



Contributions aux méthodes numériques pour les problèmes de complémentarité et problèmes d'optimisation sous contraintes de complémentarité

Tangi Migot

► To cite this version:

Tangi Migot. Contributions aux méthodes numériques pour les problèmes de complémentarité et problèmes d'optimisation sous contraintes de complémentarité. Analyse numérique [math.NA]. INSA de Rennes, 2017. Français. NNT : 2017ISAR0026 . tel-01791526

HAL Id: tel-01791526

<https://theses.hal.science/tel-01791526>

Submitted on 14 May 2018

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

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

Thèse

UNIVERSITE
BRETAGNE
LOIRE

THESE INSA Rennes
sous le sceau de l'Université Bretagne Loire
pour obtenir le titre de
DOCTEUR DE L'INSA RENNES
Spécialité : Mathématiques et leurs Interactions

présentée par

Tangi MIGOT

ECOLE DOCTORALE : MATHSTIC
LABORATOIRE : IRMAR

Contributions aux
méthodes numériques
pour les problèmes
de complémentarité et
problèmes d'optimisation
sous contraintes de
complémentarité

Christian KANZOW

Professeur, University of Würzburg (Allemagne) / Rapporteur

Olivier LEY

Professeur des Universités, INSA de Rennes / Examineur / Invité

Thèse soutenue le 06.10.2017

devant le jury composé de :

Jean Charles GILBERT

Directeur de Recherche, INRIA Paris / Président

Samir ADLY

Professeur des Universités, Université de Limoges / Rapporteur

Jocelyne ERHEL

Directrice de Recherche, INRIA Rennes / Examinatrice

Claudia SAGASTIZABAL

Directrice de Recherche, IMPA Rio de Janeiro (Brésil) / Examinatrice

Jean-Pierre DUSSAULT

Professeur, Université de Sherbrooke (Canada) / Co-encadrant

Mounir HADDOU

Professeur des Universités, INSA de Rennes / Directeur de thèse

Contributions aux méthodes numériques pour les problèmes de complémentarité et problèmes d'optimisation sous contraintes de complémentarité

Tangi Migot



En partenariat avec



Résumé

Dans cette thèse, nous avons étudié les méthodes de régularisation pour la résolution numérique de problèmes avec équilibres.

Dans une première partie, nous nous sommes intéressés aux problèmes de complémentarité au travers de deux applications : les équations en valeur absolue et les problèmes de parcimonie.

Dans une seconde partie, nous avons étudié les problèmes d'optimisation sous contraintes de complémentarité. Après avoir défini des conditions d'optimalité pour ces problèmes nous avons proposé une nouvelle méthode de régularisation appelée méthode des papillons. À partir d'une étude de la résolution des sous-problèmes de la régularisation nous avons défini un algorithme avec des propriétés de convergence forte.

Tout au long de ce manuscrit nous nous sommes concentrés sur les propriétés théoriques des algorithmes ainsi que sur leurs applications numériques. La dernière partie de ce document est consacrée aux résultats numériques des méthodes de régularisation.

Abstract

In this thesis, we studied the regularization methods for the numerical resolution of problems with equilibria.

In the first part, we focused on the complementarity problems through two applications that are the absolute value equation and the sparse optimization problem.

In the second part, we concentrated on optimization problems with complementarity constraints. After studying the optimality conditions of this problem, we proposed a new regularization method, so-called butterfly relaxation. Then, based on an analysis of the regularized sub-problems we defined an algorithm with strong convergence property.

Throughout the manuscript, we concentrated on the theoretical properties of the algorithms as well as their numerical applications. In the last part of this document, we presented numerical results using the regularization methods for the mathematical programs with complementarity constraints.

Remerciements

Cette thèse n'aurait pu être achevée sans l'aide, les encouragements et le soutien d'un grand nombre de personnes à qui j'adresse ma plus sincère gratitude et considération.

Tout d'abord, je souhaiterais exprimer ma reconnaissance et toute mon amitié à mon Directeur de thèse Mounir Haddou. Je n'aurai pu espérer personne plus disponible et sympathique pour m'encadrer. J'ai beaucoup appris à ses côtés grâce à nos échanges et discussions. Les mots sont faibles pour exprimer ma gratitude pour ses encouragements et son soutien constant au cours de ces trois années.

Au cours de ma deuxième année de thèse, j'ai eu le plaisir et l'honneur de rencontrer Jean-Pierre Dussault qui est par la suite devenu co-encadrant de la thèse. Je le remercie de son soutien, des discussions éclairantes que nous avons eu et de m'avoir fait l'honneur de m'inviter à Sherbrooke, séjours qui ont été très enrichissants tant sur le plan mathématiques qu'humain.

Mes remerciements vont aussi à Samir Adly et Christian Kanzow, qui m'ont fait l'honneur de rapporter le manuscrit, ainsi qu'à Claudia Sagastizabal, Jocelyne Erhel, Jean Charles Gilbert et Olivier Ley qui ont accepté de faire partie de mon jury de thèse.

Je remercie la composante IRMAR-Insa qui m'a accueilli et m'a permis de réaliser les travaux présents dans ce manuscrit dans de bonnes conditions. Mes remerciements vont à tous les membres de la composante ainsi que son personnel administratif dont le contact a été à la fois source d'apprentissage et d'épanouissement.

Une partie des résultats obtenus au cours la thèse a été motivée par des discussions avec d'autres chercheurs que mes encadrants. Je remercie Lina Abdallah pour son invitation au Liban, son soutien et nos discussions. Mes remerciements vont également vers Jérémy Omer pour ses précieux conseils et nos échanges qui ont bien souvent été source d'inspiration et de motivation.

Durant ces trois années, j'ai profité de la présence de beaucoup de personnes qui ont fait de cette thèse une succession de moments agréables. L'Insa a été un lieu de travail motivant, mais également riche en rencontres. Je tiens à remercier Tuyen, Viet, Samuel, Trinh, Khang, Duong, Kien pour les discussions et pauses café très agréables. Au cours de ces années mon vietnamien ne s'est pas amélioré, mais un grand merci à Tuan qui m'a supporté dans son bureau et dans sa ville natale à Da Nang. Je ne pourrai oublier mes deux co-bureaux Emilie et Audrey qui en plus de me supporter sont amies, colocataires et co-voyageuses !

J'ai eu le plaisir de passer une partie de cette thèse à Sherbrooke et je remercie sincèrement le BISOUS et les bisounours pour leur accueil, les 5 à 7, les parties de Mølky, ... Merci à

Max', Luc, Jo', Sam et Mathieu.

Évidemment je ne saurai oublier tous ceux qui en dehors des universités étaient toujours là quand le besoin de passer un bon moment se faisait sentir. J'ai le bonheur d'être entouré de personnes toujours souriantes, positives et optimistes ! Mes remerciements vont ici aux joueurs d'échecs, Bastien, Alexandre, Corina, Pierre, Raph, Mathilde, Remel, qui m'ont accompagné sur les chemins obscurs des 64 cases. Un grand merci aux belettes pour leur sourire, leur chaleur et leur gentillesse, ma petite sœur Fleur, Lucie et Vio. Je garderai toujours un excellent souvenir des moments passés à DDL et des vacances. Merci également à Cédric qui est là depuis le début et toujours disponible.

Plus important que tout, je remercie vivement ma famille. Leur soutien a été indéfectible depuis toujours et ce même dans les moments difficiles. Merci à mon frère, Jildaz, et Alice pour nos retrouvailles à Rennes, leur soutien et les bons moments. Enfin, un très grand merci à mes parents, Martine et Alain, qui ont toujours été présents pour m'aider, me réconforter et célébrer les bons moments. Il n'y a pas de mots pour leur exprimer toute ma gratitude. Ce manuscrit leur est dédié.

À Rennes, octobre 2017
Tangi Migot

Travaux en lien avec la thèse

Les résultats présentés dans ce manuscrit de thèse représentent les travaux réalisés au cours des trois années de durée de la thèse. Au cours de cette période une majeure partie des résultats obtenus ont été rendus accessibles en ligne, sur HAL¹ ou optimization-online², et soumis à des revues spécialisées.

La première partie qui s'intéresse aux problèmes de parcimonie et de complémentarité est constituée de deux articles [95] et [2] :

- *A Smoothing Method for Sparse Optimization over Polyhedral Sets*, M. Haddou & T. Migot, 2015 ;
- *Solving Absolute Value Equation using Complementarity and Smoothing Functions*, Journal of Computational and Applied Mathematics (accepted), L. Abdallah, M. Haddou & T. Migot, 2017.

La seconde partie est centrée autour des problèmes d'optimisation sous contraintes de complémentarité, de leurs conditions d'optimalité et des méthodes de régularisation. Dans cette partie nous proposons, une nouvelle méthode de régularisation introduite dans [64] :

- *The New Butterfly Relaxation Methods for Mathematical Program with Complementarity Constraints*, J.-P. Dussault, M. Haddou & T. Migot, 2016.

L'étude plus fine des sous-problèmes régularisés pour les méthodes de régularisation est réalisée dans [148] :

- *How to Compute the Local Minimum of the MPCC*, J.-P. Dussault, M. Haddou, A. Kadrani & T. Migot, 2017.

Au delà des travaux présentés ici, une étude numérique sur les méthodes de points intérieurs pour les problèmes de complémentarité linéaires a été réalisé dans [96] et est disponible en annexe de ce document.

¹*hal.archives-ouvertes.fr*

²*optimization-online.org*

Résumé de la thèse

Motivation

En sciences, plus particulièrement en mathématiques appliquées et en optimisation, la modélisation d'équilibres est un sujet fondamental pour les chercheurs depuis de nombreuses années. Des problèmes d'équilibres sont naturellement rencontrés dans la réalité, par exemple lors de changement de phases au cours d'une réaction chimique pour ne citer qu'un exemple parmi d'autres dans divers domaines comme en physique (problème de contact), en économie (équilibres de Nash),... Ces problèmes interviennent également de façon naturelle en optimisation et en recherche opérationnelle. Par exemple les conditions nécessaires d'optimalité d'un problème d'optimisation ou la modélisation de variables binaires sont très souvent reformulées avec des équilibres.

Dans le contexte des mathématiques appliquées, ces problèmes d'équilibres sont appelés problèmes de complémentarité. Soit un cône K , c'est-à-dire que si $x \in K$ alors $\tau x \in K$ pour tout $\tau \geq 0$, et deux fonctions $G, H : K \rightarrow \mathbb{R}^q$ avec $\mathbb{N} \ni q \leq n$. Le problème de complémentarité consiste à trouver $x \in \mathbb{R}^n$ qui satisfait

$$K \ni G(x) \perp -H(x) \in K^\circ,$$

où la notation \perp signifie perpendiculaire et K° est le cône polaire de K défini comme

$$K^\circ := \{d \in \mathbb{R}^n \mid v^T d \leq 0, \forall v \in K\};$$

c'est-à-dire que K° est l'ensemble des vecteurs qui forment un angle obtu avec chaque vecteur de K . Il est commun dans la littérature de se limiter à proposer des méthodes numériques où $K = \mathbb{R}_+^n$, étant donné qu'ils couvrent une large majorité des applications qui intéressent les ingénieurs et chercheurs. Dans ce cas, le problème de complémentarité consiste à trouver $x \in \mathbb{R}^n$ qui satisfait :

$$0 \leq G(x) \perp H(x) \geq 0. \quad (\text{CP})$$

La condition de complémentarité peut ainsi être réécrite comme $G_i(x)H_i(x) = 0$ pour tout $i = 1, \dots, q$.

Bien que ce problème ne soit pas un problème d'optimisation, mais simplement un problème de réalisabilité il est d'un grand intérêt pour l'optimisation. En effet, les conditions nécessaires d'optimalité de nombreux problèmes d'optimisation peuvent être représentées sous la forme (CP). Soit le problème d'optimisation non linéaire (parfois appelé problème

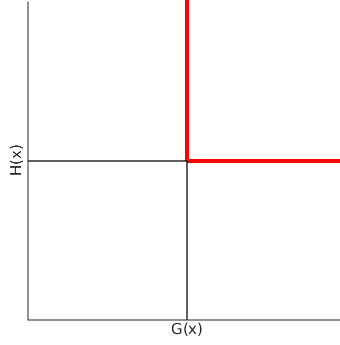


Figure 1: Ensemble des solutions du problème (CP).

de programmation non linéaire) défini par une fonction objectif $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et des fonctions de contraintes $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tel que

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

où f, g, h seront supposées continûment différentiable. La résolution globale de ce problème (i.e trouver le minimum global de la fonction f qui respecte les contraintes) est un problème très difficile. Un domaine entier de l'optimisation, appelé optimisation globale, y est exclusivement affecté.

L'objectif plus raisonnable est donc de trouver un minimum local de ce problème, c'est-à-dire un point x^* tel que pour tout $x \in V_\epsilon(x^*) \cap C$ vérifie

$$f(x) \geq f(x^*),$$

où $V_\epsilon(x^*)$ est un voisinage centré en x^* de rayon ϵ et C l'ensemble des contraintes du problème. On en déduit alors que x^* vérifie aussi la relation

$$\nabla f(x^*)^T d \geq 0 \forall d \in \mathcal{T}_{\mathcal{F}},$$

où $\mathcal{T}_{\mathcal{F}} := \{d \in \mathbb{R}^n \mid \exists t_k \geq 0 \text{ and } x^k \rightarrow_{\mathcal{F}} x^* \text{ s.t. } t_k(x^k - x^*) \rightarrow d\}$ est appelé le cône tangent à \mathcal{F} en x^* et \mathcal{F} est une notation pour l'ensemble des x réalisable de (NLP) (c.-à-d. les points qui satisfont les contraintes de (NLP)). La recherche d'un minimum local (c.-à-d. trouver le minimum dans un certain voisinage de la fonction f qui respecte les contraintes) est un problème numériquement coûteux sans hypothèse de convexité sur les fonctions en jeu. C'est pourquoi dans un objectif numérique efficace l'essentiel des algorithmes, y compris la plupart des solveurs commerciaux, résout des conditions nécessaires d'optimalité.

En 1951, Kuhn et Tucker énoncent des conditions nécessaires du premier ordre pour qu'un point x^* soit un minimum local de (NLP). Plus tard, les chercheurs découvrirent que ces conditions furent historiquement introduites par Karush dans son mémoire de maîtrise en 1939, [47]. Si les contraintes $g(x) \leq 0, h(x) = 0$ donnent une "bonne" description de

l'ensemble des points réalisables au voisinage d'un point x^* (ils vérifient une condition de qualification), alors ce point x^* est un point critique s'il vérifie

$$\begin{aligned} -\nabla f(x^*) &= \sum_{i=1}^m \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^p \lambda_i^h \nabla h_i(x^*), \\ h(x^*) &= 0, \quad 0 \leq -g(x) \perp \lambda^g \geq 0. \end{aligned} \tag{KKT}$$

On dit dans ce cas que le point x^* vérifie les conditions de Karush-Kuhn-Tucker. Il apparaît clairement que ces conditions KKT forment un problème de complémentarité.

Les difficultés majeures pour résoudre le problème de complémentarité (CP) viennent de deux aspects essentiellement géométriques. D'une part, l'ensemble des solutions de ce problème n'est en général pas convexe et pas connexe. D'autre part, l'intérieur relatif de l'ensemble des solutions est vide, c'est-à-dire qu'il n'existe pas de x^* solution de (CP) tel que $G(x^*) > 0, H(x^*) > 0$.

Diverses méthodes numériques existent pour résoudre ce problème. Parmi celles-ci on peut citer les méthodes de reformulation qui transforment (CP) comme un système d'équations sans contraintes ou encore les méthodes d'activation de contraintes qui utilisent une procédure combinatoire pour déterminer les contraintes actives. Au vu des difficultés géométriques énoncées plus haut, une approche naturelle est d'utiliser des techniques de relâchement, autrement appelées techniques de régularisation. Ces techniques relâchent les contraintes du problème pour le rendre plus simple, puis tentent de se rapprocher du problème initial. Ce processus mène bien souvent à des méthodes itératives. Ce sont ces méthodes qui sont au coeur de ce manuscrit. Parmi les méthodes de régularisation les plus connues, on peut citer les méthodes de point-intérieur et les méthodes de pénalisations ou de fonctions de mérites. Ces dernières transforment le problème de complémentarité (CP) comme un problème d'optimisation avec une fonction objectif qui incite à faire respecter les contraintes du problème de complémentarité. La méthode des points-intérieurs peut aussi être interprétée comme une reformulation avec une pénalité logarithmique.

Une généralisation naturelle du problème (CP) est de considérer la résolution d'un problème d'optimisation avec un problème de complémentarité inclus dans les contraintes. On appelle problème d'optimisation sous contrainte de complémentarité le problème qui consiste à minimiser une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.à} & g(x) \leq 0, h(x) = 0, \\ & 0 \leq G(x) \perp H(x) \geq 0, \end{aligned} \tag{MPCC}$$

pour des fonctions de contraintes $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^q$. De nombreuses applications utilisent le problème (MPCC) par exemple en contrôle optimal, en physique ou encore en recherche opérationnelle.

Considérons le problème d'optimisation bi niveaux suivant qui consiste à minimiser une fonction objectif f_0 contrainte à la résolution d'un autre problème d'optimisation, c'est-à-dire

$$\begin{aligned} \min_{x,y \in \mathbb{R}^{n_0} \times \mathbb{R}^{n_1}} f_0(x,y) \\ \text{s.à. } g_0(x,y) \leq 0, h_0(x,y) = 0, \\ y \in \arg \min_{y \in \mathbb{R}^{n_1}} \{f_1(x,y) \text{ s.à. } g_1(y) \leq 0, h_1(y) = 0\}, \end{aligned} \quad (\text{BP})$$

où $f_0 : \mathbb{R}^{n_0} \times \mathbb{R}^{n_1} \rightarrow \mathbb{R}$, $f_1 : \mathbb{R}^{n_1} \times \rightarrow \mathbb{R}$ et les fonctions de contraintes $g_0 : \mathbb{R}^{n_0} \times \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{m_0}$, $h_0 : \mathbb{R}^{n_0} \times \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{p_0}$ et $g_1 : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{m_1}$, $h_1 : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{p_1}$.

Dans ce cas, en remplaçant le problème d'optimisation "intérieur" par ses conditions (KKT), on obtient alors un problème d'optimisation sous contraintes de complémentarité. De nombreux problèmes d'économie sont notamment reformulés comme (MPCC) via cette technique.

Les difficultés géométriques mentionnées pour les problèmes de complémentarité ne font qu'accentuer les difficultés pour la résolution numérique de (MPCC). Bien que (MPCC) soit un problème d'optimisation non linéaire de la forme de (NLP), les contraintes de complémentarité font que d'une façon générique les qualifications de contraintes ne sont en général pas vérifiées et donc nous ne pouvons utiliser la théorie basée sur les conditions (KKT).

D'une façon similaire aux problèmes de complémentarité une approche naturelle est de relâcher les contraintes de complémentarité de (MPCC) afin d'obtenir des problèmes d'optimisation non linéaire et ainsi utiliser les techniques très performantes développées sur ces problèmes ces 60 dernières années.

Contributions de la thèse

Divisée en trois parties cette thèse comporte 9 chapitres. Tout au long de ce document, nous nous intéresserons aux techniques de régularisation pour les problèmes de complémentarité et les problèmes d'optimisation sous contraintes de complémentarité. Ces techniques de régularisation ont notamment permis de développer différentes méthodes qui seront abordées dans chacun des chapitres qui composent ce manuscrit.

Le premier chapitre est une introduction à l'optimisation et aux problèmes de complémentarité. Dans ce chapitre nous définirons ou rappellerons différentes notions qui sont largement utilisées dans ce manuscrit et en optimisation en général. En première partie, nous nous intéresserons en particulier à une technique de régularisation introduite dans [26] et [94] et ses applications en parcimonie et aux équations en valeur absolue. La seconde partie définit de façon détaillée les conditions d'optimalité du premier ordre et les conditions de qualifications de contraintes accessibles aux problèmes d'optimisation sous contraintes de complémentarité. Par la suite, nous discuterons des méthodes de régularisation pour ces problèmes. En particulier, nous présenterons une nouvelle méthode de régularisation appelée "méthode des papillons". Puis, nous présenterons pour un large spectre de méthodes de régularisation des résultats théoriques dans des conditions plus réalistes que celles

présentés dans la littérature jusqu'à présent. La troisième partie est centrée sur l'application numérique des méthodes de régularisation pour les problèmes d'optimisation sous contraintes de complémentarité. En particulier, nous discuterons d'une comparaison de ces méthodes et d'une implantation en JULIA de la méthode des papillons.

Nous nous concentrerons pour chacune des méthodes présentées ici à réaliser une étude théorique complète. Par ailleurs, un soin particulier sera apporté à l'implantation des méthodes via des codes informatiques et à l'application de ces méthodes.

Partie I : Problèmes de complémentarité

Cette première partie se concentre sur de nouvelles méthodes numériques pour résoudre le problème de complémentarité linéaire qui est une spécialisation de (CP) défini en introduction pour G et H affines. Ce problème consiste à trouver $x \in \mathbb{R}^n$ tel que

$$0 \leq Mx + q \perp x \geq 0, \quad (\text{LCP})$$

pour une matrice M d'ordre n et un vecteur $q \in \mathbb{R}^n$. Cette formulation pourrait être plus générale en considérant le membre de droite comme une fonction affine également. Sans perdre de généralité, nous nous conformons ici à cette formulation qui est la plus populaire dans la littérature des problèmes de complémentarité.

Des travaux récents de M.Haddou dans [94] ont inspiré une technique de régularisation pour les problèmes de complémentarité non linéaire. Nous présentons ici une interprétation originale de cette technique. Le (LCP) est équivalent à

$$y = Mx + q \geq 0, x \geq 0, \|y_i\|_0 + \|x_i\|_0 \leq 1 \quad \forall i = \{1, \dots, n\}, \quad (1)$$

où $\|\cdot\|_0$ indique la "norme ℓ_0 " d'un vecteur, c'est-à-dire que pour un vecteur $x \in \mathbb{R}^n$ la norme ℓ_0 de ce vecteur est donnée par

$$\|x\|_0 = \sum_{i \in \{i | x_i \neq 0\}} 1. \quad (\ell_0)$$

Il est à noter ici que même si ℓ_0 est appelée norme dans la littérature, ce n'est pas une norme, car elle ne vérifie pas la propriété d'homogénéité. Il est clair que ce nouveau problème n'est pas plus simple que le problème initial essentiellement, car la norme ℓ_0 n'est pas continue. Pour pallier à cette difficulté, nous introduisons une famille de fonctions θ qui approchent la norme ℓ_0 . Soit $\theta : \mathbb{R} \rightarrow]-\infty, 1]$ qui satisfait les propriétés suivantes

1. θ est deux fois continûment différentiable ;
2. $\theta(0) = 0$;
3. θ est une fonction croissante et concave ;
4. θ est négatif sur \mathbb{R}_- , c.-à-d. $\theta(< 0) < 0$;

5. θ est majorée par 1 et en particulier $\lim_{x \rightarrow \infty} \theta(x) = 1$.

Ensuite, en introduisant une technique de mise à l'échelle similaire aux fonctions de perspectives en analyse convexe nous obtenons une famille de fonction $\theta_r(x) := \theta(x/r)$ pour un paramètre de relaxation $r > 0$. Les propriétés énoncées plus haut pour les fonctions θ restent inchangées, la dernière propriété peut être précisée en fonction de r , ainsi

6. $\lim_{r \rightarrow 0} \theta_r(x) = 1 \ \forall x > 0$.

Ces fonctions sont par la suite étendues comme des fonctions de \mathbb{R}^n dans \mathbb{R}^n composante par composante, c'est-à-dire que pour un vecteur $x \in \mathbb{R}^n$ on a $\theta_r(x) = (\theta_r(x_i))_{1 \leq i \leq n}$. Cette famille de fonctions θ généralise un certain nombre de fonctions de densité ou d'entropie utilisées dans divers domaines scientifiques, nous donnons ici quelques exemples de fonctions θ_r sur \mathbb{R} :

- $\theta_r^1(x) = \frac{x}{x+r}$;
- $\theta_r^{Wk}(x) = 1 - \exp^{(-x/r)^k}$ (fonctions de densité des distributions de Weibull) ;
- $\theta^{log} = \frac{\log(1+x)}{\log(1+x+r)}$.

La fonction θ^1 jouera un rôle particulier dans notre étude et est bien souvent utilisée comme "minimum" de cette classe de fonctions. En utilisant cette classe de fonctions, le problème (1) est régularisé pour $r > 0$ comme

$$y = Mx + q \geq 0, x \geq 0, \theta_r(y) + \theta_r(x) \leq 1. \quad (2)$$

D'après les propriétés des fonctions θ_r , lorsque r tend vers 0, ce problème régularisé devrait être équivalent à (LCP). Cette partie divisée en 3 chapitres se consacre à l'étude de ces méthodes de régularisation. Le premier chapitre étudie l'application des fonctions θ_r en parcimonie. Le chapitre 2 considère la suite des problèmes (2) pour résoudre l'équation en valeur absolue, qui est un problème équivalent à (LCP).

Chapitre 2 : Méthode de régularisation et application en parcimonie

Le problème de parcimonie consiste à minimiser la norme ℓ_0 d'un vecteur x sujet à un système linéaire sous-déterminé. Autrement dit, soit A une matrice d'ordre $n \times m$ avec $m < n$ et un vecteur $b \in \mathbb{R}^m$, on cherche à résoudre le problème :

$$\min \|x\|_0 \text{ s.à } x \in F := \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\}. \quad (P_{\ell_0})$$

Il est à noter que prendre $x \geq 0$ n'est en rien une perte de généralité, mais simplifie l'exposé. Ce problème a reçu récemment un intérêt tout particulier dans la communauté scientifique en particulier due à ses applications en apprentissage statistique et en acquisition comprimée (compressed sensing en anglais). De la façon présentée en introduction, on utilise les fonctions

θ_r pour approcher la norme ℓ_0 . Ainsi pour des valeurs positives de r on obtient le problème suivant :

$$\min_{x \in F} \sum_{i=1}^n \theta_r(x_i). \quad (P_r)$$

Cette approche généralise plusieurs techniques utilisées dans la littérature en considérant diverses fonctions concaves dans [82, 167, 152, 44]. Le théorème suivant montre que la séquence des problèmes (P_r) pour $r \downarrow 0$ approche le problème initial.

Théorème (Theorem 2.2.1, p.31). *Chaque point d'accumulation d'une séquence $\{x_r\}$ de solutions de (P_r) pour $r \downarrow 0$ est une solution de (P_{ℓ_0}) .*

Par ailleurs, la Proposition 2.2.1 (p.32) ajoute que cette convergence se fait en temps fini grâce à la structure polyédrale des contraintes. Le problème (P_{ℓ_0}) étant difficile à résoudre une approche plus simple très utilisée dans la littérature, [45], consiste à résoudre le problème convexe en norme ℓ_1 , à savoir

$$\min_{x \in F} \|x\|_1. \quad (P_{\ell_1})$$

où $\forall x \in \mathbb{R}^n$, $\|x\|_1 = \sum_{i=1}^n |x_i|$. Cette approche peut-être interprétée comme une convexification de (P_{ℓ_0}) , puisque la norme ℓ_1 est l'enveloppe convexe de la norme ℓ_0 pour $x \in [-1, 1]^n$. Nous montrons dans le théorème suivant que cette approche peut servir d'initialisation à notre schéma algorithmique, car lorsque r tend vers l'infini les solutions de (P_r) sont solutions de (P_{ℓ_1}) .

Théorème (Theorem 2.2.2, p.32). *Chaque point d'accumulation de la séquence $\{x_r\}_r$, tel que x_r est solution de (P_r) , pour $r \uparrow \infty$ est une solution optimale de (P_{ℓ_1}) .*

L'idée de la preuve est d'utiliser une mise à l'échelle du problème (P_r) en considérant le problème suivant :

$$\min_{x \in F} \sum_{i=1}^n r \theta_r(x_i).$$

Ces deux théorèmes nous indiquent un schéma algorithmique qui consiste à initialiser le problème en utilisant (P_{ℓ_1}) , puis résoudre les problèmes (P_r) pour des valeurs de r décroissantes. On peut noter ici qu'à une mise à l'échelle près, il existe une monotonie des solutions comme indiqué par le Théorème 2.2.3 (p.34). Partant de ces observations notre objectif principal est que la séquence de problèmes (P_r) permette d'améliorer le résultat obtenu par la résolution du problème convexe, (P_{ℓ_1}) .

La question qui reste pour ce schéma est de savoir quand la convergence peut-être atteinte. Une étude de l'erreur nous a permis d'obtenir le résultat suivant, qui est un résultat de pénalité exacte pour la séquence des problèmes (P_r) . Ce type d'approche est nouveau par rapport aux méthodes concaves proposées dans la littérature.

Théorème (Theorem 2.3.1, p.39). *On considère les fonctions θ telles que $\theta \geq \theta^1$. Soit $\mathbb{N} \ni k = \|x^*\|_0 < n$ la valeur optimale du problème (P_{ℓ_0}) et $x^r \in S_r^*$. Alors, on a*

$$\theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1} \implies x^r \in S_{\|\cdot\|_0}^*.$$

Dans le résultat précédent, on utilise la plus petite des composantes non nulles de x^r . On peut s'attendre à ce que la séquence $\{\min_{x_i^r \neq 0} x_i^r\}_r$ soit décroissante par rapport à r comme indiqué par le Lemme 2.3.3 (p.39). En pratique, on peut utiliser une borne plus grande, mais indépendante de la valeur inconnue k en remarquant que pour tout $x_r \in F$

$$\frac{\|x_r\|_0}{\|x_r\|_0 + 1} \geq \frac{k}{k + 1}.$$

Ce dernier résultat nous permet de conduire un schéma numérique complet. Dans le chapitre 2, nous proposons une implantation de schéma où à chaque itération on cherche un point stationnaire du problème concave en utilisant une technique de linéarisation appelée SLA, [140]. Les résultats numériques en fin de chapitre 2 montrent que dans la majorité des cas, cette heuristique permet d'améliorer les résultats obtenus par l'approximation convexe, (P_{ℓ_1}) . Ces résultats montrent la validité de notre approche.

Chapitre 3 : Régularisation du problème de complémentarité linéaire et application aux équations en valeur absolue

Nous nous intéressons désormais à la résolution d'équations en valeur absolue, autrement dit on cherche $x \in \mathbb{R}^n$ qui satisfait l'équation non linéaire suivante

$$Ax - |x| = b. \quad (\text{AVE})$$

Dans le cas "simple" où la matrice A possède des valeurs singulières toutes différentes de 1 et -1, on peut appliquer des méthodes de type Newton qui se sont avérées très efficaces. Malheureusement dans un cadre plus général, ces méthodes ne peuvent être appliquées et d'autre part il est bien souvent difficile de résoudre ce problème. Dans des travaux récents de Mangasarian qui traitent de ce problème, des formulations par fonctions de mérite bilinéaires et concaves ont été proposées. Nous proposons ici une adaptation de notre méthode de régularisation à ce problème (et ainsi au problème (LCP)).

En utilisant une décomposition de la valeur absolue, on se ramène facilement à un problème de complémentarité. Soit $x^+ = \max(0, x)$ et $x^- = \max(0, -x)$, il vient que $x = x^+ - x^-$ et que $|x| = x^+ + x^-$ pour tout $x \in \mathbb{R}^n$ si x^+ et x^- sont orthogonaux. D'où (AVE) est équivalent au problème de réalisabilité suivant

$$A(x^+ - x^-) - (x^+ + x^-) = b, \quad 0 \leq x^+ \perp x^- \geq 0. \quad (\text{AVE})$$

Maintenant, considérons la méthode de régularisation (2) pour $r > 0$ il vient

$$A(x^+ - x^-) - (x^+ + x^-) = b, \quad x^+ \geq 0, \quad x^- \geq 0, \quad \theta_r(x^+) + \theta_r(x^-) \leq 1.$$

Puis, nous traitons ce problème par une formulation sous forme de problème d'optimisation

$$\begin{aligned} \min_{x^+, x^-} \quad & \sum_{i=1}^n \theta_r(x_i^+) + \theta_r(x_i^-) \\ \text{s.à} \quad & A(x^+ - x^-) - (x^+ + x^-) = b, \\ & x^+ \geq 0, \quad x^- \geq 0, \\ & x^+ + x^- \geq g(r), \end{aligned}$$

où $g(r)$ est une fonction qui va moins vite que r vers 0, par exemple $g(r) = r^\alpha$ avec $\alpha \in (0, 1)$. Cette dernière contrainte $x^+ + x^- \geq g(r)$ peut paraître surprenante, mais elle évite des phénomènes de compensation dans la fonction objectif.

Il est clair que la contrainte d'égalité de ce problème risque d'être difficile à satisfaire et donc nous proposons de relâcher également cette contrainte, en considérant le problème d'optimisation suivant pour $r > 0$:

$$\begin{aligned} \min_{x^+, x^-} \quad & \sum_{i=1}^n \theta_r(x_i^+) + \theta_r(x_i^-) \\ \text{s.à} \quad & -g(r)|A|e - g(r)e \leq A(x^+ - x^-) - (x^+ + x^-) - b \leq g(r)|A|e + g(r)e, \quad (P_r) \\ & M \geq x^+ \geq 0, \quad M \geq x^- \geq 0, \\ & x^+ + x^- \geq g(r), \end{aligned}$$

où e est le vecteur dont toutes les composantes sont 1 et $|A|$ est la matrice dont toutes les composantes sont les valeurs absolues des composantes de la matrice A . La méthode que nous étudions ici est la méthode itérative qui consiste à résoudre (P_r) pour différentes valeurs de r qui tendent vers 0. À partir de maintenant, nous supposons que l'ensemble des solutions de (AVE) notées $S_{(AVE)}^*$ est non vide et notons $S_{(P_r)}^*$ l'ensemble optimal de (P_r) . Afin de simplifier les notations, nous noterons $x \in S_{(P_r)}^*$ quand $(x^+, x^-) \in S_{(P_r)}^*$ avec $x = x^+ - x^-$ et $x^+ = \max(x, 0)$, $x^- = \max(-x, 0)$. Soit M une constante strictement positive telle que

$$M \geq \|x^*\|_\infty,$$

où x^* est un élément de $S_{(AVE)}^*$. Cette hypothèse est suffisante pour garantir l'existence de solutions du problème (P_r) comme prouvé dans le Théorème 3.3.1. On peut noter que le problème peut admettre plusieurs solutions. Le résultat suivant prouve la convergence de $\{x^r\}_{r>0}$, où $x^r := x^{r+} - x^{r-}$ avec (x^{r+}, x^{r-}) solution optimale de (P_r) , vers un élément de $S_{(AVE)}^*$.

Théorème (Theorem 3.3.2, p.49). *Chaque point d'accumulation de la séquence $\{x^r\}$ tel que $x^r \in S_{(P_r)}$ pour $r > 0$ est une solution de (AVE).*

Nous avons aussi établi une estimation de l'erreur en fonction de r dans le résultat suivant.

Théorème (Theorem 3.4.1, p.52). *Soit (\bar{x}^+, \bar{x}^-) un point limite de la séquence $\{x^{r+}, x^{r-}\}$, où on note $\bar{x} = \bar{x}^+ - \bar{x}^-$ et $x^r = x^{r+} - x^{r-}$. Alors, pour r suffisamment petit on a :*

$$d_{S_{(AVE)}^*}(x^r) = O(g(r)).$$

L'idée principale de la preuve de cette borne d'erreur est d'observer que le problème (AVE) est un problème affine par morceau et d'utiliser le Lemme classique d'Hoffman [99].

Ces résultats théoriques permettent de définir un schéma algorithmique pour résoudre le problème (AVE). Dans le chapitre 3 sont définis les outils nécessaires à une implantation de ce schéma à savoir : trouver un point initial, heuristique pour résoudre les sous-problèmes, une

technique hybride d'amélioration de la convergence et la gestion des différents paramètres. Par ailleurs, cette implantation a été testée sur différents types d'exemples pour une validation de l'approche.

Le cas qui nous intéresse le plus consiste à résoudre des équations sans aucune hypothèse sur l'ensemble des solutions à priori. Nous comparons notre approche aux méthodes proposées dans [136] et [139] sur un ensemble de problèmes générés aléatoirement. Les résultats de cette comparaison montrent que notre méthode réussit à résoudre un plus grand nombre de problèmes.

