On approximate stationary points of the regularized mathematical program with complementarity constraints

Jean-Pierre Dussault, Mounir Haddou, Abdeslam Kadrani, Tangi Migot

JOTA manuscript No.
(will be inserted by the editor)

On Approximate Stationary Points of the Regularized Mathematical Program with Complementarity Constraints

Dussault, Jean-Pierre · Haddou, Mounir · Kadrani, Abdeslam · Migot, Tangi

This paper has been accepted for publication in *JOTA*, find the published version here: https://link.springer.com/article/10.1007/s10957-020-01706-w.

Abstract We discuss the convergence of regularization methods for the mathematical program with complementarity constraints with approximate sequence of stationary points. It is now well accepted in the literature that, under some tailored constraint qualification, the genuine necessary optimality condition for this problem is the M-stationarity condition. It has been pointed out*, in Kanzow, C., and Schwartz, A., Mathematics of Operations Research (2015),* that relaxation methods with approximate stationary points fail to ensure convergence to M-stationary points. We define a new strong approximate stationarity concept and we prove that a sequence of strong approximate stationary points always converges to an M-stationary solution. We also prove under weak assumptions the existence of strong approximate stationary points in the neighborhood of an M-stationary solution.

Keywords non-linear programming · MPCC · regularization methods · constraint qualification · optimization model with complementarity constraints

Mathematics Subject Classification (2000) 90C30 · 90C33 · 49M37 · 65K05

Dussault, Jean-Pierre
Département d'informatique, Université de Sherbrooke, Sherbrooke, Canada.
Jean-Pierre.Dussault@USherbrooke.ca

Haddou, Mounir
Univ Rennes, INSA Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France.
Mounir.Haddou@insa-rennes.fr

Kadrani, Abdeslam
INSEA, Laboratoire SI2M, Rabat-Instituts, Rabat, Maroc.
kadrania@gmail.com

Tangi Migot, Corresponding author
Department of Mathematics and Statistics, University of Guelph, Guelph, Canada
tangi.migot@gmail.com

1 Introduction

The Mathematical Program with Complementarity Constraints (MPCC) is a non-linear optimization problem including a complementarity problem in the constraints. It plays a very important role in many fields, such as engineering design, economic equilibria, transportation science, multilevel game, and mathematical programming itself. However, this kind of problems is generally difficult to deal with because it fails to satisfy the standard Mangasarian-Fromovitz constraint qualification (MFCQ) at any feasible point [1]. See the monographs [2,3] for details about the basic theory, effective algorithms, and various applications of the MPCC.

During the past two decades, many researchers introduced necessary optimality conditions such as the Clarke (C-), Mordukhovich (M-), strong (S-), and Bouligand (B-) stationarity conditions for the MPCC; see, e.g., [1,4–9]. Among these stationarities, the B-stationarity is known to be a good candidate for optimality, but since it is computationally difficult, it is rarely used in algorithmic analysis; the S-stationarity is the strongest and equivalent to the KKT conditions (see, e.g., [10,11]), but its interest is reduced since it does not always hold for the MPCC. The M-stationarity, which has already widely been investigated (see, e.g.,[1,4,5,7–9,12]), is the most relevant concept since it is the weakest necessary condition holding, under suitable constraint qualifications, at any local minimizer of the MPCC and is computationaly tractable.

The feasible set of the MPCC involves a complementarity constraint equivalent to G(x)=0 OR H(x) = 0. This is a *thin* set exhibiting some irregularity when G(x)=0 AND H(x)=0. It is this thinness that makes constraint qualifications fail at any feasible point. In view of the constraint qualifications issues that plague the MPCC the regularization methods provide an intuitive answer. The complementarity constraint is regularized using a parameter so that the new feasible domain is not thin anymore. Regularization methods drive this parameter to zero so that the feasible set of the regularized problem converges to the feasible set of (3) similar to an homotopy technique; see [13–18]. In the corresponding papers it has been shown that, under suitable conditions providing convergence of the methods, they still might converge to some spurious points, called C-stationary points, the convergence to M-stationary points being guaranteed only under some additional second-order conditions, strict complementarity-type condition or conditions on the sequence of iterates that cannot be guaranteed a priori. The only exception are the methods KDB and KS proposed respectively in [17] and [18]. In those references, the authors prove convergence of their regularization scheme to an M-stationary point under a classical constraint qualification. However, in [19], Kanzow and Schwartz discuss convergence of the methods considering a sequence of approximate stationary points, that is points that satisfy approximately the KKT conditions. They illustrate the fact that, for such sequences of approximate stationary points of the regularized subproblems, the method may converge to spurious weak-stationary points. Our motivation is to deal with this issue.

We focus in this paper on the KDB regularization method introduced in [17]. By defining a new kind of approximate stationary point called strong epsilon-stationary point we prove that any cluster point of a sequence of strong epsilon-stationary points of the KDB regularization is an M-stationary points. Moreover, we analyze existence of the strong epsilon-stationary points in the neighborhood of a solution. Previous studies of regularization methods require strong assumption to guarantee the existence of stationary point [17,18]. We show that strong epsilon-stationary points have a good behavior with this respect by proving the existence in the neighborhood of any M-stationary point without the need of constraint qualification.

The results presented in this paper are theoretical and present the advantages of strong epsilon-stationary points. In particular, we cover two important questions in the study of regularization methods for the MPCC namely the global convergence of strongly epsilon-stationary points to M-stationary solutions and weak assumptions allowing to prove the existence of such approximate stationary points. This study paves the way to further studies on an algorithm able to compute strong epsilon-stationary points.

The rest of the paper is organized as follows. In the next section, we sum up some stationarity concepts and constraint qualifications for the standard non-linear program and the MPCC. In Section 3, we recall the regularization method from [17] and motivate the difficulty arising when dealing with approximate stationary points of the regularized subproblems. We introduce our new strong epsilon-stationary point and prove convergence of the method to an M-stationary point in Section 4. In Section 5, after introducing the regularization method with slack variables, we prove our main existence result of strong epsilon-stationary point. Finally, we present some perspectives of this work in Section 6.

2 Preliminaries

2.1 Non-Linear Programming

Let a general non-linear program be

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

with $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^p$ and $f : \mathbb{R}^n \to \mathbb{R}$. Denote $\mathcal{F}$ the feasible region of (1), the set of active indices $\mathcal{I}_g(x) := \{i \in \{1, ..., m\} \;:\; g_i(x) = 0\}$ at x and the Lagrangian $\mathcal{L}(x, \lambda) := f(x) + g(x)^T\lambda^g + h(x)^T\lambda^h$ with $\lambda = (\lambda^g, \lambda^h)$. A fundamental tool to address such problems is a necessary optimality condition, the best known being the Karush-Kuhn-Tucker (KKT) condition. Let $x^* \in \mathcal{F}$ be a local minimum of (1) that satisfies a constraint qualification, then there exists $\lambda \in \mathbb{R}^m \times \mathbb{R}^p$ such that (x^*, λ) satisfy the equation $\nabla_x\mathcal{L}(x^*, \lambda) = 0$, with the conditions $\lambda^g \geq 0$, $\text{supp}(\lambda^g) := \{i : \lambda^g_i \neq 0\} \subseteq \mathcal{I}_g(x^*)$. A couple (x^*, λ) satisfying the above conditions is called a KKT point; it is a stationary point

of the Lagrangian function, its gradient vanishes. The stationary condition is often expressed as follows.

$$\begin{array}{l}\nabla_x \mathcal{L}(x^*, \lambda) = 0, \\ g_i(x^*) \leq 0,\ \lambda_i^g \geq 0,\ \lambda_i^g g_i(x^*) = 0,\ \forall i \in \{1, \ldots, m\}, \\ h_i(x^*) = 0,\ \forall i \in \{1, \ldots, p\}.\end{array} \tag{2}$$

Algorithms to solve (1) aim at computing a stationary point. In a practical context, it can be difficult, even impossible, to compute such stationary points, therefore approximate, or epsilon-stationary points are usually considered. We refer the reader for instance to [19] or [20] for a deeper motivation on this condition.

Definition 2.1 Let $\epsilon \geq 0$. We say that x^* is an ϵ-stationary point of (1), if there exists $\lambda \in \mathbb{R}^m \times \mathbb{R}^p$ satisfying

$$\begin{array}{l}\|\nabla_x \mathcal{L}(x^*, \lambda)\|_\infty \leq \epsilon, \\ g_i(x^*) \leq \epsilon,\ \lambda_i^g \geq 0,\ |\lambda_i^g g_i(x^*)| \leq \epsilon,\ \forall i \in \{1, \ldots, m\}, \\ |h_i(x^*)| \leq \epsilon,\ \forall i \in \{1, \ldots, p\}.\end{array}$$

At $\epsilon = 0$ we get the usual definition (2).

