPVC)" or decision of infeasibility or unboundedness;</p>

Algorithm 1: Relaxation Algorithm for the MPVC.

Step 4 is usually computed through an iterative method that may use the previous iterate x^k as an initial point.

This generic algorithm raises two important questions:

1. Is it possible to guarantee that the limit point is an M-stationary point ?
2. Do we have guarantees of existence of stationary point of the relaxed sub-problem ?

The first question is our main focus in the sequel and will be treated successfully. Although, existence of Lagrange multipliers will be discussed in Section 7.

We conclude this section by an example that shows that the butterfly relaxation may improve the relaxation [19], which is the only other method with proved convergence to M-stationary point. Indeed, it illustrates an example where there are no sequences of stationary points that converge to any undesirable point. This phenomenon will be generalized in Theorem 6.2 (page 15) to a large class of problem.

Example 5.1.

$$\min_{x \in \mathbb{R}^2} -x_2 \text{ s.t. } x_2 \leq 1, x_2 \geq 0, x_1 x_2 \leq 0.$$

In this example, there are two stationary points: an S -stationary point $(0, 1)$ that is the global minimum and an M -stationary point $(0, 0)$, which is not a local minimum. Indeed, we see that $x^* = (0, 0)^T$ is M -stationary, since computing the gradient of the MPVC-Lagrangian function yields to

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\lambda^H \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda^G \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

so $\lambda^H = -1$ and $\lambda^G = 0$.

For the relaxation [19] the gradient of the Lagrangian equal to 0 gives

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\lambda^{H,k} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda^{\Phi,k} \begin{pmatrix} x_2^k - t_2^k \\ x_1^k \end{pmatrix}.$$

Here, for $t_{2,k} = \frac{1}{k}$ a sequence $x^k = (2t_{2,k} \ t_{2,k})^T$, with $\lambda^{\Phi,k} = k$, converge to $(0, 0)$ as k goes to infinity. However, there are no sequences of stationary point that converges to this undesirable point with the butterfly relaxation.

6 Convergence of the Butterfly Relaxation

As in [19], sequences of stationary points of (B_t) converge to an M -stationary point. Moreover, we prove here that the new MPVC-CRSC is sufficient for this result, while MPVC-CPLD has been considered in [19].

Theorem 6.1. *Given a sequence $\{t_k\}$ of positive parameters satisfying (6) and decreasing to zero as k goes to infinity. Let $\{x^k, \lambda^k\}$ be a sequence of stationary points of (B_t) for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$, where MPVC-CRSC is satisfied in x^* . Then, x^* is an M -stationary point of (MPVC).*

Proof. First, we identify the expressions of the multipliers of the vanishing constraints in Definition 3.1 in function of the stationary points of (B_t) .

Let $\{x^k, \lambda^{g,k}, \lambda^{h,k}, \lambda^{G,k}, \lambda^{H,k}, \lambda^{\Phi,k}\}$ be a sequence of stationary points of (B_t) for all $k \in \mathbb{N}$. The representation of Φ^B immediately gives $\nabla \Phi_i^B(G(x^k), H(x^k); t_k) = 0, \forall i \in \mathcal{I}_{GH}^{00}(x^k; t_k)$ for all $k \in \mathbb{N}$. Thus, we can write

$$-\nabla f(x^k) = \sum_{i=1}^p \lambda_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i(x^k) + \sum_{i=1}^q \eta_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^q \eta_i^{H,k} \nabla H_i(x^k),$$

where

$$\eta_i^{G,k} = \begin{cases} -\lambda_i^{\Phi,k} t_{2k} \theta'_{t_{1k}}(G_i(x^k)) G_i(x^k), & \text{if } i \in \mathcal{I}_{GH}^{+0}(x^k; t_k), \\ \lambda_i^{\Phi,k} F_i(x^k; t_k), & \text{if } i \in \mathcal{I}_{GH}^{0+}(x^k; t_k), \\ 0, & \text{otherwise,} \end{cases}$$

$$\eta_i^{H,k} = \begin{cases} \lambda_i^{H,k}, & \text{if } i \in \mathcal{I}_H(x^k; t_k), \\ -\lambda_i^{\Phi,k} G_i(x^k), & \text{if } i \in \mathcal{I}_{GH}^{+0}(x^k; t_k), \\ 0, & \text{otherwise.} \end{cases}$$

Noticing that whenever $i \in \{i = 1, \dots, q \mid G_i(x^k) = 0\}$, $i \in \mathcal{I}_{GH}^{0+}(x^k; t_k)$ or symmetrically $i \in \{i \mid F_i(x^k; t_k) = 0\}$ implies that $i \in \mathcal{I}_{GH}^{+0}(x^k; t_k)$ by concavity of θ and $t_{2k} \theta'(0) \leq t_{1k}$ for all k sufficiently large.

