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

Jean-Pierre Dussault, Mounir Haddou, Tangi Migot

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

Dussault, J.-P.* Haddou, M.† Migot, T.‡

February 5, 2018

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. We obtain the new MPVC-wGCQ and prove that it is the weakest constraint qualification for MPVC. We also introduce a new constraint qualification, MPVC-CRSC, that is sufficient to guarantee the convergence of the new method. Under this weak condition, we also provide an error bound and an exact penalty result for the MPVC.

Keywords: non-linear programming - MPCC - MPEC - MPVC - relaxation methods - butterfly relaxation - stationary point - constraint qualification - CRSC

AMS Subject Classification: 90C30, 90C33, 49M37, 65K05

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 $f:\mathbb{R}^n\to\mathbb{R}, g:\mathbb{R}^n\to\mathbb{R}^p, h:\mathbb{R}^n\to\mathbb{R}^m, G,H:\mathbb{R}^n\to\mathbb{R}^q$ 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, 13], but also for robots motion planning [23, 24]. 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 [13].

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

*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".

cannot be applied here due to the degeneracy of the vanishing constraints. It has been shown in [2] that most of the classical constraint qualifications used in non-linear programming does not hold here in a generic way. From a geometrical point of view, this can be explained by the fact that the feasible domain is in general non-convex (possibly non-connected see Example 3) and has an empty relative interior. Thus, enhanced stationary conditions and their corresponding constraint qualifications have been defined in the literature in [2, 9, 14, 15, 16, 19, 20, 22].

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, 17, 18, 22].

We mention here that other methods have also been proposed: a Newton with active-set strategy in [21], and a SQP method in [23] that consider sequences of quadratic programs with vanishing constraints. However, both methods require stronger assumptions that the ones used for regularization method in general and in particular in this paper.

The main aim of this paper focus on relaxation methods to solve the MPVC and in particular on methods with guarantees of convergence to an M-stationary point, which is the strongest property in the literature. A first step in this path has been proposed in [18] and we continue the discussion by introducing the new butterfly relaxation method.

One of the main advantages of the relaxation methods is that they require very weak constraint qualification to be well-defined. In this paper, we also focus on very weak assumptions to ensure convergence of the relaxation scheme. This analyze leads us to introduce new constraint qualifications with algorithmic applications: error bound, exact penalty, convergence of primal-dual sequences. We also introduce a new condition, which is proved to be the weakest MPVC constraint qualification.

In Section 2, we introduce classical definitions and results from non-linear programming. In Section 3, we consider optimality conditions for MPVCs and introduce new constraint qualifications, whose applications are studied in Section 4. In Section 5, we define extension of the butterfly relaxation. In Section 6 and 7, we prove theoretical results on convergence and existence of the multiplier of the relaxed sub-problems. We prove that the butterfly method has similar properties as the best methods in the literature. Finally, in section 8, we validate our approach on classical examples.

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)

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

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 or so-called KKT-points of (1). We call x^* a stationary point (or a KKT-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 $\lambda^g_i = 0$ 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

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

and linearized cone

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

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

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

where the polar cone of a cone K is given as $\{d \in \mathbb{R}^n \mid d^T y \leq 0,\ \forall y \in K\}$. It is also well-known since Gould and Tolle in [11], 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 $\mathcal{T}(x^*) = \mathscr{L}(x^*)$. 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 $\{\nabla g_i(x^*)\ (i \in \mathcal{I}_g(x^*)),\ \nabla h_i(x^*)\ (i = 1, ..., m)\}$ 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 such that $\|x - x^*\| \leq \delta$.

(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 $J_- := \{i \in \mathcal{I}_g(x^*) \mid -\nabla g_i(x^*) \in \mathscr{L}(x^*)^\circ\}$.

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

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

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 $\mathscr{L}(x^*)^\circ$. 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.

MPCC-LICQ $\Longrightarrow$ MPCC-MFCQ $\Longrightarrow$ MPCC-CRSC

Figure 1: Relations between the MPCC constraint qualifications.