2.2 The Mathematical Program with Complementarity Constraints

We consider the Mathematical Program with Complementarity Constraints (MPCC)

$$\begin{array}{l}\min\limits_{x \in \mathbb{R}^n} f(x) \text{ s.t. } g(x) \leq 0,\ h(x) = 0, \\ \qquad\qquad 0 \leq G(x) \perp H(x) \geq 0,\end{array} \tag{3}$$

with $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^p$ and $G, H : \mathbb{R}^n \to \mathbb{R}^q$. All these functions are assumed to be continuously differentiable through this paper. The notation $0 \leq u \perp v \geq 0$ for two vectors u and v in $\mathbb{R}^q$ is a shortcut for $u \geq 0,\ v \geq 0$ and $u^T v = 0$. Let $\mathcal{Z}$ be the set of feasible points of (3). Given $x \in \mathcal{Z}$, we denote

$$\begin{array}{rl}\mathcal{I}^{+0}(x) := \{i \in \{1, \ldots, q\} & : \ G_i(x) > 0 \text{ and } H_i(x) = 0\}, \\ \mathcal{I}^{0+}(x) := \{i \in \{1, \ldots, q\} & : \ G_i(x) = 0 \text{ and } H_i(x) > 0\}, \\ \mathcal{I}^{00}(x) := \{i \in \{1, \ldots, q\} & : \ G_i(x) = 0 \text{ and } H_i(x) = 0\}.\end{array}$$

We define the MPCC-Lagrangian function of (3) as

$$\mathcal{L}_{MPCC}(x, \lambda) := f(x) + g(x)^T \lambda^g + h(x)^T \lambda^h - G(x)^T \lambda^G - H(x)^T \lambda^H,$$

with $\lambda := (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$. We introduce more stationary concepts as in [8,9, 7,12,6,21,2]. Those concepts are needed for two reasons:

- unless assuming a restrictive constraint qualification, a local minimizer x^* may fail to be a KKT point, so that (2) needs to be weakened in order to obtain a necessary condition;

– when analyzing cluster points of algorithms, other weak stationarity conditions appear naturally.

Definition 2.2 $x^* \in \mathcal{Z}$ is said

– W-stationary if there exists $\lambda \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q$ such that

$$\begin{array}{l} \nabla_x \mathcal{L}_{MPCC}(x^*, \lambda) = 0, \\ \lambda^g \geq 0, \lambda_i^g = 0, \forall i \notin \mathcal{I}_g(x^*), \\ \lambda_i^G = 0, \forall i \in \mathcal{I}^{+0}(x^*), \text{ and, } \lambda_i^H = 0, \forall i \in \mathcal{I}^{0+}(x^*); \end{array}$$

– C-stationary, if it is W-stationary and $\lambda_i^G \lambda_i^H \geq 0$ for all $i \in \mathcal{I}^{00}(x^*)$;
– M-stationary, if it is W-stationary and either $\lambda_i^G > 0, \lambda_i^H > 0$ or $\lambda_i^G \lambda_i^H = 0$ for all $i \in \mathcal{I}^{00}(x^*)$;
– S-stationary, if it is W-stationary and $\lambda_i^G \geq 0,\ \lambda_i^H \geq 0$ for all $i \in \mathcal{I}^{00}(x^*)$.

As pointed out in [10], strong stationarity is equivalent to the standard KKT conditions of an MPCC. In order to guarantee that a local minimum x^* of (1) is a stationary point in any of the previous senses, one needs to assume that a constraint qualification (CQ) is satisfied in x^*. Since most standard CQs are violated at any feasible points of (3), many MPCC-analogues of these CQs have been developed. Here, we mention only those needed later.

Definition 2.3 A feasible point x^* of (3), i.e. $x^* \in \mathcal{Z}$, is said to satisfy

1. MPCC-linear independence CQ (MPCC-LICQ), if the gradients

$$\begin{array}{ll} \nabla g_i(x^*)\ (i \in \mathcal{I}_g(x^*)), & \nabla h_i(x^*)\ (i = 1, \ldots, p), \\ \nabla G_i(x^*)\ (i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{0+}(x^*)), & \nabla H_i(x^*)\ (i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{+0}(x^*)), \end{array}$$

   are linearly independent.
2. MPCC-Mangasarian Fromovitz CQ (MPCC-MFCQ), if the only solution to the equation

$$\sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^{p} \lambda_i^h \nabla h_i(x^*) + \sum_{i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{0+}(x^*)} \lambda_i^G \nabla G_i(x^*) + \sum_{i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{+0}(x^*)} \lambda_i^H \nabla H_i(x^*) = 0$$

   with $\lambda_i^g \geq 0$ is the trivial solution, i.e. $\lambda^g = \lambda^h = \lambda^G = \lambda^H = 0$.

Note that each of these CQs implies that a local minimum is M-stationary, see [22,8], but only MPCC-LICQ is sufficient to guarantee strong stationarity of a local minimum, see [10,2,23]. The MPCC-LICQ is among the first MPCC-tailored constraint qualifications and may already be found in [2] and [6], the MPCC-MFCQ was introduced in [6] and presented in the form above in [24].

The following example due to Scheel and Scholtes [6] exhibits a situation where the global minimizer is not an S-stationary point but an M-stationary point. We will return to this example later on.

Example 2.1

$$\begin{aligned}\min_{x\in\mathbb{R}^3} x_1 + x_2 - x_3 \text{ s.t. } & -4x_1 + x_3 \leq 0,\\ & -4x_2 + x_3 \leq 0,\\ & 0 \leq x_1 \perp x_2 \geq 0.\end{aligned}$$

The global solution is $(0,0,0)^T$ *but is not an S-stationary point. Indeed, the gradient of the Lagrangian equal to zero yields*

$$0 = \begin{pmatrix}1\\1\\-1\end{pmatrix} + \lambda_1^g \begin{pmatrix}-4\\0\\1\end{pmatrix} + \lambda_2^g \begin{pmatrix}0\\-4\\1\end{pmatrix} - \lambda^G \begin{pmatrix}1\\0\\0\end{pmatrix} - \lambda^H \begin{pmatrix}0\\1\\0\end{pmatrix},$$

and since $\lambda_1^g + \lambda_2^g = 1$ *(third line), summing the first two lines yields* $2 - 4(\lambda_1^g + \lambda_2^g) - \lambda^G - \lambda^H = 0$ *and therefore* $\lambda^G + \lambda^H = -2$*; both cannot be non-negative.*

All in all, unless MPCC-LICQ is assumed, devising algorithms to provably reach stationary points (equivalent to S-stationary point) is not possible in general, and we must satisfy ourselves in devising algorithms reaching M-stationary points.

3 The KDB Regularization Method and Its Convergence

Any complementarity constraint represents a thin geometric locus troublesome with respect to CQ. A regularization introduces a parametric way to thicken these thin constraints and hopefully the regularized subproblem will satisfy a CQ.

Given a map $\Phi : \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}^q$ regularization methods consider the following parametric non-linear program $R_{t,\bar{t}}$ parametrized by two positive parameters t and $\bar{t}$:

$$\begin{aligned}\min_{x\in\mathbb{R}^n} f(x) \text{ s.t. } & g(x) \leq 0,\ h(x) = 0,\\ & G(x) \geq -\bar{t}\mathbb{1},\ H(x) \geq -\bar{t}\mathbb{1},\ \Phi(G(x), H(x); t) \leq 0.\end{aligned} \tag{4}$$

In [17], Kadrani, Dussault and Benchakroun introduced a method which enjoys the desired goal to converge to an M-stationary point without additional assumptions such as second-order conditions. Their method considers an approximation of the complementarity constraints defined for all $i \in \{1, \ldots, q\}$ as

$$\Phi_i(G(x), H(x); t) := (G_i(x) - t)(H_i(x) - t). \tag{5}$$

The only difference here compared to the initial presentation of the method in [17] is that $\bar{t}$ might be different from t, i.e. thicken asymetrically the complementarity constraints. In the sequel we use the following notations. The Lagrangian function of (4) is defined as

$$\begin{aligned}\mathcal{L}_{R_{t,\bar{t}}}(x,\nu) := f(x) + g(x)^T\nu^g + h(x)^T\nu^h - (G(x) + \bar{t}\mathbb{1})^T\nu^G\\ -(H(x) + \bar{t}\mathbb{1})^T\nu^H + \Phi(G(x), H(x); t)^T\nu^\Phi,\end{aligned}$$

with $\nu := (\nu^g, \nu^h, \nu^G, \nu^H, \nu^\Phi) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$.