In order to prove boundedness of the sequence $\eta^k := \{\lambda^{g,k}, \lambda^{h,k}, \eta^{G,k}, \eta^{H,k}\}$, we check that all the assumptions of Theorem 4.4 are satisfied and apply Corollary 4.1.

Let $i \in \mathcal{I}^{-+} \cup \mathcal{I}^{-0} \cup \mathcal{I}^{+0}$, we verify that

$$\lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} = 0. \quad (7)$$

Obviously, it holds for k sufficiently large that $G_i(x^k) \neq 0$ and so either $i \in \mathcal{I}_H(x^k; t_k)$ or $i \in \mathcal{I}_{GH}^{+0}(x^k; t_k)$. In the former case, $\eta^{G,k} = 0$ and so (7) holds true. Thus, let us consider $i \in \mathcal{I}_{GH}^{+0}(x^k; t_k)$. In this case, it clearly holds that

$$\lim_{k \rightarrow \infty} \frac{|\eta_i^{G,k}|}{|\eta_i^{H,k}|} = \lim_{k \rightarrow \infty} \frac{\lambda_i^{\Phi,k} t_{2k} \theta'_{t_{1k}}(G_i(x^k)) G_i(x^k)}{\lambda_i^{\Phi,k} G_i(x^k)} \geq \frac{|\eta_i^{G,k}|}{\|\eta^k\|_\infty} = 0,$$

which proves (7).

It follows in a straightforward way from the definition of $\eta_i^{H,k}$ that $\eta^{H,k} \geq 0$ for all $i \in \mathcal{I}^{-0}$ and for all $i \in \mathcal{I}^{+}$

$$\lim_{k \rightarrow \infty} \frac{\eta_i^{H,k}}{\|\eta^k\|_\infty} = 0. \quad (8)$$

Finally, we verify that for all $i \in \mathcal{I}^{00}$ we get

$$\lim_{k \rightarrow \infty} \frac{\eta_i^{G,k} \eta_i^{H,k}}{\|\eta^k\|_\infty^2} = 0 \text{ and } \lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} \geq 0. \quad (9)$$

The case, where $\lambda_i^{\Phi,k} = 0$ for k sufficiently large holds trivially, thus assume that $i \in \mathcal{I}_{GH}^{+0}$ and so by definition of $\eta_i^{G,k}$ and $\eta_i^{H,k}$ we have

$$0 \leq \lim_{k \rightarrow \infty} \frac{\eta_i^{G,k} \eta_i^{H,k}}{\|\eta^k\|_\infty^2} = \lim_{k \rightarrow \infty} \frac{(\lambda^{\Phi,k} G_i(x^k))^2 t_{2k} \theta'_{t_{1k}}(G_i(x^k))}{\|\eta^k\|_\infty^2} \leq \lim_{k \rightarrow \infty} t_{2k} \theta'_{t_{1k}}(G_i(x^k)) = 0,$$

by the assumption that $t_{2k} = o(t_{1k})$. The second part of (9) is also a consequence of the previous equation since for $i \in \mathcal{I}_{GH}^{0+}$ it always holds and for $i \in \mathcal{I}_{GH}^{+0}$ it holds that $\lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} = 0$.

Thus, all the assumptions of Corollary 4.1 are verified and so x^* is an M-stationary point. \square

This theorem attains our goal defined in Theorem 4.1 to compute an M-stationary point of the (MPVC).

Removing the assumption $t_2 = o(t_1)$, made in 6, may lead the sequence of stationary points to spurious stationary points as illustrated by the following example.

Example 6.1. Consider the following two-dimensional example

$$\min_{x_1, x_2 \in \mathbb{R}^2} x_1 - x_2 \text{ s.t. } x_2 \geq 0, x_1 x_2 \leq 0,$$

and the relaxed sub-problems of the butterfly relaxation with $t_{2k} = 1/\theta'_{t_{1k}}(0)$ denoted as

$$\min_{x_1, x_2 \in \mathbb{R}^2} x_1 - x_2 \text{ s.t. } x_2 \geq -t_{2,k}, x_1(x_2 - t_{2,k} - t_{2,k} \theta'_{t_{1,k}}(x_1)) \leq 0.$$

The origin is a weak-stationary point with multipliers $\lambda^G = 1$ and $\lambda^H = 1$. Let $x^k = (1/k, t_{2,k} + t_{2,k} \theta'_{t_{1,k}}(1/k))$ (so $i \in \mathcal{I}_{GH}^{+0}(x^k; t_k)$) be a sequence of stationary point of the butterfly relaxed sub-problem with multipliers $\lambda^{G,k} = \lambda^{H,k} = 0$ and $\lambda^{\Phi,k} = k$. Then, $x^k \rightarrow_{k \rightarrow \infty} x^* = (0, 0)^T$ since $\lim_{k \rightarrow \infty} t_{2k} \theta'_{t_{1k}}(0) = 1$ for $t_{2k} = 1/\theta'_{t_{1k}}(0)$.