3 MPVC: Definitions, Stationarity and Constraint Qualifications

This section defines the stationary conditions and their related 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

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

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^G and λ^H for indices $i \in \mathcal{I}^{00}$.

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

(a) W-stationary [22], 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

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

(b) T-stationary [9], if x^* is W-stationary and $\lambda_i^G \lambda_i^H \leq 0$ for all $i \in \mathcal{I}^{00}$.

(c) M-stationary [15], 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 $\Longrightarrow$ M-stationary $\Longrightarrow$ T-stationary $\Longrightarrow$ 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 [9] 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 $\mathcal{I}^{00}(x^*) = A \cup B$ such that x^* is a stationary point of the following non-linear program, denoted $MNLP_{A,B}(x^*)$,*

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

Based on this observation, we can derive constraint qualifications for MPVC simply 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 [17] 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 $\mathcal{I}^{00}(x^*) = A \cup B$, GCQ holds at x^* for $(MNLP_{A,B}(x^*))$.*

In the sequel, we show that this is a weaker constraint qualification than MPVC-GCQ, and in Section 4.1, we prove that it is the weakest constraint qualification for MPVC.

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 $\mathscr{P}_M$ be defined as*

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

Then, it holds true that

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

Proof. By straightforward computation, we can give an explicit form of the polar cone of $\mathscr{L}_{MNLP_{A,B}}$ for all A, B. Then, computing the union of this cones yields to

$$\begin{aligned} \cup_{A \in \mathcal{P}(\mathcal{I}^{00}), B=\mathcal{I}^{00} \setminus A} \mathscr{L}_{MNLP_{A,B}}(x^*)^\circ = \{ d \in \mathbb{R}^n \mid \exists (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \text{ such that } d = \mathcal{L}_{MPVC}(x^*, \lambda), \\ \text{with } \lambda_i^g \geq 0 \, (i \in \mathcal{I}_g), \\ \lambda_i^H = 0 \, (i \in \mathcal{I}^{\cdot +}), \\ \lambda_i^H \geq 0 \, (i \in \mathcal{I}^{-0}), \\ \lambda_i^G = 0 \, (i \in \mathcal{I}^{-+} \cup \mathcal{I}^{-0} \cup \mathcal{I}^{+0}), \\ \lambda_i^G \geq 0 \, (i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}), \\ \lambda_i^G \lambda_i^H = 0 \, (i \in \mathcal{I}^{00}) \}. \end{aligned} \tag{2}$$

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

MPVC-wGCQ is different from the MPVC-GCQ used for instance in [15] or in [16]. First, let us define the MPVC-linearized cone

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

Note that $\mathscr{L}_{MPVC}(x^*)$ is, in general, a non-convex cone and then, differs from $\mathscr{L}(x^*)$ applied to (MPVC). However, due to Corollary 2.5 of [16] it holds

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

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

These definitions have been introduced earlier in [16]. The following lemma gives a relation of $\mathscr{L}_{MPVC}(x^*)^\circ$ with the set $\mathscr{P}_M$ used in Corollary 3.1.

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

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

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

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

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

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

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

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

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

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

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

Now, computing the intersection in the equation above leads to

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

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

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

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

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

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

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 [16]. Under this condition, it has been proved in [2] that the MPVC Lagrange multipliers are unique. The MPVC-CPLD introduced in [17] is very popular in the context of relaxation methods to prove their convergence, see [17, 18]. 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 $\mathcal{I}^{00}(x^*) = A \cup B$, 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 [10] 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 CCP constraint qualification 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*

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

is outer semicontinuous at x^*, that is

$$\limsup_{x \to x^*} K_{MPCC}(x) \subset K_{MPCC}(x^*).$$

In this context, the outer limit is taken in the sense of Kuratowski-Painlevé corresponding to the Definition 5.4 given in [27]. By construction, it holds that MPVC-CRSC implies MPVC-CCP, which itself implies MPVC-wGCQ.