Denote by $\mathcal{I}_\Phi(x;t) := \{i : \Phi_i(G(x), H(x); t) = 0\}$ the set of active indices of the regularization constraint.

The method has been shown to converge to an M-stationary point as recalled in the following theorem from [24] (the original convergence result from [17] being stated with MPCC-LICQ). The extension of [24, Theorem 3.5] to the case $\bar{t} \neq t$ is straightforward.

Theorem 3.1 *Let $\{t_k\} \downarrow 0$, $\{\bar{t}_k\} \downarrow 0$, and $\{x^k\}$ be a sequence of stationary points of (4). If $x^k \to x^*$ and* MPCC − MFCQ *holds in x^*, then x^* is an M-stationary point of (3).*

This result was a clear improvement over other existing methods such as [13–16], which only converge to a C-stationary point with similar assumptions. An additional necessary second order condition is often required to ensure convergence to an M-stationary point. Only the method developed in [18], extending the KDB method as a regularization, shares the same strong convergence property.

However, as shown in [19], the situation is different when considering sequence of approximate stationary points. We recall here the corresponding result, where the extension to the case $\bar{t} \neq t$ is again straightforward.

Theorem 3.2 (Theorem 9 and 10 [19]) *Let $\{t_k\} \downarrow 0$, $\{\bar{t}_k\} \downarrow 0$,$\{\epsilon_k\} \downarrow 0$, $\{x^k\}$ be a sequence of ϵ_k-stationary points of (4). If $x^k \to x^*$ and* MPCC − MFCQ *holds in x^*, then x^* is a W-stationary point of (3).*

Suppose further that $\epsilon_k = o(\bar{t}_k)$ and there is a constant $c > 0$ such that, for all $i \in \mathcal{I}^{00}(x^*)$ and all k sufficiently large, the iterates $(G_i(x^k), H_i(x^k))$ satisfy

$$(G_i(x^k), H_i(x^k)) \notin \ ](1-c)t_k, (1+c)t_k[^2 \ \setminus \ ](1-c)t_k, t_k[^2. \tag{6}$$

Then x^* is an M-stationary point.

Note that the authors also gave similar type of conditions to get convergence to a C- or an S-stationary point. The drawback of condition (6), however, is that it cannot be checked a priori and the strong convergence result of the method no longer holds.

In [18, Example 3], the authors provide an example satisfying MPCC-LICQ and exhibiting convergence to a W-stationary point. One of the main reason for this behavior being that if we perturb the relation $|\nu_i^\Phi \Phi_i(x_1, x_2; t)| \leq \epsilon$ (leaving the other conditions $\nu^\Phi \geq 0$, $\Phi(x_1, x_2; t) \leq 0$), ν^Φ may be positive when the constraint $\Phi(x_1, x_2; t)$ is not active.

Now, if the complementarity constraint is satisfied approximately, but an approximate complementarity condition of the form $\nu_i^\Phi(\Phi_i(x_1, x_2; t) - \epsilon) = 0$ is guaranteed, convergence may occur to C-stationary points as shown in the following example.

Example 3.1 *Consider the problem*

$$\min_{x \in \mathbb{R}^2} \tfrac{1}{2}((x_1 - 1)^2 + (x_2 - 1)^2) \text{ s.t. } 0 \leq x_1 \perp x_2 \geq 0.$$

We specialize the relations in Definition 2.1 for (4) *as the following*

$$\begin{array}{l}
\left\| \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} - \nu^G \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \nu^H \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \nu^\Phi \begin{pmatrix} x_2 - t \\ x_1 - t \end{pmatrix} \right\|_\infty \leq \epsilon, \\
0 \leq \nu^G, \qquad (x_1 + \bar{t}) \geq 0, \qquad \nu^G (x_1 + \bar{t}) \leq \epsilon, \\
0 \leq \nu^H, \qquad (x_2 + \bar{t}) \geq 0, \qquad \nu^H (x_2 + \bar{t}) \leq \epsilon, \\
0 \leq \nu^\Phi, \qquad (x_1 - t)(x_2 - t) \leq \epsilon, \qquad \nu^\Phi \left[(x_1 - t)(x_2 - t) - \epsilon \right] = 0.
\end{array}$$

The points $(t + \sqrt{\epsilon}, t + \sqrt{\epsilon})^T$ *together with* $\nu^G = \nu^H = 0$ *and* $\nu^\Phi = \frac{1 - t - \sqrt{\epsilon}}{\sqrt{\epsilon}}$ *satisfy the above relations. The limit point when* $t, \epsilon \longrightarrow 0$ *is the origin, which is a C-stationary point with* $\lambda^G = \lambda^H = -1$.

Epsilon-stationary points have two main drawbacks considering our goal of computing an M-stationary point. The difficulties may come from the approximation of the complementarity condition as shown by [18, Example 3] or from the approximation of the feasibility of the regularized constraint as illustrated in Example 3.1.

We focus in this paper on studying approximate stationary points. Note that an alternative but complementary approach would be to study sequential optimality conditions. This approach first designed for nonlinear programming [20] has recently been extended to the MPCC in [25] or in [26].

4 Convergence of Strong Epsilon-Stationary Sequences

We now address the convergence of sequences of epsilon–stationary points. The previous section motivates the definition of a new kind of epsilon-stationary point called strong epsilon-stationary point, which is more stringent regarding the complementarity constraint.

Definition 4.1 Let $\epsilon \geq 0$. We say that x is a **strong ϵ-stationary point** for (4) if there exists $\nu \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$ such that

$$\begin{array}{l}
\left\| \nabla \mathcal{L}_{R_{t,\bar{t}}}(x, \nu) \right\|_\infty \leq \epsilon, \\
g_i(x) \leq \epsilon,\ \nu_i^g \geq 0,\ |g_i(x) \nu_i^g| \leq \epsilon,\ \forall i \in \{1, \ldots, m\}, \\
|h_i(x)| \leq \bar{t} + 3\epsilon,\ \forall i \in \{1, \ldots, p\}, \\
G_i(x) + \bar{t} \geq -\epsilon,\ \nu_i^G \geq 0,\ \left| \nu_i^G (G_i(x) + \bar{t}) \right| \leq \epsilon,\ \forall i \in \{1, \ldots, q\}, \\
H_i(x) + \bar{t} \geq -\epsilon,\ \nu_i^H \geq 0,\ \left| \nu_i^H (H_i(x) + \bar{t}) \right| \leq \epsilon,\ \forall i \in \{1, \ldots, q\}, \\
\Phi_i(G(x), H(x); t) \leq \mathbf{0},\ \nu_i^\Phi \geq 0,\ \left| \nu_i^\Phi \Phi_i(G(x), H(x); t) \right| = \mathbf{0},\ \forall i \in \{1, \ldots, q\}.
\end{array}$$

Note that we are less demanding on the equality constraint. This will become clearer when considering slack variables to the MPCC to get our result on existence of strong epsilon-stationary point in Section 5.

Since a strong epsilon-stationary point is, in particular, an epsilon-stationary point, we can deduce from the first part of Theorem 3.2 the following result.

Lemma 4.1 *Let* $\{t_k\} \downarrow 0$, $\{\bar{t}_k\} \downarrow 0$,$\{\epsilon_k\} \downarrow 0$, $\{x^k\}$ *be a sequence of strong* ϵ_k*-stationary points of* (4). *If* $x^k \to x^*$ *and* MPCC − MFCQ *holds in* x^*, *then* x^* *is a W-stationary point of* (3).

We can now move to our main convergence result.

Theorem 4.1 *Let* $\{t_k\} \downarrow 0$, $\{\bar{t}_k\} \downarrow 0$, $\epsilon_k = o(\bar{t}_k)$, *and* $\{x^k\}$ *be a sequence of strong* ϵ_k*-stationary points of* (4). *If* $x^k \to x^*$ *and* MPCC − MFCQ *holds in* x^*, *then* x^* *is an M-stationary point of* (3).