The following result motivated by Example 5.1 shows that the butterfly relaxation may improve its behavior in some specific cases. Example 5.1 also shows that this cannot be expected with the other relaxations.

Theorem 6.2. *Given a sequence $\{t_k\}$ of positive parameters satisfying (6) and decreasing to zero as k goes to infinity. Let $\{x^k, \lambda^{g,k}, \lambda^{h,k}, \lambda^{G,k}, \lambda^{H,k}, \lambda^{\Phi,k}\} \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^{3q}$ be a sequence of stationary points of (B_t) for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$ such that MPVC-LICQ holds at x^* . Then, x^* is an S-stationary point.*

Proof. Theorem 6.1 already proves that x^* is an M-stationary point. Assume by contradiction that x^* is not an S-stationary point.

We already mention in the proof of Theorem 6.1 that for all k it holds

$$-\nabla f = \sum_{i=1}^p \lambda_i^{g,k} \nabla g_i + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i + \sum_{i=1}^q \eta_i^{G,k} \nabla G_i - \sum_{i=1}^q \eta_i^{H,k} \nabla H_i,$$

where we omit the dependence in k in the expression of the gradients, since they are constant by linear/affine assumption. Now, by continuity linear independence of these gradients holds in a neighborhood of x^* . So, for k sufficiently large we get

$$\lambda^{g,k} = \lambda^{g,*}, \lambda^{h,k} = \lambda^{h,*}, \eta^{G,k} = \eta^{G,*}, \eta^{H,k} = \eta^{H,*}. \quad (10)$$

Let $i \in \mathcal{I}^{00} \cap \text{supp}(\lambda^{\Phi,*})$, where we remind that $\text{supp}(\lambda^{\Phi,*}) \subset \text{supp}(\lambda^{\Phi,k}) \subset \mathcal{I}_{GH}(x^k; t_k)$. If no such index exists, then for all k sufficiently large $\lambda^{\Phi,k}$ is zero and x^* is S-stationary.

Moreover, by definition of η^G and η^H , it holds that $\eta_i^{G,*} = 0$ for all i such that $G_i(x^*) = H_i(x^*) = 0$ since $F_i(x^k; t_k) \rightarrow_{\|t_k\| \downarrow 0} H(x^*)$, and $\eta^H \geq 0$ since $\lambda^{H,*} \geq 0$ and $G_i(x^*) = 0$. So, x^* is a S-stationary point. \square

7 Existence of Lagrange Multipliers of the Relaxed Sub-Problems

The convergence analysis relies on the fact that Lagrange multipliers exist for the relaxed sub-problems. In this part, we study some regularity properties of the relaxed non-linear programs. The butterfly relaxation satisfies Guignard CQ as stated in Theorem 7.1, which is equivalent in terms of regularity to the relaxation from [19].

Theorem 7.1. *Let $x^* \in \mathcal{Z}$ such that MPVC-LICQ holds at x^* . Then, there exists $t^* > 0$ and a neighborhood $U(x^*)$ of x^* such that:*

$$\forall t \in \mathcal{B}(t^*) \setminus \{0\} : x \in U(x^*) \cap \mathcal{X}_t^B \implies \text{standard GCQ for } (B_t) \text{ holds in } x.$$

Proof. First we note that it always holds that $\mathcal{L}_{\mathcal{X}_t^B}(x)^\circ \subseteq \mathcal{T}_{\mathcal{X}_t^B}(x)^\circ$. So, it is sufficient to show the reverse inclusion.

The linearized cone of R_t^B is given by

$$\begin{aligned} \mathcal{L}_{\mathcal{X}_t^B}(x) &= \{d \in \mathbb{R}^n \mid \nabla g_i(x)^T d \leq 0 \ (i \in \mathcal{I}_g(x)), \nabla h_i(x)^T d = 0 \ (i = 1, \dots, m), \\ &\quad \nabla H_i(x)^T d \geq 0 \ (i \in \mathcal{I}_H(x; t)), \\ &\quad \nabla \Phi_i^B(G(x), H(x); t)^T d \leq 0 \ (i \in \mathcal{I}_{GH}^{0+}(x; t) \cup \mathcal{I}_{GH}^{+0}(x; t))\}. \end{aligned}$$

Let us compute the polar of the tangent cone.