As we already remarked the new conditions defined here (MPCC-wGCQ, MPCC-CCP and MPCC-CRSC) are very 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 [17], MPVC-MFCQ defined in [2] and MPVC-GMFCQ first introduced in [19]. Besides in the following section, we show that these new weak conditions have a theoretical and numerical impact.

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.

4.1 MPVC-wGCQ as the Weakest Constraint Qualification for MPVC

The following theorem is a key result, since it defines in some sense the best necessary optimality condition that we can expect to compute with first order methods.

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, it holds that

$$-\nabla f(x^*) \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ \quad \Longrightarrow \quad -\nabla f(x^*) \in \cup \mathscr{P}_M, \tag{5}$$

$$\Longleftrightarrow \quad x^* \text{ M-stationary}, \tag{6}$$

where the last equivalence comes from Corollary 3.1. □

A similar result as been proved under MPVC-GCQ in [16]. 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.

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 [11] and Theorem 3.1.

Theorem 4.2. *MPVC-wGCQ is the weakest constraint qualification for MPVC.*

Before moving to the proof, let us introduce the definition of (g, h, G, H) MPVC-regular point that is an extension of a definition from [11].

Definition 4.1. *(g, h, G, H) is said MPVC-regular at x^*, if for all function f that admits a local constrained minimum at x^*, x^* is an M-stationary point.*

It is very important to notice here that a constraint qualification must be independent of the objective function, since it describes only the feasible set. Considering Corollary 3.1, we can prove the following result.

Lemma 4.1. *Assume that (g, h, G, H) is MPVC-regular at $x^* \in \mathcal{Z}$. Then, for all function f that admits a local constrained minimum at x^* it holds that $-\nabla f(x^*) \in \mathscr{P}_M(x^*)$.*

We also give an additional theorem from [11]. 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 "$\Longleftarrow$" part is given by Theorem 4.1.

So, let us consider the "$\Longrightarrow$" part. Assume that (g, h, G, H) is MPVC-regular at $x^* \in \mathcal{Z}$ and prove that for any $y \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ$, $y \in \mathscr{P}_M(x^*)$. By Theorem 4.3, for any $y \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ$ there exists a function f such that $y = -\nabla f(x^*)$.

Since we assume that (g, h, G, H) is MPVC-regular at $x^* \in \mathcal{Z}$ it follows by Lemma 4.1 that for all functions f such that $-\nabla f(x^*) \in \mathcal{T}_{\mathcal{Z}}(x^*)^\circ$ we have $-\nabla f(x^*) \in \mathscr{P}_M(x^*)$. □

4.2 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 4.2. *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 [26].

We can now derive a enhanced MPVC-error bound result. Error bound and exact penalty results have been first derived in [19] 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 4.4. *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*

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

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 $\mathcal{I}^{00}(x^*) = A \cup B$. Thus, applying Theorem 5.1 of [4] yields that for A, B there exists a neighborhood $U_{A,B}$ and a constant $\alpha_{A,B}$ such that

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

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

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

This concludes the proof. □

By Clarke's exact penalty principle [8], we get the following exact penalty result.

Corollary 4.1. *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*

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

4.3 Convergence of Primal-Dual Sequences under MPVC-CRSC and MPVC-CCP

Another application of these new constraint qualifications is the convergence of sequences that are classically obtained by relaxation methods.

Theorem 4.5. *Let x^* be in $\mathcal{Z}$ such that MPVC-CRSC holds at x^*. Given two sequences $\{x^k\}$,$\{\lambda^k\}$ such that $x^k \to x^*$ and that satisfies*

$$\nabla f(x^k) + \sum_{i=1}^{p} \lambda_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^{m} \lambda_i^{h,k} \nabla h_i(x^k) + \sum_{i=1}^{q} \lambda_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^{q} \lambda_i^{H,k} \nabla H_i(x^k) \to 0, \tag{7}$$