Partie II : Problèmes d'optimisation sous contraintes de complémentarité

Cette seconde partie continue l'étude des problèmes de complémentarité réalisée dans les chapitres précédents dans un cadre plus général qui est l'étude des problèmes d'optimisation sous contraintes de complémentarité définie dans l'introduction et que nous rappelons ici. On cherche à minimiser une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tel que

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.à} & g(x) \leq 0, h(x) = 0, \\ & 0 \leq G(x) \perp H(x) \geq 0, \end{aligned} \tag{MPCC}$$

pour des contraintes $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^q$. En particulier notre objectif est d'obtenir des méthodes efficaces et rapides qui n'utilisent que les dérivées premières des fonctions en jeu pour obtenir des points critiques de ce problème. La condition de différentiabilité implique que nous supposons que toutes les fonctions de ce problème sont continûment différentiables.

L'étude de ce problème sera divisée ici en 4 chapitres. Le premier chapitre, Chapitre 4, couvre l'extension des conditions (KKT) et des qualifications de contraintes aux (MPCC). Par la suite, nous nous intéresserons aux méthodes de régularisation pour (MPCC). Dans le Chapitre 5, nous ferons un état de l'art sur les méthodes existantes et discuterons de la technique de régularisation proposée dans la première partie. Puis, nous introduirons la nouvelle méthode dite "des papillons" dans le Chapitre 6. Le Chapitre 7 se concentrera sur la convergence d'une large famille de méthodes de relaxation, incluant celle des papillons, dans un contexte proche des implantations pratiques. L'étude numérique de ces méthodes et des problématiques soulevées dans cette partie sera le sujet de la troisième partie de ce manuscrit.

Chapitre 4 : Stationarité et qualification de contraintes pour les problèmes d'optimisation sous contraintes de complémentarité

Nous nous intéresserons maintenant à l'extension aux problèmes d'optimisation sous contraintes de complémentarité des notions de point stationnaire et de qualification de contraintes bien connues en optimisation non linéaire. Bien que ce problème soit un problème

d'optimisation non linéaire, les conditions de (KKT) ne peuvent être utilisées directement pour ces problèmes qui d'une façon générique ne vérifient pas les conditions de qualification de contrainte. Il est essentiel de noter ici que les difficultés surviennent aux points où pour certains indices i on a $G_i(x) = H_i(x) = 0$.

La contrainte de complémentarité rend le problème non convexe en général et donc le cône tangent ne sera pas convexe. Cette observation a motivé l'introduction d'une "linéarisation" du cône tangent spécifique aux (MPCC), [181, 72, 159], que nous notons ici \mathcal{L}_{MPCC} :

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

où les ensembles d'indices correspondent à

$$\begin{aligned}\mathcal{I}^{+0}(x^*) &:= \{i \in \{1, \dots, q\} \mid G_i(x^*) > 0 \text{ and } H_i(x^*) = 0\}, \\ \mathcal{I}^{0+}(x^*) &:= \{i \in \{1, \dots, q\} \mid G_i(x^*) = 0 \text{ and } H_i(x^*) > 0\}, \\ \mathcal{I}^{00}(x^*) &:= \{i \in \{1, \dots, q\} \mid G_i(x^*) = 0 \text{ and } H_i(x^*) = 0\}, \\ \mathcal{I}_g(x^*) &:= \{i \in \{1, \dots, p\} \mid g_i(x^*) = 0\}.\end{aligned}$$

De façon similaire au cas non linéaire, on peut ainsi définir des conditions de Guignard spécifique aux (MPCC).

Definition. Soit x^* un point réalisable de (NLP). MPCC-Guignard CQ est vérifié en x^* si $\mathcal{T}^\circ(x^*) = \mathcal{L}_{MPCC}^\circ(x^*)$.

De nombreuses conditions d'optimalité ont été définies dans la littérature qui relâchent la condition de signe sur les multiplicateurs des indices de \mathcal{I}^{00} . Ces conditions sont connues sous le nom de faible-stationnarité, A-stationnarité, C-stationnarité et M-stationnarité. C'est notamment cette dernière qui concentre notre attention. En effet, en 2005 Flegel et Kanzow dans [71] montrent le résultat suivant qui indique que les conditions d'optimalité sont celles que l'on doit espérer calculer en pratique.

Théorème (Flegel et Kanzow, [71]). Soit x^* un minimum local de (NLP). Si MPCC-GCQ

est vérifié en x^* alors il existe $(\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ tel que x^* vérifie les conditions suivantes

$$\begin{aligned}
-\nabla f(x^*) &= \sum_{i=1}^p \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x^*) - \sum_{i=1}^q \lambda_i^G \nabla G_i(x^*) - \sum_{i=1}^q \lambda_i^H \nabla H_i(x^*), \\
h(x^*) &= 0, \quad 0 \leq -g(x) \perp \lambda^g \geq 0, \\
0 &\leq G(x^*) \perp H(x^*) \geq 0, \\
\lambda_i^G(x^*) &= 0 \quad \forall i \in \mathcal{I}^{+0}(x^*), \\
\lambda_i^H(x^*) &= 0 \quad \forall i \in \mathcal{I}^{0+}(x^*), \\
\text{either } \lambda_i^G > 0, \lambda_i^H > 0 &\text{ either } \lambda_i^G \lambda_i^H = 0 \quad \forall i \in \mathcal{I}^{00}(x^*).
\end{aligned}
\tag{M-stationnarité}$$

Dans le cadre (MPCC), on appellera ainsi MPCC-CQ une condition qui assure qu'un minimum local est un point M-stationnaire indépendamment du choix de la fonction objectif f . La plupart des CQ pour les problèmes (NLP) ont été étendues aux MPCC-CQ. Nous avons continué cette approche en considérant le diagramme présenté en Figure 2.

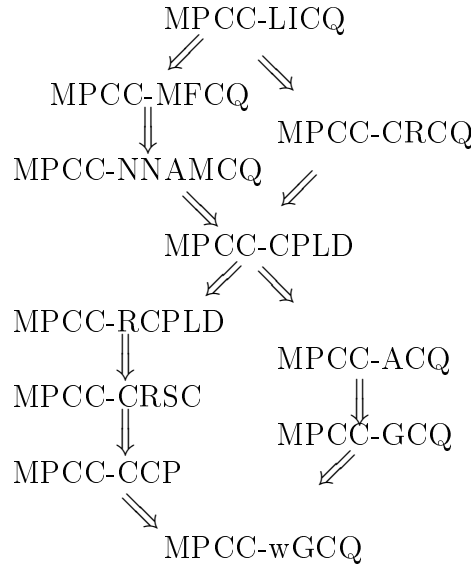


Figure 2: Qualifications de contrainte pour (MPCC).

À propos de la plus faible MPCC-CQ

Une question naturelle est de savoir si la condition MPCC-GCQ est la plus faible que l'on puisse espérer pour garantir les conditions de M-stationnarité en un minimum local. La réponse est négative comme l'illustre l'exemple suivant.

Exemple. Soit l'exemple suivant en dimension 1, défini pour une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ arbitraire.

$$\min_{x \in \mathbb{R}} f(x) \text{ s.à } 0 \leq x \perp -x^2 \geq 0.$$

Il est clair que l'ensemble des points réalisables est réduit à l'origine $\mathcal{Z} = \{0\}$. Ainsi, quel que soit le choix de la fonction f , le point $x^* = 0$ est un minimum local. Par ailleurs, il est aisé de voir qu'il existe des multiplicateurs (λ^G, λ^H) tel que x^* est un point M -stationnaire. En effet, le gradient du Lagrangien égal à zéro donne

$$-f'(0) = -\lambda^G.$$

Vérifions maintenant que MPCC-GCQ échoue en x^* . Étant donné que l'ensemble réalisable est réduit à un singleton, on a que le cône tangent est donné par $\mathcal{T}_{\mathcal{Z}}(x^*) = \{x^*\}$ et ainsi le polaire du cône tangent est $\mathcal{T}_{\mathcal{Z}}^\circ(x^*) = \mathbb{R}$. Calculons maintenant le cône \mathcal{L}_{MPCC} ,

$$\begin{aligned} \mathcal{L}_{MPCC}(x^*) &= \{d \in \mathbb{R} \mid \nabla G(x^*)^T d \geq 0, \nabla H(x^*)^T d \geq 0, \\ &\quad (\nabla G(x^*)^T d)(\nabla H(x^*)^T d) = 0\}, \\ &= \mathbb{R}_+. \end{aligned}$$

Par définition du polaire d'un cône, on obtient $\mathcal{L}_{MPCC}^\circ(x^*) = \mathbb{R}_-$. Donc MPCC-GCQ échoue au point x^* .

Partant de cette observation, nous avons défini une nouvelle condition de qualification que nous avons prouvé être la plus faible. On appelle MPCC-faible Guignard CQ (ou MPCC-wGCQ) la condition suivante.

Définition. Soit x^* un point réalisable de (NLP). MPCC-wGCQ est vérifié en x^* si $\mathcal{T}^\circ(x^*) \subset \mathcal{P}_M$ où

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

Cette condition plus faible que MPCC-GCQ est une condition de qualification de contrainte pour les MPCC et de plus est la plus faible.

Théorème. MPCC-wGuignard CQ est vérifié en x^* si et seulement si pour toute fonction f continument différentiable et qui admet un minimum contraint x^* , il existe $(\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ tel que x^* vérifie les conditions de M -stationnarité.

MPCC-CRSC et application numérique

Les conditions MPCC-GCQ et MPCC-wGCQ ont le défaut d'être difficiles à vérifier en pratique. Les conditions plus forte, mais plus simple à vérifier définies pour les (NLP) sont étendues aux (MPCC), en tenant compte du fait que les signes des multiplicateurs des indices de \mathcal{I}^{00} doivent être soit positif soit de produit nul.

Dans la majorité des cas, le bon comportement des algorithmes d'optimisation est garanti pour des CQ plus forte que Guignard et l'extension aux (MPCC) ne déroge pas à cette règle. Parmi la collection de MPCC-CQ disponible, la condition MPCC-CRSC a retenu notre attention pour une application algorithmique.

Definition. Soit $x^* \in \mathcal{Z}$. MPCC-CRSC tient en x^* si pour n'importe qu'elle partition $\mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00} = \mathcal{I}^{00}(x^*)$ tel que

$$\begin{aligned} & \sum_{i \in \mathcal{I}_g} \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x^*) - \sum_{i \in \mathcal{I}^{0+}(x^*) \cup \mathcal{I}_{++}^{00}} \lambda_i^G \nabla G_i(x^*) - \sum_{i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}_{++}^{00}} \lambda_i^H \nabla H_i(x^*) \\ & + \sum_{i \in \mathcal{I}_{-0}^{00}} \lambda_i^G \nabla G_i(x^*) + \sum_{i \in \mathcal{I}_{0-}^{00}} \lambda_i^H \nabla H_i(x^*) = 0, \end{aligned}$$

avec $\lambda_i^g \geq 0$ ($i \in \mathcal{I}_g(x^*)$), λ_i^G et $\lambda_i^H \geq 0$ ($i \in \mathcal{I}_{++}^{00}$), $\lambda_i^G > 0$ ($i \in \mathcal{I}_{-0}^{00}$), λ_i^H ($i \in \mathcal{I}_{0-}^{00}$) > 0 , il existe $\delta > 0$ tel que la famille de gradients

$$\{\nabla g_i(x) \ (i \in \mathcal{I}_1), \ \nabla h_i(x) \ (i = 1, \dots, m), \ \nabla G_i(x) \ (i \in \mathcal{I}_3), \ \nabla H_i(x) \ (i \in \mathcal{I}_4)\}$$

a le même rang pour tout $x \in \mathcal{B}_\delta(x^*)$, où

$$\begin{aligned} \mathcal{I}_1 &:= \{i \in \mathcal{I}_g(x^*) \mid -\nabla g_i(x^*) \in \mathcal{P}_M(x^*)\}, \\ \mathcal{I}_3 &:= \mathcal{I}^{0+}(x^*) \cup \{i \in \mathcal{I}_{++}^{00} \mid \nabla G_i(x^*) \in \mathcal{P}_M(x^*)\} \cup \mathcal{I}_{-0}^{00}, \\ \mathcal{I}_4 &:= \mathcal{I}^{+0}(x^*) \cup \{i \in \mathcal{I}_{++}^{00} \mid \nabla H_i(x^*) \in \mathcal{P}_M(x^*)\} \cup \mathcal{I}_{0-}^{00}. \end{aligned}$$

Cette condition de qualification nous permet de démontrer que des séquences qui convergent vers un point M-stationnaire (parfois appelée MPCC-AKKT) ont des multiplicateurs bornés. Cette application algorithmique s'avère essentielle pour démontrer la convergence de certaines méthodes de régularisation.

Corollaire. Soit $x^* \in \mathcal{Z}$ tel que MPCC-CRSC tient en x^* . Soit la séquence $\{x^k\}, \{\lambda^k\}$ tel que $x^k \rightarrow x^*$ et λ^k tend vers une certaine limite $\lambda^* \in \mathbb{R}_+^p \times \mathbb{R}^{m+2q}$ (possiblement infinie) qui satisfait

$$\begin{aligned} & \nabla f(x^k) + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i(x^k) + \sum_{j \in \mathcal{I}_g(x^*)} \lambda_j^{g,k} \nabla g_j(x^k) - \sum_{j \in \mathcal{I}^{0+}(x^*) \cup \mathcal{I}^{00}(x^*)} \lambda_j^{G,k} \nabla G_j(x^k) - \sum_{j \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*)} \lambda_j^{H,k} \nabla H_j(x^k) \rightarrow 0, \\ & \lambda_i^g = 0 \ \forall i \notin \mathcal{I}_g, \ \lambda_{\mathcal{I}^{0+}}^{G,*} = 0, \ \lambda_{\mathcal{I}^{+0}}^{H,*} = 0 \text{ either } \lambda_i^{G,*} \lambda_i^{H,*} = 0 \text{ either } \lambda_i^{G,*} > 0, \lambda_i^{H,*} > 0 \text{ for } i \in \mathcal{I}^{00}(x^*), \\ & \text{et les gradients des multiplicateurs non-nuls sont linéairement indépendants.} \end{aligned}$$

Alors, la séquence $\{\lambda^k\}$ est bornée.

Chapitre 5 : Méthodes de régularisation pour les problèmes d'optimisation sous contraintes de complémentarité

Il a été souligné que les notions bien connues en programmation non linéaire ne peuvent être appliquées directement au problème (MPCC). La motivation des méthodes de régularisation est d'introduire un paramètre pour définir une séquence de sous-problèmes qui peuvent être résolus par des solveurs classiques. À nouveau, résoudre signifie ici trouver un point stationnaire. Cette stratégie est ensuite intégrée dans un schéma itératif qui vise à réduire le paramètre de relaxation pour se rapprocher du (MPCC). Ce genre de méthode s'apparente à des méthodes homotopiques.

Le problème (R_{t_k, \bar{t}_k}) de paramètre t_k est défini génériquement de la façon suivante :

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

La régularisation de la contrainte de complémentarité est définie par l'application $\Phi_{t_k} : \mathbb{R}^q \times \mathbb{R}^q \rightarrow \mathbb{R}^q$ et vérifie la propriété suivante

$$\lim_{t_k \rightarrow 0} \Phi(a, b; t_k) = ab.$$

La plupart des méthodes existantes dans la littérature considèrent $t_k \in \mathbb{R}$, mais on peut très bien utiliser un vecteur de paramètres comme dans le cas de la méthode des papillons présentée dans le chapitre suivant. Le paramètre \bar{t}_k bien souvent pris comme t_k doit tendre vers 0 lorsque t_k tend vers 0. Ces méthodes ont été suggérées dès 2000 par Scheel et Scholtes, [181], en prenant $\forall i \in \{1, \dots, q\}$

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

Cette approche naturelle a été plus tard étendue par Demiguel, Friedlander, Nogales et Scholtes dans [52] en relâchant aussi les contraintes de positivité : $G(x) \geq -t$, $H(x) \geq -t$. Dans [185], les auteurs considèrent une régularisation concentrée uniquement sur la zone qui pose problème, i.e $G(x) = H(x) = 0$ de la façon suivante

$$\Phi_i^{SU}(G(x), H(x); t) = G_i(x) + H_i(x) - \begin{cases} |G_i(x) - H_i(x)| & \text{si } |G_i(x) - H_i(x)| \geq t \\ t\psi\left(\frac{G_i(x) - H_i(x)}{t}\right) & \text{sinon} \end{cases}, \tag{SU}$$

où ψ appartient à une certaine famille de fonctions introduite dans [185], par exemple $\psi(z) = \frac{2}{\pi} \sin(\frac{\pi}{2}z + \frac{3\pi}{2}) + 1$. Une approche différente qui considère une approximation du domaine a été proposée dans [111] par Kadrani, Dussault et Benchakroun en 2009, en considérant $\forall i \in \{1, \dots, q\}$

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

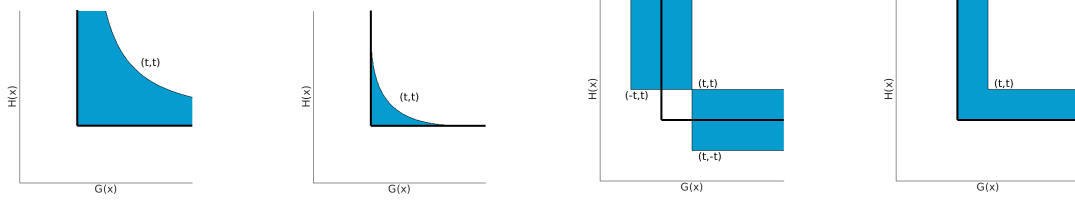


Figure 3: Ensemble réalisable des régularisations, (SS), (SU), (KDB) et (KS) dans leur forme originale, c'est-à-dire avec ou sans régularisation des contraintes de positivité.

Cette méthode a par la suite été étendue comme une régularisation en utilisant une fonction NCP, ϕ dans [112]: $\forall i \in \{1, \dots, q\}$

$$\Phi_i^{KS}(G(x), H(x); t) = \phi(G_i(x) - t, H_i(x) - t). \quad (\text{KS})$$

Le domaine réalisable de la condition de complémentarité relâchée est présenté dans la Figure 4.2. En première partie de ce document, nous avons présenté une technique de régularisation pour les problèmes de complémentarité, en considérant

$$\Phi_i^\theta(G(x), H(x); t) = \theta_r(G_i(x)) + \theta_r(H_i(x)) - 1. \quad (\text{AH})$$

Cette méthode a été introduite pour les (MPCC) par Abdallah et Haddou dans [93]. Le domaine réalisable introduit par ces relaxations peut être similaire à la méthode (SS), en observant le résultat suivant.

Lemme. Soit $\theta_r^1(x) = x/(x+r)$. Alors, on a pour $(a, b) \in \mathbb{R}^2$

$$\theta_r^1(a) + \theta_r^1(b) = 1 \iff ab = r^2.$$

Dans chacun des cas présentés ci-dessus on peut s'intéresser à deux propriétés principales :

- Convergence des méthodes vers un point stationnaire pour (MPCC) ;
- Qualification de contrainte du sous-problème régularisé, (R_{t_k, \bar{t}_k}) .

Ces deux propriétés nécessitent dans chacun des cas la satisfaction de MPCC-CQ à la solution. La première propriété de convergence est spécifiquement intéressante dans notre cas et notre objectif est de considérer des méthodes qui convergent vers des points M-stationnaires. C'est le cas des méthodes (KDB) et (KS). En effet, les méthodes (SS), (SU) et (θ) peuvent converger vers des points indésirables appelés C-stationnaires.

Chapitre 6 : La méthode des papillons

Dans ce chapitre nous introduisons une nouvelle méthode de régularisation pour (MPCC) appelée méthode des papillons qui utilise deux paramètres (t, r) . Pour $i \in \{1, \dots, q\}$, notons

$$\Phi_i^B(G(x), H(x); t, r) := \begin{cases} F_{1i}(x; r, t)F_{2i}(x; t, r), & \text{if } F_{1i}(x; r, t) + F_{2i}(x; t, r) \geq 0, \\ -\frac{1}{2}(F_{1i}(x; r, t)^2 + F_{2i}(x; t, r)^2), & \text{if } F_{1i}(x; r, t) + F_{2i}(x; t, r) < 0. \end{cases} \quad (\text{Bu.})$$

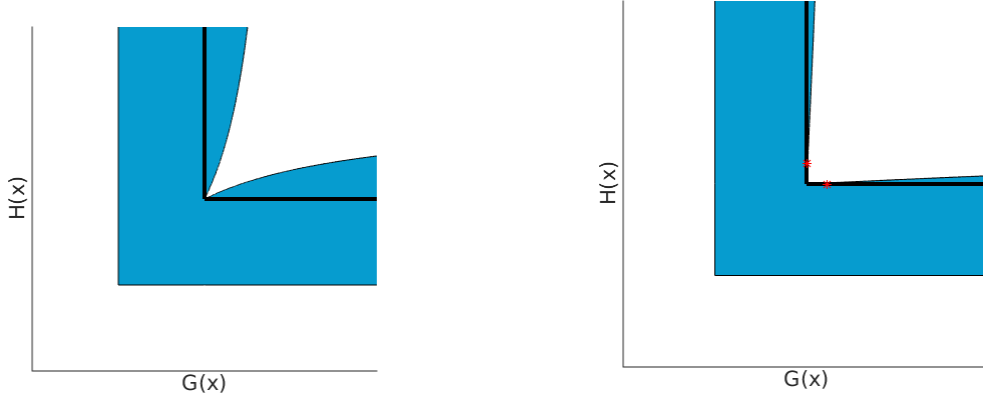


Figure 4: Domaine réalisable de la régularisation papillon pour $\theta_r(z) = \frac{z}{z+r}$ avec de gauche à droite : $t = 2r$ et $t = r^{3/2}$.

avec

$$\begin{aligned} F_{1i}(x; t, r) &:= (H_i(x) - t\theta_r(G_i(x))), \\ F_{2i}(x; t, r) &:= (G_i(x) - t\theta_r(H_i(x))). \end{aligned}$$

où $\theta_r : \mathbb{R} \rightarrow]-\infty, 1]$ est une fonction comme celles définies précédemment. Notons, $R_{(t,r)}^B$ le sous-problème régularisé (R_{t_k, \bar{t}_k}) en utilisant (Bu.). La Figure 6.1 présente le domaine réalisable de la contrainte de complémentarité pour différents choix de paramètres. Cette méthode peut être vue comme une généralisation des méthodes (KDB) et (KS). L'exemple suivant illustre même un exemple où cette nouvelle méthode évite de converger vers un point indésirable.

Exemple.

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

Dans cet exemple, il y a deux points stationnaires, le point $(1,0)$ qui est le minimum global du problème et le point $(0,0)$. Contrairement aux régularisations (KDB) et (KS) où pour $t_k = \frac{1}{k}$ une séquence $x^k = (t_k \ t_k)^T$ peut converger vers $(0,0)$, il n'existe pas de séquence de ce type pour la méthode des papillons.

Le résultat suivant montre que la méthode des papillons converge toujours vers au moins un point A-stationnaire et que pour un choix judicieux de paramètre, elle converge vers un point M-stationnaire.

Théorème. Soient $\{t_k\}$ et $\{r_k\}$ deux séquences qui décroissent vers 0. Soit $\{x^k, \lambda_k\}$ une séquence de points stationnaires de $R_{(t,r)}^B$ tel que $x^k \rightarrow x^*$.

1. Supposons que MPCC-CRCQ est vérifiée en x^* . Alors, x^* est un point A-stationnaire.

2. Supposons que MPCC-CRSC est vérifiée en x^* et $t_k = o(r_k)$. Alors, x^* est un point M-stationnaire.

Ainsi pour $t = o(r)$ la méthode des papillons atteignent notre but. Cette nouvelle méthode converge vers des points M-stationnaires. L'existence d'une séquence de points stationnaire est un autre résultat important pour la définition d'une méthode de relaxation.

Théorème. Soit $x^* \in \mathcal{Z}$ tel que MPCC-LICQ tient en x^* . Alors, il existe $\bar{t} > 0$ et un voisinage $U(x^*)$ de x^* tel que pour tout $t \in (0, \bar{t}]$, si $x \in U(x^*) \cap \mathcal{X}_{t,\bar{t}}^B$, Guignard CQ pour $R_{t,\bar{t}}^B$ tient en x .

Ce résultat d'existence de point stationnaire est classique pour ces méthodes. Malgré tout, ce résultat est légèrement décevant, car MPCC-LICQ est équivalent à Guignard CQ pour le problème (MPCC). Le résultat suivant montre néanmoins que la difficulté est située sur des points très ciblés.

Théorème. Soit $x^* \in \mathcal{Z}$ tel que MPCC-LICQ tient en x^* . Alors, il existe $\bar{t} > 0$ et un voisinage $U(x^*)$ de x^* tel que pour tout $t \in (0, \bar{t}]$, si $x \in U(x^*) \cap \mathcal{X}_{t,\bar{t}}^B$ et $\mathcal{I}_{GH}^{00}(x; \hat{t}) = \emptyset$, alors LICQ pour $R_{t,\bar{t}}^B$ tient en x .

Ces résultats d'existence de point stationnaire et de convergence permettent d'assurer la bonne définition de la méthode. Dans un contexte numérique, il est virtuellement impossible de calculer un point stationnaire, on considère bien souvent des points ϵ -stationnaires.

Définition. Soit un problème d'optimisation non linéaire (NLP) et $\epsilon \geq 0$. On dit que $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^{p+m}$ est un point ϵ -stationnaire (ou un point ϵ -KKT) s'il vérifie

$$\begin{aligned} \left\| \nabla f(x) + \sum_{i=1}^p \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) \right\|_{\infty} &\leq \epsilon, \\ |h_i(x)| &\leq \epsilon, \quad \forall i \in \{1, \dots, m\}, \\ g_i(x) &\leq \epsilon, \quad \lambda_i^g \geq 0, \quad |\lambda_i^g g_i(x)| \leq \epsilon \quad \forall i \in \{1, \dots, p\}. \end{aligned}$$

Dans le cas de convergence de séquence de points ϵ -stationnaires des relaxations, les résultats théoriques de convergence peuvent s'en trouver fortement détériorés comme observé par Kanzow et Schwarz dans [115]. La méthode des papillons en l'état n'échappe pas à la règle, comme l'indique le résultat suivant.

Théorème. Soient $\{t_k\}$ et $\{r_k\}$ deux séquences qui décroissent vers 0 tel que $t_k = o(r_k)$. Soit $\{x^k, \lambda_k\}$ une séquence de points ϵ -stationnaires de $R_{(t,r)}^B$ tel que $x^k \rightarrow x^*$.

1. Supposons que MPCC-CRSC tient en x^* et $\epsilon_k = o(\bar{t}_k)$. Alors, x^* est un point faiblement stationnaire.
2. Si en plus, on suppose que $\epsilon_k = \max(G(x^k), H(x^k))$. Alors, x^* est un point M-stationnaire.

Cette difficulté pratique des méthodes de régularisation en général et de celle des papillons en particulier est la motivation de l'étude réalisée dans le chapitre suivant. Une comparaison des différentes méthodes de régularisation est présentée dans le Chapitre 8 et montre tout l'intérêt de la méthode des papillons.

Chapitre 7 : Étude théorique de la résolution des sous-problèmes régularisés

Dans cette section, nous proposons une nouvelle approche pour pallier aux difficultés de convergence des séquences approchées de points stationnaires, pour une famille de régularisation. On considère des relaxations de la forme (R_{t_k, \bar{t}_k}) où l'application $\Phi; \mathbb{R}^q \times \mathbb{R}^q \rightarrow \mathbb{R}^q$ vérifie les hypothèses suivantes :

- $\Phi(G(x), H(x); t)$ est une application réelle continûment différentiable étendue composante par composante, telle que $\Phi_i(G(x), H(x); t) := \Phi(G_i(x), H_i(x); t)$.
- $\Phi(G(x), H(x); t)$ peut aussi être écrite comme une fonction de deux variables $G(x)$ et $H(x)$. Ainsi, le gradient par rapport à x pour $i \in \{1, \dots, q\}$ de $\Phi_i(G(x), H(x); t)$ quelque soit $x \in \mathbb{R}^n$ est donné par

$$\nabla_x \Phi_i(G(x), H(x); t) = \nabla G_i(x) \alpha_i^H(x; t) + \nabla H_i(x) \alpha_i^G(x; t),$$

où $\alpha^H(x; t)$ et $\alpha^G(x; t)$ sont des applications continues, qui satisfont $\forall x \in \mathcal{Z}$

$$\lim_{\|t\| \rightarrow 0} \alpha^H(x; t) \rightarrow H(x) \text{ et } \lim_{\|t\| \rightarrow 0} \alpha^G(x; t) \rightarrow G(x). \quad (\text{H2})$$

- À la limite, quand $\|t\|$ tend vers 0, l'ensemble réalisable du problème non linéaire paramétrique (R_{t_k, \bar{t}_k}) doit converger vers l'ensemble réalisable du (MPCC). Autrement dit, soit $\mathcal{F}(t)$ l'ensemble réalisable de (R_{t_k, \bar{t}_k}) , alors

$$\lim_{\|t\| \rightarrow 0} \mathcal{F}(t) = \mathcal{Z}, \quad (\text{H3})$$

où on considère une limite point par point.

- À la frontière de l'ensemble réalisable pour la relaxation de la contrainte de complémentarité, il est vérifié que

$$\Phi(G(x), H(x); t) = 0 \iff F_G(x; t) = 0 \text{ ou } F_H(x; t) = 0 \quad (\text{H4})$$

où

$$\begin{aligned} F_G(x; t) &= G(x) - \psi(H(x); t), \\ F_H(x; t) &= H(x) - \psi(G(x); t), \end{aligned}$$

et ψ est une fonction à valeur réelle continûment différentiable étendue composante par composante. On peut noter que la fonction ψ utilisée pour définir F_G pourrait être différente de celle utilisée pour définir F_H tant qu'elles vérifient les hypothèses ci-dessous.

Ces fonctions $\psi(H(x); t)$, $\psi(G(x); t)$ sont positives pour tout $x \in \{x \in \mathbb{R}^n \mid \Phi(G(x), H(x); t) = 0\}$ et satisfont $\forall z \in \mathbb{R}^q$

$$\lim_{\|t\| \rightarrow 0} \psi(z; t) = 0. \quad (\text{H5})$$

- Lorsque $\|t\|$ tend vers 0, la dérivée de ψ par rapport à la première variable satisfait

$$\lim_{\|t\| \rightarrow 0} \left. \frac{\partial \psi(x; t)}{\partial x} \right|_{x=z} = 0 \quad \forall z \in \mathbb{R}^q. \quad (\text{H6})$$

Nous appellerons méthodes-UF, les méthodes qui satisfont ces hypothèses et on peut montrer que (KDB), (KS) et (Bu.) font partie de cette famille. On peut également ajouter une nouvelle méthode qui vérifie ces hypothèses, la régularisation "asymétrique" : Soit I_G et I_H deux ensembles d'indices tels que $I_G \cup I_H = \{1, \dots, q\}$ et $I_G \cap I_H = \emptyset$. Alors, la régularisation asymétrique est définie par

$$\Phi_i(G(x), H(x); t) = \begin{cases} (G(x) - t)H(x) & \text{pour } i \in I_G, \\ G(x)(H(x) - t) & \text{pour } i \in I_H. \end{cases}$$

Pour assurer le bon comportement de ces méthodes, nous proposons de considérer des séquences de points stationnaires approchés appelés points fortement ϵ -stationnaire

Définition. x^k est un point fortement ϵ_k -stationnaire pour (R_{t_k, \bar{t}_k}) avec $\epsilon_k \geq 0$ s'il existe $\nu^k \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{3q}$ tel que

$$\|\nabla \mathcal{L}_{R_t}^1(x^k, \nu^k; t_k)\|_\infty \leq \epsilon_k$$

et

$$\begin{aligned} |h(x^k)| &\leq \bar{t}_k + O(\epsilon_k) \quad \forall i \in \{1, \dots, m\}, \\ g_i(x^k) &\leq \epsilon_k, \quad \nu_i^{g,k} \geq 0, \quad |g_i(x^k) \nu_i^{g,k}| \leq \epsilon_k \quad \forall i \in \{1, \dots, p\}, \\ G_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu^{G,k}_i \geq 0, \quad |\nu^{G,k}_i (G_i(x^k) + \bar{t}_k)| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\ H_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu^{H,k}_i \geq 0, \quad |\nu^{H,k}_i (H_i(x^k) + \bar{t}_k)| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\ \Phi_i(G(x^k), H(x^k); t_k) &\leq 0, \quad \nu_i^{\Phi,k} \geq 0, \quad \left| \nu_i^{\Phi,k} \Phi_i(G(x^k), H(x^k); t_k) \right| = 0 \quad \forall i \in \{1, \dots, q\}. \end{aligned}$$

La différence avec les points ϵ -stationnaires est qu'un point fortement ϵ -stationnaire vérifie exactement la réalisabilité et la condition de complémentarité. Cette définition peut en partie être motivée par le fait que la contrainte de complémentarité a déjà été relaxée par les paramètres t et ajouter un ϵ serait une relaxation supplémentaire.

Par ailleurs, on voit sur la Figure 5, qu'en relâchant la satisfaction de la contrainte de complémentarité régularisée on obtient un domaine similaire à la méthode (SS). Le résultat suivant montre l'intérêt de considérer des séquences de points fortement ϵ -stationnaires.

Théorème. Soit $\{t_k\}$ et $\{\epsilon_k\}$ deux séquences de paramètres qui décroissent vers zéro lorsque $k \in \mathbb{N}$ tend vers l'infini. Supposons que $\epsilon_k = o(\bar{t}_k)$. Par ailleurs, soit $\{x^k, \nu^k\}_k$ une séquence de points fortement ϵ_k stationnaires de (R_{t_k, \bar{t}_k}) avec $x^k \rightarrow x^*$ et que MPCC-CRSC est vérifiée en x^* . Alors, x^* est un point M -stationnaire de (MPCC).

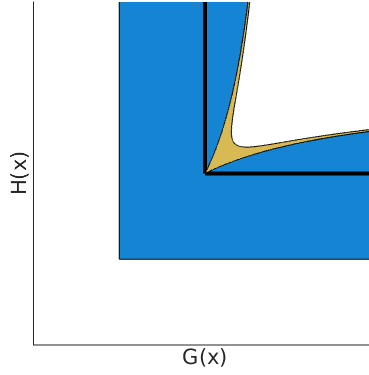


Figure 5: Domaine réalisable de la complémentarité relâchée satisfaite à ϵ près.

Nous nous sommes ensuite intéressés à l'existence des points fortement ϵ_k stationnaires au voisinage d'un point M-stationnaire. Pour montrer ce type de résultat nous ferons les hypothèses supplémentaires suivantes sur la fonction ψ pour tout $t > 0$ et $x \in \mathbb{R}^q$:

•

$$\frac{\partial \psi(x; t)}{\partial t} > 0. \quad (\text{A1})$$

•

$$\frac{\partial \psi(x; t)}{\partial x} \geq 0. \quad (\text{A2})$$

• Pour un paramètre s strictement positif, ψ doit en plus satisfaire

$$\psi(\psi(\|t\|_\infty; t); t) \leq s, \quad (\text{A3})$$

$$\psi(-\|t\|_\infty; t) \leq s. \quad (\text{A4})$$

Le résultat d'existence que nous donnons ici concerne une version équivalente de (R_{t_k, \bar{t}_k}) formulée avec des variables d'écarts de la façon suivante :

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

avec $\bar{t} \downarrow_{\|t\| \rightarrow 0} 0$ et l'application de régularisation $\Phi(s_G, s_H; t) : \mathbb{R}^q \rightarrow \mathbb{R}^q$ définie en remplaçant $G(x)$ et $H(x)$ par s_G et s_H dans l'application $\Phi(G(x), H(x); t)$. L'existence de points fortement ϵ_k stationnaires au voisinage d'un point M-stationnaire est donné par le résultat qui ne nécessite aucune condition de qualification de contraintes.

Théorème. Soit $x^* \in \mathcal{Z}$ un point M -stationnaire et $\epsilon > 0$ arbitrairement petit. Par ailleurs, supposons que les hypothèses sur ψ et la relaxation faites plus haut sont vérifiées. Alors, il existe des constantes strictement positives c, \bar{t}^* avec $\bar{t}^* > c\epsilon$ et un voisinage $U(x^*)$ de $(x^*, G(x^*), H(x^*))^T$ tel que pour tous $t \in (0, \bar{t}^*)$ et toutes $\bar{t} \in (0, \bar{t}^*)$ il existe $(x, s)^T \in U(x^*)$, qui est un point fortement ϵ_k stationnaire de $(R_t^s(x, s))$.

Dans le chapitre 7, nous présentons un exemple (Exemple 7.5.2) qui montre que sans les variables d'écart ce type de résultat n'est pas évident. Nous présentons au Chapitre 9 une stratégie de pénalisation-activation de contraintes qui grâce à l'utilisation des variables d'écart permet de calculer des points fortement ϵ -stationnaires.