Proof The proof of [19, Theorem 9] shows that the sequence $\{\eta^{g,k}, \eta^{h,k}, \eta^{G,k}, \eta^{H,k}\}$ is bounded with $\eta^{g,k} := \nu^{g,k}$, $\eta^{h,k} := \nu^{h,k}$ and for all i

$$\begin{aligned} \eta_i^{G,k} &:= \nu_i^{G,k} - \nu_i^{\Phi,k}(H_i(x^k) - t_k), \\ \eta_i^{H,k} &:= \nu_i^{H,k} - \nu_i^{\Phi,k}(G_i(x^k) - t_k). \end{aligned} \tag{7}$$

Up to a subsequence, the limit point of this sequence is a mutliplier, denoted $(\eta^{g,*}, \eta^{h,*}, \eta^{G,*}, \eta^{H,*})$, such that x^* is a W-stationary point. We now consider indices $i \in \mathcal{I}^{00}(x^*)$. Our aim here is to prove that either $\eta_i^{G,*} > 0, \eta_i^{H,*} > 0$ or $\eta_i^{G,*}\eta_i^{H,*} = 0$ by showing that $\eta_i^{G,*} < 0 \Longrightarrow \eta_i^{H,*} = 0$.

Assume that $\eta_i^{G,*} < 0$, then for k sufficiently large it holds that

$$\eta_i^{G,k} < 0 \Longleftrightarrow \nu_i^{G,k} - \nu_i^{\Phi,k}(H_i(x^k) - t_k) < 0 \Longrightarrow \nu_i^{\Phi,k} > 0 \text{ and } H_i(x^k) > t_k,$$

since $\nu^{G,k}$ and $\nu^{\Phi,k}$ are non-negative. Using the complementarity condition $\left|\nu_i^{\Phi,k}(G_i(x^k) - t_k)(H_i(x^k) - t_k)\right| = 0$ it follows that $G_i(x^k) = t_k$. Moreover, using that $H_i(x^k) > t_k$, we obtain

$$\epsilon_k \geq \left|\nu_i^{H,k}(H_i(x^k) + \bar{t}_k)\right| \geq \left|\nu_i^{H,k}\bar{t}_k\right|,$$

therefore $\nu_i^{H,k} \downarrow 0$ as $\epsilon_k = o(\bar{t}_k)$. Finally, since $G_i(x^k) = t_k$ and $\nu_i^{H,k} \downarrow 0$ it holds that $\eta_i^{H,*} = 0$. □

Further work may focus on weakening the constraint qualification necessary to prove this convergence theorem. In the literature, MPCC-RCPLD [27] or MPCC-CCP [26] have been used to prove convergence results involving M-stationary points.

5 Existence of Strong Epsilon-Stationary Points for the Relaxation with Slack Variables

5.1 On Lagrange Multipliers of the Relaxation

Beyond the convergence of a sequence of (approximate) stationary points of the regularization, another important question is wether such points exist in a neighborhood of an M-stationary point.

Following Example 2.1, we can encounter a situation where the regularized subproblems have no stationary point.

Example 5.1 *The* KDB *regularized problem for Example 2.1 is*

$$\min_{x\in\mathbb{R}^3} x_1 + x_2 - x_3 \text{ s.t. } -4x_1 + x_3 \leq 0, -4x_2 + x_3 \leq 0,$$
$$x_1 \geq -\bar{t},\ x_2 \geq -\bar{t}, (x_1 - t)(x_2 - t) \leq 0.$$

The point $(t, t, 4t)^T$ *is feasible so that the minimum value of this program is* $\leq -2t$*. Moreover, whenever* $x_1 > t$*, we must have* $x_2 \leq t$ *to satisfy* $(x_1 - t)(x_2 - t) \leq 0$*. This allows to conclude that* $(t, t, 4t)^T$ *is the global minimum of the regularized problem.* $\nu^G = \nu^H = 0$ *and the gradient of the Lagrangian equal to zero yields*

$$0 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \nu_1^g \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} + \nu_2^g \begin{pmatrix} 0 \\ -4 \\ 1 \end{pmatrix} + \nu^\Phi \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

which cannot be satisfied.

This last example seems to contradict Theorem 4.6 in [18], but MPCC-LICQ is not satisfied at the origin.

It has been pointed out earlier that a practical algorithm may not be able to compute stationary points of the regularized subproblem, but only some approximate stationary point. In the previous section, in particular, we underline the interest for strong epsilon-stationary points. An intuitive idea would be that a weaker constraint qualification (like MPCC-MFCQ) would guarantee existence of such points in the neighborhood of an M-stationary point. However, the following one-dimensional example shows that things are not that simple.

Example 5.2 *Consider the one-dimensional* MPCC

$$\min_{x\in\mathbb{R}} -x \text{ s.t. } 0 \leq x \perp x \geq 0.$$

The constraint set involves only linear functions, and the objective is linear. There is only one feasible point, $x^* = 0$*. Moreover, it is not hard to verify that this point is an M-stationary point.*

Given $1 > \epsilon > 0$*, let us now verify that there is no strong* ϵ*-stationary point for this problem. The condition on the gradient of the Lagrangian yields*

$$| - 1 - \nu^G - \nu^H + 2\nu^\Phi(x - t)| \leq \epsilon, \tag{8}$$

since in this case $\nabla\Phi(G(x), H(x); t) = 2(x - t)$*. Since* x *must be feasible, we get* $(x - t)^2 \leq 0$*, therefore* $x = t$*. Hence, there is no* x *satisfying* (8) *as* ν^G *and* ν^H *are non-negative.*

The previous example shows that existence of a strong epsilon-stationary point is not ensured in the neighborhood of an M-stationary point even so a constraint qualification holds. However, by reformulating (3) with slack variables things could be slightly different as shown by the continuation of the example.

Example 5.3 *(Example 5.2 continued) Given $\epsilon > 0$ and $t \leq \bar{t}$. We now verify that there is a strong epsilon-stationary point of the approximation* KDB *written with slack variables for this example,*

$$\min_{x\in\mathbb{R}} -x \text{ s.t. } s_G = x,\ s_H = x\ , s_G \geq -\bar{t},\ s_H \geq -\bar{t},\ (s_G - t)(s_H - t) \leq 0.$$

Let δ be a constant satisfying $2\epsilon \geq \delta > 0$. Consider the point $x = 0$ and $(s_G, s_H) = (t, t+\delta)$ with Lagrange multiplier $(\nu^{s_G}, \nu^{s_H}, \nu^G, \nu^H, \nu^\Phi) = (-1, 0, 0, 0, \frac{1}{\delta})$. Let us verify that this point is a strong epsilon-stationary point of the regularization:

- *Condition on the gradient of the Lagrangian:*

$$\begin{aligned}
|\nabla_x \mathcal{L}(x, s, \nu)| &= |-1 - \nu^{s_G} - \nu^{s_H}| = 0,\\
|\nabla_{s_G} \mathcal{L}(x, s, \nu)| &= |\nu^{s_G} - \nu^G + \nu^\Phi(s_H - t)| = 0,\\
|\nabla_{s_H} \mathcal{L}(x, s, \nu)| &= |\nu^{s_H} - \nu^H + \nu^\Phi(s_G - t)| = 0.
\end{aligned}$$

- *Condition on the feasibility: $|x - s_G| = t \leq \bar{t} + 2\epsilon$, $|x - s_H| = t + \delta \leq \bar{t} + 2\epsilon$, $s_G + \bar{t} = t + \bar{t} \geq -\epsilon$, $s_H + \bar{t} = t + \bar{t} + \delta \geq -\epsilon$, and $(s_G - t)(s_H - t) = 0$.*
- *Condition on the complementarity: $|(s_G + t)\nu^G| = 0$, $|(s_H + t)\nu^H| = 0$, and $|(s_G - t)(s_H - t)\nu^\Phi| = 0$.*

This example shows that the formulation with slack variables has a strong epsilon-stationary point.

5.2 The Regularization with Slack Variables

Consider the following MPCC reformulated with slacks variables:

$$\begin{aligned}
\min_{(x,s)\in\mathbb{R}^n\times\mathbb{R}^{2q}} f(x) \text{ s.t. } & g(x) \leq 0,\ h(x) = 0,\\
& s_G = G(x),\ s_H = H(x),\\
& 0 \leq s_G \perp s_H \geq 0.
\end{aligned} \tag{9}$$

Consider now the following parametric non-linear program with $t, \bar{t} \geq 0$:

$$\begin{aligned}
\min_{(x,s)\in\mathbb{R}^n\times\mathbb{R}^{2q}} f(x) \text{ s.t. } & g(x) \leq 0,\ h(x) = 0,\\
& s_G = G(x),\ s_H = H(x),\\
& s_G \geq -\bar{t}\mathbb{1},\ s_H \geq -\bar{t}\mathbb{1},\ \Phi(s_G, s_H; t) \leq 0.
\end{aligned} \tag{10}$$

The Lagrangian function of (10) is defined as

$$\begin{aligned}
\mathcal{L}_{R^s_{t,\bar{t}}}(x, s, \nu) := f(x) + g(x)^T\nu^g + h(x)^T\nu^h - (G(x) - s_G)^T\nu^{s_G} - (H(x) - s_H)^T\nu^{s_H}\\
-(s_G + \bar{t}\mathbb{1})^T\nu^G - (s_H + \bar{t}\mathbb{1})^T\nu^H + \Phi(s_G, s_H; t)^T\nu^\Phi.
\end{aligned}$$

The following result is a direct corollary of Theorem 4.1 stating that the reformulation with slack variables does not alter the convergence result.