$$\lim_{k\to\infty} \frac{\lambda_i^g}{\|\lambda^k\|_\infty} = 0 \ \forall i \notin \mathcal{I}_g, \ \lim_{k\to\infty} \frac{\lambda_i^H}{\|\lambda^k\|_\infty} \geq 0 \ \forall i \in \mathcal{I}^{-0}, \tag{8}$$

$$\lim_{k\to\infty} \frac{\lambda_i^{G,k}}{\|\lambda^k\|_\infty} = 0 \ \forall i \in \mathcal{I}^{-+} \cup \mathcal{I}^{-0} \cup \mathcal{I}^{+0} \ \textit{and} \ \lim_{k\to\infty} \frac{\lambda_i^{H,k}}{\|\lambda^k\|_\infty} = 0 \ \forall i \in \mathcal{I}^{\cdot +}, \tag{9}$$

$$\lim_{k\to\infty} \frac{\lambda_i^{G,k}\lambda_i^{H,k}}{\|\lambda^k\|_\infty^2} = 0 \ \forall i \in \mathcal{I}^{00}, \tag{10}$$

$$\lim_{k\to\infty} \frac{\lambda_i^{G,k}}{\|\lambda^k\|_\infty} \geq 0 \ \forall i \in \mathcal{I}^{0+} \cup \mathcal{I}^{00}, \tag{11}$$

and the family of gradients of non-vanishing multipliers in (7) *are linearly independent. Then, the sequence* $\{\lambda^k\}$ *is bounded.*

According to Lemma 7.1 of [28] we may assume without loss of generality that the gradients corresponding to non-vanishing multipliers in equation (7) are linearly independent for all $k \in \mathbb{N}$ (note that this may change the multipliers, but a previously positive multiplier will stay at least non-negative and a vanishing multiplier will remain zero).

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

$$w^k := \sum_{i\in\mathcal{I}_g} \lambda_i^{g,k}\nabla g_i(x^k) + \sum_{i=1}^m \lambda_i^{h,k}\nabla h_i(x^k) + \sum_{i\in\mathcal{I}^{0+}\cup\mathcal{I}^{00}} \lambda_i^{G,k}\nabla G_i(x^k) - \sum_{i\in\mathcal{I}^{-0}\cup\mathcal{I}^{00}} \lambda_i^{H,k}\nabla H_i(x^k). \tag{12}$$

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

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

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 (7) to (11) give that $\lim_{k\to\infty} w^k = \lim_{k\to\infty} -\nabla f(x^k)/\|\lambda^k\|_\infty = 0$.

Dividing by $\|\lambda^k\|_\infty$ and passing to the limit in (12),

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

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 $i \in \mathcal{I}^{\cdot +}$ by assumptions (8) to (11).

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 $\mathrm{supp}(\bar{\lambda}) \subseteq \mathrm{supp}(\lambda^k)$. Consequently, we get a contradiction and thus the sequence $\{\lambda^k\}$ is bounded. □

A direct consequence of this theorem is that the sequence herein converge to an M-stationary point. Indeed, by assumptions of previous theorem and since $\lambda < \infty$, then any limit point of the sequence $\{x^k, \lambda^k\}$ satisfies conditions from Definition 3.1.

Corollary 4.2. *Let x^* be in $\mathcal{Z}$ such that MPVC-CRSC holds at x^*. Given two sequences $\{x^k\}$,$\{\lambda^k\}$ that satisfy the assumptions of Theorem 4.5. Then, x^* is an M-stationary point.*

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

Theorem 4.6. *Let $x^* \in \mathcal{Z}$ such that MPVC-CCP holds at x^*. Given two sequences $\{\delta^k\}$,$\{x^k\}$ such that $\delta^k \to 0$, $x^k \to x^*$ 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

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

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

The gap between Theorem 4.5 and Theorem 4.6 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 [10] 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 problem