Perspectives pour la partie II :

L'étude des méthodes de régularisation a des conséquences sur la résolution numérique de plusieurs problèmes d'optimisation réputés difficiles. D'une part, les méthodes de régularisations peuvent être adaptées aux problèmes d'optimisation avec contraintes évanescentes et aux problèmes d'optimisation avec contraintes de cardinalité. D'autre part, la résolution efficace des (MPCC) peut être utilisée pour résoudre les problèmes d'optimisation bi niveaux. Le Chapitre 7.7.2 donne plus de détails concernant la formulation de ces problèmes ainsi que l'utilisation de la méthode des papillons pour les résoudre.

Partie III : Numérique

Nous avons présenté dans la partie 2, une étude approfondie d'une part des conditions d'optimalité pour les (MPCC), et d'autre part des méthodes de régularisation pour les résoudre. En particulier, nous avons introduit la nouvelle méthode de régularisation "papillon".

Cette partie s'intéresse maintenant à l'application numérique de ces méthodes et sera articulée en 2 chapitres. Le Chapitre 8 propose une comparaison numérique des différentes méthodes de régularisation discutées tout au long de ce document. Ensuite, le Chapitre 9 discute d'une stratégie de pénalisation et d'activation de contraintes qui permet de calculer des points fortement ϵ -stationnaires. Cette stratégie fait l'objet d'une implémentation en JULIA d'un solveur pour résoudre (MPCC).

Chapitre 8 : Comparaison numérique des méthodes de régularisation

Nous nous sommes intéressés ici à une comparaison numérique entre les différentes méthodes de régularisation en prenant en compte les méthodes proposées dans la littérature en particulier les méthodes (SS) et (KS) ainsi que trois versions de la méthode des papillons :

1. $B_{(t=r)} : (s = 0, t = r) ;$
2. $B_{(t=r^{3/2})} : (s = 0, t = r^{3/2}) ;$
3. $B_{(s=t, 2t=r)} : (s = t, 2t = r).$

Il est à noter que la troisième version utilise un nouveau paramètre s en considérant une version étendue de la méthode définie par

$$\begin{aligned} F_{1(s,t,r)_i}(x) &= (H_i(x) - s - t\theta_r(G_i(x) - s)), \\ F_{2(s,t,r)_i}(x) &= (G_i(x) - s - t\theta_r(H_i(x) - s)). \end{aligned}$$

Si nous ne considérons que les méthodes (SS) et (KS); ce choix a été décidé après différentes simulations numériques préliminaires qui ont montré que ces deux méthodes étaient les plus efficaces sur nos notre large jeu de problèmes tests.

La comparaison est réalisée sur la collection de problèmes MacMPEC, [123], composée de 101 problèmes, d'où ont été omis les problèmes excédant 300 variables ou 300 contraintes et ceux avec des problèmes d'évaluation des contraintes ou de la fonction objectif. Pour résoudre les sous-problèmes, nous utilisons AMPL à partir de Matlab. En particulier, nous avons utilisé les solveurs d'optimisation non linéaire : SNOPT, MINOS et IPOPT. L'algorithme s'arrête si la valeur des paramètres est trop petite (ici 10^{-15}) ou si le point courant est considéré comme ϵ -solution (ici $\epsilon = 10^{-7}$) de (MPCC). Une solution approchée est définie telle que :

$$\min_local(x) := \max(\nu_f(x), \nu_{comp}(x), \nu_c(x)) \leq \epsilon, \quad (3)$$

où

- a) Réalisabilité de sous-problèmes régularisés : $\nu_f(x) := \max(-g(x), |h(x)|, -\Phi(x))$,
- b) Satisfaction des contraintes de complémentarité : $\nu_{comp}(x) := \min(G(x), H(x))^2$,
- (c) La complémentarité entre les multiplicateurs de Lagrange et les contraintes :

$$\begin{aligned} \nu_c(x) &:= \max(\|g(x) \circ \lambda^g\|_\infty, \|h(x) \circ \lambda^h\|_\infty, \|G(x) \circ \lambda^G\|_\infty, \\ &\quad \|H(x) \circ \lambda^H\|_\infty, \|\Phi^B(G(x), H(x); \hat{t})(x) \circ \lambda^\Phi\|_\infty). \end{aligned}$$

Il est délicat d'être plus exigeant sur la contrainte de complémentarité ici écrite comme le minimum composante par composante tandis que la formulation relâchée considère le produit des deux. Par ailleurs, la réalisabilité duale (i.e gradient du Lagrangien égal à zéro) n'est pas prise en compte ici, car elle n'est pas un critère d'arrêt pour SNOPT et MINOS.

Afin de mesurer l'impact des paramètres sur ces méthodes nous considérons une étude de sensibilité avec plusieurs choix de valeurs initiales (7 valeurs) et plusieurs stratégies pour mettre à jour les paramètres de régularisations (5 valeurs). Le tableau 8.4 présente les résultats en pourcentage de problèmes résolus pour un jeu de paramètre (paramètre initial et mise à jour). Un problème est résolu s'il vérifie le critère défini en (3). Ces résultats numériques illustrent dans l'ensemble l'intérêt des méthodes de régularisation et en particulier celui des régularisations papillon.

Solver SNOPT	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
meilleur	92.08	94.06	96.04	96.04	97.03	96.04
moyenne	92.08	90.78	91.17	92.08	90.04	92.33
pire	92.08	83.17	86.14	87.13	82.18	87.13
écart-type	0	3.15	2.59	2.45	2.86	2.77
Solveur MINOS	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
meilleur	85.15	94.06	93.07	88.11	94.06	87.13
moyenne	85.15	90.94	90.18	81.92	90.04	80.11
pire	85.15	87.13	86.14	76.23	85.15	74.26
écart-type	0	1.50	1.62	2.65	2.31	2.95
Solveur IPOPT	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
meilleur	91.09	93.07	93.07	94.06	93.07	94.06
moyenne	91.09	91.82	89.84	89.05	88.80	89.02
pire	91.09	90.10	86.14	84.16	84.16	81.19
écart-type	0	1.14	2.19	3.09	2.72	3.86

Table 1: Étude de sensibilité sur la collection de problèmes MacMPEC avec l’optimalité de (MPCC) comme critère de succès. Les résultats sont des pourcentages de succès. meilleur : pourcentage de succès avec le meilleur jeu de paramètre. pire : pourcentage de succès avec le pire jeu de paramètre. moyenne et écart-type indiquent l’écart-type et la moyenne de pourcentage de succès.

Chapitre 9 : Implémentation d’une méthode de régularisation-pénalisation-activation de contrainte

Dans le Chapitre 7, il a été défini une version spécifique d’approximation de points stationnaire pour les sous-problèmes régularisés de (MPCC). Nous présentons ici une stratégie qui permet de calculer numériquement ces points dits fortement ϵ -stationnaires. Nous cherchons donc à résoudre les sous-problèmes régularisés écrits avec des variables d’écart, $(R_t^s(x, s))$. Ce problème régularisé est transformé en un problème pénalisé où sont laissées en contrainte celles qui font l’objet d’une restriction d’après la définition de points fortement ϵ -stationnaire.

$$\begin{aligned}
\min_{x,s} \Psi_\rho(x, s) &:= f(x) + \frac{1}{2\rho} \phi(x, s) \\
\text{s.à. } s_G &\geq -\bar{t}, \quad s_H \geq -\bar{t}, \quad \Phi(s_G, s_H; t) \leq 0,
\end{aligned} \tag{P(x, s)}$$

où ϕ est une fonction de pénalité, par exemple :

$$\phi(x, s) := \|\max(g(x), 0), h(x), G(x) - s_G, H(x) - s_H\|^2.$$

Une adaptation du théorème de convergence du Chapitre 7 donne le résultat suivant qui valide l’approche pénalisée.

Théorème. Soit $\{\rho_k\}$ et $\{\epsilon_k\}$ deux séquences de paramètres qui décroissent vers zéro lorsque $k \in \mathbb{N}$ tend vers l'infini. Supposons que $\epsilon_k = o(\bar{t}_k)$. Par ailleurs, soit $\{x^k, \nu^k\}$ une séquence de points fortement ϵ_k -stationnaires de $(R_t^s(x, s))$ avec $x^k \rightarrow x^*$ où MPCC-CRSC tient. Alors, si x^* est réalisable pour $(R_t^s(x, s))$, x^* est un point M -stationnaire de (MPCC).

Les contraintes restantes dans $(P(x, s))$ sont gérées par une stratégie d'activation de contrainte. Les contraintes de bornes ne posent aucun problème par contre pour la régularisation de la contrainte de complémentarité il est essentiel de noter que par hypothèse sur la régularisation,

$$\Phi(s_G, s_H; t) = 0 \iff s_H = \psi(s_G; t) \text{ or } s_G = \psi(s_H; t).$$

Ainsi, lorsque cette contrainte est active certaines variables sont "fixées" et on utilise une technique de substitution pour travailler dans le sous-espace des variables libres. Le calcul du gradient de la fonction objectif dans ce sous-espace se fait au travers des règles de composition de dérivées. L'algorithme complet de la résolution d'un sous-problème régularisé suit itérativement les étapes suivantes :

1. Projection du point initial sur l'ensemble réalisable s'il n'est pas réalisable ;
2. Soit $\mathcal{W}(s; t, \bar{t})$ l'ensemble des contraintes actives parmi les contraintes

$$s_G \geq -\bar{t}, s_H \geq -\bar{t}, \Phi(s_G, s_H; t) \leq 0 ;$$

3. Substitution des variables fixées par les contraintes actives ;
4. Minimisation du problème $(P(x, s))$
 - Calcul du gradient dans le sous-espace de travail en utilisant les règles de composition de dérivées ;
 - Restreindre le pas pour satisfaire les contraintes ;
5. Calcul exact des multiplicateurs de Lagrange ;
6. Relaxation (si nécessaire) de certaines contraintes actives ;
7. Réduction (si nécessaire) du paramètre de pénalité ρ .

Les étapes 1 et 4 garantissent que les contraintes du problème $(R_t^s(x, s))$ sont toujours satisfaites. L'étape 5 assure que la condition de complémentarité entre les multiplicateurs de Lagrange et les contraintes est toujours satisfaite. Ces deux missions étant celles requises pour calculer un point fortement ϵ -stationnaire.

Notations

We consider here classical notations:

- \mathbb{R}^n : the n -dimensional real Euclidian vector space;
- For two vectors $x, y \in \mathbb{R}^n$: $x^T y$ is the scalar product of x and y ;
- \emptyset denotes the empty set;
- For $x \in \mathbb{R}^n$, $\text{supp}(x) := \{i \in \{1, \dots, n\} \mid x_i \neq 0\}$;
- e the vector whose components are all one.

We also use asymptotic Landau notations:

- $f(x) = O(g(x))$ as $x \rightarrow a$ if and only if there exists positive numbers δ and M such that $|f(x)| \leq M|g(x)|$ for all $|x - a| \leq \delta$;
- $f(x) = \Omega(g(x))$ as $x \rightarrow a$ if and only if there exists positive numbers δ and M such that $|f(x)| \geq M|g(x)|$ for all $|x - a| \leq \delta$;
- $f(x) = o(g(x))$ as $x \rightarrow a$ if and only if for all positive constant M there exists a positive number δ such that $|f(x)| \leq M|g(x)|$ for all $|x - a| \leq \delta$, in other words $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$;
- $f(x) = \omega(g(x))$ as $x \rightarrow a$ if and only if for all positive constant M there exists a positive number δ such that $|f(x)| \geq M|g(x)|$ for all $|x - a| \leq \delta$;
- $f(x) \sim_K (g(x))$ as $x \rightarrow a$ if and only $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = K$ with K a positive finite constant and in a classical way $f(x) \sim (g(x))$ when $K = 1$.

Acronyms

- NLP: Non-linear program;
- MPCC: Mathematical program with complementarity constraints;
- MPVC: Mathematical program with vanishing constraints;
- OPCC: Optimisation problem with cardinality constraints;
- BP: Bilevel program;
- CP: Complementarity problem;
- LCP: Linear complementarity problem;
- NCP: Non-linear complementarity problem;
- VI: Variational inequality;
- AVE: Absolute value equation;
- KKT: Karush-Kuhn Tucker;
- AKKT: Approximate Karush-Kuhn Tucker;
- CQ: Constraint qualification;
- MPCC CQ: MPCC constraint qualification;
- LICQ: Linear independence constraint qualification;
- MFCQ: Mangasarian Fromowitz constraint qualification;
- NNAMCQ: Non-abnormal multiplier constraint qualification;
- CRCQ: Constant rank constraint qualification;
- CPLD: Co-positive linear dependence;
- RCPLD: Relaxed co-positive linear dependence;

- CRSC: Constant rank in the subspace of components;
- CCP: Cone continuity property;
- ACQ: Abadie constraint qualification;
- GCQ: Guignard constraint qualification;

Contents

1	Introduction to optimisation and complementarity	1
1.1	Convex analysis and cones	2
1.2	Optimisation and variational inequalities	10
1.3	Non-linear programming	13
1.4	On the θ 's functions	19
I	Complementarity Problems	24
2	Relaxation methods and application to sparse optimisation	26
2.1	A smoothing method	29
2.2	Convergence	31
2.3	Error estimate	37
2.4	Algorithm	40
2.5	Numerical simulations	41
2.6	Conclusion and outlook	42
3	Relaxation methods to solve the absolute value equation	44
3.1	Introduction	45
3.2	AVE as a sequence of concave minimization programs	46
3.3	Convergence	48
3.4	Error estimate	50
3.5	Algorithm	54
3.6	Numerical simulations	55
II	Mathematical Programs with Complementarity Constraints	62
4	Stationary and constraint qualifications for MPCC	65
4.1	Stationary conditions for MPCC	65
4.2	Constraint qualifications for MPCC	70
4.3	On the weakest MPCC CQ	73
4.4	Algorithmic applications of the MPCC CQs	78

4.5	Diagram of relationship between the MPCC CQs	82
5	Relaxation methods for MPCC	85
5.1	Regular relaxations	86
5.2	A local relaxation	91
5.3	The paradigm of M-relaxations	92
6	The new butterfly relaxation method	97
6.1	The butterfly relaxation method	97
6.2	Convergence	101
6.3	Existence of Lagrange multipliers of the relaxed sub-problems	108
6.4	Convergence of the epsilon-stationary points	115
7	Theoretical study of the relaxed sub-problems	121
7.1	A unified framework for relaxation/approximation methods	121
7.2	Existing methods under the unified framework	123
7.3	Motivations on epsilon-solution to the regularised sub-problems	127
7.4	Convergence of strong epsilon-stationary sequences	129
7.5	On Lagrange multipliers of the regularisation	133
7.6	The MPCC with slack variables	135
7.7	Existence of strong epsilon-stationary points for the regularisation with slack variables	136
	Notes and extensions of Part II	144
III	Numerics	152
8	Numerical comparison between the relaxation methods	154
8.1	On the implementation of the butterfly relaxation	154
8.2	Comparison of the relaxation methods	155
8.3	Benchmarking MPCCs solvers	157
8.4	An example of numerical difficulties	162
9	How to compute strong epsilon-stationary points	165
9.1	A penalized formulation	165
9.2	Active-set method for the penalised problem	166
9.3	Numerical results	170
	Concluding remarks	172
A	A generalized direction in interior point method for monotone linear com- plementarity problems	177

A.1	Introduction	178
A.2	Preliminaries and Problem Settings	179
A.3	Polynomial Complexity	180
A.4	Conclusions	191
A.5	Numerics	192
B	Formations complémentaires et participations aux conférences	194

Chapter 1

Introduction to optimisation and complementarity

In this section we present basic results from convex analysis, variational analysis and non-linear programming that are used through the following chapters. Among the major references that have been used while studying these topics, we may cite some important books such as [170] for convex analysis, [171] for variational analysis, [65] for variational inequalities and complementarity problems and finally [21] for various subjects on optimisation. A special care has been taken to make the manuscript understandable for any undergraduate student in applied mathematics.

This chapter is divided in 4 sections. Section 1.1 presents essential notions from convex analysis. In particular, this section gives definitions and results on convex sets, cones and polyhedral sets, convex functions and briefly discuss error bounds. Section 1.2 discusses the link between optimization problems, variational inequalities and complementarity problems. Section 1.3 introduces optimality conditions for non-linear problems and gives several definitions of constraint qualifications. The experienced reader may skip Section 1.1, 1.2 and 1.3. The final section of this chapter, Section 1.4, introduces a family of concave functions and some of their properties that will be extensively used through this manuscript.

1.1 Convex analysis and cones

In this section, we introduce elementary notions about convex sets, convex functions and cones.

1.1.1 Convex sets

Definition 1.1.1. A subset C of \mathbb{R}^n is convex if $\forall x, y \in C, \forall t \in [0, 1]$

$$tx + (1 - t)y \in C.$$

A geometrical interpretation of this definition is that a set is convex if any segment joining two points of this set also belongs to this set as illustrated on Figure 1.1.

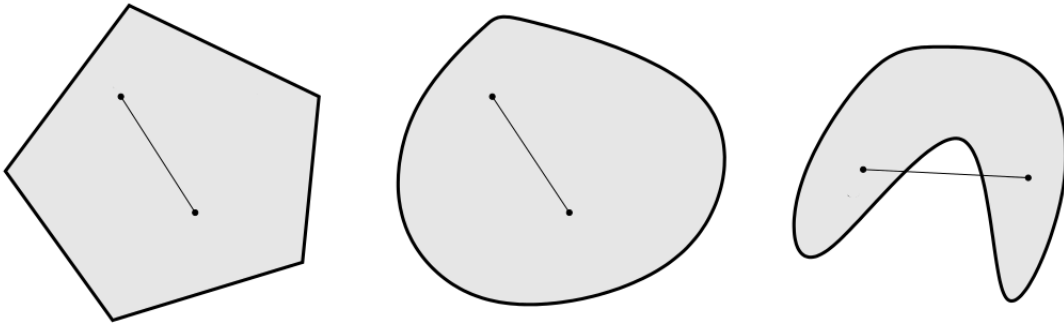


Figure 1.1: Some examples of two convex sets and a non-convex set.

Some elementary operations that preserve the convexity of a set are collected in the following proposition.

Proposition 1.1.1. For any collection $\{C_i \mid i \in I\}$ of convex sets $C_i \subset \mathbb{R}^{n_i}$ we have:

1. $C_1 \times \cdots \times C_m$ is convex in $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$;
2. $\cap_{i \in I} C_i$ is convex, here with $n_i = n$ for all i ;
3. The finite sum $\sum_{i=1}^m C_i$ is convex, with $n_i = n$ for all i .

For a set $C \subset \mathbb{R}^n$, the convex hull of C , denoted by $\text{conv } C$, is the intersection of all convex sets containing C .

Proposition 1.1.2. For a set $C \subset \mathbb{R}^n$, it holds that $\text{conv } C$ is the set of all convex combinations of elements of C , i.e.,

$$\text{conv } C = \left\{ \sum_{i=1}^m t_i x_i \mid x_i \in C, t_i \geq 0, \sum_{i=1}^m t_i = 1 \right\}.$$

A fundamental characterisation of convex sets is provided by Caratheodory's theorem.

Theorem 1.1.1. *For any $C \subset \mathbb{R}^n$, any element of $\text{conv } C$ can be represented as a convex combination of no more than $(n + 1)$ elements of C .*

We now give a lemma that we extensively use to prove convergence results later in this thesis. This result may be interpreted as a Caratheodory kind lemma. Proof of this result can be found in [12], Lemma A.1 [185] or Lemma 7.1, [184]. We present here the proof of the latter mentioned.

Lemma 1.1.1. *Let $\{a_i \mid i = 1, \dots, p\}$, $\{b_i \mid i = 1, \dots, m\}$ and c be vectors in \mathbb{R}^n and $\alpha \in \mathbb{R}_+^p$, $\beta \in \mathbb{R}^m$ multipliers such that*

$$\sum_{i=1}^p \alpha_i a^i + \sum_{i=1}^m \beta_i b^i = c.$$

Then there exist multipliers $\alpha^ \in \mathbb{R}_+^p$ and β^* with $\text{supp}(\alpha^*) \subseteq \text{supp}(\alpha)$, $\text{supp}(\beta^*) \subseteq \text{supp}(\beta)$ and*

$$\sum_{i=1}^p \alpha_i^* a^i + \sum_{i=1}^m \beta_i^* b^i = c$$

such that the vectors

$$\{a^i \mid i \in \text{supp}(\alpha^*)\} \cup \{b^i \mid i \in \text{supp}(\beta^*)\}$$

are linearly independent.

Proof. If the vectors

$$\{a^i \mid i \in \text{supp}(\alpha)\} \cup \{b^i \mid i \in \text{supp}(\beta)\}$$

are already linearly independent, we can choose $a^* = a$, $b^* = b$ and are done. Otherwise, there are scalar δ_i , $i \in \text{supp}(a)$ and τ_i , $i \in \text{supp}(b)$ not all equal to zero such that

$$\sum_{i \in \text{supp}(a)} \delta_i a^i + \sum_{i \in \text{supp}(b)} \tau_i b^i = 0.$$

If all δ_i are equal to zero, we can choose an arbitrary $i^* \in \text{supp}(\tau) \subset \text{supp}(\beta)$ and define

$$\bar{\alpha} := \alpha, \text{ and } \bar{\beta} := \begin{cases} \beta_i - \frac{\beta_{i^*}}{\tau_{i^*}} \tau_i, & \text{if } i \in \text{supp}(\beta), \\ 0 & \text{else.} \end{cases}$$

Otherwise, we can assume without loss of generality that there is at least one $\alpha_i > 0$ and choose i^* as an index with

$$\frac{\alpha_{i^*}}{\delta_{i^*}} = \min \left\{ \frac{\alpha_i}{\delta_i} \mid i \in \text{supp}(\alpha), \delta_i > 0 \right\}.$$

In this case, we define the new multipliers as

$$\bar{\alpha} := \begin{cases} \alpha_i - \frac{\alpha_{i^*}}{\delta_{i^*}} \delta_i, & \text{if } i \in \text{supp}(\alpha), \\ 0 & \text{else,} \end{cases} \quad \text{and } \bar{\beta} := \begin{cases} \beta_i - \frac{\alpha_{i^*}}{\tau_{i^*}} \tau_i, & \text{if } i \in \text{supp}(\beta), \\ 0 & \text{else.} \end{cases}$$

In both cases we have $\bar{\alpha} \geq 0$ and $\text{supp}(\bar{\alpha}, \bar{\beta}) \subsetneq \text{supp}(\alpha, \beta)$. Additionally, these multipliers still have the property that

$$\sum_{i=1}^p \bar{\alpha}_i a^i + \sum_{i=1}^m \bar{\beta}_i b^i = c.$$

If the vectors

$$\{a^i \mid i \in \text{supp}(\bar{\alpha})\} \cup \{b^i \mid i \in \text{supp}(\bar{\beta})\}$$

are linearly independent, we can finish here. Otherwise, we have to repeat the procedure above. Since the support of $(\bar{\alpha}, \bar{\beta})$ decreases each time, after a finite number of iterations either the vectors corresponding to non-vanishing multipliers are linearly independent or $\text{supp}(\bar{\alpha}, \bar{\beta}) = \emptyset$, in which case the assertion is trivially satisfied. \square

1.1.2 Cones and polyhedral sets

Cones are fundamental geometric objects associated with sets. They play a key role in several aspects of mathematics.

Definition 1.1.2. *A set $K \subset \mathbb{R}^n$ is called a cone if $tx \in K$ for all $x \in K$ and for all $t > 0$.*

It can be observed that if K is a non-empty closed cone then $0 \in K$. Examples of convex cones include linear subspaces of \mathbb{R}^n and the non-negative orthant $\mathbb{R}_+^n := \{x \mid x_i \geq 0, i = 1, \dots, n\}$. Other cones playing an important role in convex optimisation problems are the cone of symmetric real positive semi-definite matrices of order n and the cone of Lorentz. More examples of cones that are classical in the context of optimisation are given in the rest of this section.

In the study of convergence of sequences involved in an optimisation context, one may be interested to handle situations where a given sequence $\{x^k\} \subset \mathbb{R}^n$ is unbounded. To derive some convergence properties we are led to consider direction $d = x^k / \|x^k\|$ with $x^k \neq 0$. This motivates the study of the asymptotic cone of a set C .

Definition 1.1.3. *Let C be a non-empty set of \mathbb{R}^n . Then, the asymptotic cone of C , denoted C^∞ , is the set of vectors $d \in \mathbb{R}^n$ that are limits in direction of the sequences $\{x^k\} \subset C$, that is*

$$C^\infty := \{d \in \mathbb{R}^n \mid \exists t_k \rightarrow \infty, \exists x^k \in C \text{ with } \lim \frac{x^k}{t_k} = d\}.$$

This cone is sometimes called recession cone in the literature as in [170]. From the definition it is clear that C^∞ is a closed cone. One characterization of this cone in the convex case is given in the following theorem.

Theorem 1.1.2. [170, Theorem 8.4] Let C be non-empty convex set. The asymptotic cone is a convex cone containing the origin. It is the same as the set of vectors y such that $C + y \subset C$.

One property of importance that motivates the use of asymptotic cone is the following.

Proposition 1.1.3. [21, Prop. 2.1.2] A set $C \subset \mathbb{R}^n$ is bounded if and only $C^\infty = \{0\}$.

These property will be useful to study existence of optimal solution of optimisation problems as in Chapter 2.

We now study two more examples of cones that are the tangent cone and the normal cone. Beforehand, we introduce the notion of polar cone.

Definition 1.1.4. Given a cone $K \subset \mathbb{R}^n$, the polar of K is the cone defined by

$$K^\circ := \{y \in \mathbb{R}^n \mid y^T x \leq 0, \forall x \in K\}.$$

The bipolar is the cone $K^{\circ\circ} := (K^\circ)^\circ$.

Proposition 1.1.4. For a cone $K \subset \mathbb{R}^n$, the polar cone is closed and convex, and $K^{\circ\circ} = cl(conv K)$. If K is also closed and convex, one then has $K^{\circ\circ} = K$.

An important and useful object in variational problems is the tangent cone of C at \bar{x} , given by

$$\mathcal{T}_C(\bar{x}) := \{d \in \mathbb{R}^n \mid \exists t \geq 0 \text{ and } X \ni x^k \rightarrow \bar{x} \text{ s.t. } t(x^k - \bar{x}) \rightarrow d\}.$$

An example of the tangent cone and its polar are given in the following example.

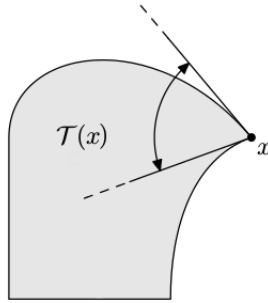


Figure 1.2: The tangent cone at x .

Example 1.1.1. $C = \{(a, b) \in \mathbb{R}^2 \mid a \geq 0, b \geq 0, ab \leq 0\}$. In this case, $\mathcal{T}_C(\bar{x}) = \{d \in \mathbb{R}_+^2 \mid d_1 d_2 = 0\}$ and $\mathcal{T}_C(\bar{x})^\circ = \mathbb{R}_-^2$.

A cone that is very often related to the tangent cone in the literature is the normal cone of a convex set C , defined as

$$\mathcal{N}_C(\bar{x}) := \{d \in \mathbb{R}^n \mid d^T(y - \bar{x}) \leq 0, \forall y \in C\}.$$

In the case of a convex set C , the tangent cone and the normal cone are polar to each others.

Proposition 1.1.5. *Let C be a convex subset of \mathbb{R}^n and let $x \in C$ be arbitrary. It holds that*

$$\mathcal{T}_C(x)^\circ = \mathcal{N}_C(x).$$

Proof. First, It is straightforward from the definition to see that $\mathcal{N}_C(x) \subset \mathcal{T}_C(x)^\circ$.

Now, we show the converse. Let d be a vector from $\mathcal{T}_C(x)^\circ$ and let $y \in C$. By convexity of the set C , it holds that $y - x$ is tangent to C at x . Therefore, by the choice of d it follows that $d^T(y - x) \leq 0$ and so $d \in \mathcal{N}_C(x)$. \square

The proof of this result is rather classical and can be found for instance in [65, Prop. 1.3.2]. We can notice here that the convexity of the set C is only required in one inclusion. In the non-convex case, there are various definitions of normal cones, which will not be developed here. The interested reader may refer to [171] for more details.

Some useful operations on polar cones are summarised below.

Proposition 1.1.6. [21, Prop. 1.1.16] *For cones K_i of $\mathbb{R}^n, i = 1, 2$ one has*

1. $K_1 \subset K_2 \implies K_2^\circ \subset K_1^\circ$ and $K_1^{\circ\circ} \subset K_2^{\circ\circ}$;
2. $K = K_1 + K_2 \implies K^\circ = K_1^\circ \cap K_2^\circ$;
3. $K = K_1 \cap K_2$ with K_i closed $\implies K^\circ = K_1^\circ \cap K_2^\circ$. The closure operation can be removed if $0 \in \text{int}(K_1 - K_2)$;
4. For a family of cones $\{K_i \mid i \in I\}$ in \mathbb{R}^n ,

$$K = \cup_{i \in I} K_i \implies K^\circ = \cap_{i \in I} K_i^\circ.$$

A cone $K \subset \mathbb{R}^n$ is said to be finitely generated if it can be written as

$$K = \left\{ \sum_{i=1}^p t_i a_i \mid a_i \in \mathbb{R}^n, t_i \geq 0, i = 1, \dots, p \right\}.$$

A set $P \subset \mathbb{R}^n$ is called polyhedral if it has the form

$$P = \{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, i = 1, \dots, p\},$$

where $a_i \in \mathbb{R}^n, b_i \in \mathbb{R}, i = 1, \dots, p$. When $b_i = 0, \forall i$, then P is called a polyhedral cone.

Theorem 1.1.3 (Minkowski-Weyl Theorem). *A cone K is polyhedral if and only if it is finitely generated.*

We conclude this section by an example of a cone that will be used in Part II.

Example 1.1.2. *Given two continuously differentiable functions $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\mathcal{I}_g(x^*) \subset \{1, \dots, p\}$, the linearised cone \mathcal{L} at a point $x^* \in \mathbb{R}^n$ is defined as*

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

This cone is polyhedral by the Minkowski-Weyl Theorem and therefore convex.

The polar of the linearised cone is a very useful tool for deriving optimality conditions of non-linear problems as in Section 1.3.

Lemma 1.1.2. *[23, Theorem 3.2.2] Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be two continuously differentiable functions and $\mathcal{I}_g(x^*) \subset \{1, \dots, p\}$. The polar cone of*

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

is given by

$$\mathcal{L}(x)^\circ = \{v \in \mathbb{R}^n \mid v = \sum_{i \in \mathcal{I}_g(x)} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) : \lambda^g \geq 0\}.$$

1.1.3 Convex functions

Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$. An important and useful set associated with a function f is the epigraph defined by

$$\text{epi } f := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid \alpha \geq f(x)\}.$$

The epigraph is thus a subset of \mathbb{R}^{n+1} that consists of all points of \mathbb{R}^{n+1} lying on or above the graph of f . An optimisation problem can thus be expressed equivalently in terms of its epigraph as

$$\inf f = \inf \{\alpha \mid (x, \alpha) \in \text{epi } f\}.$$

For an extended real-valued function $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, f is convex if and only if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \ \forall x, y \in \mathbb{R}^n, \ \forall t \in (0, 1).$$

The function is called strictly convex if the above inequality is strict for all $x, y \in \mathbb{R}^n$ with $x \neq y$ and $t \in (0, 1)$. A function is concave whenever $-f$ is convex.

A characterisation of the convexity of a function f can also be done using the epigraph of the function.

Definition 1.1.5. *A function $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$ with $f \neq \infty$ is called convex if $\text{epi } f$ is a non-empty convex set.*

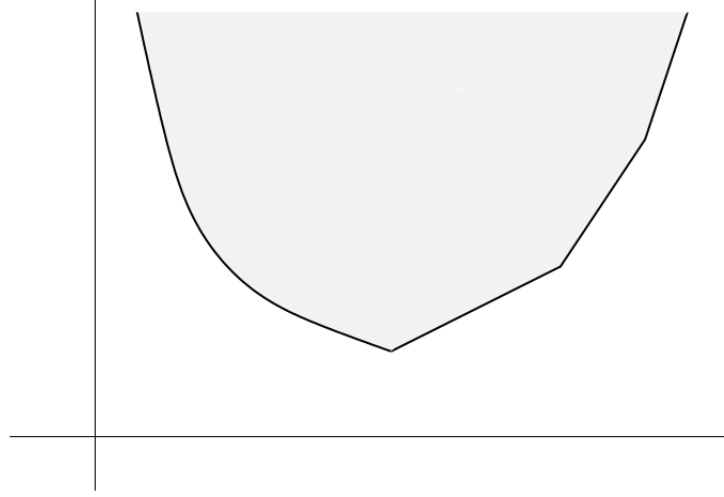


Figure 1.3: Epigraph of a convex function.

An illustration is given on Figure 1.3 and now we give some examples of convex and concave functions.

Example 1.1.3. *Examples of convex functions are:*

$$x \in \mathbb{R} \mapsto |x|, \quad x \mapsto \|x\|_2, \quad x \in \mathbb{R} \mapsto \exp(x).$$

Examples of concave functions are:

$$x \in \mathbb{R}_+ \mapsto \sqrt{x}, \quad x \in \mathbb{R}_+ \mapsto \frac{x}{x+1}, \quad x \in \mathbb{R} \mapsto \exp(-x).$$

Examples of functions that are neither convex nor concave:

$$x \in \mathbb{R} \mapsto \sin x, \quad x \in \mathbb{R} \mapsto \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases} \quad \text{and } x \in \mathbb{R} \mapsto \frac{1}{x}.$$

In the case of a minimisation problem, we do not necessarily need continuity of the function but some weaker condition that we define here. Recall that

$$\liminf_{x \rightarrow y} f(x) := \sup_{r > 0} \inf_{x \in \mathcal{B}_r(y)} f(x),$$

where $\mathcal{B}_r(y)$ denotes the ball centred in y of radius r . Lower limits are characterised via

$$\liminf_{x \rightarrow y} f(x) = \min\{\alpha \in \mathbb{R} \cup \{-\infty, \infty\} \mid \exists x_n \rightarrow y \text{ with } f(x_n) \rightarrow \alpha\}.$$

Note that one always has $\liminf_{x \rightarrow y} f(x) \leq f(y)$.

Definition 1.1.6. The function $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$ is lower semi-continuous at x if

$$f(x) = \liminf_{y \rightarrow x} f(y),$$

and lower semi-continuous on \mathbb{R}^n if this holds for every $x \in \mathbb{R}^n$.

Example 1.1.4. The function $x \in \mathbb{R} \mapsto \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases}$ is not continuous but lower semi-continuous.

1.1.4 Error bounds

An error bound is an estimation of the distance from a given point to some set. For a given norm $\|\cdot\|$ on \mathbb{R}^n (not necessarily the euclidean norm), we look for the distance defined by

$$d_P(x) := \inf_{y \in P} \|y - x\|.$$

The simplest is the set $P := \{\bar{x}\}$, where \bar{x} is the unique solution of the linear system $Ax = b$ with A any invertible matrix. For any x , it holds that $x - \bar{x} = A^{-1}(Ax - b)$. Thus, it follows

$$d_P(x) \leq \|A^{-1}\| \|Ax - b\|.$$

Convex polyhedron

Let $P := \{x \in \mathbb{R}^n : Ax \leq b\}$ be a convex polyhedron, where $A \in \mathbb{R}^{m \times n}$ is a real matrix and $b \in \mathbb{R}^m$. We denote $\mathcal{B}(A)$ the convex cone of vectors $b \in \mathbb{R}^m$ such that $P \neq \emptyset$. For $v \in \mathbb{R}^m$, we denote v^+ the vector from \mathbb{R}^m such that the i -th component is equal to $\max(0, v_i)$.

Lemma 1.1.3 (Hoffman Lemma [99]). *There exists a constant h independent of A , $\|\cdot\|$ and $\|\cdot\|'$ such that*

$$\forall b \in \mathcal{B}(A), \quad \forall x \in \mathbb{R}^n : \quad d_P(x) \leq h \|(Ax - b)^+\|'.$$

Convex constraints

Whenever the considered set is given by

$$P := \{x \in \mathbb{R}^n : c(x) \leq 0\},$$

where $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is convex, in general, we cannot have a Hoffman-type error bound, that is

$$\exists \beta > 0, \quad \forall x \in P : \quad d_P(x) \leq \beta \|c(x)^+\|,$$

without additional hypothesis. Indeed, it is necessary to have a constraint qualification hypothesis to have a linear error bound or to take a non-linear error bound of the form $\beta \|c(x)^+\|^\alpha$, with $\alpha > 0$. An example is given by Robinson error bound [169], for a convex set defined by a convex function and satisfying Slater constraint qualification.