Corollary 5.1 *Let* $\{t_k\} \downarrow 0$, $\{\bar{t}_k\} \downarrow 0$, $\epsilon_k = o(\bar{t}_k)$, $\{x^k, s^k\}$ *be a sequence of strong epsilon-stationary points of* (10). *If* $x^k \to x^*$ *and* MPCC − MFCQ *holds at* x^*, *then* x^* *is an M-stationary point of* (3).

Proof Let $\tilde{h}(x) : \mathbb{R}^n \to \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q$ be such that $\tilde{h}(x) := (h(x), s_G - G(x), s_H - H(x))$ and $\tilde{x} := (x, s_G, s_H)$. It is clear that the non-linear program (10) fall under the formulation (4). Moreover, MPCC-MFCQ for (9) is equivalent to MPCC-MFCQ applied to (3) with slack variables. Therefore, we can apply Theorem 4.1 to conclude this proof. □

The following lemma giving an explicit form of the gradient of the Lagrangian function of (10) can be deduced through direct computations.

Lemma 5.1 *The gradient of* $\mathcal{L}_{R^s_{t,\bar{t}}}(x, s, \nu)$ *is given by*

$$\begin{aligned}
\nabla_x \mathcal{L}_{R^s_{t,\bar{t}}}(x, s, \nu) &= \nabla \mathcal{L}_{MPCC}(x, \nu^g, \nu^h, \nu^{s_G}, \nu^{s_H}), \\
\nabla_{s_G} \mathcal{L}_{R^s_{t,\bar{t}}}(x, s, \nu) &= \nu^{s_G} - \nu^G + \nabla_{s_G} \Phi(s_G, s_H; t)^T \nu^\Phi, \\
\nabla_{s_H} \mathcal{L}_{R^s_{t,\bar{t}}}(x, s, \nu) &= \nu^{s_H} - \nu^H + \nabla_{s_H} \Phi(s_G, s_H; t)^T \nu^\Phi.
\end{aligned}$$

There are two direct consequences of this result. First, it is easy to see from this lemma that computing a stationary point of $\mathcal{L}_{R^s_{t,\bar{t}}}(x, s, \nu)$ is equivalent to computing a stationary point of $\mathcal{L}_{R_{t,\bar{t}}}(x, \nu)$. Secondly, a stationary point of $\mathcal{L}_{R^s_{t,\bar{t}}}(x, s, \nu)$ satisfies $\nabla \mathcal{L}_{MPCC}(x, \nu^g, \nu^h, \nu^{s_G}, \nu^{s_H}) = 0$, which is one of the conditions of W-stationarity for (3).

5.3 Existence

The following lemma proves that given any M-stationary point x^* there is a neighborhood such that for all x in this neighborhood there exists s such that (x, s) is a strong epsilon-stationary point for the regularized problem.

Lemma 5.2 *Assume that* f, g, h, G, H *have Lipschitz continuous gradients. Let* x^* *be an M-stationary point of* (3) *and* $\epsilon > 0$ *be small enough. Then, there exist positive constants* $c, \delta, \bar{t}^* \geq c\epsilon, t^*$ *such that, for all* $\bar{t} \in]c\epsilon, \bar{t}^*[$ *and* $t \in]0, \min(t^*, \bar{t} + c\epsilon)[$, *it holds that*

$$\begin{aligned}
&\forall x \in \mathbb{B}_\delta(x^*), \exists \hat{s} \in \mathbb{B}_\epsilon((G(x^*), H(x^*))^T), \\
&(x, \hat{s}) \text{ is a strong epsilon-stationary point of } (10)\ .
\end{aligned}$$

Proof The proof is conducted in two steps. First, taking any $\hat{x} \in \mathbb{B}_\delta(x^*)$ we construct a point that is a candidate to be strong epsilon-stationary. Then, we verify that this candidate is actually a strong epsilon-stationary point.

Let x^* be an M-stationary point. Therefore, x^* is feasible for (3) and there exists $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ such that:

$$\begin{aligned}
&\nabla_x \mathcal{L}_{MPCC}(x^*, \lambda) = 0, \\
&\lambda^g \geq 0, \lambda^g_i = 0 \ (i \notin \mathcal{I}_g(x^*)),\ \lambda^G_i = 0 \ (i \in \mathcal{I}^{+0}(x^*)),\ \lambda^H_i = 0 \ (i \in \mathcal{I}^{0+}(x^*)), \\
&\text{either } \lambda^G_i > 0,\ \lambda^H_i > 0 \text{ or } \lambda^G_i \lambda^H_i = 0 \ (i \in \mathcal{I}^{00}(x^*)).
\end{aligned}$$

Let c be the positive constant defined as follows

$$c := \sup_{i\in\mathrm{supp}(\lambda^G), j\in\mathrm{supp}(\lambda^H), k\in\mathrm{supp}(\lambda^g)} \frac{1}{|\lambda_i^G|} + \frac{1}{|\lambda_j^H|} + \frac{1}{|\lambda_k^g|},$$

with the convention that $c := 0$ if $\mathrm{supp}(\lambda^G) \cup \mathrm{supp}(\lambda^H) \cup \mathrm{supp}(\lambda^g) = \emptyset$.

Besides, all of the functions involved here are assumed to have Lipschitz continuous gradients, thus there exist constants L^1 and L^2 such that

$$L^1 := \sup_{i,j,k,l} \|\nabla f(x^*)\|_\infty + \|\nabla g_i(x^*)\|_\infty + \|\nabla h_j(x^*)\|_\infty + \|\nabla G_k(x^*)\|_\infty + \|\nabla H_l(x^*)\|_\infty,$$

and for $\hat{\nu}$ depending of x^* and λ defined below

$$L^2 := \|\nabla^2_{xx} \mathcal{L}_{R^s_{t,\bar{t}}}(x^*, G(x^*), H(x^*), \hat{\nu})\|_\infty.$$

Let $\delta > 0$ be a constant satisfying

$$\max\left(L^1, L^1\|\lambda\|_\infty, L^2\right)\delta + o(\delta) \le \epsilon. \tag{11}$$

In other words, $\delta = O(\epsilon)$.

Construction of the point $(\hat{x}, \hat{s}, \hat{\nu})$ For any $\hat{x} \in \mathbb{B}_\delta(x^*)$, let us construct a point $(\hat{x}, \hat{s}, \hat{\nu})$ that satisfies Definition 4.1. Set

$$\xi := \hat{x} - x^*,\ \hat{\nu}^g := \lambda^g,\ \hat{\nu}^h := \lambda^h,\ \hat{\nu}^{s_G} := \lambda^G,\ \hat{\nu}^{s_H} := \lambda^H,$$

where $\|\xi\|_2 < \delta$. Using the following notations

$$\begin{aligned}
&\mathcal{I}^{00}_{-0} := \{i \in \mathcal{I}^{00} : \lambda_i^G \le 0, \lambda_i^H = 0\}, \mathcal{I}^{00}_{0-} := \{i \in \mathcal{I}^{00} : \lambda_i^G = 0, \lambda_i^H \le 0\},\\
&\mathcal{I}^{+0}_{-} := \{i \in \mathcal{I}^{+0} \colon \lambda_i^G = 0, \lambda_i^H < 0\}, \mathcal{I}^{+0}_{+} := \{i \in \mathcal{I}^{+0} \colon \lambda_i^G = 0, \lambda_i^H \ge 0\},\\
&\mathcal{I}^{0+}_{-} := \{i \in \mathcal{I}^{0+} \colon \lambda_i^G < 0, \lambda_i^H = 0\}, \mathcal{I}^{0+}_{+} := \{i \in \mathcal{I}^{0+} \colon \lambda_i^G \ge 0, \lambda_i^H = 0\},\\
&\mathcal{I}^{00}_{++} := \{i \in \mathcal{I}^{00} : \lambda_i^G > 0, \lambda_i^H > 0\},
\end{aligned}$$