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

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

$$\Phi^B_i(G(x), H(x); t) := G_i(x) F_i(x; t),$$

and

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

$\theta : \mathbb{R} \to]-\infty, 1]$ 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 $\lim_{t_1 \to 0} \theta_{t_1}(x) = 1 \ \forall x \in \mathbb{R}_{++}$ 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 $\theta^1_{t_1}(x) = \frac{x}{x+t_1}$ and $\theta^2_{t_1}(x) = 1 - \exp^{-\frac{x}{t_1}}$. These functions have already been used in the context of complementarity constraints in [10, 12, 25]. This new relaxation handles two parameters t_1 and t_2 chosen such that

$$t_2 = o(t_1). \tag{13}$$

A figure illustrating the feasible set of the (MPVC) and the relaxed feasible set is given in Figure 2. This method is an extension of the method introduced in [18]. A typical choice to satisfy (13) 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

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

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

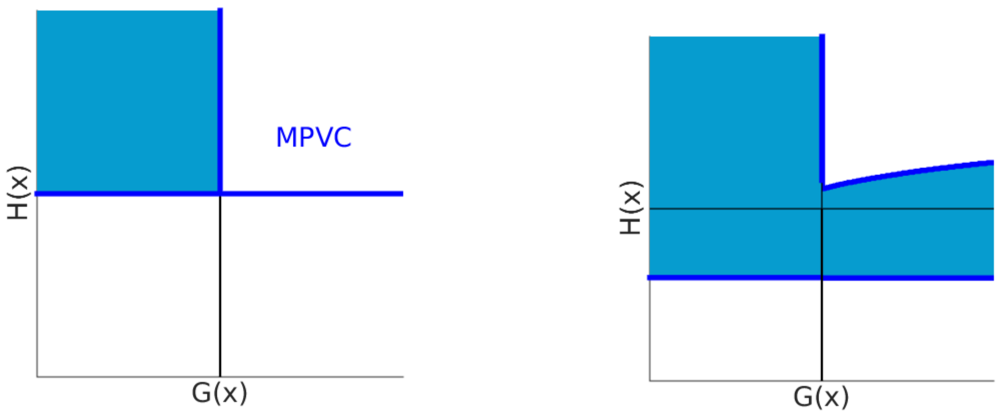

Figure 2: Feasible set of the MPVC and the butterfly relaxation for $\theta_{t_1}(z) = \frac{z}{z+t_1}$ with $t_2 = {t_1}^{3/2}$.

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

1. $\mathcal{X}_0^B = \mathcal{Z}$;
2. $\mathcal{X}_{t_a}^B \subset \mathcal{X}_{t_b}^B$ *for all* $0 < \frac{t_{2a}}{t_{1a}} < \frac{t_{2b}}{t_{1b}}$;
3. $\cap_{t \geq 0} \mathcal{X}_t^B = \mathcal{Z}$.

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 (Bu.) already belongs to $\mathcal{Z}$, then it is a local minimum of the (MPVC).

Finally, the following lemma from [10] gives more information on the asymptotic convergence of the θ's functions.

Lemma 5.2. *Given two sequences $\{t_{1k}\}$ and $\{t_{2k}\}$, which converge to 0 as k goes to infinity and $\forall k \in \mathbb{N}$, $(t_{1k}, t_{2k}) \in \mathbb{R}^2_{++}$. Then, for any $z \in \mathbb{R}_+$ it holds*

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

and

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

Since we assume that $t_{2k} = o(t_{1k})$ in (13), it implies that $t_{2k} \theta'_{t_{1k}}(z_k) \to 0$.

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

Data: x^0 an initial point, $(t_0, \bar{t}_0)$ initial parameters, $\sigma_t \in (0,1)$ parameters update;

Set $k := 0$, $(t_k, \bar{t}_k) := (t_0, \bar{t}_0)$;
repeat
  $(t_{k+1}, \bar{t}_{k+1}) = \sigma_t (t_k, \bar{t}_k)$;
  x^{k+1} := stationary point of (Bu.) with x^k initial point;
  $k := k + 1$;
until x^{k+1} *is a "stationary point of* (MPVC)*" or decision of infeasibility or unboundedness*;

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 ?