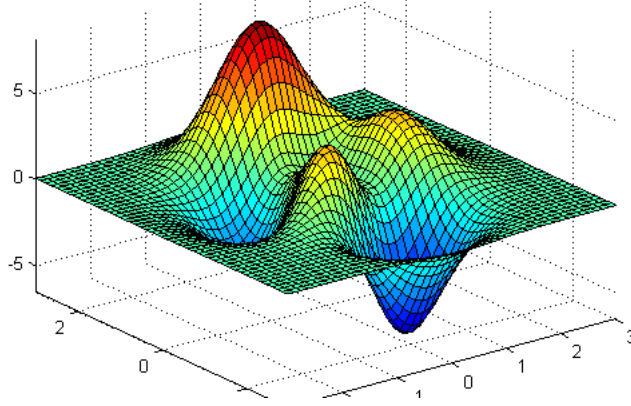


Figure 1.4: How to find the minimum of a function ?

1.2 Optimisation and variational inequalities

We discuss in this section the definition of an optimisation problem, a local minimum and some related problems called variational inequalities and complementarity problems.

1.2.1 Optimisation problems and local minima

Consider the problem of minimising a continuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ over a compact set $C \subset \mathbb{R}^n$ denoted by

$$\min_{x \in C} f(x).$$

Existence of minimisers is given by the classical Weierstrass Theorem.

Theorem 1.2.1. *Let $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuous function defined on a compact set C . Then, there exists a global minimiser $x^* \in C$ of f on C , that is,*

$$f(x^*) \leq f(x), \forall x \in C.$$

It is to be mentioned here that the continuity hypothesis on f may be reduced to lower semi-continuity. Minimising a non-linear smooth function over an arbitrary compact set is already a very hard problem. An example is given in Figure 1.4. A more realistic and more accessible goal for numerical methods is to compute a local minimum. A point $x^* \in C$ is a local minimum of f over C if there exists $\epsilon > 0$ such that for all $x \in V_\epsilon(x^*) \cap C$ it holds that

$$f(x^*) \leq f(x),$$

where $V_\epsilon(x^*)$ denotes a neighbourhood centred in x^* of radius ϵ . A consequence of the definition of a local minimum is that a local minimum may also be characterized by the fact that there is no feasible descent direction, that is

$$\nabla f(x^*)^T d \geq 0, \forall d \in \mathcal{T}_C(x^*).$$

Using the polar of the tangent cone, this may also be written as

$$-\nabla f(x^*) \in \mathcal{T}_C^\circ. \quad (1.1)$$

Unfortunately those definitions are virtually impossible to use in practice, since checking whether a point is a local minimum would require to compare its value with an infinity of other points. Thus, mathematical analysis usually focus on critical (or stationary) points that satisfies some necessary conditions to be a local minimum.

In the special case of minimising a convex function over a compact convex set, then every local minimum is also a global minimum. This observation makes the study of convex optimisation problem especially useful. However, in the case of minimising a concave function (or maximising a convex function) over a compact convex set minimising may be a hard task, since there may be many local maxima besides the global maximum.

Example 1.2.1. $f(x) = |x - a|$ with $a \in \mathbb{R}^2$ over a triangle C in \mathbb{R}^2 . Minimise f over C can be anywhere in C . Maximising is one of the three vertices but local maxima may well occur at these vertices.

1.2.2 Variational inequalities and properties

Variational inequalities are intimately connected with optimisation problems. First, let us defined a variational inequality.

Definition 1.2.1. Given $K \subset \mathbb{R}^n$ and a mapping $F : K \rightarrow \mathbb{R}^n$, the variational inequality, denoted (VI), is to find a vector $x \in K$ such that

$$(y - x)^T F(x) \geq 0, \quad \forall y \in K. \quad (\text{VI})$$

Using the normal cone for K convex, we can give a more geometric interpretation of this problem as finding $x \in K$ such that

$$-F(x) \in \mathcal{N}_K(x).$$

The condition to be a local minimum given in (1.1) is therefore an instance of the (VI). A special case of this problem is the complementarity problem, defined below.

Definition 1.2.2. Given a cone K and a mapping $F : K \rightarrow \mathbb{R}^n$, the complementarity problem, denoted (CP), is to find a vector $x \in \mathbb{R}^n$ satisfying the following conditions:

$$K \ni x \perp F(x) \in -K^\circ, \quad (\text{CP})$$

where the notation \perp means that $x^T F(x) = 0$.

The complementarity problem, (CP), is actually equivalent to (VI), whenever K is a cone.

Proposition 1.2.1. [65, Prop. 1.1.3] *Let K be cone in \mathbb{R}^n . A vector x solves (VI) if and only if x solves (CP)*

Many special cases of the complementarity problem are very important in modelling. We introduce two of the most important ones. When K is the non-negative orthant we can notice that $K = -K^\circ$ and this leads us to the non-linear complementarity problem.

Definition 1.2.3. *Given a mapping $F : \mathbb{R}_+^n \rightarrow \mathbb{R}^n$, the (NCP) is to find a vector $x \in \mathbb{R}^n$ satisfying*

$$0 \leq x \perp F(x) \geq 0. \quad (\text{NCP})$$

In the special case of F being an affine function given by:

$$F(x) = Mx + q$$

for some vector $q \in \mathbb{R}^n$ and matrix $M \in \mathbb{R}^{n \times n}$, we get the linear complementarity problem.

Definition 1.2.4. *Given a vector $q \in \mathbb{R}^n$ and a matrix $M \in \mathbb{R}^{n \times n}$, the (LCP) is to find a vector $x \in \mathbb{R}^n$ satisfying*

$$0 \leq x \perp Mx + q \geq 0. \quad (\text{LCP})$$

Variational inequalities and complementarity problems arise from a variety of interesting sources. Foremost among these sources are differentiable constrained optimisation problems. We already point out the link between the geometric interpretation of (VI) and (1.1) for $F = \nabla f$. Thus, it is not surprising to see the (LCP) as optimality condition of the quadratic program for a polyhedron K

$$\min_{x \in K} q^T x + \frac{1}{2} x^T M x.$$

Many more applications than constrained optimisation may be a source of complementarity problems among other we may cite geochemical model [149], Nash equilibrium, obstacle problems, frictional contact problems,...

1.3 Non-linear programming

Let a general non-linear program (NLP) be

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f(x) \\ & \text{s.t. } g(x) \leq 0, \quad h(x) = 0, \end{aligned} \tag{NLP}$$

with $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that h, g and f are continuously differentiable functions. We present in this section optimality conditions for (NLP). After introducing essential notations in Sect. 1.3.1, we introduce the classical Karush Kuh Tucker (KKT) optimality conditions in Sect. 1.3.2. These optimality conditions require some hypothesis on the set $\{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\}$ that are called constraint qualifications. A short review of some of these constraint qualifications is presented in Sect. 1.3.3.

1.3.1 Notations

Denote \mathcal{F} the feasible region of (NLP),

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

The set of active indices $\mathcal{I}_g(x)$ is given by

$$\mathcal{I}_g(x) := \{i \in \{1, \dots, p\} \mid g_i(x) = 0\}.$$

Let the generalised Lagrangian $\mathcal{L}^r(x, \lambda)$ be

$$\mathcal{L}^r(x, \lambda) := rf(x) + g(x)^T \lambda^g + h(x)^T \lambda^h,$$

where $\lambda = (\lambda^g, \lambda^h)$ is the vector of Lagrange multipliers.

We remind that the tangent cone of a set X at $x^* \in X$ is a closed cone defined by

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

Another useful tool for our study is the linearised cone of (NLP) at $x^* \in \mathcal{F}$ defined by

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

1.3.2 Karush-Kuhn-Tucker optimality conditions

By definition, λ is an index r multiplier for (NLP) at a feasible point x if

$$(r, \lambda) \neq 0 \text{ and } \nabla_x \mathcal{L}^r(x, \lambda) = 0, \lambda^g \geq 0, g(x)^T \lambda^g = 0.$$

An index 0 multiplier is also called singular multiplier, [30], or an abnormal multiplier, [46]. We call a KKT point or a stationary point a couple (x, λ) with λ an 1-index multiplier at x .

The motivation to compute KKT point in non-linear programming is based on the following fundamental theorem from [120] that is well-known under the name KKT theorem, [47]. This result requires that some constraint qualification hold at a local minimum. The original result from [116] uses a stronger condition than Guignard CQ that will be considered later.

Theorem 1.3.1. *Assume x^* is a local minimum of (NLP), where the functions f , $\{g_i\}$ and $\{h_i\}$ are once differentiable. Furthermore assume that a constraint qualification (CQ) holds at x^* . Then, there exists a KKT point (or stationary point) at x^* .*

The proof of this result, which can also be found in [69], relies on some duality argument. These optimality conditions are stronger than other optimality conditions known as Fritz-John conditions [108] that do not require any constraint qualification. These weaker optimality conditions that we do not formally state here give that there exists an index-r multiplier at any local minimum. This result, however, suffer from the drawback that at any point there always exists an index-r multiplier, the trivial solution.

Constraint qualifications are independent of the objective function and measure the way the feasible set is described by the inequality and equality constraints not the geometry of the feasible set itself. It is to be noted that only active constraints among the inequality constraints at x^* are taking into account since others are not playing any role at this point.

During the past 60 years, a wide variety of constraint qualifications has been derived. Beyond their utility in the KKT theorem, they are also used in a quantity of algorithmic applications. Indeed, a large number of algorithms for (NLP) including a majority of commercial solvers solves the KKT conditions to determine a stationary point. To give some examples, constraint qualifications may give properties on the set of multipliers, existence of a local error bound, metric regularity, differentiability in the central path,...

In practice, it is very difficult to find a point that conforms exactly to the KKT conditions. Hence, an algorithm may stop when such conditions are satisfied approximately. We mention here two approaches. A first approach is to consider epsilon-stationary points. It can be noticed here that for $\epsilon = 0$ we recover the classical stationary points.

Definition 1.3.1. *Given a general non-linear program (NLP) and $\epsilon \geq 0$. We say that $\lambda \in \mathbb{R}^n \times \mathbb{R}^{p+m}$ is an epsilon-stationary point (or an epsilon-KKT point) if it satisfies*

$$\left\| \nabla f(x) + \sum_{i=1}^p \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) \right\|_{\infty} \leq \epsilon,$$

with

$$\begin{aligned} |h_i(x)| &\leq \epsilon, \quad \forall i \in \{1, \dots, m\}, \\ g_i(x) &\leq \epsilon, \quad \lambda_i \geq 0, \quad |\lambda_i^g g_i(x)| \leq \epsilon \quad \forall i \in \{1, \dots, p\}. \end{aligned}$$

Another way to deal with this problem is to give necessary optimality conditions of the point and its neighbourhood in the form of sequential optimality conditions. The most popular among those conditions is the Approximate KKT (AKKT) conditions introduced in [11].

Definition 1.3.2. *AKKT holds at a feasible point x^* if there are sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^{g,k}\} \subset$*

\mathbb{R}_+^p and $\{\lambda^{h,k}\} \subset \mathbb{R}^m$ not necessarily feasible such that

$$\begin{aligned} \lim_{k \rightarrow \infty} x^k &= x^*, \\ \lim_{k \rightarrow \infty} \nabla f(x^k) + \nabla g(x^k)^T \lambda^{g,k} + \nabla h(x^k)^T \lambda^{h,k} &= 0, \\ \lambda_i^{g,k} &= 0 \text{ for } i \notin \mathcal{I}_g(x^*). \end{aligned}$$

The attractiveness of sequential optimality conditions is associated to three properties. First, they are genuine necessary optimality conditions, independently of the fulfilment of a CQ. Second, they are strong, in the sense that they imply the classical first order optimality conditions. Third, there are many practical algorithms that generate sequences whose limit points satisfy them. Many practical optimisation algorithms, such as augmented Lagrangian methods, some SQP algorithms, interior-point methods and inexact restoration methods.

1.3.3 Constraint qualifications for NLP

We now define some of the classical constraint qualifications. Note that there exists a wide variety of such conditions and we define here only those that are essential for our purpose. In Definition 1.3.3 the Linear Independence CQ (LICQ) and Constant Rank CQ (CRCQ) are presented. Both are very classical the latter being defined first in [107].

Definition 1.3.3 (LICQ and its Relaxations). *Let $x^* \in \mathcal{F}$.*

(a) *LICQ holds at x^* if the family of gradients*

$$\{\nabla g_i(x^*) \ (i \in \mathcal{I}_g(x^*)), \ \nabla h_i(x^*) \ (\forall i = 1, \dots, m)\}$$

is linearly independent.

(b) *CRCQ holds at x^* if there exists $\delta > 0$ such that for any subsets $I_1 \subseteq \mathcal{I}_g(x^*)$ and $I_2 \subseteq \{1, \dots, m\}$ the family of gradients*

$$\{\nabla g_i(x) \ (i \in I_1), \ \nabla h_i(x) \ (i \in I_2)\}$$

has the same rank for all $x \in \mathcal{B}_\delta(x^)$.*

We now introduce the definition of positive-linearly dependent vectors, which helps up building constraint qualifications since it takes into account the sign of some multipliers.

Definition 1.3.4. *A finite set of vectors $\{a^i | i \in I_1\} \cup \{b^i | i \in I_2\}$ is said to be positive-linearly dependent if there exist scalars α_i ($i \in I_1$) and β_i ($i \in I_2$), not all of them being zero, with $\alpha_i \geq 0$ for all $i \in I_1$ and*

$$\sum_{i \in I_1} \alpha_i a^i + \sum_{i \in I_2} \beta_i b^i = 0.$$

Otherwise, we say that these vectors are positive-linearly independent. In a usual way the unsigned vectors are denoted with double brackets, that is $\{a^i | i \in I_1\} \cup \{\{b^i | i \in I_2\}\}$.

Another family of constraint qualifications that take into account the non-negativity of the Lagrange multiplier of inequality constraints can now be derived using this notion.

Definition 1.3.5 (MFCQ and its Relaxations). *Let $x^* \in \mathcal{F}$.*

(a) *PLICQ holds at x^* if the family of gradients*

$$\{\nabla g_i(x^*) \mid i \in \mathcal{I}_g(x^*)\} \cup \{\{\nabla h_j(x^*) \mid j = 1, \dots, m\}\}$$

is positively linearly independent.

(b) *CPLD holds at x^* if, for each $I \subseteq \mathcal{I}_g(x^*)$ and $J \subseteq \{1, \dots, m\}$, whenever*

$$\{\nabla g_i(x^*) \mid i \in I\} \cup \{\{\nabla h_j(x^*) \mid j \in J\}\}$$

is positively linearly dependent, there exists $\delta > 0$ such that for every $x \in \mathcal{B}_\delta(x^)$,*

$$\{\nabla g_i(x), \nabla h_j(x) \mid i \in \mathcal{I}_g(x^*), j = 1, \dots, m\}$$

is linearly dependent.

(c) *RCPLD holds at x^* if we assume $J \subseteq \{1, \dots, m\}$ be such that $\{\nabla h_j(x^*)\}_{j \in J}$ is a basis for $\text{span}\{\nabla h_j(x^*)\}_{j=1}^m$ and if there exists $\delta > 0$ such that*

(i) *$\{\nabla h_j(x)\}_{j=1}^m$ has the same rank $\forall x \in \mathcal{B}_\delta(x^*)$;*

(ii) *for each $I \subseteq \mathcal{I}_g(x^*)$, if*

$$\{\nabla g_i(x^*) \mid i \in I\} \cup \{\{\nabla h_j(x^*) \mid j \in J\}\}$$

is positively linearly dependent, then

$$\{\nabla g_i(x), \nabla h_j(x) \mid i \in \mathcal{I}_g(x^*), j \in J\}$$

is linearly dependent $\forall x \in \mathcal{B}_\delta(x^)$.*

(d) *CRSC holds at x^* if there exists $\delta > 0$ such that the family of gradients*

$$\{\nabla g_i(x^*), \nabla h_j(x^*) \mid i \in J_-, j = 1, \dots, m\}$$

has the same rank for every $x \in \mathcal{B}_\delta(x^)$, where $J_- = \{i \in \mathcal{I}_g \mid -\nabla g_i(x^*) \in \mathcal{L}(x^*)^\circ\}$.*

The Positive Linear Independence CQ is equivalent to Mangasarian Fromowitz CQ and can also be called NNAMCQ (no non-zero abnormal multiplier) or BCQ (basic constraint qualification). The usual definition of MFCQ given in the literature is given in the following definition. Equivalence with PLICQ relies on a duality argument, for instance using Motzkin's alternative lemma.

Definition 1.3.6. Let $x^* \in \mathcal{F}$. MFCQ holds at x^* if there exists a vector $d \in \mathbb{R}^n$ that satisfies

$$\nabla h_i(x^*)^T d = 0 \forall i = 1, \dots, m \text{ and } \nabla g_i(x^*)^T d < 0 \forall i \in \mathcal{I}_g(x^*).$$

The constant positive linear dependence defined by [162] has been proved to be a constraint qualification in [15]. Constant rank of the subspace component, CRSC, was introduced recently in [13]. 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.

As pointed out earlier in equation (1.1), a local minimum is characterised by the fact that there is no feasible descent direction for the objective function of (NLP), that is

$$-\nabla f(x^*) \in \mathcal{T}_{\mathcal{F}}(x^*)^\circ.$$

From the other side the KKT conditions build ∇f using a linearisation of the active constraints. This motivates the following CQs defined as early as 1969 in [88] for GCQ and in [1] for ACQ.

Definition 1.3.7. A point $x^* \in \mathcal{F}$ is said to satisfy Guignard CQ if $\mathcal{T}_{\mathcal{F}}(x^*)^\circ = \mathcal{L}^\circ(x^*)$ and satisfy Abadie CQ if $\mathcal{T}_{\mathcal{F}}(x^*) = \mathcal{L}(x^*)$.

As proved in [86], Guignard CQ is the weakest constraint qualification that ensures that a local minimum satisfies the KKT conditions independently of the objective function. It is easy to see that for all x it holds that $\mathcal{T}_{\mathcal{F}}(x) \subseteq \mathcal{L}(x)$ and so $\mathcal{L}^\circ(x) \subseteq \mathcal{T}_{\mathcal{F}}(x)^\circ$ by Proposition 1.1.6. The fact that Abadie CQ holds at x^* implies that Guignard CQ also holds at x^* is also a consequence of Proposition 1.1.6.

The supplementary condition necessary such that AKKT implies KKT is sometimes called Strict Constraint Qualification, [14]. Motivation behind SCQ is to confirm that an AKKT sequence is approaching a true minimiser. This motivation is well illustrated by the following example. For instance, ACQ is a CQs that is not SCQ.

Example 1.3.1 (AKKT but not minimizer).

$$\min_{(x_1, x_2) \in \mathbb{R}^2} x_2 \text{ s.t. } x_1 \geq 0, x_1 x_2 = 0.$$

the feasible point $x^* = (0, 1)$ is neither a minimiser nor a KKT point, but it satisfies AKKT. Computing the tangent cone at x^* gives $\mathcal{T}_{\mathcal{F}}(x^*) = \{0\} \times \mathbb{R}$ and the linearised cone gives $\mathcal{L}(x^*) = \{d \in \mathbb{R}^2 | d_1 \geq 0, x_2^* d_1 + x_1^* d_2 = 0\} = \mathcal{T}_{\mathcal{F}}(x^*)$. Thus, ACQ holds at x^* .

This has been a motivation to the definition of the CCP condition in [14] that is the weakest constraint qualification that ensures that AKKT is actually a first order optimality condition.

Definition 1.3.8. We say that a point $x^* \in \mathcal{F}$ satisfies the Cone-Continuity Property (CCP) if the set-valued mapping $\mathbb{R}^n \ni x \Rightarrow K(x)$ such that

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

is outer semi-continuous at x^* , that is

$$\limsup_{x \rightarrow x^*} K(x) := \{u | \forall x^\nu \rightarrow x^*, \exists u^\nu \rightarrow u \text{ with } u^\nu \in K(x^\nu)\} \subset K(x^*).$$

The definition of outer semi-continuity and related properties can be found for instance in [171, Definition 5.4]. It is to be noted here that $K(x)$ depends on x^* , since it considers only active constraints at x^* . Clearly, $K(x^*)$ is a closed convex cone and coincides with the polar linearised cone $\mathcal{L}(x^*)^\circ$. Moreover, $K(x)$ is always inner semi-continuous, i.e. $K(x^*) \subset \liminf_{x \rightarrow x^*} K(x)$, due to the continuity of the gradients and the definition of $K(x)$. For this reason, outer semi-continuity is sufficient for the continuity of $K(x)$ at x^* .

Theorem 1.3.2. [14, Theorem 3.1] CCP is the weakest property under which AKKT implies KKT, independently of the objective function.

A corollary of this result is that CCP is a CQ. Besides, it has been shown in [14] that CCP is strictly stronger than ACQ and weaker than CRSC.

We would like to insist on an essential characteristic of CQs that they measure the way the feasible set is described and not its geometry as illustrated by the following simple example.

Example 1.3.2. Consider the following one-dimensional minimisation problem :

$$\min_{x \in \mathbb{R}} x \text{ s.t. } x^2 \leq 0.$$

Obviously, the origin is the only feasible point and thus is the global minimum. Besides, since $\mathcal{F} = \{0\}$ it follows that $\mathcal{T}_{\mathcal{F}}(0) = \{0\}$, while the linearised cone is given by $\mathcal{L}(0) = \mathbb{R}$. Thus, ACQ does not hold at the origin and neither does Guignard CQ, since the definition of the polar of a cone yields $\mathcal{L}(0)^\circ = \{0\}$ and $\mathcal{T}_{\mathcal{F}}(0)^\circ = \mathbb{R}$.

We sum up the various CQs defined in this section by giving the diagram in Figure 1.5.

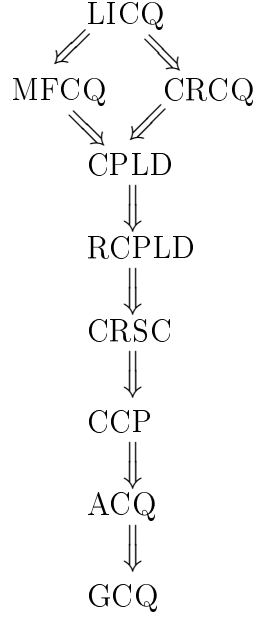


Figure 1.5: Relations between the constraint qualifications for (NLP).

1.4 On the θ 's functions

We consider a family of smooth functions that are extensively used in the various numerical methods introduced in this manuscript. This family has already been used in the context of complementarity [93, 94] and in image restoration [26]. These functions are non-decreasing continuous smooth concave functions such that

$$\theta : \mathbb{R} \rightarrow]-\infty, 1[\text{ with } \theta(x) < 0 \text{ if } x < 0, \theta(0) = 0 \text{ and } \lim_{x \rightarrow +\infty} \theta(x) = 1.$$

Example 1.4.1. *Examples of this family are $\theta^1(x) = x/(x+1)$ if $x \geq 0$ and $\theta^1(x) = x$ if $x < 0$, $\theta^2(x) = 1 - e^{-x}$ with $x \in \mathbb{R}$.*

Then using a scaling technique similar to the perspective functions we define $\theta_r(x) := \theta\left(\frac{x}{r}\right)$ for $r > 0$ and we get

$$\theta_r(0) = 0 \quad \forall r > 0 \text{ and } \lim_{r \rightarrow 0} \theta_r(x) = 1 \quad \forall x > 0.$$

In order to simplify the presentation we sometimes used the notation $\theta_r(x)$ to denote $\theta(x, r)$.

Example 1.4.2. *For the previous examples of this family and $x \geq 0$ we have $\theta_r^1(x) = x/(x+r)$, $\theta_r^2(x) = 1 - e^{-x/r}$.*

The function $\theta_r^1(x)$ will be extensively used in this paper and is illustrated in Figure 1.6 for several values of r . The function θ^2 is a special case of the Weibull distribution defined for $k \leq 1$ as $\theta_r^W(x) := (1 - e^{-(\frac{x}{r})})^k$. Functions that possess these properties have been used in several domains in the literature. This chapter and the whole manuscript in general study

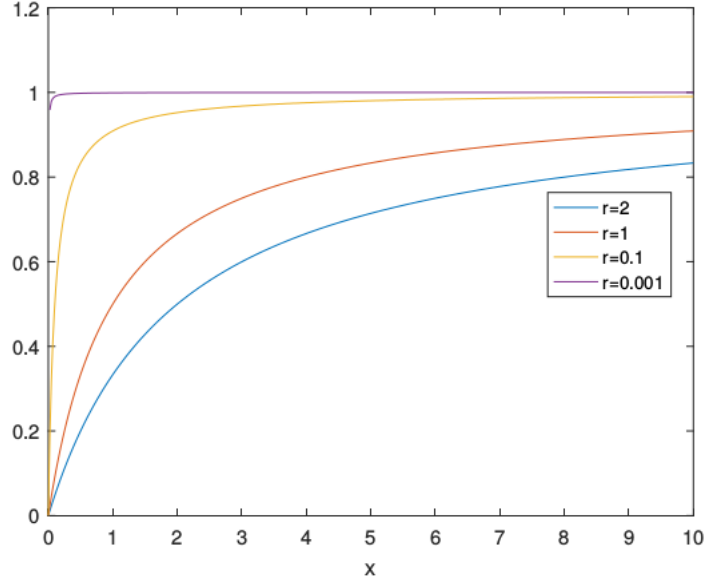


Figure 1.6: Function θ^1 for several values of r .

these functions in a general way, therefore allowing to extend existing results and derive new ones.

It can be seen on the Figure 1.6 that the functions θ_r behave as a step function when r becomes small. This observation has been the motivation of the study of Chapter 2. Another asymptotic property of these functions can be obtained when r grows as given in the following proposition, whose proof will be given in Theorem 2.2.2 (p.32).

Proposition 1.4.1. *Let x be a non-negative vector. Then, it holds that*

$$\lim_{r \rightarrow \infty} r\theta_r(x) = \theta'(0)x.$$

This property in particular has been very useful in Chapter 2 to find an initial point of the regularisation scheme. We point out here that by composition of the derivative it always holds that $\theta'_r(0) = \theta'(0)/r$. Through the manuscript, $\theta'_r(x)$ denotes the derivative of $\theta_r(x)$ with respect to x , while the derivative with respect to r is explicitly noted.

A classical property shared by all concave functions that vanish in zero is the sub-additivity.

Lemma 1.4.1. *θ_r is sub-additive for non-negative values, i.e. given $x, y \geq 0$ it holds that*

$$\theta_r(x) + \theta_r(y) \geq \theta_r(x + y).$$

Proof. Since θ is concave, we obtain

$$\forall x \neq y \in \mathbb{R}, \forall t \in (0, 1), t\theta_r(x) + (1 - t)\theta_r(y) \leq \theta_r(tx + (1 - t)y),$$

with equality if $t = 0$ or 1 and if $x = y$. Considering $y = 0$ and $\theta_r(0) = 0$ yields

$$\theta_r(tx) = \theta_r(tx + (1-t)y) \geq t\theta_r(x) \quad \forall t \in (0, 1),$$

with equality if $t = 0, 1$ or $x = 0$. Let $i \in \{1, \dots, n\}$ and suppose that $x_i + y_i \neq 0$ (the case $x_i = y_i = 0$ stay true)

$$\begin{aligned} \theta_r(x_i) + \theta_r(y_i) &= \theta_r((x_i + y_i) \frac{x_i}{x_i + y_i}) + \theta_r((x_i + y_i) \frac{y_i}{x_i + y_i}), \\ &\geq \frac{x_i}{x_i + y_i} \theta_r((x_i + y_i)) + \frac{y_i}{x_i + y_i} \theta_r((x_i + y_i)), \\ &= \theta_r(x_i + y_i), \end{aligned}$$

with equality if and only if $x_i = 0$ or $y_i = 0$. This concludes the proof. \square

Through this manuscript, we use the functions θ_r to regularise the complementarity problem. The following lemma, whose proof will be given in Lemma 3.2.1 (p.46), provides an intuition of the motivation behind such technique.

Lemma 1.4.2. *Given $x, y \in \mathbb{R}_+$, we have*

$$x \perp y \iff \lim_{r \searrow 0} \theta_r(x) + \theta_r(y) \leq 1.$$

1.4.1 How to build θ 's functions

In [94], the authors present a technique to construct functions close to the functions θ_r as a solution of some differential equation. We offer here a different approach. One way to build θ functions is to consider non-increasing probability density functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and then take the corresponding cumulative distribution function

$$\forall x \geq 0, \theta(x) = \int_0^x f(y) dy \text{ and } \forall x < 0, \theta(x) < 0.$$

By definition of f we can verify that

$$\lim_{x \rightarrow +\infty} \theta(x) = \int_0^{+\infty} f(y) dy = 1 \text{ and } \theta(0) = \int_0^0 f(y) dy = 0.$$

The non-increasing hypothesis on f gives the concavity of θ .

1.4.2 Useful results

We conclude this section by giving several lemmas that are used in the sequel of this manuscript. The first two results give more insight on the asymptotic behaviour of the functions θ_r .

Lemma 1.4.3. *Given $\{r_k\}$ a sequence of positive parameters decreasing to zero. Then, it holds true that for all $x \in \mathbb{R}_{++}$*

$$\lim_{k \rightarrow \infty} \theta'_{r_k}(x) = 0.$$

Proof. By definition of the θ_r functions as perspective function, we can write

$$\theta'_{r_k}(x) = \frac{1}{r_k} \theta' \left(\frac{x}{r_k} \right).$$

Thus, it is sufficient to prove that

$$\lim_{k \rightarrow \infty} x \theta'_{r_k}(x) = \lim_{k \rightarrow \infty} \frac{x}{r_k} \theta' \left(\frac{x}{r_k} \right) = \lim_{z \rightarrow \infty} z \theta'(z) = 0.$$

Let z and z' be two positive numbers such that $z > z'$. By the mean value theorem it follows that there exists a constant $c \in [z', z]$ such that

$$\theta(z) - \theta(z') = \theta'(c)(z - z') \geq 0,$$

where the positivity comes from the non-decreasing hypothesis on θ . By concavity it follows that

$$\theta(z) - \theta(z') \geq \theta'(z)(z - z') \geq 0,$$

since $\theta' \geq 0$. Taking $z' = z/2$ yields

$$2(\theta(z) - \theta(z/2)) \geq \theta'(z)z \geq 0.$$

Using that $\lim_{z \rightarrow \infty} \theta(z) = 1$ and passing to the limit gives the result. \square

The following lemma concerns also the asymptotic behavior of θ' . We mention here that this result should also be true without the assumption that $\theta_{r_k}(y^k) = 1 - \theta_{r_k}(x^k)$ for all k .

Lemma 1.4.4. *Given $\{r_k\}$ a sequence of positive parameters decreasing to zero. Given a sequence $\{x^k\} \subset \mathbb{R}$ such that $x^k \rightarrow x \in \mathbb{R}_{++}$. Then, it holds for all sequence $\{y^k\} \subset \mathbb{R}$ with $y^k \rightarrow 0$ and $\theta_{r_k}(y^k) = 1 - \theta_{r_k}(x^k)$ that*

$$\theta'_{r_k}(x^k) = o(\theta'_{r_k}(y^k)).$$

Proof. In order to prove that $\theta'_{r_k}(x^k) = o(\theta'_{r_k}(y^k))$ we show that

$$\lim_{k \rightarrow \infty} \frac{\theta'_{r_k}(x^k)}{\theta'_{r_k}(y^k)} = 0.$$

By Lemma 1.4.3, it holds that $\lim_{k \rightarrow \infty} \theta'_{r_k}(x^k) = 0$. Now, if $\lim_{k \rightarrow \infty} \theta'_{r_k}(y^k) \neq 0$, then the proof is complete.

So, let us assume that $\lim_{k \rightarrow \infty} \theta'_{r_k}(y^k) = 0$. We show that this leads to a contradiction. A direct consequence of this assumption is that necessarily $\lim_{k \rightarrow \infty} y^k/r_k \neq 0$. Therefore, it follows that $\lim_{k \rightarrow \infty} \theta(y^k/r_k) \neq 0$. However, this is a contradiction with the assumption that y_k satisfies $\theta_{r_k}(y^k) = 1 - \theta_{r_k}(x^k)$ for all k . Indeed, as k grows to infinity $\theta_{r_k}(x^k)$ goes to 1 and so $\theta_{r_k}(y^k)$ must go to 0.

Thus, $\lim_{k \rightarrow \infty} \theta'_{r_k}(y^k) \neq 0$ and the result follows. \square

The following result illustrates the behaviour of functions θ_r and their derivatives when t and r are going through zero.

Lemma 1.4.5. *Given two sequences $\{r_k\}$ and $\{t_k\}$, which converge to 0 as k goes to infinity and $\forall k \in \mathbb{N}$, $(r_k, t_k) \in \mathbb{R}_{++}^2$. Then, for any $x \in \mathbb{R}_+$*

$$\lim_{k \rightarrow \infty} t_k \theta_{r_k}(x) = 0.$$

Furthermore, let $\{x^k\}$ be such that $\lim_{k \rightarrow \infty} x^k = 0$. Then, either $x^k = O(r_k)$ and so there exists a constant $C_\theta \in [0, \theta'(0)]$ such that,

$$\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(x_k) = \lim_{k \rightarrow \infty} C_\theta \frac{t_k}{r_k},$$

otherwise, i.e $x^k = \omega(r_k)$, then

$$\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(x_k) \leq \lim_{k \rightarrow \infty} \theta'(1) \frac{t_k}{r_k}.$$

Proof. First part of the lemma follows from the definition of functions θ_r . Indeed, it holds for all $x \in \mathbb{R}_+$ that $\theta_r(x) \in [0, 1]$. Therefore, $\lim_{k \rightarrow \infty} t_k \theta_{r_k}(x_k) = 0$.

Second part of the lemma uses the fact that functions θ_r are defined as perspective functions, that is for all $x^k \in \mathbb{R}_+$

$$\theta_{r_k}(x^k) = \theta\left(\frac{x^k}{r_k}\right),$$

and so, computing the derivative gives

$$t_k \theta'_{r_k}(x^k) = \frac{t_k}{r_k} \theta'\left(\frac{x^k}{r_k}\right).$$

So, either $x^k = o(r_k)$ and by $0 < \theta'(0) < \infty$

$$\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(x^k) = \lim_{k \rightarrow \infty} \frac{t_k}{r_k} \theta'\left(\frac{x^k}{r_k}\right) = \lim_{k \rightarrow \infty} \frac{t_k}{r_k} \theta'(0).$$

Either there exists a constant $C > 0$ such that $x^k = Cr_k$ and so

$$\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(x^k) = \lim_{k \rightarrow \infty} \frac{t_k}{r_k} \theta'\left(\frac{Cr^k}{r_k}\right) = \lim_{k \rightarrow \infty} \frac{t_k}{r_k} \theta'(C).$$

Otherwise for k sufficiently large $r_k \leq x^k$ and by concavity of θ_r

$$0 \leq \lim_{k \rightarrow \infty} t_k \theta'_{r_k}(x^k) \leq \lim_{k \rightarrow \infty} t_k \theta'_{r_k}(r_k) = \lim_{k \rightarrow \infty} \frac{t_k}{r_k} \theta'(1).$$

□

Part I

Complementarity Problems

This part focus on numerical methods to solve the linear complementarity problem, which consists in finding $x \in \mathbb{R}^n$ such that

$$0 \leq Mx + q \perp x \geq 0, \quad (\text{LCP})$$

with M a matrix of order n and a vector $q \in \mathbb{R}^n$. Recent works by M. Haddou and P. Maheux [94], inspired a regularisation technique to deal with the complementarity problems. Indeed, the (LCP) is equivalent to

$$y = Mx + q \geq 0, x \geq 0, \|y\|_0 + \|x\|_0 \leq 1, \quad (1.2)$$

where $\|\cdot\|_0$ denotes the ℓ_0 norm of vectors, that is for a vector $x \in \mathbb{R}^n$

$$\|x\|_0 = \sum_{i \in \text{supp}(x)} 1. \quad (\ell_0)$$

It is to be noted here that even so ℓ_0 is very often called a norm in the literature, it is not one since it does not satisfy the homogeneity property. This equivalent formulation is not easier than the initial problem since the ℓ_0 is not continuous. Now, using the class of functions θ used in [94] and defined in Section 1, we consider a regularisation of (1.2) for $r > 0$ by

$$y = Mx + q \geq 0, x \geq 0, \theta_r(y) + \theta_r(x) \leq 1. \quad (1.3)$$

By construction of the functions θ_r we should recover the initial problem when r decrease to 0. An equivalent formulation to (LCP) is the absolute value equation, denoted AVE, that has also received some interest in the literature.