Consider the non-linear program $(\text{NLP}_{t,I}(x))$ with $I \subset \mathcal{I}_{GH}^{00}(x; t)$, defined by

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} f(x) \\ &\text{s.t. } g(x) \leq 0, \ h(x) = 0, \ H(x) \geq -t_2, \\ &\quad \Phi_i^B(G(x), H(x); t) \leq 0, \ i \notin \mathcal{I}_{GH}^{00}(x; t), \\ &\quad G_i(x) \leq 0, \ F_i(x; t) \geq 0, \ i \in I, \\ &\quad G_i(x) \geq 0, \ F_i(x; t) \leq 0, \ i \in I^c, \end{aligned} \quad (\text{NLP}_{t,I}(x))$$

where $I^c \cup I = \mathcal{I}_{GH}^{00}(x; t)$ and $I \cap I^c = \emptyset$. By construction of $U(x^*)$ and t^* , the gradients $\{\nabla g_i(x^*) (i \in \mathcal{I}_g), \nabla h_i(x^*) (i = 1, \dots, m), \nabla G_i(x^*) (i \in \mathcal{I}^{00} \cup \mathcal{I}^{0+}), \nabla H_i(x^*) (i \in \mathcal{I}^0)\}$ remain linearly independent for all $x \in U(x^*)$ by continuity of the gradients in a neighborhood and

$$\begin{aligned} \mathcal{I}_g(x) &\subseteq \mathcal{I}_g, \quad \mathcal{I}_H(x; t) \subseteq \mathcal{I}^0, \\ \mathcal{I}_{GH}^{00}(x; t) \cup \mathcal{I}_{GH}^{0+}(x; t) &\subseteq \mathcal{I}^{00} \cup \mathcal{I}^{0+}, \\ \mathcal{I}_{GH}^{00}(x; t) \cup \mathcal{I}_{GH}^{+0}(x; t) &\subseteq \mathcal{I}^{+0} \cup \mathcal{I}^{00}. \end{aligned} \tag{11}$$

Therefore, we can apply Lemma A.1 of [11] that gives that MFCQ holds for $(\text{NLP}_{t,I}(x))$ at x . Furthermore, by Lemma 8.10 of [29] and since MFCQ in particular implies Abadie CQ it follows

$$\mathcal{T}_{\mathcal{X}_t^B}(x) = \bigcup_{I \subseteq \mathcal{I}_{GH}^{00}(x; t)} \mathcal{T}_{\text{NLP}(t,I)}(x) = \bigcup_{I \subseteq \mathcal{I}_{GH}^{00}(x; t)} \mathcal{L}_{\text{NLP}(t,I)}(x).$$

By [7, Theorem 3.1.9], passing to the polar yields

$$\mathcal{T}_{\mathcal{X}_t^B}(x)^\circ = \bigcap_{I \subseteq \mathcal{I}_{GH}^{00}(x; t)} \mathcal{L}_{\text{NLP}(t,I)}(x)^\circ,$$

and by [7, Theorem 3.2.2]

$$\begin{aligned} \mathcal{L}_{\text{NLP}(t,I)}(x)^\circ &= \{v \in \mathbb{R}^n \mid v = \sum_{i \in \mathcal{I}_g(x)} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) - \sum_{i \in \mathcal{I}_H(x; t)} \lambda_i^H \nabla H_i(x) \\ &\quad + \sum_{i \in \mathcal{I}_{GH}^{+0}(x; t) \cup \mathcal{I}_{GH}^{0+}(x; t)} \lambda_i^\Phi \nabla \Phi_i^B(G(x), H(x); t) \\ &\quad + \sum_{i \in I} \lambda_i^G \nabla G_i(x) - \sum_{i \in I^c} \lambda_i^G \nabla G_i(x) \\ &\quad - \sum_{i \in I} \nabla H_i(x) + \sum_{i \in I^c} \nabla H_i(x) : (\lambda^g, \lambda^G, \lambda^H, \lambda^\Phi) \geq 0\}. \end{aligned}$$

Taking $v \in \mathcal{T}_{\mathcal{X}_t^B}(x)^\circ$ implies $v \in \mathcal{L}_{\text{NLP}(t,I)}(x)^\circ$ for all $I \subseteq \mathcal{I}_{GH}^{00}(x; t)$. If we fix such I , then there exists some multipliers λ^h and $(\lambda^g, \lambda^G, \lambda^H, \lambda^\Phi) \geq 0$ so that

$$\begin{aligned} v &= \sum_{i \in \mathcal{I}_g(x)} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) - \sum_{i \in \mathcal{I}_H(x; t)} \lambda_i^H \nabla H_i(x) \\ &\quad + \sum_{i \in \mathcal{I}_{GH}^{+0}(x; t) \cup \mathcal{I}_{GH}^{0+}(x; t)} \lambda_i^\Phi \nabla \Phi_i^B(G(x), H(x); t) \\ &\quad + \sum_{i \in I} \lambda_i^G \nabla G_i(x) - \sum_{i \in I^c} \lambda_i^G \nabla G_i(x) - \sum_{i \in I} \nabla H_i(x) + \sum_{i \in I^c} \nabla H_i(x). \end{aligned}$$