Both questions will be treated successfully in the following two sections.

We conclude this section by an example that shows that the butterfly relaxation may improve the relaxation [18], 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.

Example 1.

$$\min_{x\in\mathbb{R}^2} -x_2 \ s.t\ x_2 \leq 1, x_2 \geq 0,\ x_1x_2 \leq 0.$$

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 [18] 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* $\lambda^{\Phi,k} = k$*, may 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. The gradient of the Lagrangian for the relaxation equal to 0 gives

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\lambda^{H,k} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda^{\Phi,k} \begin{pmatrix} (x_2^k - t_2^k - t_2^k\theta_{t_{1,k}}(x_1^k - t_{2,k})) - t_2^k\theta_{t_{1,k}}(x_1^k - t_{2,k})x_1^k \\ x_1^k \end{pmatrix}.$$

Now, evaluating this expression at the point $x^k = (2t_{2,k}\ t_{2,k})^T$ *yields that* $\lambda^{\Phi,k} = 0$ *and* $\lambda^{H,k} = -1$*, which contradicts the sign constraints in the KKT condition.*

6 Convergence of the Butterfly Relaxation

As in [18], sequences of stationary points of (Bu.) 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 [18].

Theorem 6.1. *Given a sequence* $\{t_k\}$ *of positive parameters satisfying* (13) *and decreasing to zero as* k *goes to infinity. Let* $\{x^k, \lambda^k\}$ *be a sequence of points that are stationary points of* (Bu.) *for all* $k \in \mathbb{N}$ *with* $x^k \to x^*$*, 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 (Bu.). Let $\{x^k, \lambda^{g,k}, \lambda^{h,k}, \lambda^{G,k}, \lambda^{H,k}, \lambda^{\Phi,k}\}$ be a sequence of KKT points of (Bu.) 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 rewrite the equation above as

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

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 $k \in \mathbb{N}$.

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 4.5 are satisfied and apply Corollary 4.2.

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

$$\lim_{k\to\infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} = 0. \tag{14}$$

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 (14) holds true. Thus, let us consider $i \in \mathcal{I}_{GH}^{+0}(x^k;t_k)$. In this case, it clearly holds that

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

which proves (14).

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 $i \in \mathcal{I}^{\cdot+}$

$$\lim_{k\to\infty} \frac{\eta_i^{H,k}}{\|\eta^k\|_\infty} = 0. \tag{15}$$

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

$$\lim_{k\to\infty} \frac{\eta_i^{G,k}\eta_i^{H,k}}{\|\eta^k\|_\infty^2} = 0 \text{ and } \lim_{k\to\infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} \geq 0. \tag{16}$$

The case, where $\lambda^{\Phi,k} = 0$ 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

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

by the assumption that $t_{2k} = o(t_{1k})$. The second part of (16) 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 $\lim_{k\to\infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} = 0$.

Thus, all the assumptions of Corollary 4.2 are verified and so x^* is an M-stationary point. □

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 13, may lead the sequence of stationary points to spurious stationary points as illustrated by the following example.

Example 2. *Consider the following two-dimensional example*

$$\min_{x_1,x_2\in\mathbb{R}^2} x_1 - x_2 \text{ s.t. } x_2 \geq 0, x_1x_2 \leq 0,$$

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

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

The origin is a weak-stationary point with multipliers $\lambda^G = 1$ and $\lambda^H = 1$. Let $x^k = (1/k, t_k + t_{2k}\theta'_{t_{1k}}(1/k))$ (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, $x^k \to_{k\to\infty} x^* = (0,0)^T$ since $\lim_{k\to\infty} t_{2k}\theta'_{t_{1k}}(0) = 1$ for $t_{2k} = \theta'_{t_{1k}}(0)$.

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. The following theorem illustrates this assumption.