This part is divided in two chapters that focus on the study of this regularisation technique. First chapter considers the functions θ_r to approximate the ℓ_0 norm and their applications in sparse optimisation. Second chapter deal with the sequence of problems (1.3) to solve the AVE.

Chapter 2

Relaxation methods and application to sparse optimisation

This chapter is a paper submitted to RAIRO entitled:

A Smoothing Method for Sparse Optimization over Polyhedral Sets

Authors:

M. Haddou, IRMAR-INSA Rennes

T. Migot, IRMAR-INSA Rennes

Abstract: In this paper, we investigate a class of heuristic schemes to solve the NP-hard problem of minimizing ℓ_0 -norm over a polyhedral set. A well-known approximation is to consider the convex problem of minimizing ℓ_1 -norm. We are interested in finding improved results in cases where the problem in ℓ_1 -norm does not provide an optimal solution to the ℓ_0 -norm problem. We consider a relaxation technique using a family of smooth concave functions depending on a parameter. Some other relaxations have already been tried in the literature and the aim of this paper is to provide a more general context. This motivation allows deriving new theoretical results that are valid for general constraint set. We use a homotopy algorithm, starting from a solution to the problem in ℓ_1 -norm and ending in a solution of the problem in ℓ_0 -norm. We show the existence of the solutions of the subproblem, convergence results, a kind of monotonicity of the solutions as well as error estimates leading to an exact penalization theorem. We also provide keys for implementing the algorithm and numerical simulations.

Mathematics Subject Classification. 90-08 and 65K05

Keywords : smoothing functions ; sparse optimization ; concave minimization ; ℓ_0 -norm

Introduction

Consider a polyhedron F defined by linear inequalities, $F = \{x \in \mathbb{R}^n \mid Ax \leq b\} \cap \mathbb{R}_+^n$ for some $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$, which we suppose non-empty and not reduced to a singleton. Although, we consider a polyhedron here, most of the results presented in this article can be generalized as F being a closed convex set. One should note that the hypothesis of considering polyhedron in the non-negative orthant is not restrictive. It is only assumed to simplify the presentation and to avoid the absolute value in the definition of the problem.

We are interested in finding the sparsest point over this polyhedron, which is equivalent to minimize the ℓ_0 -norm, i.e.

$$\min_{x \in F} \|x\|_0, \quad (P_{\ell_0})$$

where

$$\forall x \in \mathbb{R}^n, \|x\|_0 = \sum_{i=1}^n s(|x_i|), \text{ where for } t \in \mathbb{R}, s(t) = \{0 \text{ if } t = 0 ; 1 \text{ otherwise}\}.$$

Note that the ℓ_0 -norm is not a norm as it does not have the homogeneity property. (P_{ℓ_0}) is an NP-hard problem as shown in [156].

This problem has several applications and received a considerable interest recently. Sparsity is involved in several domains including signal and image processing [160, 85, 59, 133, 33], statistics [80, 189, 186], machine learning [31, 140, 142]. The compressed sensing [45, 61, 62, 58, 40, 39, 41] has been the most popular application involving sparsity and creating cross-disciplinary attention in recent years and stimulates a plethora of new applications of sparsity. For more details about applications in image and signal modelling as well as a review on related questions see [33] or [188].

The problem (P_0) being difficult to solve, a classical approximation consists in solving the convex problem in ℓ_1 -norm. The ℓ_1 -norm is denoted by

$$\forall x \in \mathbb{R}^n, \|x\|_1 = \sum_{i=1}^n |x_i|. \quad (2.1)$$

The convex problem in ℓ_1 -norm is defined by

$$\min_{x \in F} \|x\|_1. \quad (P_1)$$

It can be seen as a convexification of (P_0) , because the absolute value of x is the convex envelope of $s(x)$ for $x \in [-1, 1]$. Furthermore, (P_1) has the benefits that it can be reformulated as a linear program. This approach has been extensively studied in [60, 38, 41, 45, 63, 81, 187] and in particular with inequality constraints. Moreover, several criteria have been found which guarantee that solving (P_1) will also solve (P_0) under various assumptions involving the coefficients of the matrix A . These criteria show the efficiency of this convex approximation to solve (P_{ℓ_0}) .

A more sophisticated version of this convex formulation could be to consider a reweighted- ℓ_1 problem, [207]. For more details about applications in image and signal modelling as well as a review on related questions see reviews in [33] or [188]. It is clear from these references that the study of the convex problem (P_1) to solve (P_{ℓ_0}) is of great importance.

Also formulation (P_1) does not solve all the time the initial problem. Consider for instance the following example in two dimensions.

Example 2.0.3. *Given a matrix $A \in \mathbb{R}^{n \times n}$ and a vector $b \in \mathbb{R}^n$ such that*

$$A = \begin{pmatrix} -0.1 & -1 \\ -10 & -1 \end{pmatrix} \text{ and } b = \begin{pmatrix} -1 \\ -10 \end{pmatrix}.$$

Geometrical observation allows to conclude that the solution of problem (P_1) is $(\frac{10}{11}, \frac{10}{11})^T$, while solution of problem (P_{ℓ_0}) are of the form $(0, 10 + \epsilon)^T$ and $(10 + \epsilon, 0)^T$ with $\epsilon \geq 0$.

Nonconvex optimization has been one of the main approach to tackle this problem [197, 68, 28, 191, 122]. For instance, in [68, 28], the authors proposed a reformulation of the problem as a mathematical program with complementarity constraints. Thresholding algorithms have also some recent popularity in [191, 51, 29, 157, 24, 132]. A Difference of Convex (DC) decomposition of the ℓ_0 -norm combined with DC Algorithm has been used in [122]. We are interested here in nonconvex methods to improve the solution we get by solving (P_1) in the general case where this approach does not solve the initial problem. In this aim, several concave relaxations of $\|\cdot\|_0$ have been tried in the literature.

An intuitive approach trying to bridge the gap between the ℓ_1 -norm and the ℓ_0 -norm has been to study homotopy methods based on the ℓ_p -norm for $0 \leq p \leq 1$. This approach has been initiated in [87] and later analyzed in [82, 44, 78, 121, 83], where the authors prove the link between (P_{ℓ_0}) and (P_1) as well as conditions involving the coefficients of A to show a sufficient convergence condition, so that p does not have to decrease to 0 but only to some small value. The homotopy method considers non-convex subproblems and solving the problem in ℓ_p is not a trivial task. In [82], the authors study a linearization algorithm, while in [83] the authors consider an interior-point method to solve the subproblems. Besides, the problem of minimizing the ℓ_p -norm might lead to numerical difficulties due to the non-differentiability at the origin, in [121] the authors consider a smoothing of the ℓ_p -norm to circumvent this problem.

Following the progress made during the last decade in the study of reweighted ℓ_1 -norm and ℓ_p -norm, we study here smooth regularizations. In [152] and related works the authors present a general family of smoothing function including the gaussian family and propose a homotopy method starting from the ℓ_2 -norm solutions.

Approximating the ℓ_0 -norm by smooth functions through an homotopy method starting from the ℓ_1 -norm has been studied in the PhD thesis [167] and in [168, 166, 129]. In these works, the authors consider a selection of minimization problems using smooth functions such that $(t+r)^p$ with $r > 0$ and $0 < p < 1$, $-(t+r)^{-p}$ with $r > 0$ and $1 < p$, $\log(t+r)$ with $0 < r \ll 1$ or $1 - e^{-rt}$ with $r > 0$ and $p \in \mathbb{N}$. The subproblems of the homotopy algorithm are solved using a Frank and Wolfe approach [168], also called SLA in [82], and this method

is further studied in [128].

The aim of this paper is to pursue the study of smooth concave approximation of the ℓ_0 -norm by offering a more general theoretical context for this study. Focusing on concave functions is a logical choice considering that the ℓ_p -norm is itself concave. The motivation here is to keep the good properties of the method from [168] and related work, a homotopy method between the ℓ_1 -norm and the ℓ_0 -norm problems, and smoothness at the origin. In particular, such a theoretical study has not been done in the literature.

The method considered here is a homotopy method with a parameter r such that the method recovers the ℓ_1 -norm problem for r large and the ℓ_0 -norm problem for r small. We provide here a complete analysis of the convergence of the algorithm as well as a monotonicity study of the objective function during the iterations of the homotopy scheme. We also prove the existence of the solutions of the subproblems without any boundedness assumption on the constraints.

For the convex problem of minimizing the ℓ_1 -norm, we already pointed out that several criteria involving the coefficients of the matrix A guarantee that solving the problem is sufficient to compute a solution to (P_{ℓ_0}) . Such a result guarantees the good behavior of the method. Considering our homotopy algorithm, we show a similar result independently of the constraints that state that it is not necessary to tend r to zero to compute a solution of (P_{ℓ_0}) . It can be seen as an exact penalty result. This property is a key to ensure the interest of the method.

Most of the theoretical results presented here are valid for any non-empty closed convex set F , which make them valid for several smoothing functions but also for several formulations of the problem.

In order to validate our approach, we give technical details and some numerical results on a Frank and Wolfe method to solve the subproblems of the homotopy scheme. In particular, these results show that we manage to improve the results given by the ℓ_1 norm, which shows the validity of our approach.

This document is organized as follows. Section 2.1 introduces a general formulation of the relaxation methods using concave functions. Section 2.2 discusses convergence and monotonicity results leading to a homotopy method. Section 2.3 proves error estimates and an exact penalization theorem. Finally, Section 2.4 presents the algorithm with several remarks concerning its implementation and numerical results can be found in Section 2.5.

2.1 A smoothing method

We consider a family of smooth functions designed to approximate the ℓ_0 -norm. This family has already been used in the different context of complementarity [94] and image restoration [26]. These functions are smooth non-decreasing concave functions such that

$$\theta : \mathbb{R} \rightarrow]-\infty, 1[\text{ with } \theta(t) < 0 \text{ if } t < 0, \theta(0) = 0 \text{ and } \lim_{t \rightarrow +\infty} \theta(t) = 1.$$

One way to build θ functions is to consider non-increasing probability density functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and then take the corresponding cumulative distribution function

$$\forall t \geq 0, \theta(t) = \int_0^t f(x)dx \text{ and } \forall t < 0, \theta(t) < 0.$$

By definition of f we can verify that

$$\lim_{t \rightarrow +\infty} \theta(t) = \int_0^{+\infty} f(x)dx = 1 \text{ and } \theta(0) = \int_0^0 f(x)dx = 0.$$

The non-increasing hypothesis on f gives the concavity of θ .

Examples of this family are $\theta^1(t) = t/(t+1)$ if $t \geq 0$ and $\theta^1(t) = t$ if $t < 0$, $\theta^2(t) = 1 - e^{-t}$ with $t \in \mathbb{R}$.

Then using a scaling technique similar to the perspective functions in convex analysis we define $\theta(t, r) := \theta\left(\frac{t}{r}\right)$ for $r > 0$ and we get

$$\theta(0, r) = 0 \quad \forall r > 0 \text{ and } \lim_{r \rightarrow 0} \theta(t, r) = 1 \quad \forall t > 0.$$

For the previous examples of this family and $t \geq 0$ we have $\theta^1(t, r) = t/(t+r)$, $\theta^2(t, r) = 1 - e^{-t/r}$. The function $\theta^1(t, r)$ will be extensively used in this paper.

Throughout this paper we will consider the concave optimization problem for $r > 0$

$$\min_{x \in F} \sum_{i=1}^n \theta(x_i, r). \quad (P_r)$$

Before moving to the proofs of convergence, we give a result on the existence of solutions of (P_r) . The proof relies on an argument similar to the use of asymptotic cones and directions as introduced in Chapter 1.

Theorem 2.1.1. *Let $F \subset \mathbb{R}_+^n$ be a non-empty closed convex set. The optimal set of (P_r) for $r > 0$ is non-empty.*

Proof. Since $f(x) := \sum_{i=1}^n \theta(x_i, r)$ is bounded below on the closed set F it admits an infimum. Now, assume by contradiction that there exists an unbounded sequence $\{x_n\}$ such that $x_n \in F$, $\forall n$ and

$$\lim_{n \rightarrow \infty} f(x_n) = \inf_{x \in F} f(x) < f(x_0).$$

Let $\{d_n\}$ be the sequence defined for all n by

$$d_n := \frac{x_n - x_0}{\|x_n\|}.$$

This sequence is bounded, therefore it converges, up to a subsequence, to some limit, $\lim_{n \rightarrow \infty} d_n = \lim_{n \rightarrow \infty} x_n / \|x_n\| = d \in F^\infty$, where F^∞ denotes the cone of asymptotic directions of F , Definition 1.1.3. Since F is a closed convex set, it holds for all $x \in F$ that

$$x + \alpha d \in F, \quad \forall \alpha \geq 0.$$

Then, since $F \subset \mathbb{R}_+^n$, we obtain that $d \geq 0$.

Using component-wise monotonicity and continuity assumption on θ gives

$$\lim_{n \rightarrow \infty} f(x + \alpha_n d_n) \geq f(x), \quad \forall x \in F$$

as long as $\alpha_n > 0$ for all n and the sequence $\{\alpha_n d_n\}$ admits some limit. Choosing $x = x_0$, $\alpha_n = \|x_n\|$ and d_n as defined above, we obtain

$$\lim_{n \rightarrow \infty} f(x_0 + \alpha_n d_n) = \lim_{n \rightarrow \infty} f(x_n) \geq f(x_0),$$

which is a contradiction with our initial assumption. This completes the proof. \square

2.2 Convergence

In this section, we will show the link between problems (P_{ℓ_0}) , (P_1) and (P_r) . We denote $S_{\|\cdot\|_0}^*$ the set of solutions of (P_{ℓ_0}) , $S_{\|\cdot\|_1}^*$ the set of solutions of (P_1) and S_r^* the set of solutions of (P_r) .

Our aim is to illustrate that for r sufficiently large (P_r) is close to (P_1) (see Theorem 2.2.2), and for r sufficiently small (P_r) is close to (P_{ℓ_0}) (see Theorem 2.2.1). In this way, we define an homotopy method starting from r large and decreasing r step by step. Thus, we use the convex approximation (P_1) and come closer and closer to the problem we want to solve. A monotonicity-kind result of the sequence computed by the homotopy scheme is proved in Theorem 2.2.3. Finally, Theorem 2.2.4 shows that this formulation may also be of interest for more complicated objective function than the one in (P_{ℓ_0}) .

Theorem 2.2.1 gives convergence of (P_r) to (P_0) for r decreasing to 0.

Theorem 2.2.1 (Convergence to ℓ_0 -norm). *Let $F \subset \mathbb{R}_+^n$ be a non-empty closed convex set. Every limit point of any sequence $\{x_r\}_r$, such that $x_r \in S_r^*$ and $r \downarrow 0$, is an optimal solution of (P_{ℓ_0}) .*

Proof. Given \bar{x} the limit of the sequence $\{x_r\}_r$, up to a subsequence, and $x^* \in S_{\|\cdot\|_0}^*$. Since F is a closed set one has $\bar{x} \in F$. Furthermore, we have for all r in this subsequence that

$$\sum_{i \in \text{supp}(\bar{x})} \theta(x_{r,i}, r) \leq \sum_{i=1}^n \theta(x_{r,i}, r) \leq \sum_{i=1}^n \theta(x_i^*, r) \leq \|x^*\|_0. \quad (2.2)$$

Moreover the definition of $\theta(\cdot, r)$ functions, for $r > 0$ and $t \in \mathbb{R}^n$ give

$$\sum_{i=1}^n \lim_{r \downarrow 0} \theta(t_i, r) = \|t\|_0.$$

Passing to the limit into (2.2) we get

$$\|\bar{x}\|_0 \leq \|x^*\|_0,$$

and thanks to the definition of x^*

$$\|\bar{x}\|_0 = \|x^*\|_0.$$

□

We now give another convergence result from [82], which adds that the convergence appears in a finite number of iteration.

Proposition 2.2.1. *Given a non-empty polyhedron $F \subset \mathbb{R}_+^n$. Then there exists a \bar{r} such that for all $r \leq \bar{r}$ a vertex of F is an optimal solution of (P_{ℓ_0}) and $(P_{r \leq \bar{r}})$.*

Proof. (P_r) is a problem of minimizing a concave function over a polyhedron F . We can use Corollary 32.3.4 of [170], since there is no half-line in F such that $\theta(., r)$ is unbounded below, so the infimum over F is attained and it is attained at one of the extremal point of F .

Given that there is a finite number of extremal point, one vertex, say x' , will repeatedly solve (??) for some increasing infinite sequence $R = (r_0, r_1, r_2, \dots)$. Moreover the objective function of (??) is non-increasing and bounded below by the infimum of ℓ_0 -norm, so

$$\sum_{i=1}^n \theta(x'_i, r_j) = \min_{x \in F} \sum_{i=1}^n \theta(x_i, r_j) \leq \inf_{x \in F} \|x\|_0. \quad (2.3)$$

Going through the limit in R for $j \rightarrow \infty$ and as the concave function is continuous and $x' \in F$, we have the results. □

Theorem 2.2.2 and Proposition 2.2.1 show that the scheme converge to (P_{ℓ_0}) as r decreases to zero.

The next theorem shows that for r sufficiently large the solutions of (??) are the same than solutions of (P_1) . This will be especially useful as an initialization of the homotopy scheme.

Theorem 2.2.2 (Convergence to ℓ_1 -norm). *Let $F \subset \mathbb{R}_+^n$ be a non-empty closed convex set. Every limit point of any sequence $\{x_r\}_r$, such that $x_r \in S_r^*$ for $r \uparrow \infty$, is an optimal solution of (P_1) .*

Proof. As $r > 0$, we can use a scaling technique for $S_r^{*(2)} = \arg \min_{x \in F} \sum_{i=1}^n r \theta(x_i, r)$

$$\min_{x \in F} \sum_{i=1}^n \theta(x_i, r) \iff \min_{x \in F} \sum_{i=1}^n r \theta(x_i, r), \quad (2.4)$$

$$S_r^* = S_r^{*(2)}. \quad (2.5)$$

So, it is sufficient to show that every limit point of any sequence $\{x_r\}_r$, such that $x_r \in S_r^*$ for $r \uparrow \infty$, is an optimal solution of (P_1) .

Given $x^r \in S_r^{*(2)}$ and $\bar{x} \in S_{\|\cdot\|_1}^*$. We use the first order Taylor's theorem for $\theta(t)$ in 0,

$$\theta(t) = t\theta'(0) + g(t), \text{ where } \lim_{t \rightarrow 0} \frac{g(t)}{t} = 0. \quad (2.6)$$

By concavity of the functions θ , it holds that $\theta'(0) > 0$.

By definition of \bar{x} , we get

$$\sum_{i=1}^n r\theta(x_i^r, r) \leq \sum_{i=1}^n r\theta(\bar{x}_i, r). \quad (2.7)$$

Now, using (2.6) yields

$$\sum_{i=1}^n x_i^r \theta'(0) + rg\left(\frac{x_i^r}{r}\right) \leq \sum_{i=1}^n x_i^r \theta'(0) + rg\left(\frac{\bar{x}_i}{r}\right), \quad (2.8)$$

$$\sum_{i=1}^n x_i^r - \sum_{i=1}^n \bar{x}_i \leq \frac{r}{\theta'(0)} \sum_{i=1}^n g\left(\frac{\bar{x}_i}{r}\right) - \frac{r}{\theta'(0)} \sum_{i=1}^n g\left(\frac{x_i^r}{r}\right), \quad (2.9)$$

$$\leq \frac{1}{\theta'(0)} \left| \sum_{i=1}^n \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} \bar{x}_i \right| + \frac{1}{\theta'(0)} \left| \sum_{i=1}^n \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} x_i^r \right|, \quad (2.10)$$

$$\leq \frac{1}{\theta'(0)} \left(\sum_{i=1}^n \left| \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} \right| \right) \left(\sum_{i=1}^n \bar{x}_i \right), \quad (2.11)$$

$$+ \frac{1}{\theta'(0)} \left(\sum_{i=1}^n \left| \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} \right| \right) \left(\sum_{i=1}^n x_i^r \right), \quad (2.12)$$

$$\sum_{i=1}^n x_i^r \leq \left(\sum_{i=1}^n \bar{x}_i \right) \frac{1 + \frac{1}{\theta'(0)} \left(\sum_{i=1}^n \left| \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} \right| \right)}{1 - \frac{1}{\theta'(0)} \left(\sum_{i=1}^n \left| \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} \right| \right)}. \quad (2.13)$$

Then, we show that the right-hand side in previous equation goes to 1, when passing to the limit.

It holds true that

$$\lim_{r \rightarrow +\infty} \frac{\bar{x}}{r} = 0. \quad (2.14)$$

Besides, by definition of x^r yields

$$\sum_{i=1}^n \theta(x_i^r, r) \leq \sum_{i=1}^n \theta(\bar{x}_i, r) \quad (2.15)$$

$$\lim_{r \rightarrow +\infty} \sum_{i=1}^n \theta(x_i^r, r) \leq \lim_{r \rightarrow +\infty} \sum_{i=1}^n \theta(\bar{x}_i, r) \leq 0. \quad (2.16)$$

so, we get

$$\lim_{r \rightarrow +\infty} \sum_{i=1}^n \theta(x_i^r, r) = 0. \quad (2.17)$$

By definition of functions θ , it is true that $\theta(x, r) := \theta(x/r)$ and $\theta^{-1}(0, r) = 0$. Thus, by previous equation we obtain

$$\lim_{r \rightarrow +\infty} \frac{x_i^r}{r} = 0 \quad \forall i. \quad (2.18)$$

Using (2.14) and (2.18) it follows

$$\lim_{r \rightarrow +\infty} \frac{\bar{x}_i}{r} = 0 \implies \lim_{r \rightarrow +\infty} \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} = 0, \quad (2.19)$$

$$\lim_{r \rightarrow +\infty} \frac{x_i^r}{r} = 0 \implies \lim_{r \rightarrow +\infty} \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} = 0. \quad (2.20)$$

Then going to the limit in (2.13) yields

$$\lim_{r \rightarrow +\infty} \sum_{i=1}^n x_i^r \leq \sum_{i=1}^n \bar{x}_i. \quad (2.21)$$

However, by definition of \bar{x} , it always hold that $\sum_{i=1}^n \bar{x}_i \leq \sum_{i=1}^n x_i$ for all x feasible for (P_1) . Since, this is true for the limit point of the sequence $\{x^r\}_r$, the inequality in (2.21) is actually an equality. So, the limit point of the sequence $\{x^r\}_r$ is also a solution of (P_1) . This proves the result. \square

The next theorem gives a monotonicity result, which illustrates the relations between the three problems (P_{ℓ_0}) , (P_1) and (P_r) . By monotonicity, we mean that for a given feasible point we want a relation of monotony in r for the objective function of (P_1) , (P_r) and (P_{ℓ_0}) . As the components of the ℓ_0 -norm and the $\theta_r(t)$ are in $[0, 1[$, it is necessary to put the components of the ℓ_1 -norm in a similar box, which explains the change of variable in the theorem.

Remark 2.2.1. *In the following theorem we use the hypothesis that θ functions are convex in r . This is not so restrictive as we think several functions verify it. If we take the three examples of θ functions given in the introduction, θ^1 and $\theta^{\log} := \log(1+x)/\log(1+x+r)$ are convex in r but not θ^2 .*

Theorem 2.2.3 (Monotonicity of solutions). *Let $F \subset \mathbb{R}_+^n$ be a non-empty closed convex set. Given $x \in F$, let $y = x/(||x||_\infty + \epsilon)$ where $\epsilon > 0$, so that $y \in [0, 1]^n$. Let a function $\Psi(t, r) : [0, 1] \rightarrow [0, 1]$ be defined as*

$$\Psi(t, r) = \frac{\theta(t, r)}{\theta(1, r)},$$

where $\theta(t, r)$ is the smooth function described in the introduction, which we will consider here as convex in r . For r and \bar{r} such that $0 < \bar{r} < r < +\infty$, then one has

$$||y||_1 \leq \sum_{i=1}^n \Psi(y_i, r) \leq \sum_{i=1}^n \Psi(y_i, \bar{r}) \leq ||y||_0.$$

Proof. The proof is divided in three step regarding the three inequalities.

The functions θ are sub-additive functions, since they concave and $\theta(0) = 0$. Then, it follows

$$\theta(y_i, r) \geq y_i \theta(1, r).$$

Therefore, we get

$$\begin{aligned} \sum_{i=1}^n \Psi(y_i, r) - \|y\|_1 &= \sum_{i=1}^n \left(\frac{\theta(y_i, r)}{\theta(1, r)} - y_i \right), \\ &\geq 0, \end{aligned}$$

which leads to the first inequality

$$\|y\|_1 \leq \sum_{i=1}^n \Psi(y_i, r).$$

We continue with the second inequality showing that $\Psi(y, r)$ functions are non-increasing in r , i.e.

$$\sum_{i=1}^n \Psi(y_i, r) \leq \sum_{i=1}^n \Psi(y_i, \bar{r}). \quad (2.22)$$

The function $\Psi(y, r)$ is non-increasing in r if its derivative with respect to r

$$\frac{\partial}{\partial r} \Psi(y, r) = \frac{(\frac{\partial}{\partial r} \theta(y, r))\theta(1, r) - (\frac{\partial}{\partial r} \theta(1, r))\theta(y, r)}{\theta(1, r)^2}, \quad (2.23)$$

is negative. Since $\theta(y, r)$ is a non-decreasing function in y we have

$$\frac{\theta(y, r)}{\theta(1, r)} \leq 1,$$

and

$$\frac{\partial}{\partial r} \theta(y, r) = -\frac{1}{r^2} \frac{\partial}{\partial y} \theta(y, r) \leq 0.$$

So, $\theta(y, r)$ is a non-increasing function in r . Using convexity of $\theta(y, r)$ in r it follows

$$\frac{\frac{\partial}{\partial r} \theta(y, r)}{\frac{\partial}{\partial r} \theta(1, r)} = \frac{\frac{\partial}{\partial r} \theta(1, r/y)}{\frac{\partial}{\partial r} \theta(1, r)} \geq 1.$$

Then in (2.23) the derivative with respect to r is negative and we have (2.22). Finally, since $\theta(y, r)$ is non-decreasing in y and $y \in [0, 1]^n$ one has

$$\|y\|_0 - \sum_{i=1}^n \Psi(y_i, \bar{r}) = \sum_{i=1; y_i \neq 0}^n 1 - \frac{\theta(y_i, \bar{r})}{\theta(1, \bar{r})} \geq 0,$$

which gives the last inequality and completes the theorem. \square

Remark 2.2.2. Both choices of scaling parameter in Theorem 2.2.2 and Theorem 2.2.3 are linked. In the former, we set that $\lim_{r \rightarrow +\infty} r \sum_{i=1}^n \theta(x_i, r) = \sum_{i=1}^n x_i$, so evaluating in one dimension and $x = 1$ we have $\lim_{r \rightarrow +\infty} r\theta(1, r) = 1$ and then we see that r and $1/\theta(1, r)$ have the same behaviour for r sufficiently large.

All this results lead us to the general behaviour of the method. First, we start from one solution of (P_1) then by decreasing parameter r the solution of (P_r) becomes closer to a solution of (P_{ℓ_0}) .

Another approach would be to define a new problem which selects one solution of the possibly many optimal solution of (P_{ℓ_0}) . We consider the following problem which is a selective version of (P_r)

$$\min_{x \in F} \sum_{i=1}^n \theta(x_i, r) + \sum_{i=1}^n r^{\frac{i}{2n+1}} x_i. \quad (P_{r-sel})$$

We use a lexicographic norm and we note

$$\|y\|_{lex} < \|x\|_{lex} \iff \exists i \in \{1, \dots, n\}, y_i < x_i \text{ and } \forall 1 \leq j < i, y_j = x_j.$$

In the next theorem we want to choose the solution of (P_{ℓ_0}) which has the smallest lexicographic norm. From the previous equation it is clear that this optimal solution is unique.

Theorem 2.2.4. [Convergence of the selective concave problem] We use functions θ such that $\theta \geq \theta^1$. Given $\{x^r\}_r$ a sequence of solutions of (P_{r-sel}) and \bar{x} the limit point of this sequence. Then, \bar{x} is the unique solution of $S_{\|\cdot\|_0}^*$ such that $\forall y \in S_{\|\cdot\|_0}^*, \|\bar{x}\|_{lex} \leq \|y\|_{lex}$.

Proof. Given x^* an optimal solution of (P_{ℓ_0}) such that $\forall y \in S_{\|\cdot\|_0}^*, \|x^*\|_{lex} \leq \|y\|_{lex}$ and \bar{x} the limit of a sequence of $\{x^r\}_r$ solution of (P_{r-sel}) . For all r , x_r is the minimum of a concave function over a bounded polyhedron thus by Theorem 32.3 of [170] it follows that the minimum is attained at one of the finitely many extreme point of F . Since, \bar{x} is assumed to be limit point of the sequence $\{x^r\}_r$, we can extract a subsequence such that for all r in this subsequence it holds that

$$\sum_{i=1}^n \theta(\bar{x}_i, r) + \sum_{i=1}^n r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=1}^n \theta(x_i^*, r) + \sum_{i=1}^n r^{\frac{i}{2n+1}} x_i^* \leq \|x^*\|_0 + \sum_{i=1}^n r^{\frac{i}{2n+1}} x_i^*. \quad (2.24)$$

Going to the limit for $r \downarrow 0$ we have

$$\|\bar{x}\|_0 \leq \|x^*\|_0,$$

which is an equality by definition of x^* and prove the first part of the theorem. Now we need to verify the selection of the solution. Using that for all $x \in \mathbb{R}$ $\theta(x, r) \leq 1$ in (2.24) one has

$$\sum_{i=1}^n (\theta(\bar{x}_i, r) - 1) + \sum_{i=1}^n r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=1}^n r^{\frac{i}{2n+1}} x_i^*.$$

Consider $\theta \geq \theta^1$, r sufficiently small ($\min_{\{i|x_i \neq 0\}} \bar{x}_i \geq \sqrt{r}$) and using that $\|\bar{x}\|_0 = k$ we have

$$-k \frac{r}{r + \sqrt{r}} + \sum_{i=1}^n r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=1}^n r^{\frac{i}{2n+1}} x_i^*. \quad (2.25)$$

Dividing by $r^{\frac{1}{2n+1}}$, we get

$$-k \frac{r}{r^{\frac{1}{2n+1}}(r + \sqrt{r})} + x_1 + \sum_{i=2}^n r^{\frac{i}{2n+1}} \bar{x}_i \leq x_1^* + \sum_{i=2}^n r^{\frac{i}{2n+1}} x_i^*.$$

Therefore, going to the limit for $r \downarrow 0$ one has

$$\bar{x}_1 \leq x_1^*,$$

which is an equality by hypothesis on x^* being the smallest $\|\cdot\|_{lex}$ solution of (P_{ℓ_0}) . So, as $\bar{x}_1 = x_1^*$ in (2.25) one has

$$-k \frac{r}{r + \sqrt{r}} + \sum_{i=2}^n r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=2}^n r^{\frac{i}{2n+1}} x_i^*.$$

By induction we get $\bar{x}_i = x_i^*$, $\forall i \in \{1, \dots, n\}$ and so $\bar{x} = x^*$, because we have

$$\lim_{r \rightarrow 0} \frac{r}{r^{\frac{j}{2n+1}}(r + \sqrt{r})} = 0, \quad \forall j \in \{1, \dots, n\}.$$

Finally we have the results as \bar{x} is the optimal solution which has the smallest lexicographic norm. \square

Remark 2.2.3. *If we try to get an equivalent result as in Theorem 2.2.2 for this selection problem, it is clear that for r sufficiently large we will solve the ℓ_1 -norm problem but with a reversed lexicographical order than the one we are looking for, i.e for a non-decreasing sequence of r*

$$\bar{x} = \lim_{r \rightarrow \infty} \{x^r\}_r \text{ with } x^r \in S_{r_j\text{-sel}}^* \implies \bar{x} \in S_{\|\cdot\|_1}^* \text{ and } \bar{x} = \arg \max_{y \in S_{\|\cdot\|_1}^*} \|y\|_{lex}.$$

This will definitely prevent us of any kind of monotonicity result such as Theorem 2.2.3. So, unless $S_{\|\cdot\|_0}^$ admits only one solution, the initial point as a solution of (P_1) has no chance of being a good initial point. This argument and the fact that this problem looks numerically not advisable lead us not to follow the study of this selective problem.*

2.3 Error estimate

In this section, we focus on what happen when r becomes small. We denote $\text{card}(I)$ the number of elements in a set I . Note that the following results are given for functions $\theta \geq \theta^1$.

Lemma 2.3.1. Consider θ functions where $\theta \geq \theta^1$. Let $\mathbb{N} \ni k = \|x^*\|_0 < n$ be the optimal value of problem (P_{ℓ_0}) and $I(x, r) := \{i | x_i \geq kr\}$. Then one has

$$x^r \in \arg \min_{x \in F} \sum_{i=1}^n \theta(x_i, r) \Rightarrow \text{card}(I(x^r, r)) \leq k.$$

Proof. We use a proof by contradiction. Consider that $\text{card}(I(x^r, r)) \geq k + 1$ and we have $x^r \in \arg \min_{x \in F} \sum_{i=1}^n \theta(x_i, r)$, then

$$\sum_{i=1}^n \theta(x_i^r, r) \geq (k+1)\theta(kr, r) \geq (k+1)\theta^1(kr, r) = (k+1)\frac{kr}{kr+r} = k,$$

which is a contradiction with the definition of x^r . \square

This lemma gives us a theoretical stopping criterion for the decrease of r , as for $r < \bar{r} = \min_{x_i^r \neq 0} x_i^r/k$, x^r becomes an optimal solution. In the following lemma we look at the consequences in the evaluation of θ .

Lemma 2.3.2. Consider θ functions where $\theta \geq \theta^1$. Let $\mathbb{N} \ni k = \|x^*\|_0 < n$ be the optimal value of problem (P_{ℓ_0}) and

$$\bar{r} := \min_{x_i^r \neq 0} x_i^r/k.$$

Then one has

$$r \leq \bar{r} \iff \theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1}.$$

Proof. We first show the equivalence in (??) for θ^1 . Assume that

$$\theta^1(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1}. \quad (2.26)$$

Using the expression of θ^1 , it follows

$$\theta^1(\min_{x_i^r \neq 0} x_i^r, r) = \frac{\min_{x_i^r \neq 0} x_i^r}{\min_{x_i^r \neq 0} x_i^r + r} \geq \frac{k}{k+1} \quad (2.27)$$

$$\iff \min_{x_i^r \neq 0} x_i^r(k+1) \geq k(\min_{x_i^r \neq 0} x_i^r + r) \quad (2.28)$$

$$\iff \min_{x_i^r \neq 0} x_i^r \geq kr \quad (2.29)$$

$$\iff \bar{r} = \frac{\min_{x_i^r \neq 0} x_i^r}{k} \geq r. \quad (2.30)$$

Considering the functions θ such that $\theta \geq \theta^1$, the equivalence follows in the exact same way. This proves the result. \square

Both previous lemmas lead us to the following theorem, which is an exact penalization result for our method.

Theorem 2.3.1 (Exact Penalization Theorem). *Consider θ functions where $\theta \geq \theta^1$. Let $\mathbb{N} \ni k = \|x^*\|_0 < n$ be the optimal value of problem (P_{ℓ_0}) and $x^r \in S_r^*$. Then one has*

$$\theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1} \implies x^r \in S_{\|\cdot\|_0}^*.$$

Proof. By Lemma 2.3.2 and with $\bar{r} = \min_{x_i^r \neq 0} x_i^r / k$ one has

$$\theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1} \iff r \leq \bar{r}.$$

Then by Lemma 2.3.1 and using $x_r \in S_r^*$ we have

$$x^r \in \arg \min_{x \in F} \sum_{i=1}^n \theta(x_i, r) \implies \text{card}(I(x_r, r)) \leq k.$$

Finally, using $r \leq \bar{r}$ and that k is the optimal value of problem in ℓ_0 -norm we have the result. \square

We use in the previous result the minimum non-zero component of x^r , which is logical as we expect that for r sufficiently small the sequence of $\{\min_{x_i^r \neq 0} x_i^r\}_r$ should be increasing. The following lemma gives us a clue on this behaviour.