Now, it also holds that $v \in \mathcal{L}_{\text{NLP}(t, I^c)}(x)^\circ$ and so there exists some multipliers λ^h and $(\lambda^g, \lambda^G, \lambda^H, \lambda^\Phi) \geq 0$ such that

$$\begin{aligned} v &= \sum_{i \in \mathcal{I}_g(x)} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) - \sum_{i \in \mathcal{I}_H(x; t)} \lambda_i^H \nabla H_i(x) \\ &\quad + \sum_{i \in \mathcal{I}_{GH}^{+0}(x; t) \cup \mathcal{I}_{GH}^{0+}(x; t)} \lambda_i^\Phi \nabla \Phi_i^B(G(x), H(x); t) \\ &\quad + \sum_{i \in I} \lambda_i^G \nabla G_i(x) - \sum_{i \in I^c} \lambda_i^G \nabla G_i(x) - \sum_{i \in I} \nabla H_i(x) + \sum_{i \in I^c} \nabla H_i(x). \end{aligned}$$

By the construction of t^* and $U(x^*)$ the gradients involved here are linearly independent and so the multipliers in both previous equations must be equal. Thus, the multipliers λ_i^G and λ_i^H with indices i in $I \cup I^c$ vanish. Therefore, $v \in \mathcal{L}_{\mathcal{X}_t^B}(x)^\circ$ and as v has been chosen arbitrarily then $\mathcal{T}_{\mathcal{X}_t^B}(x)^\circ \subseteq \mathcal{L}_{\mathcal{X}_t^B}(x)^\circ$. The result follows since it always holds that $\mathcal{L}_{\mathcal{X}_t^B}(x)^\circ \subseteq \mathcal{T}_{\mathcal{X}_t^B}(x)^\circ$. \square

It is disappointing to require MPCC-LICQ to obtain only GCQ, but when \mathcal{I}_{GH}^{00} is empty we obtain the stronger LICQ.

Theorem 7.2. *Let $x^* \in \mathcal{Z}$, satisfying MPVC-LICQ. Then, there exists $t^* > 0$ and a neighborhood $U(x^*)$ of x^* such that*

$$\forall t \in \mathcal{B}(t^*) : x \in U(x^*) \cap \mathcal{X}_t^B \text{ and } \mathcal{I}_{GH}^{00}(x; t) = \emptyset \implies \text{LICQ holds at } x \text{ for } (B_t).$$

Proof. Let $x \in U(x^*) \cap \mathcal{X}_t^B$ and t sufficiently small. We prove that the gradients of the constraints involved in R_t^B are linearly independent, by verifying that the trivial solution is the only solution to the following equation

$$\begin{aligned} 0 = & \sum_{i \in \mathcal{I}_g(x)} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) - \sum_{i \in \mathcal{I}_H(x; \bar{t})} \lambda_i^H \nabla H_i(x) \\ & - \sum_{i \in \mathcal{I}_{GH}^{+0}(x; t)} \lambda_i^\Phi G_i(x) t_2 \theta'_{t_1}(G_i(x)) \nabla G_i(x) + \sum_{i \in \mathcal{I}_{GH}^{+0}(x; t)} \lambda_i^\Phi F_i(x; t) \nabla G_i(x) + \sum_{i \in \mathcal{I}_{GH}^{+0}(x; t)} \lambda_i^\Phi G_i(x) \nabla H_i(x). \end{aligned}$$

MPVC-LICQ and the inclusions (11) gives that the only solution is the trivial one. \square

8 Numerical Examples

In this section, we present two examples of numerical tests of the butterfly relaxation on MPVC that are motivated by bar-truss topology problems. The two examples have been used previously in the literature in [18] and are described in depth in the thesis [14]. We remind here that models for topology optimization was the motivation to introduce MPVC in [2]. Our motivation here is to validate our method on classical instances of MPVC, additionally to its strong theoretical properties.

8.1 Academic Example

Let us consider the following two dimensional MPVC from [18].

Example 8.1.

$$\begin{aligned} \min_{x \in \mathbb{R}^2} & 4x_1 + 2x_2 \\ \text{s.t.} & x_1 \geq 0, x_2 \geq 0, \\ & (5\sqrt{2} - x_1 - x_2)x_1 \leq 0, \\ & (5 - x_1 - x_2)x_2 \leq 0. \end{aligned}$$

The feasible set of this example is given in Figure 2. As the geometry indicates, numerical methods based on feasible descent concepts generally converge to the point $\hat{x} = (0, 5\sqrt{2})^T$. The unique global solution to the problem is the point $x^* = (0, 0)^T$. In practical application this point must be excluded by an additional constraint. Finally, the point $\bar{x} = (0, 5)^T$ is a local minimum.