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:*

$$\forall t \in (0, \bar{t}] : \; x \in U(x^*) \cap \mathcal{X}_t^B \Longrightarrow \textit{standard GCQ for (Bu.) holds in } x.$$

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

The linearized cone of R_t^B is given by

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

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

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

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

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

Therefore, we can apply Lemma A.1 of [10] that gives that MFCQ holds for $(\text{NLP}_{t,I}(x))$ at x. Furthermore, by Lemma 8.10 of [28] and since MFCQ in particular implies Abadie CQ it follows

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

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

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

and by [7, Theorem 3.2.2]

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

Taking $v \in \mathcal{T}^\circ_{\mathcal{X}_t^B}(x)$ implies $v \in \mathscr{L}^\circ_{NLP(t,I)}(x)$ for all $I \subseteq \mathcal{I}^{00}_{GH}(x;t)$. 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

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

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

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

By the construction of $\bar{t}$ 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, $v \in \mathscr{L}^\circ_{\mathcal{X}_t^B}(x)$ and as v has been chosen arbitrarily then $\mathcal{T}^\circ_{\mathcal{X}_t^B}(x) \subseteq \mathscr{L}^\circ_{\mathcal{X}_t^B}(x)$.
The result follows since it always holds that $\mathscr{L}^\circ_{\mathcal{X}_t^B}(x) \subseteq \mathcal{T}^\circ_{\mathcal{X}_t^B}(x)$. □

This result gives a similar conclusion to the one from [18] and shows that, in general, there exist Lagrange multipliers of the relaxed sub-problem.

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 [17] and are described in depth in the thesis [13]. 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 [17].

Example 3.

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

The feasible set of this example is given in Figure 3. 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 3 using an initial point inside the feasible domain $x^0 = (6,6)^T$. Results are presented in Table 1 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.

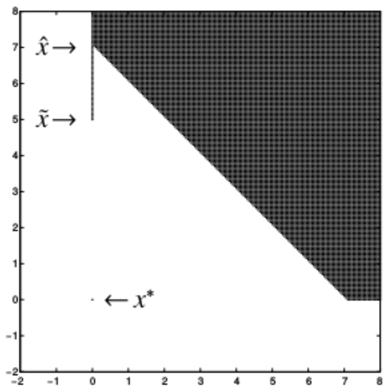

Figure 3: The feasible set of Example 3 from [13].

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

Table 1: Butterfly (MPVC) relaxation with $t_2 = {t_1}^{3/2}$ on Example 3 with initial point $(6,6)^T$.

8.2 Application in bar-truss topology

In [17], they also consider a general model for bar-truss well known from engineering literature.

$$\begin{aligned} \min_{a\in\mathbb{R}^n, u\in\mathbb{R}^{L\times d}} \quad & \sum_{i=1}^{n} l_i a_i \\ \text{s.t} \quad & K(a)u = f_l,\ l = 1,\dots,L \\ & f_l^T u_l \le c,\ l = 1,\dots,L \\ & a_i \le \bar{a}\ ,\ i = 1,...,n \\ & a_i \ge 0\ ,\ i = 1,...,n \\ & (\sigma_{il}(a,u)^2 - \bar{\sigma}^2)a_i \le 0\ ,\ i = 1,...,n,\ l = 1,...,L \end{aligned} \tag{17}$$

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 $\|f\| = 1$. 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 4 (Ten-bar truss problem)**.** $L = 1, N = 10, d = 8$. *The bar lengths are* $l_i = 1$ *for* $i \in \{1,3,5,6,8,10\}$

sovler	x^*	$f(x^*)$	last value of t
SNOPT	$(2.82, 0, 0.85, 2.00, 0.70, 0.70, 0.50, 0, 0.50, 1.24\epsilon_1)^T$	9.34	1
IPOPT	$(2.82, -7.63\epsilon_2, 0.85, 2.00, 0.70, 0.70, 0.50, -5.62\epsilon_2, 0.50, -5.76\epsilon_2)^T$	9.34	1

Table 2: Butterfly (MPVC) relaxation with $t_2 = {t_1}^{3/2}$ on Example 4 with initial point $a = u = 1$. $\epsilon_1 = 10^{-14}$, $\epsilon_2 = 10^{-9}$.