Lemma 2.3.3. *Consider θ functions where $\theta \geq \theta^1$. Let $x^* \in S_{\|\cdot\|_0}^*$, $\|x^*\|_0 = k$ and*

$$r^* := \frac{1}{k} \min_{x_i^* \neq 0} x_i^*.$$

Then one has

$$\forall r \leq r^*, x^r \in S_r^* \implies \min_{x_i^r \neq 0} x_i^r \leq \min_{x_i^* \neq 0} x_i^*.$$

Proof. Suppose that $\min_{x_i \neq 0} x_i > \min_{x_i^* \neq 0} x_i^*$. Since $x^r \in S_r^*$ we have

$$\begin{aligned} \sum_{i=1}^n \theta(x_i^r, r) &\geq \sum_{i=1}^n \theta(x_i^r, r^*), \\ &> (k+1) \theta(\min_{x_i^* \neq 0} x_i^*, r^*), \\ &> (k+1) \frac{kr^*}{kr^* + r^*}, \\ &= k, \end{aligned}$$

which is in contradiction with the definition of x^r . \square

2.4 Algorithm

The previous results allow us to build a generic algorithm

$$[\text{Thetal0}] \begin{cases} \{r^k\}_{k \in \mathbb{N}}, r^0 > 0 \text{ and } \lim_{k \rightarrow +\infty} r^k = 0, \\ \text{find } x^k : x^k \in \arg \min_{x \in F} \sum_{i=1}^n \theta(x_i, r^k). \end{cases}$$

Now, several questions remain to be answered such as initialization, choice of the sequence $\{r^k\}$ and the method used to solve the concave minimization problems. In Sect. 2.3, we have shown an exact penalization result, which help us building a stopping criterion. We make a few remarks about these questions. Note that interesting related remarks can be found in [152].

Remark 2.4.1 (On the behaviour of θ functions). *These concave functions are acting as step function for r sufficiently small. That is one has the following behaviour*

$$\theta(t, r) \simeq \begin{cases} 1 & \text{if } t \gg r, \\ 0 & \text{if } t \ll r. \end{cases}$$

which gives us a strategy to update r . Let x^k be our current iterate and r^k the corresponding parameter. We divide our iterate into two sets, those with indices in $I = \{i \mid x_i^k \geq r^k\}$ and the others with indices in $\bar{I} = \{i \mid x_i^k < r^k\}$. We can see I as the set of indices of the "non-zero" components and \bar{I} as the set of indices of the "zero" components of x^k . So we will choose r^{k+1} around $\max_{i \in \bar{I}} x_i^k$ to ask whether or not it belongs to zeros and we repeat this operation until r is sufficiently small to consider \bar{I} the set of effective zeros. Also this is a general behaviour, to be sure to have decrease of r one should add a fixed parameter of minimum decrease.

Remark 2.4.2 (Initialization). *It is the main purpose of our method to start with the solution x^0 of the problem (P_1) , which is a convex problem. So, we need to find the r^0 related to x^0 . A natural, but non-trivial, way of doing this would be to find the parameter which minimizes the following problem*

$$\min_{r > 0} \left\| \sum_{i=1}^N \theta(x_i^0, r) - \|x^0\|_1 \right\|_2^2.$$

A simpler idea is to be inspired from last remark and put r^0 as a value which is just beyond the top value of x_i^0 .

Remark 2.4.3 (Stopping criterion). *It has been shown, in Sect. 2.3, an exact penalization theorem using the quantity $k/(k+1)$, which depends on the solution we are looking for. Numerically, we can make more iterations but being sure to satisfy this criterion using the fact that $\|x^0\|_0 \geq k$, which gives us the following criterion*

$$\theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{\|x^0\|_0}{\|x^0\|_0 + 1} \geq \frac{k}{k+1}.$$

Remark 2.4.4 (Algorithm for concave minimization). *In the same way as in [82] and [167] we will use a successive linearization algorithm (SLA) algorithm to solve the concave minimization problem at each iteration in r . This algorithm is a finitely timestep Frank & Wolfe algorithm, [141].*

Proposition 2.4.1 (SLA algorithm for concave minimization). *Given ϵ sufficiently small and r^k . We know x^k and we find x^{k+1} as a solution of the linear problem*

$$\min_{x \in F} x^t \nabla_x \theta(x^k, r^k),$$

with x^0 a solution of the problem (P_1) . We stop when

$$x^{k+1} \in F \text{ and } (x^{k+1} - x^k)^t \nabla_x \theta(x^k, r^k) \leq \epsilon.$$

This algorithm generates a finite sequence with strictly decreasing objective function values.

Proof. see [[141], Theorem 4.2]. □

We note that this algorithm didn't provide necessarily a global optimum as it ends in a local solution, so we don't expect global solutions in our algorithm. Also when considering the objective function of this linear program the gradient of functions θ tends to be very large as $\theta'_r(t) \approx O(1/r)$, so it can be numerically efficient to add a scaling parameter of order r .

2.5 Numerical simulations

Thanks to the previous sections we have keys for an algorithm. We will show now some numerical results. These simulations have been done using MATLAB language, [146], with the linear programming solver GUROBI, [92].

The precision in our simulations is $\epsilon = 10^{-8}$. We generate various polyhedron $F = \{x \in \mathbb{R}^n \mid b \in \mathbb{R}^m, Ax \leq b\} \cap \mathbb{R}_+^n$ with $m < n$. In the same way as in [82] we choose $n = (500, 750, 1000)$ and in each case $m = (40\%, 60\%, 80\%)$. For each pair (n, m) we choose randomly one hundred problems. We take a random matrix A of size $m \times n$ and a default sparse solution x_{init} with 10% of non-zero components. We get b by calculating the matrix-vector product $b = Ax_{init}$. In the end, we will compare the sparsity of the solution from Thetal0-algorithm using θ^1 ($\#\theta^1$), the default sparse solution ($\#\ell_0$) and the initial iterate ($\#\ell_1$). We get the initial iterate as a solution of problem (P_1) . The item $\#$ indicates the number of non-zero components in a vector.

Results are sum up in Table 2.1. The first two columns give the dimensions of the problems, Column 3 gives the number of problems where the solution of θ algorithm has at least the same sparsity as the default sparse solution. Column 4 in the same vein compare the sparsity of the solution in ℓ_1 -norm with the default sparse solution. Column 5 gives the number of problems where the solution by θ -algorithm improves strictly the solution by ℓ_1 -norm.

Table 2.1: Numerical results with random $F = \{x \in \mathbb{R}_+^n \mid b \in \mathbb{R}^m, Ax \leq b\}$, dimensions of the problem are first 2 columns. Compare a default sparse solution with 10% of non-zero components, $\#l_0$, the initial iterate solution of (P_1) , $\#l_1$, and the solution by θ -algorithm with function θ^1 , $\#\theta^1$. The item $\#$ indicates the number of non-zeros.

n	m	$\#l_0 \geq \#\theta^1$	$\#l_1 \leq \#l_0$	$\#\theta^1 < \#l_1$
1000	800	100	100	0
1000	600	100	98	2
1000	400	50	1	99
750	600	100	100	0
750	450	100	98	2
750	300	54	0	100
500	400	100	100	0
500	300	100	94	6
500	200	63	0	100

These results validate our algorithm, as in the majority of the cases it manages to find at least an equivalent solution to the default sparse solution. One may notice that in many cases the ℓ_1 -norm minimization solution solves the problem in ℓ_0 -norm, which is not surprising according to [61].

In Figure 2.1, we show the behaviour of the minimum non-zero component of the current iterate along the iterations in r for one example. We can see the increasing behaviour that is the general behaviour expected in the Remark 2.4.3.

2.6 Conclusion and outlook

We proposed a class of heuristics schemes to solve the NP-hard problem of minimizing the ℓ_0 -norm. Our method requires to find a sequence of solution from concave minimization problem, which we solved with a successive linearization algorithm. This method has the benefit that it can only improve the solution we get by solving the ℓ_1 -norm problem. We gave convergence results, an exact penalization theorem and keys to implement the method. To confirm the validity of this algorithm we gave numerical results from randomly generated problems.

Further studies can investigate the special case where the ℓ_1 -norm solves the ℓ_0 -norm problem, to find an improved stopping condition. Thanks to several studies, for instance [61], we have criteria which can help us identifying the cases where the solution we get by solving (P_1) is an optimal solution of (P_{ℓ_0}) . We can wonder if there exists a better sufficient condition than the one presented here in the case where $x^r \in S_{\|\cdot\|_1}^* \cap S_r^*$.

We can also study a very similar problem which is the one of minimizing ℓ_0 -norm with noise, see for instance [62] or [22], that is

$$(P_{0,\delta}) \min \|x\|_0 \text{ s.t. } Ax \leq b + \delta.$$

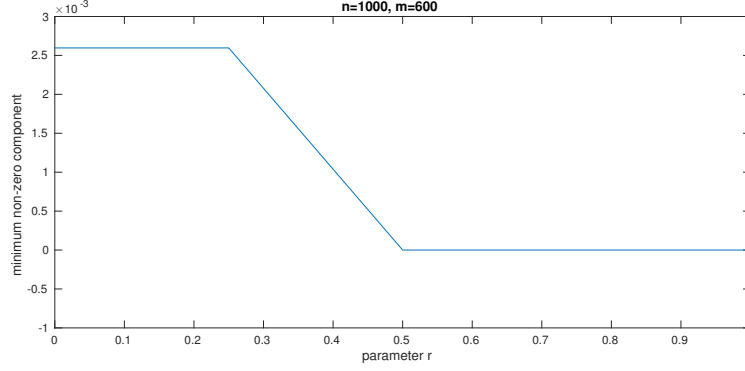


Figure 2.1: Evolution of the minimum non-zero component of x_r in function of the parameter $r \downarrow 0$.

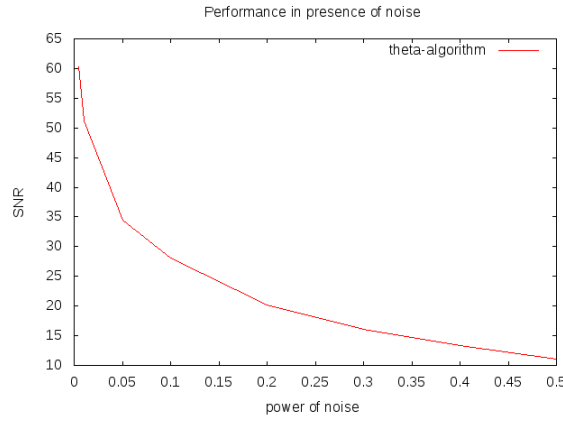


Figure 2.2: Performance of θ -algorithm in presence of noise, using function θ^1 . $n = 500$, $m = 200$. Mean of SNR for 100 random problems in function of σ^2 .

As a first step in this direction we run our heuristic schemes on some perturbed problems. We generate polyhedron in a similar way as in the previous section with noise in $b = Ax_{init} + \vartheta$, where ϑ follows $\mathcal{N}(0, \sigma^2 I_{n \times n})$. We build a signal to noise ratio (SNR) for several values of σ^2 from 0.5 to 0,

$$SNR = 20 \log\left(\frac{\|x^*\|_2}{\|x^* - x_b\|_2}\right),$$

where x^* and x_b are generated by our algorithm, the former comes from the problem without noise and the later from the perturbed problem. We choose dimensions $n = 500$ and $m = 200$. Then for one hundred randomly selected problems we compute the mean of the SNR. Results in Figure 2.2 show very logical behaviour as more noise is present more informations are lost. Further work could compare these results with existing methods and shows theoretical study, which could help building an improved algorithm.

Chapter 3

Relaxation methods to solve the absolute value equation

This chapter is a paper accepted in Journal of Computational and Applied Mathematics entitled:

Solving Absolute Value Equation using Complementarity and Smoothing Functions

Authors:

L. Abdallah, Université Libanaise (LaMa), Tripoli, Liban

M. Haddou, IRMAR-INSA, Rennes, France

T. Migot, IRMAR-INSA, Rennes, France

Abstract: In this paper, we reformulate the NP-hard problem of the absolute value equation (AVE) as a horizontal linear complementarity one and then solve it using a smoothing technique. This approach leads to a new class of methods that are valid for general absolute value equation. An asymptotic analysis proves the convergence of our schemes and provides some interesting error estimates. This kind of error bound or estimate had never been studied for other known methods. The corresponding algorithms were tested on randomly generated problems and applications. These experiments show that, in the general case, one observes a reduction of the number of failures.

AMS Subject Classification : 90C59 ; 90C30 ; 90C33 ; 65K05 ; 49M20

Keywords : smoothing function ; concave minimization ; complementarity ; absolute value equation

3.1 Introduction

We consider the absolute value equation, which consists in finding $x \in \mathbb{R}^N$ such that

$$Ax - |x| = b, \quad (\text{AVE})$$

where $A \in \mathbb{R}^{N \times N}$, $b \in \mathbb{R}^N$ and $|x| := (|x_1|, \dots, |x_N|)^T$. A slightly more general problem has been introduced in [172]

$$Ax + B|x| = b, \quad (\text{AVP})$$

where $A, B \in \mathbb{R}^{M \times N}$, $b \in \mathbb{R}^M$ and unknown $x \in \mathbb{R}^N$. Here, we focus on (AVE), which has been more popular in the literature. The recent interest in these problems can be explained by the fact that frequently occurring optimization problems such as linear complementarity problems and mixed integer programming problems can be reformulated as an (AVE), see [161, 139]. The general NP-hard linear complementarity problem can be formulated as an (AVE), which implies that it is NP-hard in general. Moreover, it has been proved in [161] that checking if an (AVE) has one or an infinite number of solutions is NP-complete.

Theoretical criteria regarding existence of solutions and unique solvability of (AVE) have been studied in [161, 143, 176, 175, 130]. An important criterion among others is that (AVE) has a unique solution if all of the singular values of the matrix A exceed 1. In the special case where the problem is uniquely solvable, a family of Newton methods has been proposed first in [134], then completed with global and quadratic convergence in [37], an inexact version in [25] and other related methods [97, 154, 204]. Also, Picard-HSS iteration methods and nonlinear HSS-like methods have been considered for instance in [180, 163, 206]. It is of a great interest to consider methods that remain valid in the general case. Most of such methods which are valid in the general case are due to Mangasarian in [136, 137, 138] by considering a concave or a bilinear reformulation of (AVE) solved by a sequence of linear programs. A hybrid method mixing Newton approach of [134] and [138] can be found in [135]. A method based on interval matrix has been studied by Rohn in [173, 174].

The special case where (AVE) is not solvable also received some interests in the literature. Prokopyev shows numerical results using a mixed integer programming solver in [161]. Theoretical study in order to correct b and A to make (AVE) feasible can be found in [117, 118].

Our aim in this paper is to pursue the study of (AVE) without additional hypothesis such as unique solvability and propose a new method, which solves a sequence of linear programs. The motivation is to diminish the number of instances where classical methods can not solve the problem. We propose a new reformulation of (AVE) as a sequence of concave minimization problems using complementarity and a smoothing technique.

The paper is organized as follows. In Section 2, we present the new formulation of (AVE) as a sequence of concave minimization problems. In Section 3, we prove convergence to a solution of (AVE) and in Section 4, we establish error estimate. Finally, Section 5 provides numerical results on simple examples and randomly generated test problems.

3.2 AVE as a sequence of concave minimization programs

We consider a reformulation of (AVE) as a sequence of concave minimization problems. First, we use a classical decomposition of the absolute value to reformulate (AVE) as an horizontal linear complementarity problem. Set $x = x^+ - x^-$, where $x^+ \geq 0$, $x^- \geq 0$ and $x^+ \perp x^-$, so that $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$. This decomposition guarantees that $|x| = x^+ + x^-$. So (AVE) is equivalent to the following complementarity problem

$$\begin{cases} A(x^+ - x^-) - (x^+ + x^-) = b, \\ x^+ \geq 0, \ x^- \geq 0, \\ x^+ \perp x^-. \end{cases} \quad (3.1)$$

Now, we reformulate this problem as a sequence of concave optimization problems using a smoothing technique. This technique has been first studied in [150, 94] and uses a family of non-decreasing continuous smooth concave functions $\theta : \mathbb{R} \rightarrow]-\infty, 1[$ that satisfies

$$\theta(t) < 0 \text{ if } t < 0, \ \theta(0) = 0 \text{ and } \lim_{t \rightarrow +\infty} \theta(t) = 1.$$

One generic way to build such functions is to consider non-increasing probability density functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and then take the corresponding cumulative distribution functions

$$\forall t \geq 0, \quad \theta(t) = \int_0^t f(x) dx.$$

By definition of f

$$\lim_{t \rightarrow +\infty} \theta(t) = \int_0^{+\infty} f(x) dx = 1 \text{ and } \theta(0) = \int_0^0 f(x) dx = 0.$$

The hypothesis on f gives the concavity of θ on \mathbb{R}_+ . We then extend the functions θ for negative values in a differentiable way, for instance taking $\theta(t < 0) = t f(0)$.

Two interesting examples of this family are $\theta^1(t) = t/(t+1)$ if $t \geq 0$ and $\theta^1(t) = t$ if $t < 0$, $\theta^2(t) = 1 - e^{-t}$ with $t \in \mathbb{R}$. In particular, θ^1 will play a special role in our analysis.

We introduce $\theta_r(t) := \theta\left(\frac{t}{r}\right)$ for $r > 0$. This definition is similar to the perspective functions in convex analysis. These functions satisfy

$$\theta_r(0) = 0 \ \forall r > 0 \text{ and } \forall t > 0, \ \lim_{r \searrow 0} \theta_r(t) = 1.$$

Previous examples lead to $\theta_r^1(t) = t/(t+r)$ if $t \geq 0$ and $\theta_r^1(t) = t/r$ if $t < 0$, $\theta_r^2(t) = 1 - e^{-t/r}$ with $t \in \mathbb{R}$.

The following lemma shows the link between this family of functions and the complementarity.

Lemma 3.2.1. *Given $s, t \in \mathbb{R}_+$, we have*

$$s \perp t \iff \lim_{r \searrow 0} \theta_r(s) + \theta_r(t) \leq 1.$$

Proof. Prove by contradiction that

$$\lim_{r \searrow 0} \theta_r(s) + \theta_r(t) \leq 1 \implies s \perp t.$$

Suppose $s, t > 0$, then

$$\lim_{r \searrow 0} (\theta_r(s) + \theta_r(t)) = \lim_{r \searrow 0} \theta_r(s) + \lim_{r \searrow 0} \theta_r(t) = 2.$$

This leads to a contradiction and therefore $s \perp t$. Conversely it is clear that $s \perp t$ implies $s = 0$ or $t = 0$ and the result follows. \square

In the case of the function θ_r^1 , it holds that

$$\theta_r^1(s) + \theta_r^1(t) = 1 \iff st = r^2.$$

Using the previous lemma, problem (3.1) can be replaced by a sequence of concave optimization problems for $r > 0$:

$$\begin{aligned} \min_{x^+, x^- \in \mathbb{R}^N} \quad & \sum_{i=1}^N \theta_r(x_i^+) + \theta_r(x_i^-) - 1, \\ & A(x^+ - x^-) - (x^+ + x^-) = b, \\ & x^+ \geq 0, \quad x^- \geq 0. \end{aligned} \tag{3.2}$$

In order to avoid compensation phenomenon and generate strictly feasible iterates we consider a relaxed version of (3.2):

$$\left\{ \begin{array}{l} \min_{x^+, x^- \in \mathbb{R}^N} \quad \sum_{i=1}^N \theta_r(x_i^+) + \theta_r(x_i^-) - 1, \\ -g(r)|A|e - g(r)e \leq A(x^+ - x^-) - (x^+ + x^-) - b \leq g(r)|A|e + g(r)e, \\ x^+ + x^- \geq g(r)e, \\ 0 \leq x^+ \leq M, \quad 0 \leq x^- \leq M, \end{array} \right. \tag{P_r}$$

where e is the unit vector, $|A|$ denotes the matrix where each element is the absolute value of the corresponding element in A , M is some positive constant to be specified later and $g : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ is a function satisfiung

$$\lim_{r \searrow 0} \frac{r}{g(r)} = 0 \text{ and } \lim_{r \searrow 0} g(r) = 0.$$

For instance, we can choose $g(r) = r^\alpha$ with $0 < \alpha < 1$.

3.3 Convergence

From now on, we assume that the set of solutions of (AVE) denoted $S_{(AVE)}^*$ is non-empty and denote $S_{(P_r)}^*$ the optimal set of (P_r) . In order to simplify the notation, we denote $x \in S_{(P_r)}^*$ when $(x^+, x^-) \in S_{(P_r)}^*$ with $x = x^+ - x^-$ and $x^+ = \max(x, 0)$, $x^- = \max(-x, 0)$. Let M be a positive constant such that

$$M \geq \|x^*\|_\infty,$$

where x^* is some solution of $S_{(AVE)}^*$. The following theorem shows that for $r > 0$, the set of solutions $S_{(P_r)}^*$ is non-empty.

Theorem 3.3.1. *(P_r) has at least one solution for any $r > 0$.*

Proof. Since $S_{(AVE)}^* \neq \emptyset$, there exists a point $\bar{x} \in S_{(AVE)}^*$. We can write $\bar{x} = \bar{x}^+ - \bar{x}^-$ with $\bar{x}^+ \perp \bar{x}^-$. It follows that $(y^{r+} := \bar{x}^+ + g(r), y^{r-} := \bar{x}^-)$ is a feasible point of (P_r) . Furthermore, we minimize a continuous function over a non-empty compact set so the objective function attains its minimum. \square

The following three lemmas will be used to prove the main convergence Theorem 3.3.2.

Lemma 3.3.1. *For $r > 0$, functions θ_r and g defined above and $x^+, x^- \in \mathbb{R}_+^N$, such that $x^+ + x^- \geq g(r)e$. It holds that*

$$\forall i \in \{1, \dots, N\}, \quad \theta_r(x_i^+) + \theta_r(x_i^-) - 1 \geq \theta_r(g(r)) - 1.$$

Proof. θ_r is concave and $\theta_r(0) = 0$ so θ_r is sub-additive on \mathbb{R}_+ . Thus, for all $i \in \{1, \dots, N\}$ it follows that

$$\theta_r(x_i^+) + \theta_r(x_i^-) - 1 \geq \theta_r(x_i^+ + x_i^-) - 1.$$

Furthermore θ_r is non-decreasing and $x^+ + x^- \geq g(r)e$ therefore

$$\theta_r(x_i^+) + \theta_r(x_i^-) - 1 \geq \theta_r(g(r)) - 1.$$

\square

Lemma 3.3.2. *Given functions θ_r and g defined above we have*

$$\lim_{r \searrow 0} \theta_r(g(r)) - 1 = 0.$$

Proof. Since $\theta_r(g(r)) = \theta_{\frac{r}{g(r)}}(1)$ and $\lim_{r \searrow 0} r/g(r) = 0$, it follows that

$$\lim_{r \searrow 0} \theta_r(g(r)) = \lim_{r \searrow 0} \theta_{\frac{r}{g(r)}}(1) = 1. \quad (3.3)$$

\square

In the special case, where many solutions of (AVE) have at least a zero component, it can be difficult to find a feasible point (x^+, x^-) of (P_r) such that $x^+ + x^- \geq g(r)e$. The following lemma explains how to build such point in this case.

Lemma 3.3.3. *Let $\bar{x} \in S_{(AVE)}^*$ and $r > 0$ be such that $g(r) < \min_{\bar{x}_i \neq 0} |\bar{x}_i|$. Then $y^r := \bar{x} + g(r)$ is a solution of the following equation*

$$Ax - |x| = b + g(r)Ae - g(r)\delta(\bar{x}), \quad ((AVE)_r)$$

where $\delta(x) \in \mathbb{R}^N$ is such that $\delta_i(x) := \begin{cases} 1 & \text{if } x_i \geq 0, \\ -1 & \text{if } x_i < 0. \end{cases}$

Proof. \bar{x} is a solution of (AVE), that is

$$A\bar{x} - |\bar{x}| = b.$$

Therefore it holds that

$$A\bar{x} + g(r)Ae - |\bar{x}| - g(r)\delta(\bar{x}) = b + g(r)Ae - g(r)\delta(\bar{x}),$$

and so

$$A(\bar{x} + g(r)e) - |\bar{x} + g(r)e| = b + g(r)Ae - g(r)\delta(\bar{x}).$$

Thus, $y^r = \bar{x} + g(r)$ is a solution of $((AVE)_r)$. \square

We now prove the convergence of $\{x^r\}_{r>0}$, where $x^r := x^{r+} - x^{r-}$ with $(x^{r+}, x^{r-}) \in S_{(P_r)}^*$, to an element of $S_{(AVE)}^*$. It is to be noted that $S_{(P_r)}^*$ is not necessarily a singleton.

Theorem 3.3.2. *Every limit point of the sequence $\{x^r\}$ such that $x^r \in S_{(P_r)}$ for $r > 0$ is a solution of (AVE).*

Proof. By Theorem 3.3.1 there exists at least one solution of (P_r) . According to Lemma 3.3.3, we can build a sequence $\{y^r\}_{r>0}$ where $y^r = y^{r+} - y^{r-}$ with $y^{r+} \perp y^{r-}$ that are solutions of $(AVE)_r$. Furthermore, for r sufficiently small (y^{r+}, y^{r-}) is a feasible point of (P_r) . Let $x^r = (x^{r+}, x^{r-})$ with $\{x^r\}_{r>0}$ be a sequence of optimal solutions of (P_r) , then

$$\sum_{i=1}^N (\theta_r(x_i^{r+}) + \theta_r(x_i^{r-}) - 1) \leq \sum_{i=1}^N (\theta_r(y_i^{r+}) + \theta_r(y_i^{r-}) - 1) \leq 0.$$

For all $i \in \{1, \dots, N\}$, it holds that

$$\theta_r(x_i^{r+}) + \theta_r(x_i^{r-}) - 1 \leq - \sum_{j=1; j \neq i}^N (\theta_r(x_j^{r+}) + \theta_r(x_j^{r-}) - 1).$$

By Lemma 3.3.1, we obtain for all $i \in \{1, \dots, N\}$, that

$$\theta_r(x_i^{r+}) + \theta_r(x_i^{r-}) \leq 1 + (N-1)(1 - \theta_r(g(r))).$$

For any limit point $\bar{x} = (\bar{x}^+, \bar{x}^-)$ of the sequence $\{x^r\}_r$, where $\bar{x}^+ = \lim_{r \searrow 0} x^{r+}$ and $\bar{x}^- = \lim_{r \searrow 0} x^{r-}$, using Lemma (3.3.2) ($\lim_{r \searrow 0} 1 - \theta_r(g(r)) = 0$) and passing to the limit, it follows that

$$\lim_{r \searrow 0} \theta_r(\bar{x}_i^+) + \theta_r(\bar{x}_i^-) \leq 1.$$

Thus, $\bar{x}^+ \perp \bar{x}^-$ by the previous inequality and Lemma 3.2.1.

Now, we verify that \bar{x} is a solution of (AVE). Let $r > 0$ and x^r be a solution of (P_r) , we have

$$b - g(r)|A|e - g(r)e \leq A(x^{r+} - x^{r-}) - (x^{r+} + x^{r-}) \leq b + g(r)|A|e + g(r)e.$$

Passing to the limit when $r \searrow 0$, we obtain

$$A(\bar{x}^+ - \bar{x}^-) - (\bar{x}^+ + \bar{x}^-) = b.$$

So, $\bar{x} = \bar{x}^+ - \bar{x}^-$ is a solution of (AVE). \square

We now continue the discussion on convergence by a characterization of the limit points in the case where $S_{(AVE)}^* = \emptyset$.

Theorem 3.3.3. *Assume that $S_{(AVE)}^*$ is empty and that (3.2) admits a feasible point whose infinite norm is bounded by M . Then, every limit point of $\{x^r\}$ such that $x^r \in S_{(P_r)}$ for $r > 0$, does not satisfy the complementarity constraint.*

Proof. A straightforward adaptation of Theorem 3.3.1 to the new assumption proves that there exists a sequence $\{x^r\}$ such that $x^r \in S_{(P_r)}$ for $r > 0$. Therefore, we can extract up to a subsequence a limit point of this sequence since by assumption it is bounded.

By contradiction, if the limit point satisfies the complementarity constraint. Then, this point will be a solution of (AVE). However, this is a contradiction with the assumption that $S_{(AVE)}^*$ is empty. \square

In the case where no feasible point of (3.2) exists. It clearly holds that for r sufficiently small the problem (P_r) becomes infeasible.

3.4 Error estimate

In this section we study, the asymptotic behaviour of the sequence $\{x^r\}_{r>0}$ for small values of r . We remind the definition of the Landau notation O often used in the context of asymptotic comparison. Given two functions f and h , one writes

$$f(x) = O_{x \rightarrow a}(h(x))$$

if $\exists C > 0$, $\exists d > 0$ such that $|f(x)| \leq C|h(x)|$ when $|x - a| \leq d$. This notation becomes $O(h(x))$ when a is 0.

The following simple lemma will be useful for the rest of this section.

Lemma 3.4.1. *Let θ_r be such that $\theta_r \geq \theta_r^1$. For $x^r \in S_{(P_r)}^*$ and r sufficiently small, we have*

$$\forall i \in \{1, \dots, N\}, \quad x_i^{r+} x_i^{r-} = O(r g(r)).$$

Proof. Set $i \in \{1, \dots, N\}$. Thanks to the convergence proof of the Theorem 3.3.2 for r sufficiently small it holds that

$$\theta_r(x_i^{r+}) + \theta_r(x_i^{r-}) - 1 \leq (N-1)(1 - \theta_r(g(r))).$$

Thus, it also holds, in particular for $\theta = \theta^1$, that

$$\theta_r^1(x_i^{r+}) + \theta_r^1(x_i^{r-}) - 1 \leq (N-1)(1 - \theta_r^1(g(r))).$$

Therefore for r sufficiently small (i.e. such that $g(r) \geq r$ and $(1 - (N-1)g(r)) > 0$) it holds that

$$\begin{aligned} \frac{x_i^{r+}}{x_i^{r+} + r} + \frac{x_i^{r-}}{x_i^{r-} + r} - 1 &\leq (N-1)\left(1 - \frac{g(r)}{g(r) + r}\right) \leq (N-1)g(r), \\ \frac{2x_i^{r+}x_i^{r-} + rx_i^{r+} + rx_i^{r-} - x_i^{r+}x_i^{r-} - rx_i^{r+} - rx_i^{r-} - r^2}{(x_i^{r+} + r)(x_i^{r-} + r)} &\leq (N-1)g(r), \\ \frac{x_i^{r+}x_i^{r-} - r^2}{(x_i^{r+} + r)(x_i^{r-} + r)} &\leq (N-1)g(r), \\ x_i^{r+}x_i^{r-} - r^2 &\leq (N-1)g(r)(x_i^{r+}x_i^{r-} + rx_i^{r+} + rx_i^{r-} + r^2), \\ x_i^{r+}x_i^{r-} &\leq r^2g(r)\frac{1 + (N-1)}{1 - (N-1)g(r)} + rg(r)\frac{(N-1)(x_i^{r+} + x_i^{r-})}{1 - (N-1)g(r)}, \end{aligned}$$

and the results follows. \square

The following proposition gives a first error estimate that concerns only the components of (x^{r+}, x^{r-}) converging to zero.

Proposition 3.4.1. *Let θ_r be such that $\theta_r \geq \theta_r^1$. Let (\bar{x}^+, \bar{x}^-) be a limit point of the sequence $\{x^{r+}, x^{r-}\}_r$ of optimal solutions of (P_r) . The convergence of the components of the variable x^{r+} or x^{r-} to the possibly zero part of the accumulation point is done in $O(r)$.*

Proof. Set $i \in \{1, \dots, N\}$. We work with one component. Assume that $\bar{x}_i^+ = 0$. The opposite case is completely similar. By assumption on x_i^{r+} , for r sufficiently small it holds that $\max(x_i^{r+}, x_i^{r-}) = x_i^{r-}$. Thus, for r sufficiently small we have

$$g(r) \leq x_i^{r+} + x_i^{r-} \leq 2x_i^{r-}.$$

Additionally, by Lemma 3.4.1 it follows

$$\begin{aligned} x_i^{r+}x_i^{r-} &= O(rg(r)), \\ \implies x_i^{r+} &= \frac{O(rg(r))}{x_i^{r-}}, \\ \implies x_i^{r+} &= \frac{O(rg(r))}{g(r)}, \\ \implies |x_i^{r+} - \bar{x}_i^+| &= O(r). \end{aligned}$$

\square

In the next theorem we provide an error estimate for the non-zero part of the solution in the couple (x^+, x^-) .

To establish this result, we use the classical Hoffman's lemma on linear inequalities.

Lemma 3.4.2 (Hoffman's lemma,[99]). *Given a convex polyhedron P such that*

$$P = \{x \in \mathbb{R}^n \mid Ax \leq b\}.$$

We set $d_P(x)$ the distance from x to P , by choosing a norm $\|\cdot\|$, where $d_P(x) = \inf_{y \in P} \|y - x\|$. There exists a constant K that only depends on A , such that

$$\forall x \in \mathbb{R}^n : d_P(x) \leq K \|(Ax - b)^+\|.$$

It is to be noted that if the constraints are given by $Ax = b$ with A a square full-rank matrix instead of $Ax \leq b$ then the polyhedron is reduced to a singleton and one can choose $K = \|A^{-1}\|$.

Theorem 3.4.1. *Given (\bar{x}^+, \bar{x}^-) a limit point of the sequence $\{x^{r+}, x^{r-}\}$, if we denote $\bar{x} = \bar{x}^+ - \bar{x}^-$ and $x^r = x^{r+} - x^{r-}$. Then, for r sufficiently small*

$$d_{S_{(AVE)}^*}(x^r) = O(g(r)). \quad (3.4)$$

We remind here that the function g can be chosen such that $g(r) = r^\alpha$ with $\alpha \in (0, 1)$. This means that (3.4) is almost a linear bound.

Proof. We split the proof in two cases, either $\min_{i \in \{1, \dots, N\}} |\bar{x}_i| \neq 0$, either $\exists i \in \{1, \dots, N\}$, $\bar{x}_i = 0$ respectively denoted as a) and b).

a) First, assume that there is no zero component in \bar{x} . Let V be a neighbourhood of \bar{x} defined as

$$V = B_\infty(\bar{x}, \alpha) = \{x \mid \max_{1 \leq i \leq N} |x_i - \bar{x}_i| \leq \alpha\},$$

where $\alpha = \min_{i \in \{1, \dots, N\}} |\bar{x}_i|/2$. For all $x \in V$, \bar{x} and x have the same signs component-wise.

Denote $D = \text{diag}(\delta(\bar{x}))$, where $\delta(x) \in \mathbb{R}^N$ with $\delta_i(x) = \begin{cases} 1 & \text{if } x_i \geq 0, \\ -1 & \text{if } x_i < 0. \end{cases}$

By taking $S^* = \{x \in \mathbb{R}^n \mid Ax - Dx = b\} \cap V$ we obtain a convex polyhedron. This set is non-empty because $\bar{x} \in S^*$. In the neighbourhood V , solving $Ax - Dx = b$ gives a solution of (AVE). Using Hoffman lemma [99] for r sufficiently small such that $x_r \in V$ there exists a constant K such that

$$\begin{aligned} d_{S^*}(x^r) &\leq K \left\| \begin{pmatrix} (A - D)x^r - b \\ (x^r - \alpha - \bar{x})^+ \\ (-x^r - \alpha + \bar{x})^+ \end{pmatrix} \right\| \leq K(\|(A - D)x^r - b\| + \|(x^r - \alpha - \bar{x})^+\| \\ &\quad + \|(-x^r - \alpha + \bar{x})^+\|), \\ &= K\|(A - D)x^r - b\|, \\ &= K\|Ax^r - |x^r| - b\|. \end{aligned}$$

Since x^r is feasible for (P_r) , it holds that

$$\begin{aligned}
\|Ax^r - |x^r| - b\| &= \|g(r)Ae - g(r)\delta(x^r)\|, \\
&= \|(Ae - \delta(x^r))g(r)\|, \\
&\leq \|Ae - \delta(x^r)\| |g(r)|, \\
&= \|Ae - \delta(x^r)\|g(r) = O(g(r)).
\end{aligned}$$

Combining both previous inequalities, we obtain

$$d_{S^*}(x^r) \leq K\|Ae - \delta(x^r)\|g(r) = O(g(r)).$$

b) Now we move to the case where $\exists i \in \{1, \dots, N\}$, $\bar{x}_i = 0$. We denote $\sigma(t) = \{i | t_i \neq 0\}$. Set $\alpha = \min_{i \in \sigma(\bar{x})} |\bar{x}_i|/2$ and a neighbourhood V of \bar{x} defined as

$$V = B_\infty(\bar{x}, \alpha) = \{x \mid \max_{i \in \sigma(\bar{x})} |x_i - \bar{x}_i| \leq \alpha\}.$$

V is non-empty because $\bar{x} \in V$. For all $x \in V$, \bar{x} and x have the same signs for the components $i \in \sigma(\bar{x})$. Furthermore, for r sufficiently small we have $x^r \in V$.