We run butterfly relaxation tailored for (MPVC) on Example 8.1 using an initial point in the interior of the feasible domain $x^0 = (6, 6)^T$. Results are presented in Table 2 with solvers SNOPT and IPOPT to compute the iterates in the relaxation method, see Algorithm 1. In one case the butterfly method manages to converge to the global optimum and in the second case it converges to the point $(0, 5)$ which is a local minimum.

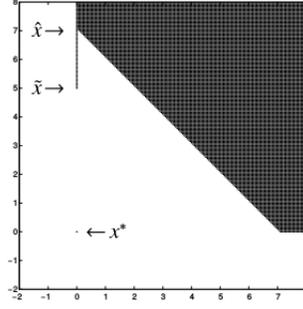


Figure 2: The feasible set of Example 8.1 from [14].

sovler	x^*	$f(x^*)$	last value of t_1
SNOPT	$(0, 5)^T$	10	0.5
IPOPT	$(0, 0)^T$	0	$4.67 \cdot 10^{-4}$

Table 2: Butterfly (MPVC) relaxation with $t_2 = t_1^{3/2}$ on Example 8.1 with initial point $(6, 6)^T$.

8.2 Application in bar-truss topology

In [18], they also consider a general model for bar-truss well known from engineering literature.

$$\begin{aligned}
 \min_{a \in \mathbb{R}^n, u \in \mathbb{R}^{L \times d}} \quad & \sum_{i=1}^n l_i a_i \\
 \text{s.t.} \quad & K(a)u = f_l, \quad l = 1, \dots, L, \\
 & f_l^T u_l \leq c, \quad l = 1, \dots, L, \\
 & a_i \leq \bar{a}, \quad i = 1, \dots, n, \\
 & a_i \geq 0, \quad i = 1, \dots, n, \\
 & (\sigma_{il}(a, u)^2 - \bar{\sigma}^2) a_i \leq 0, \quad i = 1, \dots, n, \quad l = 1, \dots, L,
 \end{aligned} \tag{12}$$

where $K(a) = \sum_{i=1}^N a_i \frac{E}{l_i} \gamma_i \gamma_i^T \in \mathbb{R}^{d \times d}$, $\sigma_{il}(a, u) = E \frac{\gamma_i^T u}{l_i}$ and $\|f\|_2 = 1$. All in all, this problem possesses $n := N + L \times d$ variables, $p := L \times d$ equality constraints, $m := L + N$ (ordinary) inequality constraints, and, formally, $N \times L$ vanishing (stress) constraints. Moreover, as a simplification in all problems below we use the setting $E := 1$ for the Young's modulus, which can be regarded as a scaling of the problem and is not essential.

Example 8.2 (Ten-bar truss problem). $L = 1, N = 10, d = 8$. The bar lengths are $l_i = 1$ for $i \in \{1, 3, 5, 6, 8, 10\}$ and $l_i = \sqrt{2}$ for $i \in \{2, 4, 7, 9\}$. Vectors γ_i are given in the following array

$i =$	1	2	3	4	5	6	7	8	9	10	
	1			$\frac{\sqrt{2}}{2}$		-1	$-\frac{\sqrt{2}}{2}$				u_{11}
		$\frac{\sqrt{2}}{2}$	1				-1	$-\frac{\sqrt{2}}{2}$			u_{21}
						1			$\frac{\sqrt{2}}{2}$		u_{31}
							$\frac{\sqrt{2}}{2}$	1			u_{41}
				$-\frac{\sqrt{2}}{2}$	-1		$-\frac{\sqrt{2}}{2}$				u_{51}
					1				$\frac{\sqrt{2}}{2}$		u_{61}
									$-\frac{\sqrt{2}}{2}$	-1	u_{71}
							$\frac{\sqrt{2}}{2}$			1	u_{81}