let $\hat{s}_G, \hat{s}_H, \hat{\nu}^G, \hat{\nu}^H$ and $\hat{\nu}^\Phi$ be as follows:

$$\hat{s}_G := \begin{cases} t, & i \in \mathcal{I}^{00}_{-0},\\ \bar{t} + c\epsilon, & i \in \mathcal{I}^{00}_{0-},\\ \frac{\epsilon/2 - \bar{t}\lambda_i^G}{\lambda_i^G}, & i \in \mathcal{I}^{00}_{++},\\ G_i(\hat{x}), & i \in \mathcal{I}^{+0},\\ t, & i \in \mathcal{I}^{0+}_{-},\\ -\bar{t}, & i \in \mathcal{I}^{0+}_{+}. \end{cases} \qquad \hat{s}_H := \begin{cases} \bar{t} + c\epsilon, & i \in \mathcal{I}^{00}_{-0},\\ t, & i \in \mathcal{I}^{00}_{0-},\\ t, & i \in \mathcal{I}^{00}_{++},\\ t, & i \in \mathcal{I}^{+0}_{-},\\ -\bar{t}, & i \in \mathcal{I}^{+0}_{+},\\ H_i(\hat{x}), & i \in \mathcal{I}^{0+}, \end{cases} \tag{12}$$

$$\hat{\nu}^G := \begin{cases} \lambda_i^G, i \in \mathcal{I}^{00}_{++} \cup \mathcal{I}^{0+}_{+},\\ 0, \quad \text{otherwise}, \end{cases} \qquad \hat{\nu}^H := \begin{cases} \lambda_i^H, i \in \mathcal{I}^{+0}_{+},\\ 0, \quad \text{otherwise}, \end{cases}$$

$$\hat{\nu}^\Phi := \begin{cases} \frac{-\lambda_i^G}{\hat{s}_{H,i} - t}, i \in \mathcal{I}^{00}_{-0} \cup \mathcal{I}^{0+}_{-},\\ \frac{-\lambda_i^H}{\hat{s}_{G,i} - t}, i \in \mathcal{I}^{00}_{0-} \cup \mathcal{I}^{00}_{++} \cup \mathcal{I}^{+0}_{-},\\ 0, \qquad \text{otherwise}. \end{cases}$$

$(\hat{x}, \hat{s}, \hat{\nu})$ *satisfy the stationarity:* We verify that $(\hat{x}, \hat{s}, \hat{\nu})$ satisfies the strong epsilon-stationarity conditions, that we divide into the next 6 parts:

I. $\left\|\nabla_x \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \hat{s}, \hat{\nu})\right\|_\infty \leq \epsilon$, $g_i(\hat{x}) \leq \epsilon$, $\hat{\nu}^g_i \geq 0$, $|g_i(\hat{x})\hat{\nu}^g_i| \leq \epsilon \ \forall i \in \{1, \ldots, m\}$ and $|h_i(\hat{x})| \leq \bar{t} + (1+c)\epsilon \ \forall i \in \{1, \ldots, p\}$;
II. $\left\|\nabla_s \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \hat{s}, \hat{\nu})\right\|_\infty \leq \epsilon$;
III. $|G_i(\hat{x}) - \hat{s}_{G,i}| \leq \bar{t} + (1+c)\epsilon$, $|H_i(\hat{x}) - \hat{s}_{H,i}| \leq \bar{t} + (1+c)\epsilon \ \forall i \in \{1, \ldots, q\}$;
IV. $\hat{s}_{G,i} + \bar{t} \geq -\epsilon$, $\left|\hat{\nu}^G_i(\hat{s}_{G,i} + \bar{t})\right| \leq \epsilon$, $\hat{s}_{H,i} + \bar{t} \geq -\epsilon$, $\left|\hat{\nu}^H_i(\hat{s}_{H,i} + \bar{t})\right| \leq \epsilon \ \forall i \in \{1, \ldots, q\}$;
V. $\Phi_i(\hat{s}_G, \hat{s}_H; t) \leq 0$, $\left|\hat{\nu}^\Phi_i \Phi_i(\hat{s}_G, \hat{s}_H; t)\right| = 0 \ \forall i \in \{1, \ldots, q\}$;
VI. $\hat{\nu}^G \geq 0, \hat{\nu}^H \geq 0, \hat{\nu}^\Phi \geq 0$.

Let us now run through these 6 conditions. Note that the cases $i \in \mathcal{I}^{00}_{-0}$ (resp. $i \in \mathcal{I}^{+0}$) and $i \in \mathcal{I}^{00}_{0-}$ (resp. $i \in \mathcal{I}^{0+}$) are symmetric, hence it is sufficient to verify these conditions for $i \in \mathcal{I}^{00}_{-0} \cup \mathcal{I}^{00}_{++} \cup \mathcal{I}^{+0}$.

I. Since $(\hat{\nu}^g, \hat{\nu}^h, \hat{\nu}^{s_G}, \hat{\nu}^{s_H}) = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ it holds that

$$\begin{array}{l} \left\|\nabla_x \mathcal{L}_{R^s_{t,\bar{t}}}(x^*, \hat{s}, \hat{\nu})\right\|_\infty = 0, \\ g_i(x^*) \leq 0, \ \hat{\nu}^g_i \geq 0, \ |g_i(x^*)\hat{\nu}^g_i| = 0 \ \forall i \in \{1, \ldots, m\}, \\ |h_i(x^*)| = 0 \ \forall i \in \{1, \ldots, p\}. \end{array}$$

Now, since $\hat{x} = x^* + \xi$ and $\|\xi\| \leq \delta$, using Taylor theorem at x^* yields

$$g_i(\hat{x}) = g_i(x^*) + \xi_i g_i'(x^*) + o(\xi_i) \leq \delta L^1 + o(\delta) \leq \epsilon,$$

where the last inequality follows by definition of δ in (11). We can proceed in the same way to prove that $|h_i(\hat{x})| \leq \epsilon$, $|g_i(\hat{x})\hat{\nu}^g_i| \leq \epsilon$ and $\|\nabla_x \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \hat{s}, \hat{\nu})\|_\infty \leq \epsilon$.

II. As pointed in Lemma 5.1, the gradient of the Lagrangian with respect to s is given by

$$\begin{array}{l} \nabla_{s_{G,i}} \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \hat{s}, \hat{\nu}) = \hat{\nu}^{s_G}_i - \hat{\nu}^G_i + \hat{\nu}^\Phi_i(\hat{s}_{H,i} - t), \\ \nabla_{s_{H,i}} \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \hat{s}, \hat{\nu}) = \hat{\nu}^{s_H}_i - \hat{\nu}^H_i + \hat{\nu}^\Phi_i(\hat{s}_{G,i} - t). \end{array} \tag{13}$$

It is sufficient to verify these conditions for $i \in \mathcal{I}^{00}_{-0} \cup \mathcal{I}^{00}_{++} \cup \mathcal{I}^{+0}_{+} \cup \mathcal{I}^{+0}_{-}$. The following tables summarize the relevant quantities to prove that the gradient (13) vanishes.

$\nabla_{s_{G,i}} \mathcal{L}_{R^s_{t,\bar{t}}}$	$\hat{\nu}^{s_G}_i$	$\hat{\nu}^G_i$	$\hat{\nu}^\Phi_i$	$(\hat{s}_{H,i} - t)$	$\nabla_{s_{G,i}} \mathcal{L}_{R^s_{t,\bar{t}}}$
$\mathcal{I}^{00}_{++}$	λ^G_i	λ^G_i	$\frac{-\lambda^H_i}{\hat{s}_{G,i} - t}$	0	0
$\mathcal{I}^{00}_{-0}$	λ^G_i	0	$\frac{-\lambda^G_i}{\hat{s}_{H,i} - t}$	$\hat{s}_{H,i} - t$	0
$\mathcal{I}^{+0}_{+}$	$\lambda^G_i = 0$	0	0	$\hat{s}_{H,i} - t$	0
$\mathcal{I}^{+0}_{-}$	$\lambda^G_i = 0$	0	$\frac{-\lambda^H_i}{\hat{s}_{G,i} - t}$	0	0

set	$\hat{\nu}_i^{s_H}$	$\hat{\nu}_i^H$	$\hat{\nu}_i^\Phi$	$(\hat{s}_{G,i}-t)$	$\nabla_{s_{H,i}}\mathcal{L}_{R^s_{t,\bar{t}}}$
$\mathcal{I}^{00}_{++}$	λ_i^H	0	$\frac{-\lambda_i^H}{\hat{s}_{G,i}-t}$	$\hat{s}_{G,i}-t$	0
$\mathcal{I}^{00}_{-0}$	$\lambda_i^H=0$	0	$\frac{-\lambda_i^G}{\hat{s}_{H,i}-t}$	0	0
$\mathcal{I}^{+0}_{+}$	λ_i^H	λ_i^H	0	$\hat{s}_{G,i}-t$	0
$\mathcal{I}^{+0}_{-}$	λ_i^H	0	$\frac{-\lambda_i^H}{\hat{s}_{G,i}-t}$	$\hat{s}_{G,i}-t$	0