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}

All in all, problem possesses 18 variables, 8 bilinear equality constraints, 21 linear inequality constraints, and 10 non-linear inequality constraints modeling the vanishing stress constraints. Furthermore, we fix $c = 10$, $\bar{a} = 100$ *and* $\bar{\sigma} = 1$*. Moreover, we set* $f_1 = (1/\sqrt{8}, \dots, 1/\sqrt{8})$ *so that it satisfies* $\|f_1\|_2 = 1$*.*

Results are presented in Table 2 with solvers SNOPT and IPOPT to compute the iterates in the relaxation method and with initial point $a = u = 1$, $t_0 = 1, \sigma_t = 0.1$. In both cases, the method manages to find a solution.

9 Conclusion

This article proposes a new family of relaxation schemes for the mathematical program with vanishing constraints. This new method is an extension of existing method with the best-known convergence property, while weakening the required assumptions. We define new and weak constraint qualifications for the MPVC, all of them with direct algorithmic applications.

Further research will focus on several aspects of the new relaxation method. First of all, we plan a numerical comparison between relaxation methods on several examples. Then, convergence of the method needs to be analyzed, when the sequence of stationary points is only computed approximately. This potential issue and some keys to solve it have already been proposed in [25] in the context of complementarity constraints.

References

[1] Wolfgang Achtziger, Tim Hoheisel, and Christian Kanzow. A smoothing-regularization approach to mathematical programs with vanishing constraints. *Computational Optimization and Applications*, 55(3):733–767, 2013.

[2] Wolfgang Achtziger and Christian Kanzow. Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications. *Mathematical Programming*, 114(1):69–99, 2008.

[3] Wolfgang Achtziger, Christian Kanzow, and Tim Hoheisel. *On a relaxation method for mathematical programs with vanishing constraints*, volume 35. Inst. of Math., nov 2012.

[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] Frank H. Clarke. *Optimization and nonsmooth analysis*. Siam, 1990.

[9] Dominik Dorsch, Vladimir Shikhman, and Oliver Stein. Mathematical programs with vanishing constraints: Critical point theory. *Journal of Global Optimization*, 52(3):591–605, 2012.

[10] Jean-Pierre Dussault, Mounir Haddou, and Tangi Migot. The new butterfly relaxation methods for mathematical program with complementarity constraint. *Optimization-Online.org*, 2016.

[11] 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.

[12] Mounir Haddou. A new class of smoothing methods for mathematical programs with equilibrium constraints. *Pacific Journal of Optimization*, 5(1), 2009.

[13] Tim Hoheisel. *Mathematical programs with vanishing constraints*. PhD thesis, 2009.

[14] 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.

[15] 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.

[16] Tim Hoheisel and Christian Kanzow. On the Abadie and Guignard constraint qualifications for mathematical programs with vanishing constraints. *Optimization*, 58(4):431–448, 2009.

[17] Tim Hoheisel, Christian Kanzow, and Alexandra Schwartz. Convergence of a local regularization approach for mathematical programmes with complementarity or vanishing constraints. *Optimization Methods and Software*, 27(3):483–512, 2012.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] Christian Kirches, Andreas Potschka, Hans Georg Bock, and Sebastian Sager. A parametric active set method for quadratic programs with vanishing constraints. *Technical Report*, 2012.

[24] Jean-Claude Latombe. *Robot motion planning.* Kluwer Academic Publishers, Norwell, MA, 1991.

[25] Tangi Migot, Jean-Pierre Dussault, Mounir Haddou, and Abdeslam Kadrani. How to compute a local minimum of the MPCC. *Optimization-Online.org*, 2017.

[26] Leonid Minchenko and Alexander Tarakanov. On error bounds for quasinormal programs. *Journal of Optimization Theory and Applications*, 148(3):571–579, 2011.

[27] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.

[28] 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.