- [4] Roberto Andreani, Gabriel Haeser, María Laura Schuverdt, and Paulo J.S. Silva. Two new weak constraint qualifications and applications. *SIAM Journal on Optimization*, 22(3):1109–1135, 2012.
- [5] Roberto Andreani, José Mário Martinez, Alberto Ramos, and Paulo J.S. Silva. A cone-continuity constraint qualification and algorithmic consequences. *SIAM Journal on Optimization*, 26(1):96–110, 2016.
- [6] Alfred Auslender and Marc Teboulle. *Asymptotic cones and functions in optimization and variational inequalities*. Springer Science & Business Media, 2006.
- [7] Mokhtar S. Bazaraa and Chitharanjan Marakada Shetty. *Foundations of optimization*. Springer Science & Business Media, 2012.
- [8] Bertsekas, Dimitri P.. *Nonlinear programming*. Belmont: Athena scientific, 1999.
- [9] Frank H. Clarke. *Optimization and nonsmooth analysis*. Siam, 1990.
- [10] Dominik Dorsch, Vladimir Shikhman, and Oliver Stein. Mathematical programs with vanishing constraints: Critical point theory. *Journal of Global Optimization*, 52(3):591–605, 2012.
- [11] Jean-Pierre Dussault, Mounir Haddou, and Tangi Migot. The new butterfly relaxation methods for mathematical program with complementarity constraint. *Optimization-Online.org*, 2016.
- [12] F. J. Gould and Jon W. Tolle. A necessary and sufficient qualification for constrained optimization. *SIAM Journal on Applied Mathematics*, 20(2):164–172, mar 1971.
- [13] Mounir Haddou. A new class of smoothing methods for mathematical programs with equilibrium constraints. *Pacific Journal of Optimization*, 5(1), 2009.
- [14] Tim Hoheisel. *Mathematical programs with vanishing constraints*. PhD thesis, 2009.
- [15] Tim Hoheisel and Christian Kanzow. First-and second-order optimality conditions for mathematical programs with vanishing constraints. *Applications of Mathematics*, 52(6):495–514, dec 2007.
- [16] Tim Hoheisel and Christian Kanzow. Stationary conditions for mathematical programs with vanishing constraints using weak constraint qualifications. *Journal of Mathematical Analysis and Applications*, 337(1):292–310, 2008.
- [17] Tim Hoheisel and Christian Kanzow. On the Abadie and Guignard constraint qualifications for mathematical programs with vanishing constraints. *Optimization*, 58(4):431–448, 2009.
- [18] Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz. Convergence of a local regularization approach for mathematical program with complementarity or vanishing constraints. *Optimization Methods and Software*, 27(3):483–512, 2012.
- [19] Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz. Mathematical programs with vanishing constraints: a new regularization approach with strong convergence properties. *Optimization*, 61(6):619–636, 2012.
- [20] Tim Hoheisel, Christian Kanzow, and Jiri V. Outrata. Exact penalty results for mathematical programs with vanishing constraints. *Nonlinear Analysis, Theory, Methods and Applications*, 72(5):2514–2526, 2010.
- [21] Qingjie Hu, Jiguang Wang, Yu Chen, and Zhibin Zhu. On an ℓ_1 exact penalty result for mathematical programs with vanishing constraints. *Optimization Letters*, 2016.
- [22] Alexey F. Izmailov and A. L. Pogosyan. Active-set Newton methods for mathematical programs with vanishing constraints. *Computational Optimization and Applications*, 53(2):425–452, 2012.

- [23] Alexey F. Izmailov and Mikhail V. Solodov. Mathematical programs with vanishing constraints: Optimality conditions, sensitivity, and a relaxation method. *Journal of Optimization Theory and Applications*, 142(3):501–532, 2009.
- [24] Christian Kirches, Andreas Potschka, Hans Georg Bock, and Sebastian Sager. A parametric active set method for quadratic programs with vanishing constraints. *Technical Report*, 2012.
- [25] Jean-Claude Latombe. *Robot motion planning*. Kluwer Academic Publishers, Norwell, MA, 1991.
- [26] Tangi Migot, Jean-Pierre Dussault, Mounir Haddou, and Abdeslam Kadrani. How to compute an M-stationary point of the MPCC. *Optimization-Online.org*, 2017.
- [27] Leonid Minchenko and Alexander Tarakanov. On error bounds for quasinormal programs. *Journal of Optimization Theory and Applications*, 148(3):571–579, 2011.
- [28] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.
- [29] Alexandra Schwartz. *Mathematical programs with complementarity constraints: Theory, methods, and applications*. PhD thesis, Ph. D. dissertation, Institute of Applied Mathematics and Statistics, University of Würzburg, 2011.

A MPVC-CRSC and Error Bound for MPVC

An interesting application of the constraint qualifications is the existence of an error bound. That is, if close to a feasible point x it is possible to estimate the distance to the feasible set \mathcal{F} using a natural measure of infeasibility. We first give the definition in the context of a non-linear program.

Definition A.1. *We say that an error bound holds in a neighborhood $U(x^*)$ of a feasible point $x^* \in \mathcal{F}$ if there exists $\alpha > 0$ such that for every $y \in U(x^*)$*

$$\min_{z \in \mathcal{F}} \|z - y\| \leq \alpha \max\{\|h(y)\|_\infty, \|g^+(y)\|_\infty\},$$

whenever $\mathcal{F} := \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\}$.