III. Using the same technique as in **I**. it holds that

$$|G_i(\hat{x})-\hat{s}_{G,i}|\leq\epsilon+|\hat{s}_{G,i}|,\forall i\in\mathcal{I}^{0+}\cup\mathcal{I}^{00},$$
$$|H_i(\hat{x})-\hat{s}_{H,i}|\leq\epsilon+|\hat{s}_{H,i}|,\forall i\in\mathcal{I}^{+0}\cup\mathcal{I}^{00},$$

while $|G_i(\hat{x})-\hat{s}_{G,i}|=0$ for $i\in\mathcal{I}^{+0}$ and $|H_i(\hat{x})-\hat{s}_{H,i}|=0$ for $i\in\mathcal{I}^{0+}$. For $i\in I^{00}_{-0}\cup I^{00}_{0-}$, $|\hat{s}^{G,i}|\leq\bar{t}+c\epsilon$ follows from $\hat{s}^{G,i}=t$ or $\hat{s}^{G,i}=\bar{t}+c\epsilon$. For $i\in I^{00}_{++}$,$|\frac{\epsilon/2-\bar{t}\lambda_i^G}{\lambda_i^G}|\leq\bar{t}+c\epsilon$.

IV. By construction, $\hat{s}_{G,i}+\bar{t}\geq 0$ and $\hat{s}_{H,i}+\bar{t}\geq 0$. Now, we verify the complementarity condition for $i\in\mathcal{I}^{00}_{++}\cup\mathcal{I}^{0+}_{+}$ as otherwise $\hat{\nu}_i^G=0$ (the other case follows by symmetry).

$$\left|\hat{\nu}_i^G(\hat{s}_{G,i}+\bar{t})\right|=\begin{cases}0,\ \text{for } i\in\mathcal{I}^{0+}_{+},\\ \left|\lambda_i^G\frac{\epsilon}{2\lambda_i^G}\right|=\frac{\epsilon}{2},\ \text{for } i\in\mathcal{I}^{00}_{++}.\end{cases}$$

V. By construction of $\hat{s}_G$ and $\hat{s}_H$, we have $\Phi_i(\hat{s}_G,\hat{s}_H;t)=0$ for $i\in\mathcal{I}^{00}\cup\mathcal{I}^{0+}_{-}\cup\mathcal{I}^{+0}_{-}$ and $\Phi_i(\hat{s}_G,\hat{s}_H;t)<0$ for the remaining indices taking t sufficiently small such that $G_i(\hat{x})>t$ for $i\in\mathcal{I}^{+0}_{+}$ and $H_i(\hat{x})>t$ for $i\in\mathcal{I}^{0+}_{+}$. For indices $\mathcal{I}^{+0}_{+}\cup\mathcal{I}^{0+}_{+}$, where $\Phi_i(\hat{s}_G,\hat{s}_H;t)\neq 0$ the complementarity condition is satisfied as $\hat{\nu}_i^\Phi=0$.

VI. By construction it is clear that $\hat{\nu}_i^G\geq 0,\hat{\nu}_i^H\geq 0$. Consider indices i $\in\mathcal{I}^{00}_{0-}\cup\mathcal{I}^{00}_{++}\cup\mathcal{I}^{+0}_{-}$, i.e. $\hat{\nu}_i^\Phi=\frac{-\lambda_i^H}{\hat{s}_{G,i}-t}$, as $i\in\mathcal{I}^{00}_{-0}\cup\mathcal{I}^{0+}_{-}$ will follow by symmetry. $\frac{-\lambda_i^H}{\hat{s}_{G,i}-t}$ is non-negative since $\hat{s}_{G,i}>t$ when $\lambda_i^H<0$ and $\hat{s}_{G,i}<t$ when $\lambda_i^H\geq 0$. For the case $\mathcal{I}^{00}_{0-}$, $\lambda_i^H\leq 0$ and $\hat{s}_{G,i}=\bar{t}+c\epsilon>t$. For the case $\mathcal{I}^{+0}_{-}$, $\lambda_i^H<0$ and $\hat{s}_{G,i}=G_i(\hat{x})>G_i(x^*)-L_1\delta-o(\delta)>G_i(x^*)-\epsilon>t$ as t is chosen sufficiently small. For the case $\mathcal{I}^{00}_{++}$, $\lambda_i^H>0$ and $\hat{s}_{G,i}=\frac{\epsilon}{2\lambda_i^G}-\bar{t}<0<t$.

□

This result cannot be extended to a general neighborhood for s due to the definition of strong epsilon-stationary point. However, we prove that s can live in a neighborhood of $\hat{s}$ (defined in Lemma 5.2) on the boundary of the feasible set.

Theorem 5.1 *Assume that f,g,h,G,H have Lipschitz continuous gradients. Let x^* be an M-stationary point of (3) and $\epsilon>0$ be small enough. Then, there exist positive constants $c,\delta,\delta',\bar{t}^*\geq c\epsilon,t^*$ such that, for all $\bar{t}\in]c\epsilon,\bar{t}^*[$ and $t\in]0,\min(t^*,\bar{t}+c\epsilon)[$, any point in the following set*

$$(x,s)\in\mathbb{B}_\delta(x^*)\times\mathbb{B}_{\delta'}(\hat{s})\cap\{(s_G,s_H)\ :\ \Phi_i(s_G,s_H;t)=0,\ \forall i\in\mathcal{I}_\Phi(x^*;t)\},$$

with $\hat{s}$ defined in (12)*, is a strong epsilon-stationary point of the regularization* (10)*.*

Proof First we stress that we keep the multipliers $\hat{\nu}$ as defined in lemma 5.2

The proof is an adaptation of the one of Lemma 5.2, where are unchanged the parts I, V, VI. Part I remains true since s only appears in $\nabla_x \mathcal{L}_{R^s_{t,\bar{t}}}$ which is actually independent on s by Lemma 5.1. Part V remains true as the active constraints remain active for all $s \in \mathbb{B}_{\delta'}((G(x^*), H(x^*))^T) \cap \{(s_G, s_H) \ : \Phi_i(s_G, s_H; t) = 0 \text{ for all } i \in \mathcal{I}_\Phi(x^*; t)\}$. Part VI remains true as we keep the definition of multipliers $\hat{\nu}^\Phi$ using $\hat{s}$.

Hence, we discuss only the parts II, III and IV.

Let δ' be such that

$$\max\left(1, 2\|\lambda\|_\infty, \frac{\|\lambda\|_\infty}{\min(\min_{i\in\mathcal{I}^{+0}}\{G_i(x^*)\} - t - \epsilon, \bar{t} + t, \bar{t} + c\epsilon - t)}\right)\delta' < \epsilon,$$

and $\tilde{s} = \hat{s} + \alpha$ with $\|\alpha\| < \delta'$.

II. In Lemma 5.2, we proved stronger than strong-epsilon stationarity since we obtained vanishing gradient of the Lagrangian with respect to s, stronger than normwise bounded by ϵ. Now, our use of $\hat{\nu}$ introduces perturbations in the second line of first table and lines one and four of the second table in the proof of II in previous Lemma 5.2. Hence, by (13) with respect to s_G (the case s_H follows in the same way), we have $\nabla_{s_{G,i}} \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \tilde{s}, \hat{\nu}) = \nabla_{s_{G,i}} \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \hat{s}, \hat{\nu}) + \hat{\nu}^\Phi_i \alpha_i$. Therefore, $\|\nabla_{s_{G,i}} \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \tilde{s}, \hat{\nu})\| < \|\nabla_{s_{G,i}} \mathcal{L}_{R^s_{t,\bar{t}}}(\hat{x}, \hat{s}, \hat{\nu})\| + |\hat{\nu}^\Phi_i|\delta' < \epsilon$.