Taking $S^* = \{x \in \mathbb{R}^n \mid Ax - Dx = b, Dx \geq 0\} \cap V$ with $D = \text{diag}(\delta(x^r))$ we obtain again a convex polyhedron. The choice of D depending on x^r is not restrictive as we can always take a subsequence of the sequence $\{x^r\}_{r>0}$, which converge to \bar{x} , with constant signs near \bar{x} . This set is non-empty because $\bar{x} \in S^*$. In the neighbourhood V , solving $Ax - Dx = b$ with the constraints $Dx \geq 0$ gives a solution of (AVE). Using Hoffman lemma [99] there exists a constant K such that

$$\begin{aligned}
d_{S^*}(x^r) &\leq K \left\| \begin{pmatrix} (A - D)x^r - b \\ (x^r - \alpha - \bar{x})^+ \\ (-x^r - \alpha + \bar{x})^+ \\ (-Dx)^+ \end{pmatrix} \right\| \leq K(\|(A - D)x^r - b\| + \|(x^r - \alpha - \bar{x})^+\| \\
&\quad + \|(-x^r - \alpha + \bar{x})^+\| + \|(-Dx)^+\|), \\
&= K\|(A - D)x^r - b\|, \\
&= K\|Ax^r - |x^r| - b\|.
\end{aligned}$$

As x^r is feasible for (P_r) , we have

$$\begin{aligned}
\|Ax^r - |x^r| - b\| &\leq \|g(r)Ae - g(r)\delta(x^r)\|, \\
&= \|(Ae - \delta(x^r))g(r)\|, \\
&\leq \|Ae - \delta(x^r)\| |g(r)|, \\
&= \|Ae - \delta(x^r)\|g(r) = O(g(r)).
\end{aligned}$$

Combining both previous inequalities, we obtain

$$d_{S^*}(x^r) \leq K\|Ae - \delta(x^r)\|g(r) = O(g(r)).$$

In both cases a) and b), the proof is complete since $S^* \subset S_{(AVE)}^*$. □

Remark 3.4.1. *We can be a bit more specific in the case where $(A - D)$ is invertible. In this case $S^* = \{\bar{x}\}$, so (3.4) becomes*

$$\|x^r - \bar{x}\| \leq \|(A - D)^{-1}\| \|Ae - \delta(x)\| g(r) = O(g(r)).$$

This case corresponds to the special cases where (AVE) has isolated solutions.

3.5 Algorithm

In the previous sections we present theoretical results about convergence and error estimate of an algorithm to compute a solution of (AVE). In this section, we focus on the algorithm and its implementation.

Consider the generic algorithm where C_k is the feasible set of (P_{r^k}) :

$$\begin{cases} \{r_k\}_{k \in \mathbb{N}}, r_0 > 0 \text{ and } \lim_{k \rightarrow +\infty} r_k = 0 \\ \text{find } x^k : x^k \in \arg \min_{x \in C_k} \sum_{i=1}^n \theta_{r_k}(x_i^+) + \theta_{r_k}(x_i^-) - 1. \end{cases} \quad (\text{TAVE})$$

In a practical implementation of (TAVE) one should probably more likely use the initial problem (3.2) with the constraint $x^+ + x^- \geq g(r)e$. The sequence of computed points will probably be infeasible but we believe that it leads to improved numerical behaviour. The constraint $x^+ + x^- \geq g(r)e$ prevents the sequence to possibly go to a local minimum with a zero component.

Algorithm TAVE requires an initial point. In a same way as in [138, 135], one can use the solution of the following linear program

$$\begin{cases} \min_{x^+, x^- \in \mathbb{R}_+^N} (x^+ + x^-)^T e, \\ A(x^+ - x^-) - (x^+ + x^-) = b. \end{cases}$$

Indeed, this program find an initial feasible point of (P_r) and the objective function may encourage this point to satisfy the complementarity condition.

In this study we put the variables in a compact set. Indeed, the functions θ_r are more efficient when their arguments live in $[0, 1]$. Besides, we use one way to express complementarity with Lemma 3.2.1 another way, which will be used in the numerical study, is to consider the following

$$\theta_r(s) + \theta_r(t) - \theta_r(s + t) = 0. \quad (3.5)$$

In this case we don't necessarily need the constraint $x^+ + x^- \geq g(r)$, since it is a reformulation of the complementarity and no longer a relaxation.

Regarding the choice of the parameters α , r_0 and the update parameter of r , it is to be noted that they are all used in the constraint $x^+ + x^- \geq g(r)e$ with $g(r) = r^\alpha$ and $0 < \alpha < 1$. Theorem 3.4.1 shows that the convergence to the zero part of the solution is a $O(g(r))$. So

it is clear that α needs to be taken as big as possible, for instance $\alpha = 0.99$. Also there is a link between the value of α and the update rule for r . We choose to select a sequence of values with an update constant T , so that $r_{k+1} = \frac{r_k}{T}$. The initial parameter r_0 can be chosen according the relation

$$\theta_r^1(s) + \theta_r^1(t) = 1 \iff st = r^2.$$

At each step in r , we solve a concave optimization problem to get the current point. The following heuristic can be rather useful to accelerate convergence and assure a good precision when we are close to the solution. After finding the current point x^k we solve if possible the linear system

$$(A - \text{diag}(\delta(x^k)))z = b. \quad (3.6)$$

If x solves (AVE), then we solved (AVE) with the same precision as we solved the linear system. However, if x does not solve (AVE), we continue the iteration in r with x^k . This idea is similar to compute a Newton iteration.

3.6 Numerical simulations

We present numerical results on two examples and some randomly generated problems. Those random problems can be divided into two different classes: problems with singular values exceeding one and problems that are completely general without any particular property. The latter class is the most interesting for us since numerical methods can fail to obtain a solution. This is the main motivation for our approach.

These simulations have been done using MATLAB, [147], with the linear programming solver GUROBI, [92]. We used the Successive Linearisation Algorithm (SLA) of [141] to solve concave minimization problems encountered at each iteration.

Proposition 3.6.1 (SLA for concave minimization). *Given ϵ sufficiently small and r^k . Denote C the feasible set of (P_{r^k}) . Given $x^k = x^{k+} - x^{k-}$, x^{k+1} is designed as a solution of the linear problem*

$$\min_{y^+, y^- \in C} (y^+)^T \nabla \theta_{r^k}(x^{k+}) + (y^-)^T \nabla \theta_{r^k}(x^{k-}),$$

with $x^0 = x^{0+} - x^{0-}$ a random point. We stop when

$$x^{k+1} \in C \text{ and } (x^{k+1} - x^k)^T \nabla \theta_{r^k}(x^k) \leq \epsilon.$$

This algorithm generates a finite sequence with strictly decreasing objective function values.

Proof. see [141, Theorem 4.2]. □

Along these simulations we use the parameters detailed in Table 3.1 for TAVE. The maximum number of iterations in r for one instance is fixed to 20 and the maximum number of linear programs for one SLA is fixed to 10. We measure the time in seconds, the number of linear programs solved and the number of linear systems solved respectively denoted by nb-LP-method and nb-lin-syst-method.

T	initial $r : r_0$	function θ_r	α
1.8	1	θ_r^2	0.99

Table 3.1: Parameters for the simulations

We consider two concrete examples and then two kinds of randomly generated problems. The first one is a second order ordinary differential equation with initial conditions and the second example is an obstacle problem. We remind that our main motivation is to consider general absolute value equations and this has been treated in subsection 3.6.4.

3.6.1 An ordinary differential equation

We consider the ordinary differential equation

$$\ddot{x}(t) - |x(t)| = 0, \quad x(0) = x_0, \quad \dot{x}(0) = \gamma.$$

We get an (AVE) by using a finite difference scheme in order to discretize this equation. We use the following second-order backward difference to approximate the second derivative

$$\frac{x_{i-2} - 2x_{i-1} + x_i}{h^2} - |x_i| = 0.$$

This equation was derived with an equispaced grid $x_i = ih$, $i = 1, \dots, N$. Neumann boundary conditions were approximated using a centred finite difference scheme

$$\frac{x_{-1} - x_1}{2h} = \gamma.$$

We compare the obtained solution by TAVE to the one of the predefined Runge-Kutta ode45 function in MATLAB, [147]. The domain is $t \in [0, 4]$, initial conditions $x_0 = -1$, $\gamma = 1$ and $N = 100$. Results are presented in Figure 3.1. TAVE solves the problem and gives consistent results.

3.6.2 Obstacle problem

In this simple obstacle problem, we try to compute a trajectory joining the boundary of a domain with an obstacle, g , and a minimal curvature, f . This can be formulated using the following equation and inequalities: find u such that

$$(\ddot{u}(x) - f(x))^T(u(x) - g(x)) = 0, \quad \ddot{u}(x) - f(x) \geq 0, \quad u(x) - g(x) \geq 0.$$

We approximate the second order derivative with a second-order centred finite difference to get a discrete version on an equispaced grid $x_i = ih$, $i = 1, \dots, N$.

$$(Du - f)^T(u - g) = 0, \quad Du - f \geq 0, \quad u - g \geq 0, \quad \text{where } D = \begin{pmatrix} \frac{2}{h^2} & \frac{-1}{h^2} & & \\ \frac{-1}{h^2} & \ddots & \ddots & \\ & \ddots & \ddots & \frac{-1}{h^2} \\ & & \frac{-1}{h^2} & \frac{2}{h^2} \end{pmatrix},$$

$g_i = g(x_i)$ and $f_i = f(x_i)$. This can be written as a linear complementarity problem by setting $z = u - g$, $M = D$ and $q = Dg - f$, that is

$$(Mz + q)^T z = 0, \quad Mz + q \geq 0, \quad z \geq 0. \quad (3.7)$$

This equation is equivalent to an (AVE) whenever 1 is not an eigenvalue of M by proposition 2 of [136]:

$$(3.7) \iff (M - I)^{-1}(M + I)x - |x| = (M - I)^{-1}q.$$

We present results for our method and for the LPM method from [138], with $g(x) = \max(0.8 - 20(x - 0.2)^2, \max(1 - 20(x - 0.75)^2, 1.2 - 30(x - 0.41)^2))$, $f(x) = 1$, $N = 50$ in Figure 3.2. Both methods give 20 points on the curve g and none below g over 50 points. Once again TAVE method gives consistent results.

3.6.3 Random easy problems with unique solution

We consider the special case where (AVE) has a unique solution. Following [205], we generate data (A, b) by the following Matlab code for $n = 100, 200, 400, 800$

```
n=input('dimension of matrix A =');
rand('state',0);
R=rand(n,n);
b=rand(n,1);
A=R'*R+n*eye(n);
```

Then, the corresponding (AVE) has a unique solution and A is a symmetric matrix. The required precision for solving (AVE) is 10^{-6} and thanks to the heuristic from Section 3.5 in equation (3.6) we get in the worst case 10^{-10} . For each n we consider 100 instances. TAVE is compared to a Newton method from [134], which is denoted GN. Results are summarized in Table 3.2, which gives for TAVE the number of linear programs solved, the time required to solve all the instances and gives for GN the number of linear systems and the time required. Note that other Newton methods like [37, 97, 154] should give similar conclusions so we do not include them in our comparison. It is to be expected that the method GN is faster than the method TAVE in particular when the dimension grows since it solves only linear systems whereas TAVE solves linear programs.

Our method solves all the problems, which once again valid our approach. We observe that the number of solved linear programs is very low. Indeed in each case, the initialization

n	nb-LP-TAVE	time TAVE	nb-lin-syst-GN	time GN
32	100	0.1841	217	0.0403
64	100	0.4702	224	0.0553
128	100	1.5880	219	0.1079
256	100	14.5161	226	0.3924
512	100	129.1686	214	2.2327

Table 3.2: TAVE and Newton method from [134], GN, for (AVE) in the case with singular values of A exceeds 1.

step has been sufficient to solve the problem. The method GN outperform in time TAVE as it was expected. Since some other methods to solve (AVE) have no problem solving these instances they are not our main focus here.

3.6.4 General problems

Now, we consider general (AVE) without any assumption on A . This kind of problems correspond to the main goal of this paper. We generate for several n and several values of the parameters one hundred instances of the problem following [136]:

"Choose a random A from a uniform distribution on $[-10, 10]$, then choose a random x from a uniform distribution on $[-1, 1]$ and set $b = Ax - |x|$."

We compare 4 methods tailored for general (AVE):

- TAVE method from Algorithm TAVE;
- TAVE2 which is the same algorithm with the different objective based on (3.5)

$$\sum_{i=1}^N \theta_r(x_i^+) + \theta_r(x_i^-) - \theta_r(x_i^+ + x_i^-);$$

- concave minimization method CMM from [136];
- successive linear programming method LPM from [138].

In this situation GN method has not been applied since it has no guarantee of convergence.

In Tables 3.3–3.6, "nnztot" gives the number of violated expressions for all problems, "nnzx" gives the maximum violated expressions for one problem, "out-iter" gives the number of iteration in r and "in-iter" gives the number of linear programs solved for all the problems. We also provide the time in seconds and the number of failures.

In each case our methods manage to reduce the number of unsolved problems. This confirm the interest of the relaxation method presented here. Also one should note that an improved number of solved problems comes with a price, since it requires more time.

Table 3.4 shows promising results for TAVE2. It is a slightly different method, since it is not a relaxation but a reformulation of the complementarity. In every case it gives the smallest number of unsolved problem in a very reasonable time.

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	0	0	74	306	0.6234	0
64	3	1	156	491	2,8173	3
128	8	1	269	841	20,9447	8
256	8	1	324	1129	281,6190	8

Table 3.3: TAVE

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	0	0	33	164	0.3137	0
64	2	1	81	221	1,1280	2
128	4	1	136	303	6,3953	4
256	4	1	131	292	56,4148	4

Table 3.4: TAVE2

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	9	1	-	485	1.0823	9
64	8	1	-	458	2,9234	8
128	10	1	-	568	18,4404	10
256	11	1	-	595	124,5728	11

Table 3.5: CMM

Conclusion and perspectives

In this paper, we propose a class of heuristic schemes to solve the NP-hard problem (AVE). A complete analysis is provided including convergence and error estimate.

Furthermore, a numerical study shows that our approach is full of interest. Indeed, our methods turn out to be consistent with real examples and problems with unique solution. We do not compare our method with those specially designed for these problems since they do not belong to the same class. Finally, the last set of generated problems concerns general

n	nnztot	nnzx	out-iter	in-iter	time	nb-failure
32	7	1	-	248	0.5546	7
64	19	4	-	342	2,5822	13
128	19	3	-	409	16,6830	13
256	29	5	-	439	143,0973	11

Table 3.6: LPM

(AVE). In comparison with some existing methods, our approach improves the number of failures. This was our main goal.

It is of interest to note that the methods presented here can also solve the linear complementarity problem using the same technique as for the obstacle problem example.

Further studies could improve the choice of parameters in order to reduce the computational time to solve the problems, especially for large instances. Promising results were shown by the modified algorithm TAVE2, which considers a different way to formulate the penalty. So we may wonder if it is possible to improve our algorithms in this case and if there exists other similar reformulations of the complementarity which can give even better results.

We are working on a hybrid algorithm which can benefit from both the minimization methods and Newton methods as in [135]. Indeed, this philosophy is fully applicable here and can lead to computational improvements.

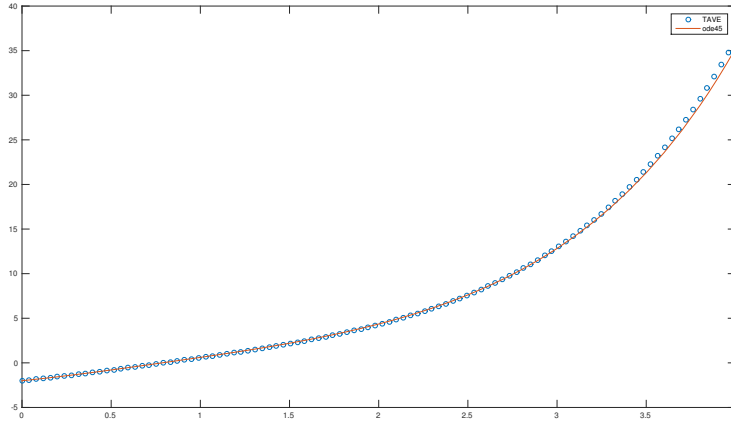


Figure 3.1: Numerical solution of (3.6.1) with ode45 and ThetaAVE.

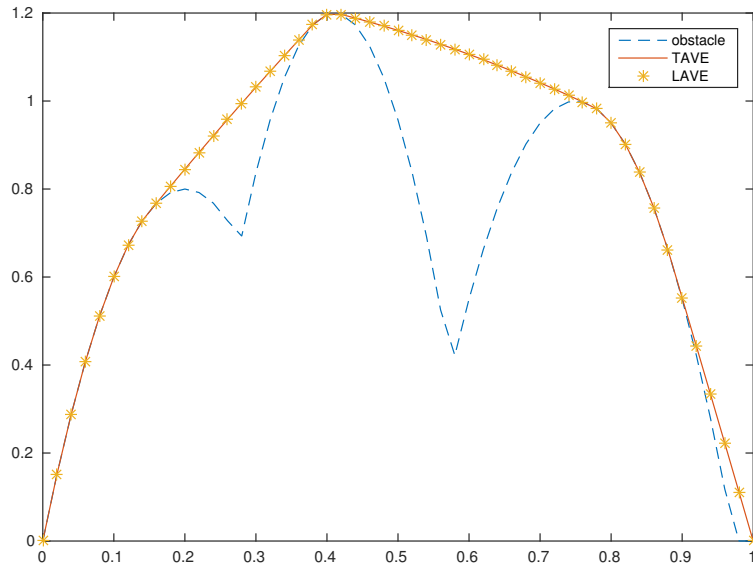


Figure 3.2: Numerical solution of the obstacle problem (3.6.2) with ThetaAVE and the method from [136]

Part II

Mathematical Programs with Complementarity Constraints

We consider the mathematical program with complementarity constraint, denoted MPCC, defined as

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

with $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ and $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^q$ that are assumed continuously differentiable. The notation $0 \leq u \perp v \geq 0$ for two vectors u and v in \mathbb{R}^q is a shortcut for $u_i \geq 0$, $v_i \geq 0$ and $u_i v_i = 0$ for all $i \in \{1, \dots, q\}$.

This problem has become an active subject in the literature in the last two decades and has been the subject of several monographs [131, 158] and PhD thesis [70, 109, 100, 190, 184, 48]. The wide variety of applications that can be cast as an MPCC is one of the reasons for this popularity. Among other we can cite truss topology optimisation [100], discrete optimisation [3], image restoration [26], optimal control [10, 98]. Otherwise, another source of problems are bilevel programming problems [53, 55], where the lower-level problem is replaced by its optimality conditions. This may lead to a more general kind of problem called Mathematical Program with Equilibrium Constraints [158] or Optimisation Problem with Variational Inequality Constraints [200]. (MPCC) is a special case of the following optimisation problem with a geometric constraint:

$$\min f(x) \text{ s.t. } F(x) \in \Lambda, \tag{3.8}$$

where $F(x) := (g(x), h(x), \Psi(x))$, $\Psi(x) := (-G_1(x), -H_1(x), \dots, -G_q(x), -H_q(x))$, $\Lambda := \{0\}^m \times]-\infty, 0]^p \times C^q$ and $C := \{(a, b) \in \mathbb{R}^2 \mid 0 \leq a \perp b \geq 0\}$. The MPCC formulation has been the most popular in the literature motivated by more accessible numerical approaches. (MPCC) is clearly a non-linear program and in general most of the functions involved in the formulation are non-convex.

In this context solving the problem means finding a local minimum. Even so this goal apparently modest is hard to achieve in general due to the degenerate nature of the MPCC. Therefore, numerical methods that consider only first order information may be expected to compute a stationary point.

The wide variety of approaches with this aim computes the KKT conditions, which require that some constraint qualifications hold at the solution to be an optimality condition. However, these constraint qualifications never hold in general for (MPCC). For instance, the classical MFCQ that is very often used to guarantee convergence of algorithms is violated at any feasible point. This is partly due to the geometry of the complementarity constraint that always has an empty relative interior.

These issues have motivated the definition of enhanced constraint qualifications and optimality conditions for (MPCC) as in [200, 198, 181, 73] to cite some of the earliest research. In 2005, Flegel & Kanzow provide an essential result that defines the right necessary optimality condition to (MPCC). This optimality condition is called M(Mordukhovich)-stationary condition. This is the subject of the following chapter, Chapter 4.

A wide range of numerical methods have been proposed to solve this problem such as relaxation methods, interior-point methods [127, 164, 124], penalty methods [131, 104, 153], SQP methods [77], elastic mode [16, 48], dc methods [145], filter methods [125] and Levenberg-Marquardt methods [89] to cite some of them. This first family of method called relaxation or regularisation method is developed in detail in this thesis. Despite the difficulties explained above, the methods reformulating (MPCC) as a non-linear program have shown to be successful in practice in [75, 76]. However, the theoretical guarantee for this approach is often far from the desired goal to compute M-stationary points.

In view of the constraint qualifications issues that plague the (MPCC), the relaxation methods provide an intuitive answer. The complementarity constraints are 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 complementarity constraints. Finally, as the relaxing parameter is reduced, convergence to the feasible set of (MPCC) is obtained similar to a homotopy technique. The interest for such methods has already been the subject of some PhD thesis in [184, 190] and is an active subject in the literature. These methods are the central theme along the rest of this part from Chapter 5 to Chapter 7. Chapter 5 presents the state of the art of the relaxation methods proposed in the literature. Then, Chapter 6 introduces a new method called Butterfly relaxation. The analysis of some of these methods in a numerical context is made in Chapter 7.

Chapter 4

Stationary and constraint qualifications for MPCC

We have introduced in Chapter 1, in the non-linear programming case, the most classical necessary optimality conditions so-called KKT conditions. These optimality conditions also called stationary conditions require some constraint qualification to hold at a local minimum. The enhanced conditions for the (MPCC) presented in this chapter follow a similar construction.

This chapter is divided in five sections. Section 4.1 introduces in a classical way the alphabet of stationary conditions for the (MPCC). We point out the most important conditions and define what is an MPCC-constraint qualification, denoted MPCC-CQ for short. Section 4.2 presents some of the various MPCC-CQs. In particular, we show an interesting generic way to define these conditions that encompass some of the existing conditions and allows new ones to be introduced. The two following sections focus on two properties of the MPCC-CQs. In Section 4.3, we discuss the weakest MPCC-CQ. In Section 4.4, we give some algorithmic applications that will be useful for the study of relaxation methods. Particular emphasis is given in the relations between these conditions. This is the subject of the final section of this chapter, Section 4.5.

4.1 Stationary conditions for MPCC

The following example due to Kanzow and Schwartz, [112], exhibits a situation where all of the functions are affines and the global minimiser is not a KKT point. We will return to this example later on.

Example 4.1.1.

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & x_1 + x_2 - x_3 \\ \text{s.t.} \quad & g_1(x) := -4x_1 + x_3 \leq 0, \\ & g_2(x) := -4x_2 + x_3 \leq 0, \\ & 0 \leq G(x) := x_1 \perp H(x) := x_2 \geq 0. \end{aligned}$$

The global solution is $(0, 0, 0)^t$ but is not a KKT point. Indeed, the gradient of the Lagrangian equal to zero yields

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

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

This example is not a contradiction with the Theorem 1.3.1 (p.14). Indeed, there is no constraint qualification holding here. We make some remarks about the use of constraint qualification in this context. The following lemma shows that MFCQ fails to hold at any feasible point of (MPCC).

Lemma 4.1.1. *Let x^* be in \mathcal{Z} . Then, MFCQ fails to hold at x^* .*

The proof is straightforward using the Definition 1.3.6 of MFCQ. Now, consider the use of ACQ on the following example, which is the simplest (MPCC).

Example 4.1.2.

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

The tangent cone, \mathcal{T} , at the origin is given by

$$\mathcal{T}_{\mathcal{Z}}(x^*) = \{d \in \mathbb{R}^2 \mid d_1 = 0, d_2 \geq 0 \text{ or } d_1 \geq 0, d_2 = 0\}.$$

The linearised cone, $\mathcal{L}_{\mathcal{Z}}(x^*)$, is given by

$$\mathcal{L}_{\mathcal{Z}}(x^*) = \mathbb{R}_+^2.$$

Thus, we see that ACQ fails to hold at $x^* = 0$.

The geometrical behaviour of the cones \mathcal{T} and \mathcal{L} in the previous example is typical for (MPCC). Indeed, due to the complementarity constraint, the tangent cone is in general non-convex, while the linearised cone is polyhedral and so convex. Thus, ACQ is very unlikely to hold.

It is clear that we cannot expect to compute usual KKT point since classical constraint qualifications in general does not hold, so we introduce weaker stationary concepts as in [181, 198]. Beforehand, let us introduce some notations. Let \mathcal{Z} be the set of feasible points of (MPCC). Given $x^* \in \mathcal{Z}$, we denote

$$\begin{aligned} \mathcal{I}^{+0}(x^*) &= \{i \in \{1, \dots, q\} \mid G_i(x^*) > 0 \text{ and } H_i(x^*) = 0\}, \\ \mathcal{I}^{0+}(x^*) &= \{i \in \{1, \dots, q\} \mid G_i(x^*) = 0 \text{ and } H_i(x^*) > 0\}, \\ \mathcal{I}^{00}(x^*) &= \{i \in \{1, \dots, q\} \mid G_i(x^*) = 0 \text{ and } H_i(x^*) = 0\}, \\ \mathcal{I}_g(x^*) &= \{i \in \{1, \dots, p\} \mid g_i(x^*) = 0\}. \end{aligned}$$

In order to derive weaker optimality conditions, we consider an enhanced Lagrangian function. Let \mathcal{L}_{MPCC}^r be the generalised MPCC-Lagrangian function of (MPCC) such that

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

with $\lambda := (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}^{p+m+2q}$. The alphabet of stationary conditions can now be defined.

Definition 4.1.1. $x^* \in \mathcal{Z}$ is said

- *Weak-stationary* if there exists $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in \mathbb{R}_+^p \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^q$ such that

$$\begin{aligned} \nabla_x \mathcal{L}_{MPCC}^1(x^*, \lambda) &= 0, \\ \lambda_i^g &= 0 \ \forall i \notin \mathcal{I}_g(x^*), \ \lambda_{\mathcal{I}^{+0}(x^*)}^G = 0, \ \lambda_{\mathcal{I}^{0+}(x^*)}^H = 0. \end{aligned}$$

- *Clarke (C)-stationary point* if x^* is weak-stationary and

$$\forall i \in \mathcal{I}^{00}(x^*), \ \lambda_i^G \lambda_i^H \geq 0.$$

- *Alternatively or Abadie (A)-stationary point* if x^* is weak-stationary and

$$\forall i \in \mathcal{I}^{00}(x^*), \ \lambda_i^G \geq 0 \text{ or } \lambda_i^H \geq 0.$$

- *Mordukhovich (M)-stationary point* if x^* is weak-stationary and

$$\forall i \in \mathcal{I}^{00}(x^*), \text{ either } \lambda_i^G > 0, \ \lambda_i^H > 0 \text{ or } \lambda_i^G \lambda_i^H = 0.$$

- *Strong (S)-stationary point* if x^* is weak-stationary and

$$\forall i \in \mathcal{I}^{00}(x^*), \ \lambda_i^G \geq 0, \ \lambda_i^H \geq 0.$$

Relations between these definitions are given in Figure 4.1 and follow in a straightforward way from the definitions. It is also to be noted that if we assume strict complementarity, i.e. for all $i \in \{1, \dots, q\}$

$$G_i(x^*) + H_i(x^*) > 0,$$

then all of the stationary conditions presented here are equivalent. For instance, this is the case for the class of binary optimisation problems, whose integer constraints are replaced by complementarity constraints. Local optimal solutions are often denoted Bouligand (B)-stationary point in the literature, but this will not be used here.

An essential question here is to know what are the strongest stationary conditions that can be expected to hold at a local minimum for a (MPCC) with non-degenerate description of the constraints (i.e. satisfies some kind of constraint qualification). A first observation based on Definition 4.1.1 is that S-stationary condition correspond to usual KKT conditions applied to (MPCC). However, as pointed out in Example 4.1.1 this condition is too stringent

Figure 4.1: Relations between the stationary definitions

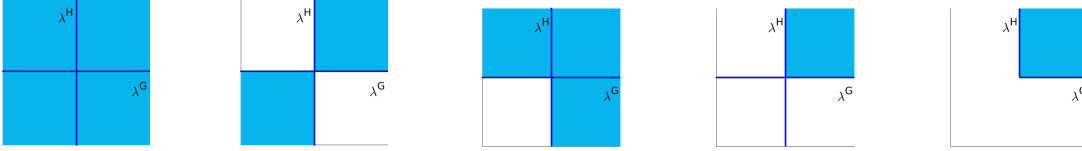
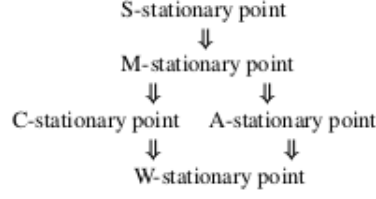


Figure 4.2: Signs of λ^G, λ^H for indices $i \in \mathcal{I}^{00}(x^*)$. From the left to the right : weak-stationary, C-stationary, A-stationary, M-stationary and S-stationary.

for (MPCC) in general. The answer to this question is given in the following key result from Flegel & Kanzow, which considers an MPCC-version of Guignard CQ. Based on Example 4.1.2 and our observation on \mathcal{L}_Z , an adaptation of this cone for MPCC, \mathcal{L}_{MPCC} , in [181, 72, 159] is defined such that:

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

This cone is not a polyhedral cone any more and consequently not necessarily convex. However due to [72], one always has the following inclusions

$$\mathcal{T}_Z(x^*) \subseteq \mathcal{L}_{MPCC}(x^*) \subseteq \mathcal{L}_Z(x^*).$$

Let us continue the study of Example 4.1.2.

Example 4.1.3. *Example 4.1.2 continued. The cone $\mathcal{L}_{MPCC}(x^*)$ at $x^* = 0$ is given by*

$$\mathcal{L}_{MPCC}(x^*) = \{d \in \mathbb{R}^2 \mid d_1 = 0, d_2 \geq 0 \text{ or } d_1 \geq 0, d_2 = 0\}.$$

Thus, for this example it holds that $\mathcal{L}_{MPCC}(x^) = \mathcal{T}_Z(x^*)$.*

The condition illustrated in the previous example is an Abadie-kind CQ. We now define MPCC-Abadie CQ and MPCC-Guignard CQ, denoted respectively MPCC-ACQ and MPCC-GCQ.

Definition 4.1.2. *Let $x^* \in \mathcal{Z}$. We say that MPCC-ACQ holds at x^* if $\mathcal{T}_{\mathcal{Z}}(x^*) = \mathcal{L}_{MPCC}(x^*)$ and MPCC-GCQ holds at x^* if $\mathcal{T}_{\mathcal{Z}}^\circ(x^*) = \mathcal{L}_{MPCC}^\circ(x^*)$.*

The following theorem is a keystone to define necessary optimality conditions for (MPCC).

Theorem 4.1.1 ([71]). *A local minimum of (MPCC) that satisfies MPCC-GCQ or any stronger MPCC CQ is an M-stationary point.*

This result is coherent with Example 4.1.1, where the point $(0, 0, 0)^T$ is an M-stationary point. Therefore, devising algorithms to reach KKT stationary points (S-stationary) is not possible in general, and we must satisfy ourselves in devising algorithms reaching M-stationary points. We already defined some MPCC CQ. We now formalise the discussion by stating an explicit definition.

Definition 4.1.3. *Let $x^* \in \mathcal{Z}$. An MPCC CQ is a condition that holds independently of the choice of the objective function and that ensures the relation*

$$x^* \text{ local minimum} + \text{MPCC CQ holds at } x^* \implies x^* \text{ M-stationary.}$$

In the context of (MPCC), MPCC-GCQ is not the weakest MPCC CQ as it will be shown in the following section. Furthermore, due to the difficulty of computing the tangent cone, a wide variety of stronger conditions, that are easier to verify, have been designed. Some of these conditions are also useful for algorithmic applications.

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

Proposition 4.1.1. *Let $x^* \in \mathcal{Z}$ be an M-stationary point of (MPCC) if and only if there exists a partition of $\mathcal{I}^{00}(x^*) = \mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00}$, with $(\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}, \mathcal{I}_{-0}^{00})$ disjointed two by two, such that $x^* \in \mathcal{Z}$ is a KKT point of the following non-linear program,*

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f(x) \\ & \text{s.t. } g(x) \leq 0, \quad h(x) = 0, \\ & \quad G_i(x) = 0, \quad \forall i \in \mathcal{I}^{0+}(x^*), \quad H_i(x) = 0, \quad \forall i \in \mathcal{I}^{+0}(x^*), \quad (MNL P_{\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}}(x^*)) \\ & \quad G_i(x) \leq 0, \quad \forall i \in \mathcal{I}_{-0}^{00}, \quad H_i(x) \leq 0, \quad \forall i \in \mathcal{I}_{0-}^{00}, \\ & \quad G_i(x) \geq 0, \quad H_i(x) \geq 0, \quad \forall i \in \mathcal{I}_{++}^{00}. \end{aligned}$$

Proof. Assuming that x^* is a KKT point of $(MNL P_{\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}}(x^*))$ implies by definition that there exists $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ such that

$$\begin{aligned} 0 &= \nabla f(x^*) + \sum_{i=1}^p \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x^*) - \sum_{i=1}^q \lambda_i^G \nabla G_i(x^*) - \sum_{i=1}^q \lambda_i^H \nabla H_i(x^*), \\ \min(-g(x^*), \lambda^g) &= 0, \quad h(x^*) = 0, \quad G_i(x^*) = 0 \quad \forall i \in \mathcal{I}^{0+}(x^*), \quad H_i(x^*) = 0 \quad \forall i \in \mathcal{I}^{+0}(x^*), \\ \min(G_i(x^*), \lambda_i^G) &= 0 \quad \forall i \in \mathcal{I}_{++}^{00}, \quad \min(-G_i(x^*), \lambda_i^G) = 0 \quad \forall i \in \mathcal{I}_{-0}^{00}, \\ \min(H_i(x^*), \lambda_i^H) &= 0 \quad \forall i \in \mathcal{I}_{++}^{00}, \quad \min(-H_i(x^*), \lambda_i^H) = 0 \quad \forall i \in \mathcal{I}_{0-}^{00}. \end{aligned}$$

Additionally, x^* feasible for (MPCC) yields

$$\begin{aligned}
0 &= \nabla f(x^*) + \sum_{i=1}^p \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x^*) - \sum_{i=1}^q \lambda_i^G \nabla G_i(x^*) - \sum_{i=1}^q \lambda_i^H \nabla H_i(x^*), \\
\min(-g(x^*), \lambda^g) &= 0, \quad h(x^*) = 0, \\
G_i(x^*) &= 0 \quad \forall i \in \mathcal{I}^{0+}(x^*) \cup \mathcal{I}^{00}(x^*), \quad H_i(x^*) = 0 \quad \forall i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*), \\
\lambda_i^G &\geq 0 \quad \forall i \in \mathcal{I}_{++}^{00}, \quad \lambda_i^G \leq 0 \quad \forall i \in \mathcal{I}_{-0}^{00}, \\
\lambda_i^H &\geq 0 \quad \forall i \in \mathcal{I}_{++}^{00}, \quad \lambda_i^H \leq 0 \quad \forall i \in \mathcal{I}_{0-}^{00}.
\end{aligned} \tag{4.1}$$

Let us now verify that this condition is equivalent to the M-stationary conditions. It is clear that assuming x^* M-stationary point, then there exists λ such that (x^*, λ) verifies (4.1). Assume that (x^*, λ) verifies (4.1). Therefore, x^* is a weak-stationary point. To conclude to M-stationary it is sufficient to see that for all $i \in \mathcal{I}^{00}(x^*)$ then either $\lambda_i^G > 0, \lambda_i^H > 0$ or $\lambda_i^G \lambda_i^H = 0$. In particular, it holds

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

This concludes the proof. □

This new interpretation of M-stationary will be especially useful for deriving constraint qualifications for (MPCC). Figure 4.3 highlight the feasible set of $(MNL P_{\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}}(x^*))$.

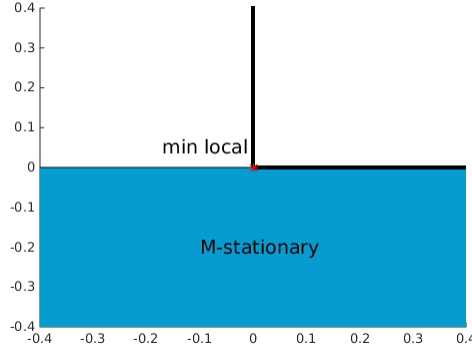


Figure 4.3: Optimality zone of an M-stationary point.