In [4], the authors show that if CRSC holds at a point x and that g, h admit second derivatives, then an error bound holds in a neighborhood of x . It is to be noted that the same result may be obtained if \mathcal{F} satisfies quasi-normality at x as shown in [27].

We can now derive an enhanced MPVC-error bound result. Error bound and exact penalty results have been first derived in [20] in the context of MPVC, where the authors proved very general results under MPVC-GMFCQ, which correspond to classical MFCQ applied to $(MNLP_{A,B}(x^*))$. The following result shows that we can, indeed, weaken this assumption.

Theorem A.1. *Given $x^* \in \mathcal{Z}$ satisfying MPVC-CRSC. Then, there exists a neighborhood $U(x^*)$ and a constant $\alpha > 0$ such that for every $y \in U(x^*)$ the following holds*

$$\min_{z \in \mathcal{Z}} \|z - y\| \leq \alpha \max\{\|h(y)\|_\infty, \|g^+(y)\|_\infty, \|G_{\mathcal{T}^0 + \cup \mathcal{T}^0}^+(y)\|_\infty, \|H_{\mathcal{T}^+ + \cup \mathcal{T}^0}^+(y)\|_\infty, \|H^-(y)\|_\infty\}.$$

Proof. We remind that definition of MPVC-CRSC states that CRSC holds for all non-linear programs of the form $(MNLP_{A,B}(x^*))$, where $\mathcal{T}^{00} = A \cup B$. Thus, applying Theorem 5.1 of [4] yields that for A, B there exists a neighborhood $U_{A,B}$ and a constant α_{AB} such that

$$\min_{z \in \mathcal{Z}} \|z - y\| \leq \alpha_{AB} \max\{\|h(y)\|_\infty, \|g^+(y)\|_\infty, \|G_{\mathcal{T}^0 + \cup A}^+(y)\|_\infty, \|H_{\mathcal{T}^+ \cup B}^+(y)\|_\infty, \|H_{\mathcal{T}^- \cup 0}^-(y)\|_\infty\},$$

for all $y \in U_{A,B}(x^*)$. Taking $\alpha := \max_{A,B} \alpha_{AB}$, $U(x^*) = \cap_{A,B} U_{A,B}(x^*)$ we get for all $y \in U(x^*)$

$$\min_{z \in \mathcal{Z}} \|z - y\| \leq \alpha \max\{\|h(y)\|_\infty, \|g^+(y)\|_\infty, \|G_{\mathcal{I}^0+ \cup A}^+(y)\|_\infty, \|H_{\mathcal{I}^0+ \cup B}(y)\|_\infty, \|H_{\mathcal{I}^0-}(y)\|_\infty\}.$$

Finally, it holds that $A, B \subset \mathcal{I}^{00}$ so we have

$$\min_{z \in \mathcal{Z}} \|z - y\| \leq \alpha \max\{\|h(y)\|_\infty, \|g^+(y)\|_\infty, \|G_{\mathcal{I}^0+ \cup \mathcal{I}^{00}}^+(y)\|_\infty, \|H_{\mathcal{I}^0+ \cup \mathcal{I}^{00}}(y)\|_\infty, \|H_{\mathcal{I}^0-}(y)\|_\infty\}.$$

Now, using that $\|H^-(y)\|_\infty \geq \|H_{\mathcal{I}^0- \cup \mathcal{I}^0+ \cup \mathcal{I}^{00}}^-(y)\|_\infty$ yields to

$$\min_{z \in \mathcal{Z}} \|z - y\| \leq \alpha \max\{\|h(y)\|_\infty, \|g^+(y)\|_\infty, \|G_{\mathcal{I}^0+ \cup \mathcal{I}^{00}}^+(y)\|_\infty, \|H_{\mathcal{I}^0+ \cup \mathcal{I}^{00}}^+(y)\|_\infty, \|H^-(y)\|_\infty\}.$$

This concludes the proof. □

By Clarke's exact penalty principle [9], we get the following exact penalty result.

Corollary A.1. *Let $x^* \in \mathcal{Z}$ be a local optimum that satisfies MPVC-CRSC. Then, there exists a constant $\kappa > 0$ such that x^* is a local optimum of*

$$\min_{x \in \mathbb{R}^n} f(x) + \kappa \max\{\|h(x)\|_\infty, \|g^+(x)\|_\infty, \|G_{\mathcal{I}^0+ \cup \mathcal{I}^{00}}^+(x)\|_\infty, \|H_{\mathcal{I}^0+ \cup \mathcal{I}^{00}}^+(x)\|_\infty, \|H^-(x)\|_\infty\}.$$