III. In Lemma 5.2, we proved that $\forall i \in \{1, \ldots, q\},\ |G_i(\hat{x}) - \hat{s}_{G,i}| \leq \bar{t} + (1 + c)\epsilon,\ |H_i(\hat{x}) - \hat{s}_{H,i}| \leq \bar{t} + (1 + c)\epsilon$. So, using that $\delta' \leq \epsilon$ gives $\forall i \in \{1, \ldots, q\},\ |G_i(\hat{x}) - \tilde{s}_{G,i}| \leq \bar{t} + (2 + c)\epsilon,\ |H_i(\hat{x}) - \tilde{s}_{H,i}| \leq \bar{t} + (2 + c)\epsilon$, which proves III.

IV. From Lemma 5.2, when $\hat{s}_{G,i} + \bar{t} \geq 0$ (resp. $\hat{s}_{H,i} + \bar{t} \geq 0$) we deduce $\tilde{s}_{G,i} + \bar{t} \geq -\epsilon$ (resp. $\tilde{s}_{H,i} + \bar{t} \geq -\epsilon$) as $\delta' < \epsilon$.

Now, we verify the complementarity condition for $i \in \mathcal{I}^{00}_{++} \cup \mathcal{I}^{0+}_{+}$:

$$\left|\hat{\nu}^G_i(\tilde{s}_{G,i} + \bar{t})\right| \leq \left|\hat{\nu}^G_i(\hat{s}_{G,i} + \bar{t})\right| + \left|\hat{\nu}^G_i \tilde{s}_{G,i}\right| = \begin{cases} \delta'\lambda^G_i \leq \frac{\epsilon}{2}, \text{ for } i \in \mathcal{I}^{0+}_{+}, \\ \frac{\epsilon}{2} + \delta'\lambda^G_i \leq \epsilon, \text{ for } i \in \mathcal{I}^{00}_{++}. \end{cases} \quad \square$$

We can now deduce the following result.

Theorem 5.2 *For any M-stationary point of* (3) *that satisfies* MPCC − MFCQ*, there exists a sequence of strong epsilon-stationary points of the regularization* (10) *that converges to that point.*

Proof Theorem 5.1 gives more relations between the parameters that are compatible with Corollary 5.1. Indeed for a chosen sequence of arbitrarily small parameters $\{\epsilon^k\}$, Corollary 5.1 requires that $\epsilon^k = o(\bar{t}_k)$ and Theorem 5.1 requires that $\bar{t}_k > c\epsilon$ and t_k must be sufficiently small.

Thus, a straightforward application of both of these results allows to conclude. $\square$

6 Conclusions

We studied in this paper a regularization method for the MPCC with approximate stationary points. Motivated by the recent discussion in [19], we introduced a new type of stationary point for the regularized subproblem, called strong epsilon-statonary point, whose sequence converges to an M-stationary point of the MPCC. Additionally, we proved that for any M-stationary point there exists a sequence of strong epsilon-statonary points in its neighborhood. These two theoretical results pave the way to a regularization method with strong properties.

The main perspective of this work is now to study an optimization algorithm able to compute a strong epsilon-stationary point for given $t, \bar{t}$ and ϵ. This algorithm must satisfy at each step exactly the feasibility of the regularized complementarity constraint as well as the complementarity condition of this constraint. We would like to mention that a possible direction is to adapt the active-set strategy studied in [28] since the active-set approach will be able to maintain the feasibility at each step.

Acknowledgements This research was partially supported by a grant from "l'Ecole des Docteurs de l'UBL" and "le Conseil Régional de Bretagne". This research was partially supported by NSERC and FRQNT grants.

References

1. Ye, J.J., Zhu, D.L., Zhu, Q.J.: Exact penalization and necessary optimality conditions for generalized bilevel programming problems. SIAM Journal on optimization **7**(2), 481–507 (1997)
2. Luo, Z.Q., Pang, J.S., Ralph, D.: Mathematical programs with equilibrium constraints. Cambridge University Press (1996)
3. Outrata, J., Kočvara, M., Zowe, J.: Nonsmooth approach to optimization problems with equilibrium constraints: theory, applications and numerical results, vol. 28. Springer Science & Business Media (2013)
4. Flegel, M.L., Kanzow, C.: A direct proof for M-stationarity under MPEC-GCQ for mathematical programs with equilibrium constraints. Springer (2006)
5. Kanzow, C., Schwartz, A.: Mathematical programs with equilibrium constraints: enhanced fritz john-conditions, new constraint qualifications, and improved exact penalty results. SIAM Journal on Optimization **20**(5), 2730–2753 (2010)
6. Scheel, H., Scholtes, S.: Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. Mathematics of Operations Research **25**(1), 1–22 (2000)
7. Ye, J.J.: Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints. SIAM Journal on Optimization **10**(4), 943–962 (2000)
8. Ye, J.J.: Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. Journal of Mathematical Analysis and Applications **307**(1), 350–369 (2005)
9. Ye, J.J., Ye, X.Y.: Necessary optimality conditions for optimization problems with variational inequality constraints. Mathematics of Operations Research **22**(4), 977–997 (1997)
10. Flegel, M.L., Kanzow, C.: On the guignard constraint qualification for mathematical programs with equilibrium constraints. Optimization **54**(6), 517–534 (2005)

11. Guo, L., Lin, G.H., Ye, J.J.: Solving mathematical programs with equilibrium constraints. Journal of Optimization Theory and Applications **166**(1), 234–256 (2015)
12. Outrata, J.V.: Optimality conditions for a class of mathematical programs with equilibrium constraints. Mathematics of Operations Research **24**(3), 627–644 (1999)
13. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM Journal on Optimization **11**(4), 918–936 (2001)
14. DeMiguel, V., Friedlander, M.P., Nogales, F.J., Scholtes, S.: A two-sided relaxation scheme for mathematical programs with equilibrium constraints. SIAM Journal on Optimization **16**(2), 587–609 (2005)
15. Lin, G.H., Fukushima, M.: A modified relaxation scheme for mathematical programs with complementarity constraints. Annals of Operations Research **133**(1-4), 63–84 (2005)
16. Steffensen, S., Ulbrich, M.: A new relaxation scheme for mathematical programs with equilibrium constraints. SIAM Journal on Optimization **20**(5), 2504–2539 (2010)
17. Kadrani, A., Dussault, J.P., Benchakroun, A.: A new regularization scheme for mathematical programs with complementarity constraints. SIAM Journal on Optimization **20**(1), 78–103 (2009)
18. Kanzow, C., Schwartz, A.: A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. SIAM Journal on Optimization **23**(2), 770–798 (2013)
19. Kanzow, C., Schwartz, A.: The price of inexactness: Convergence properties of relaxation methods for mathematical programs with complementarity constraints revisited. Mathematics of Operations Research **40**(2), 253–275 (2015)
20. Andreani, R., Haeser, G., Martìnez, J.M.: On sequential optimality conditions for smooth constrained optimization. Optimization **60**(5), 627–641 (2011)
21. Outrata, J.V.: A generalized mathematical program with equilibrium constraints. SIAM Journal on Control and Optimization **38**(5), 1623–1638 (2000)
22. Flegel, M.L., Kanzow, C.: Abadie-type constraint qualification for mathematical programs with equilibrium constraints. Journal of Optimization Theory and Applications **124**(3), 595–614 (2005)
23. Pang, J.S., Fukushima, M.: Complementarity constraint qualifications and simplified b-stationarity conditions for mathematical programs with equilibrium constraints. Computational Optimization and Applications **13**(1), 111–136 (1999)
24. Hoheisel, T., Kanzow, C., Schwartz, A.: Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. Mathematical Programming **137**(1-2), 257–288 (2013)
25. Andreani, R., Haeser, G., Secchin, L.D., Silva, P.J.S.: New sequential optimality conditions for mathematical programs with complementarity constraints and algorithmic consequences. SIAM Journal on Optimization **29**(4), 3201–3230 (2019)
26. Ramos, A.: Mathematical programs with equilibrium constraints: a sequential optimality condition, new constraint qualifications and algorithmic consequences. Optimization Methods and Software **0**(0), 1–37 (2019)
27. Andreani, R., Secchin, L., Silva, P.: Convergence properties of a second order augmented lagrangian method for mathematical programs with complementarity constraints. SIAM Journal on Optimization **28**(3), 2574–2600 (2018)
28. Kadrani, A., Dussault, J.P., Benchakroun, A.: A globally convergent algorithm for mpcc. EURO Journal on Computational Optimization **3**(3), 263–296 (2015)