4.2 Constraint qualifications for MPCC

We present in this section an extension of the CQ that has been presented in Chapter 1 to the MPCC. This does not cover all of the existing MPCC CQ, but we focus on those that

have algorithmic applications. The study of these conditions have been an active subject in the literature [183, 73, 74, 71, 199, 184, 90, 91, 201, 112, 103, 202, 165, 201, 202]. Apart from MPCC-GCQ and MPCC-ACQ that have been presented earlier, one of the most principle constraint qualifications used in the literature is the MPCC-linear independance CQ, denoted MPCC-LICQ, which is discussed in [183].

Definition 4.2.1. *MPCC-LICQ holds at $x^* \in \mathcal{Z}$ if the gradients*

$\{\nabla g_i(x^) \ (i \in \mathcal{I}_g(x^*)), \ \nabla h_i(x^*) \ (\forall i), \ \nabla G_i(x^*) \ (i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{0+}(x^*)), \ \nabla H_i(x^*) \ (i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{+0}(x^*))\}$*
are linearly independent.

The following discussion illustrates the fact that S-stationary points (i.e. classical KKT points) can be expected under MPCC-LICQ.

Example 4.2.1 (Continuation of Example 4.1.2). *The most basic MPCC*

$$\begin{aligned} & \min_{x_1, x_2} f(x) \\ & s.t. \ G(x) = x_1 \geq 0, \\ & \quad H(x) = x_2 \geq 0, \\ & \quad G(x)H(x) = x_1x_2 = 0. \end{aligned} \tag{4.2}$$

We already point out that at $x^ = (0, 0)^T$ it holds $\mathcal{T}(x^*) = \{d \in \mathbb{R}_+^2 \mid d_1d_2 = 0\}$ and $\mathcal{L}(x^*) = \mathbb{R}_+^2$. However, the polar cones coincide so GCQ holds at x^* . Besides, the gradients of G and H at x^* are obviously linearly independent. Thus, MPCC-LICQ also holds at x^* .*

This observation is not a surprise due to the following result from [74].

Proposition 4.2.1 ([74]). *Standard GCQ is always satisfied under MPCC-LICQ and not true for less.*

Example 4.1.1 is an example, where GCQ fails to hold under less than MPCC-LICQ. The following result about uniqueness of the multipliers is similar to the classical case.

Theorem 4.2.1 ([181]). *Let $x^* \in \mathcal{Z}$ be a local minimiser of (MPCC). If MPCC-LICQ holds in x^* , then there exists a unique Lagrange multiplier λ^* such that (x^*, λ^*) is S-stationary.*

Apart from LICQ another constraint qualification that ignores the sign of the multipliers is the MPCC-constant rank CQ. In a similar way we extend CRCQ as in [91].

Definition 4.2.2. *Let $x^* \in \mathcal{Z}$. MPCC-CRCQ holds at x^* if there exists $\delta > 0$ such that for any subsets $\mathcal{I}_1 \subseteq \mathcal{I}_g(x^*)$, $\mathcal{I}_2 \subseteq \{1, \dots, m\}$, $\mathcal{I}_3 \subseteq \mathcal{I}^{0+}(x^*) \cup \mathcal{I}^{00}(x^*)$, and $\mathcal{I}_4 \subseteq \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*)$, the family of gradients*

$$\{\nabla g_i(x^*) \ (i \in \mathcal{I}_1), \ \nabla h_i(x^*) \ (i \in \mathcal{I}_2), \ \nabla G_i(x^*) \ (i \in \mathcal{I}_3), \ \nabla H_i(x^*) \ (i \in \mathcal{I}_4)\}$$

has the same rank for each $x \in \mathcal{B}_\delta(x^)$.*

In [184, 112, 103], the authors defined MPCC CQ by applying the classical definition of CQ to the following non-linear program called Tighten NLP:

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f(x) \\ & \text{s.t. } g(x) \leq 0, \quad h(x) = 0, \\ & \quad G_i(x) = 0, \quad \forall i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{0+}(x^*), \\ & \quad H_i(x) = 0, \quad \forall i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{+0}(x^*). \end{aligned} \quad (TNLP(x^*))$$

In particular, they used MFCQ and CPLD applied to TNLP to prove convergence of relaxation methods for (MPCC).

Definition 4.2.3. *Let $x^* \in \mathcal{Z}$. We say that MPCC-MFCQ (resp. MPCC-TCPLD) holds at x^* if MFCQ (resp. CPLD) holds at x^* for $(TNLP(x^*))$.*

The notation MPCC-TCPLD means "Tighten CPLD" and is used to differentiate the definition of MPCC-CPLD given in [103, 112] (MPCC-TCPLD) and in [91, 203] (MPCC-CPLD defined in the definition below). The flaw of this approach is to ignore the signs of the multipliers that should have the signs of the M-stationary conditions for indices in $\mathcal{I}^{00}(x^*)$. We now use the philosophy of the Proposition 4.1.1 to define others MPCC CQ.

Definition 4.2.4. *Let $x^* \in \mathcal{Z}$. We say that MPCC-NNAMCQ (reps. MPCC-CPLD, MPCC-RCPLD, MPCC-CRSC, MPCC-CCP, MPCC-mACQ, MPCC-mGCQ) holds at x^* if for any partition of $\mathcal{I}^{00}(x^*) = \mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00}$, MFCQ (resp. CPLD, RCPLD, CRSC, CCP, ACQ, GCQ) holds at x^* for $(MNLP_{\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}}(x^*))$.*

Additionally, it can be noticed that defining MPCC-LICQ and MPCC-CRCQ in a same way would lead to the same definitions than the one given in Definitions 4.2.1 and 4.2.2. We use the notation MPCC-mACQ and MPCC-mGCQ to underline the difference with the definitions of MPCC-ACQ and MPCC-GCQ. A straightforward result from this definition is the relations between these conditions that come from their relations in the non-linear programming case, noticing that MPCC-LICQ and MPCC-CRCQ defined above are equivalent to LICQ and CRCQ applied to $(MNLP_{\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}}(x^*))$ for any partition of $\mathcal{I}^{00}(x^*) = \mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00}$.

Proposition 4.2.2. *Let $x^* \in \mathcal{Z}$. The diagram given in Figure 4.4 is true.*



Figure 4.4: Relation between the MPCC-constraint qualifications

"MPCC-CQ1 \implies MPCC-CQ2" means that MPCC-CQ1 holds at x^* implies that MPCC-CQ2 holds at x^* .

In the following section, we discuss the weakest MPCC CQ. A complete diagram continuing the one introduced in Proposition 4.2.2 is given in Section 4.5.

4.3 On the weakest MPCC CQ

In this section, we discuss the weakest constraint qualification for (MPCC). This condition is presented in the continuity of our approach from the previous section. Although, these results have been first introduced in [90], we decide to present our approach that is slightly different than the original one. In [90], the authors use the formulation (3.8).

4.3.1 MPCC-weak Guignard CQ

In order to introduce our new MPCC CQ, we give a conic interpretation of the M-stationary conditions. Let us define the cone \mathcal{P}_M at $x^* \in \mathcal{Z}$ as follows

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

This cone is closed but in general not convex. A representation of this cone is given in the following simple example.

Example 4.3.1.

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

The cone $\mathcal{P}_M(x^*)$ at $x^* = (0 \ 0)^T$ is given by

$$\mathcal{P}_M(x^*) = \{d = -(\lambda^G \ \lambda^H)^T \text{ with } \lambda^G \lambda^H = 0 \text{ or } \lambda^G > 0, \lambda^H > 0\},$$

which is illustrated in the following Figure 4.5. In this example, the polar of \mathcal{L}_{MPCC} is the negative orthant and coincide with the polar of the tangent cone.

The motivation to introduce this cone is given by the following lemma.

Lemma 4.3.1. *Let $x^* \in \mathcal{Z}$. Then, the following holds*

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

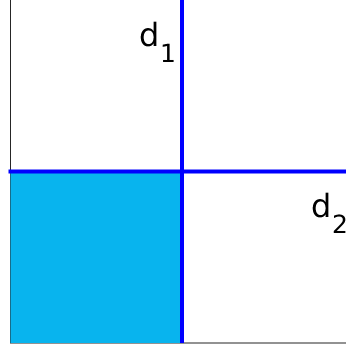


Figure 4.5: Figure of the cone \mathcal{P}_M in Example 4.1.2

Let us now introduce our new condition denoted MPCC-weak Guignard constraint qualification, or MPCC-wGCQ for short.

Definition 4.3.1. *Let $x^* \in \mathcal{Z}$. MPCC-wGCQ holds at $x^* \in \mathcal{Z}$ if*

$$\mathcal{T}_{\mathcal{Z}}(x^*)^\circ \subset \mathcal{P}_M,$$

The reversed inclusion is very unlikely to hold, since \mathcal{P}_M is not a convex cone. However, providing a sharper condition is not trivial.

The following lemma shows the relation between the widely used \mathcal{L}_{MPCC} and \mathcal{P}_M .

Lemma 4.3.2. *Let $x^* \in \mathcal{Z}$. The following inclusion holds*

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

Furthermore, the polar of \mathcal{L}_{MPCC} at x^* is given by

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

Proof. First, it is to be observed that the cone \mathcal{L}_{MPCC} can be rewritten as a union of polyhedral cones such as

$$\mathcal{L}_{MPCC}(x^*) = \cup_{I \in \mathcal{I}^{00}(x^*)} \mathcal{L}_{MPCC}(x^*, I),$$

where $\mathcal{L}_{MPCC}(x^*, I)$ is a notation for

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

Computing the polar of \mathcal{L}_{MPCC} at x^* using Proposition 1.1.6 gives

$$\mathcal{L}_{MPCC}(x^*)^\circ = \cap_{I \in \mathcal{I}^{00}(x^*)} \mathcal{L}_{MPCC}(x^*, I)^\circ.$$

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

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

Now, computing the intersection in the equation above leads to

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

This concludes the proof, since we prove the explicit formula for \mathcal{L}_{MPCC}° and the inclusion is direct by definition of \mathcal{P}_M . \square

In general the reversed inclusion never holds since \mathcal{L}_{MPCC}° is convex and not \mathcal{P}_M . A direct consequence of this lemma is that MPCC-GCQ is stronger than MPCC-wGCQ.

Proposition 4.3.1. *Let $x^* \in \mathcal{Z}$. If MPCC-GCQ holds at x^* , then MPCC-wGCQ holds at x^* .*

Proof. Since MPCC-GCQ holds at x^* it follows that $\mathcal{T}_{\mathcal{Z}}(x^*)^\circ = \mathcal{L}_{MPCC}(x^*)^\circ$. Furthermore, by Lemma 4.3.2, $\mathcal{L}_{MPCC}(x^*)^\circ \subset \mathcal{P}_M$ and so $\mathcal{T}_{\mathcal{Z}}(x^*)^\circ \subset \mathcal{P}_M$. This concludes the proof. \square

This relation is sharp, since Example 4.3.2 shows an example where MPCC-wGCQ holds and not MPCC-GCQ.

Example 4.3.2.

$$\min_{x \in \mathbb{R}} -x \text{ s.t. } 0 \leq x \perp -x^2 \geq 0.$$

It is clear that $\mathcal{Z} = \{0\}$, thus any function f admits a local minimum at $x^ = 0$. Direct computation shows that there exists multipliers (λ^G, λ^H) such that M-stationary condition hold at x^* . Indeed, the gradient of MPCC-Lagrangian function equal to zero yields to*

$$-f'(0) = -\lambda^G.$$

We now show that MPCC-GCQ does not hold at x^ . The feasible set is reduced to the singleton x^* , so the tangent cone is given by $\mathcal{T}_{\mathcal{Z}}(x^*) = \{x^*\}$ and therefore the polar of the tangent cone is $\mathcal{T}_{\mathcal{Z}}^\circ(x^*) = \mathbb{R}$. Let us compute the cone \mathcal{L}_{MPCC} ,*

$$\begin{aligned} \mathcal{L}_{MPCC}(x^*) &= \{d \in \mathbb{R} \mid \nabla G(x^*)^T d \geq 0, \nabla H(x^*)^T d \geq 0, \\ &\quad (\nabla G(x^*)^T d)(\nabla H(x^*)^T d) = 0\}, \\ &= \mathbb{R}_+. \end{aligned}$$

Furthermore, the definition of the polar cone gives $\mathcal{L}_{MPCC}(x^)^\circ = \mathbb{R}_-$. Therefore, MPCC-GCQ does not hold at x^* . Let us compute the cone \mathcal{P}_M .*

$$\begin{aligned} \mathcal{P}_M(x^*) &= \{d \in \mathbb{R} \mid \exists (\lambda^G, \lambda^H) \text{ with } \lambda^G \lambda^H = 0 \text{ or } \lambda^G > 0, \lambda^H > 0, \\ &\quad d = -\lambda^G \nabla G(x^*) - \lambda^H \nabla H(x^*)\}, \\ &= \mathbb{R}. \end{aligned}$$

In this particular case \mathcal{P}_M is a convex cone and so, $\mathcal{P}_M(x^) = \mathcal{T}_{\mathcal{Z}}^\circ(x^*) \neq \mathcal{L}_{MPCC}(x^*)^\circ$. Thus, MPCC-wGCQ holds at x^* .*

It remains to prove that MPCC-wGCQ is a constraint qualification. The following theorem is the analogous of Theorem 4.1.1 of [71], whose proof is now made very straightforward.

Theorem 4.3.1. *A local minimum of (MPCC) that satisfies MPCC-wGCQ or any stronger MPCC CQ is an M-stationary point.*

Proof. Let x^* be a local minimum of (MPCC). Therefore, using MPCC-wGCQ and Lemma 4.3.2 yields to

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

The last equivalence comes from Lemma 4.3.1 and concludes the proof. \square

4.3.2 The weakest constraint qualification for MPCC

In the previous section, we have introduced a new weak constraint qualification for (MPCC). We now prove that this new condition is the weakest MPCC CQ. In order to formalise the discussion let us introduce the definition of (g, h, G, H) MPCC-regular at a feasible point that is an extension of (g, h) regular in [86].

Definition 4.3.2. (g, h, G, H) is said MPCC-regular at x^* , if for all function f that admits a local constrained minimum at x^* , x^* is an M-stationary point.

Indeed, it is very important to notice here that a constraint qualification must be independent of the objective function, since it describes the feasible set only. Considering Lemma 4.3.1 applied to this definition yields to an analogous lemma.

Lemma 4.3.3. Assume that (g, h, G, H) is said MPCC-regular at $x^* \in \mathcal{Z}$. Then, for all function f that admits a local constrained minimum at x^* it holds that $-\nabla f(x^*) \in \mathcal{P}_M(x^*)$.

Our main result is now stated as follows.

Theorem 4.3.2. (g, h, G, H) is MPCC-regular at $x^* \in \mathcal{Z} \iff$ MPCC-wGCQ holds at x^* .

Before advancing to the proof of our main theorem, we give an additional theorem whose proof can be found in [86]. The original result is stated for local maximum, but there is no loss of generality to write it for a local minimum.

Theorem 4.3.3. For every, $y \in \mathcal{T}(x^*)^\circ$ 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 4.3.2.

Proof. The " \Leftarrow " part is given by Theorem 4.3.1.

So, let us consider the " \Rightarrow " part. Assume that (g, h, G, H) is MPCC-regular at $x^* \in \mathcal{Z}$ and prove that for any $y \in \mathcal{T}_\mathcal{Z}(x^*)^\circ$ then $y \in \mathcal{P}_M(x^*)$. By Theorem 4.3.3, 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 MPCC-regular at $x^* \in \mathcal{Z}$ it follows by Lemma 4.3.3 that for all functions f such that $-\nabla f(x^*) \in \mathcal{T}_\mathcal{Z}(x^*)^\circ$ we have $-\nabla f(x^*) \in \mathcal{P}_M(x^*)$. □

A local minimum of some (MPCC) may satisfy the M-stationary conditions in a case where MPCC-wGCQ fails to hold at x^* as illustrated in the following Example 4.3.3. However, this would not be true independently of the choice of the objective function as illustrated by Example 4.3.4. So, it does not contradict our result.

Example 4.3.3.

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

The only feasible is the origin and therefore the global minimum of this problem. We can easily verify that it is an M-stationary point for $\lambda^G = -1, \lambda^H = 0$. The polar of tangent cone is given by $\mathcal{T}_Z^\circ = \mathbb{R}^2$. Let us now verify that $\mathcal{P}_M(x^*) \neq \mathbb{R}^2$. The cone \mathcal{P}_M at x^* is given by

$$\mathcal{P}_M(x^*) = \{d \in \mathbb{R}^n \mid \exists (\lambda^G, \lambda^H) \in \mathbb{R}^2 \text{ with } \lambda^G \lambda^H = 0 \text{ or } \lambda^G > 0, \lambda^H > 0, d = -(\lambda^G \lambda^H)^T\},$$

which is obviously smaller than \mathcal{T}_Z° and thus MPCC-wGCQ fails to hold.

Example 4.3.4.

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

The only feasible is the origin and therefore the global minimum of this problem. However there is no multiplier such that $x^* = 0$ is an M-stationary point.

4.4 Algorithmic applications of the MPCC CQs

This section focusses on some algorithmic applications of MPCC-CCP and MPCC-CRSC. We remind that it has been proved earlier than the latter is a stronger condition. In particular, we focus on the convergence of MPCC-AKKT sequences that are sequences, which satisfy the M-stationary conditions only asymptotically. This definition has first been given in [165].

Definition 4.4.1. We say that $x^* \in \mathcal{Z}$ is an MPCC-AKKT point if there are sequences $\{x^k\}$ (MPCC-AKKT sequence), $\{\lambda^k\}$ such that

$$\nabla f(x^k) + \sum_{i=1}^p \lambda_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i(x^k) - \sum_{i=1}^q \lambda_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^q \lambda_i^{H,k} \nabla H_i(x^k) \rightarrow 0, \quad (4.3)$$

where $\text{supp}(\lambda^{g,k}) \subset \mathcal{I}_g(x^*)$, $\text{supp}(\lambda^{G,k}) \subset \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*)$, $\text{supp}(\lambda^{H,k}) \subset \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{0+}(x^*)$ and either $\lambda_i^{G,k} \lambda_i^{H,k} = 0$ or $\lambda_i^{G,k} > 0, \lambda_i^{H,k} > 0$ for each $i \in \mathcal{I}^{00}(x^*)$.

The interest for MPCC-AKKT optimality conditions essentially relies on the fact that it is a necessary condition to be a local minimiser that does not require any constraint qualification as shown in Theorem 3.1 of [165]. Now, studying the convergence of algorithms that compute MPCC-AKKT sequences, it is of interest to know under which condition these sequences converge to an M-stationary point. A similar question has been answered in the non-linear programming case in [14]. The extension to the MPCC is presented in [165]. We state here the result that uses the MPCC-CCP constraint qualification defined in Definition 4.2.4 and that we write explicitly here.

Definition 4.4.2. We say that a feasible point x^* satisfies the MPCC-CCP if the set-valued mapping $\mathbb{R}^n \ni x \rightrightarrows K_{MPCC}(x)$ such that

$$\begin{aligned} K_{MPCC}(x) := & \left\{ \sum_{i \in \mathcal{I}_g(x^*)} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) - \sum_{i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*)} \lambda_i^G \nabla G_i(x) - \sum_{i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*)} \lambda_i^H \nabla H_i(x) \right. \\ & \left. : \lambda_i^g \in \mathbb{R}_+ \text{ and , either } \lambda_i^G \lambda_i^H = 0 \text{ either } \lambda_i^G > 0, \lambda_i^H > 0 \text{ for } i \in \mathcal{I}^{00}(x^*) \right\} \end{aligned}$$

is outer semi-continuous at x^* , that is

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

MPCC-CCP has been proved to be an MPCC CQ in the Theorem 3.5 of [165]. The following result from [165] states that this condition is actually the weakest condition that ensures the converge of MPCC-AKKT sequences to a M-stationary point.

Theorem 4.4.1. *MPCC-CCP holds if and only if every MPCC-AKKT point is an M-stationary point.*

It is to be noted that MPCC-CCP does not require any boundedness assumption on the sequence $\{\lambda^k\}$. It may be of interest to know for which constraint qualification such property is satisfied. The answer turns out to be MPCC-CRSC that was defined in Definition 4.2.4 and that we write explicitly here.

Definition 4.4.3. *Let $x^* \in \mathcal{Z}$. MPCC-CRSC holds at x^* if for any partition $\mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00} = \mathcal{I}^{00}(x^*)$ such that*

$$\begin{aligned} & \sum_{i \in \mathcal{I}_g(x^*)} \lambda_i^g \nabla g_i(x^*) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x^*) - \sum_{i \in \mathcal{I}^{0+}(x^*) \cup \mathcal{I}_{++}^{00}} \lambda_i^G \nabla G_i(x^*) - \sum_{i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}_{++}^{00}} \lambda_i^H \nabla H_i(x^*) \\ & + \sum_{i \in \mathcal{I}_{-0}^{00}} \lambda_i^G \nabla G_i(x^*) + \sum_{i \in \mathcal{I}_{0-}^{00}} \lambda_i^H \nabla H_i(x^*) = 0, \end{aligned}$$

with $\lambda_i^g \geq 0$ ($i \in \mathcal{I}_g(x^*)$), λ_i^G and $\lambda_i^H \geq 0$ ($i \in \mathcal{I}_{++}^{00}$), $\lambda_i^G > 0$ ($i \in \mathcal{I}_{-0}^{00}$), λ_i^H ($i \in \mathcal{I}_{0-}^{00}$) > 0 , there exists $\delta > 0$ such that the family of gradients

$$\{\nabla g_i(x) \ (i \in \mathcal{I}_1), \ \nabla h_i(x) \ (i = 1, \dots, m), \ \nabla G_i(x) \ (i \in \mathcal{I}_3), \ \nabla H_i(x) \ (i \in \mathcal{I}_4)\}$$

has the same rank for every $x \in \mathcal{B}_\delta(x^*)$, where

$$\begin{aligned} \mathcal{I}_1 &:= \{i \in \mathcal{I}_g(x^*) \mid -\nabla g_i(x^*) \in \mathcal{P}_M(x^*)\}, \\ \mathcal{I}_3 &:= \mathcal{I}^{0+}(x^*) \cup \{i \in \mathcal{I}_{++}^{00} \mid \nabla G_i(x^*) \in \mathcal{P}_M(x^*)\} \cup \mathcal{I}_{-0}^{00}, \\ \mathcal{I}_4 &:= \mathcal{I}^{+0}(x^*) \cup \{i \in \mathcal{I}_{++}^{00} \mid \nabla H_i(x^*) \in \mathcal{P}_M(x^*)\} \cup \mathcal{I}_{0-}^{00}. \end{aligned}$$

It is not necessary to add that the gradients $-\nabla G_i(x^*)$ and $-\nabla H_i(x^*)$ belong to $\mathcal{P}_M(x^*)$. Indeed, we already require that λ_i^G and λ_i^H must be non-zero respectively for the indices $i \in \mathcal{I}_{-0}^{00}$ and $i \in \mathcal{I}_{0-}^{00}$ and so it implies that these gradients belong to this set. We prove that under MPCC-CRSC any MPCC-AKKT sequence is bounded. Beforehand, we prove that this result is true under MPCC-CRCQ. The following results give a characterisation of some sequences that satisfy MPCC-CRCQ and MPCC-CRSC at their limit point. Note that this result is essential for the convergence proof of relaxation methods for (MPCC) that will be studied in the next section, since it proves boundedness of approximate stationary sequences.

During the process of an iterative algorithm it is common to compute sequences $\{x^k\}$ and $\{\lambda^k\} \in \mathbb{R}_+^p \times \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^q$ such that $x^k \rightarrow x^*$ and

$$\nabla f(x^k) + \sum_{i=1}^p \lambda_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i(x^k) - \sum_{i=1}^q \lambda_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^q \lambda_i^{H,k} \nabla H_i(x^k) \rightarrow 0, \quad (4.4)$$

$$\forall i \notin \mathcal{I}_g(x^*) \quad \lim_{k \rightarrow \infty} \frac{\lambda_i^{g,k}}{\|\lambda^k\|_\infty} = 0, \quad \forall i \in \mathcal{I}^{+0} \quad \lim_{k \rightarrow \infty} \frac{\lambda_i^{G,k}}{\|\lambda^k\|_\infty} = 0 \text{ and } \forall i \in \mathcal{I}^{0+} \quad \lim_{k \rightarrow \infty} \frac{\lambda_i^{H,k}}{\|\lambda^k\|_\infty} = 0. \quad (4.5)$$

We will refer to (4.4) and (4.5) as *Assumption 1*. This condition may correspond to some kind of sequential optimality conditions.

According to Lemma 1.1.1, we can build a sequence $\{\bar{\lambda}^k\}$ that satisfies *Assumption 1*, such that the gradients corresponding to non-vanishing multipliers in equation (4.4) are linearly independent for all $k \in \mathbb{N}$. This may change the multipliers, but a previously positive multiplier will stay at least non-negative and a vanishing multiplier will remain zero.

Theorem 4.4.2. *Let x^* be in \mathcal{Z} such that MPCC-CRCQ holds at x^* . Given two sequences $\{x^k\}, \{\lambda^k\}$ that satisfies Assumption 1 and the family of gradients of non-vanishing multipliers in (4.4) are linearly independent for all $k \in \mathbb{N}$. Then, the sequence $\{\lambda^k\}$ is bounded.*

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

$$w^k := \sum_{j \in \mathcal{I}_g(x^*)} \lambda_j^{g,k} \nabla g_j(x^k) + \sum_{i=1}^m \lambda_i^{h,k} \nabla h_i(x^k) - \sum_{j \in \mathcal{I}^{0+}(x^*) \cup \mathcal{I}^{00}(x^*)} \lambda_j^{G,k} \nabla G_j(x^k) - \sum_{j \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*)} \lambda_j^{H,k} \nabla H_j(x^k). \quad (4.6)$$

We prove by contradiction that the sequence $\{\lambda^k\}$ is bounded. Assuming that λ^k is not bounded, therefore there exists a subsequence such that

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

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

Note that conditions (4.4) and (4.5) give that $\lim_{k \rightarrow \infty} w^k = \lim_{k \rightarrow \infty} -\nabla f(x^k) / \|\lambda^k\|_\infty = 0$. Dividing by $\|\lambda^k\|_\infty$ and passing to the limit in the equation above yields

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

with $\bar{\lambda}_j^g = 0$ for $j \notin \mathcal{I}_g(x^*)$, $\bar{\lambda}_j^G = 0$ for $j \in \mathcal{I}^{+0}(x^*)$ and $\bar{\lambda}_j^H = 0$ for $j \in \mathcal{I}^{0+}(x^*)$ by (4.5).

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

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

The following result is similar to Theorem 4.4.2 and focus on the case where the limit point is an M-stationary point.

Theorem 4.4.3. *Given two sequences $\{x^k\}, \{\lambda^k\}$ that satisfies Assumption 1 and the family of gradients of non-vanishing multipliers in (4.4) are linearly independent. Suppose that $x^k \rightarrow x^* \in \mathcal{Z}$, and MPCC-CRSC holds at x^* . Furthermore, assume that $\forall i \in \mathcal{I}^{00}(x^*)$*

$$\text{either } \lim_{k \rightarrow \infty} \frac{\lambda_i^{G,k}}{\|\lambda^k\|_\infty} \lim_{k \rightarrow \infty} \frac{\lambda_i^{H,k}}{\|\lambda^k\|_\infty} = 0 \text{ or } \lim_{k \rightarrow \infty} \frac{\lambda_i^{G,k}}{\|\lambda^k\|_\infty} > 0, \lim_{k \rightarrow \infty} \frac{\lambda_i^{H,k}}{\|\lambda^k\|_\infty} > 0. \quad (4.7)$$

Then, the sequence $\{\lambda^k\}$ is bounded.

Proof. The proof is completely similar to Theorem 4.4.2. Assuming that $\{\lambda^k\}$ is not bounded, we can extract a subsequence such that

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

Dividing by $\|\lambda^k\|_\infty$ and passing to the limit in the equation (4.6) yields

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

with $\bar{\lambda}_j^g = 0$ for $j \notin \mathcal{I}_g(x^*)$, $\bar{\lambda}_j^G = 0$ for $j \in \mathcal{I}^{+0}(x^*)$, $\bar{\lambda}_j^H = 0$ for $j \in \mathcal{I}^{0+}(x^*)$ and either $\bar{\lambda}_j^G \bar{\lambda}_j^H = 0$ or $\bar{\lambda}_j^G > 0, \bar{\lambda}_j^H > 0$ for $j \in \mathcal{I}^{00}(x^*)$ by (4.5) and (4.7).

It is clear that the family of gradients considered in the definition of MPCC-CRSC corresponds to the gradients with non-zero multipliers in the previous equation. Indeed, by linear dependence of the gradients at x^* any gradient whose multiplier is non-zero may be formulated as a linear combination of the other gradients.

Therefore, those gradients with non-vanishing multipliers belong to the polar of the M-linearized cone. MPCC-GCRSC guarantees that these gradients remain linearly dependent in a whole neighborhood.

In the same way as in Theorem 4.4.2, this leads to a contradiction. Thus, the sequence $\{\lambda^k\}$ is bounded. \square

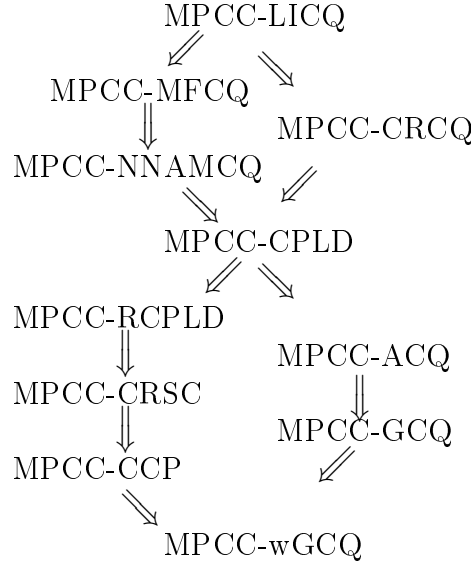


Figure 4.6: Diagram of the relationships between the MPCC-Constraint Qualifications.

4.5 Diagram of relationship between the MPCC CQs

The diagram connecting all the MPCC CQs defined in this chapter is given in Figure 4.6. Most of these relation comes from Proposition 4.2.2. Proposition 4.5.1 proved below gives that MPCC-mGCQ is equivalent to MPCC-wGCQ, therefore MPCC-GCQ implies MPCC-mGCQ by Proposition 4.3.1. Some of these relationships have been known from the literature. By definition of the cone \mathcal{L}_{MPCC} , it can be obtained that $MPCC-ACQ \implies MPCC-GCQ$, see [159, 181, 72]. MPCC-CPLD implies MPCC-ACQ was proved in [91] and [203]. MPCC-CCP and MPCC-ACQ have been proved to be independent in [165], while the same hold for MPCC-RCPLD introduced in [91]. Relations between MPCC-LICQ [181], MPCC-MFCQ [114, 181] and MPCC-NNAMCQ [114, 198] are straightforward from their definition and their independence with MPCC-CRCQ has been illustrated in the aforementioned articles.

The following result shows that the conditions defined in Definition 4.2.4 are indeed some MPCC constraint qualifications, since they imply the weakest MPCC CQ.

Proposition 4.5.1. *Let $x^* \in \mathcal{Z}$. MPCC-mGCQ holds at x^* if and only MPCC-wGCQ holds at x^* .*

Proof. We prove in Theorem 4.3.2 that

$$MPCC-wGCQ \text{ holds at } x^* \iff (g, h, G, H) \text{ is MPCC-regular at } x^*.$$

Now, by definition of MPCC-regular given in Definition 4.3.2 and by Definition of M-

stationary it follows that

$$(g, h, G, H) \text{ is MPCC-regular at } x^* \iff ((g, G_{\mathcal{I}_{++}^{00}}, H_{\mathcal{I}_{++}^{00}}, G_{\mathcal{I}_{-0}^{00}}, H_{\mathcal{I}_{-0}^{00}}), (h, G_{\mathcal{I}^{0+}(x^*)}, H_{\mathcal{I}^{0+}(x^*)})) \\ \text{regular for all partition } \mathcal{I}^{00}(x^*) = \mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00}.$$

According to [86] a regular pair (g, h) at x^* is equivalent to Guignard CQ holding at x^* for (NLP) and thus for all partition $\mathcal{I}^{00}(x^*) = \mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00}$

$$((g, G_{\mathcal{I}_{++}^{00}}, H_{\mathcal{I}_{++}^{00}}, G_{\mathcal{I}_{-0}^{00}}, H_{\mathcal{I}_{-0}^{00}}), (h, G_{\mathcal{I}^{0+}(x^*)}, H_{\mathcal{I}^{0+}(x^*)})) \text{ regular} \iff \text{GCQ holds at } x^* \\ \text{for } (MNL P_{\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}}(x^*)).$$

By definition 4.2.4, GCQ applied to each non-linear program $(MNL P_{\mathcal{I}_{++}^{00}, \mathcal{I}_{0-}^{00}}(x^*))$ for every partition $\mathcal{I}^{00}(x^*) = \mathcal{I}_{++}^{00} \cup \mathcal{I}_{0-}^{00} \cup \mathcal{I}_{-0}^{00}$ is MPCC-mGCQ. Since we use only equivalence the result follows. \square

We conclude this section by providing counter-examples illustrating the relations of the new MPCC-CRSC to other conditions. The first example shows that MPCC-CCP is strictly weaker than MPCC-CRSC. This example will be later used in Section 6.2 to illustrate convergence of relaxation methods.

Example 4.5.1. *Consider the following two-dimensional set*

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1^2 \perp x_1 + x_2^2 \geq 0\}.$$

The feasible set of this example is the set $\mathcal{Z} = \{(x_1, x_2)^T \in \mathbb{R}^2 \mid x_1 = 0\} \cup \{(x_1, x_2)^T \in \mathbb{R}^2 \mid x_1 = -x_2^2\}$. There is a unique stationary point $(0, 0)$, which is M -stationary with $(\lambda^G, \lambda^H = 0)$.

It is easy to verify that MPCC-CCP holds at this point. However, MPCC-CRSC fails to hold at any point $(0, a \in \mathbb{R})^T$ since the gradient of x_1^2 is non-zero for $x \neq 0$.

The following example proves that MPCC-CRSC is strictly weaker than MPCC-RCPLD.

Example 4.5.2. *Consider the following two-dimensional set*

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid g(x) = x_1 + x_2^2 \leq 0, 0 \leq G(x) = -x_1 + x_2 \perp H(x) = x_1 - x_2 \geq 0\}$$

at a point $x^ = (0, 0)$. The gradients of the active constraints are given by*

$$\nabla g(x^*) = (1, 0)^T, -\nabla G(x^*) = (-1, 1)^T, -\nabla H(x^*) = (1, -1)^T.$$

The family of three gradients is positively linearly dependent and span the whole space. However, moving in a neighbourhood of x^ the three gradients become linearly dependent. Thus, MPCC-RCPLD fails to hold at x^* .*

MPCC-CRSC holds at x^ , since $\mathcal{J}_- = \{2, 3\}$ and these two vectors have the same rank in a neighbourhood of x^* .*

We now consider the relation between MPCC-CRSC and MPCC-ACQ. Next example shows that MPCC-ACQ may hold even when MPCC-CRSC does not.

Example 4.5.3. *Consider the following two-dimensional set*

$$\{x \in \mathbb{R}^2 \mid 0 \leq G(x) = x_1 \perp H(x) = x_1^2 x_2^2 - x_1 \geq 0\}.$$

The feasible set of this problem can be written as $x_1 = 0$ or $x_1 = x_1^2 x_2^2$. MPCC-ACQ holds at $x^ = 0$ since $\mathcal{T}_Z = \mathcal{L}_{MPCC} = \{d \in \mathbb{R}^2 \mid d_1 = 0\}$. The two active gradients verify $\nabla G(x^*) = -\nabla H(x^*)$ and so $\mathcal{I}_3 = \{1\}$ and $\mathcal{I}_4 = \{1\}$. Moreover, in a neighbourhood of x^* the gradients are not linearly dependent and thus MPCC-CRSC fails to hold at x^* .*

The technique used in Definition 4.2.4 may very well be used to extend all the constraint qualifications and their applications from (NLP) to (MPCC). Besides, we present here only the first order constraint qualifications, but necessary and sufficient conditions have also been studies in the literature as in [91].

Chapter 5

Relaxation methods for MPCC

In this section, we consider relaxation methods to solve (MPCC) and we review most of the existing methods in the literature. We already point out in the previous section that (MPCC) is a difficult non-linear program due to the degeneracy of the complementarity constraints. A relaxation of the complementarity constraints generates a sequence of non-linear programs, which are more regular than the initial problem. Thus, we can apply the well-studied numerical methods for non-linear programming.

The sketch of such a method is described in Algorithm 1 and behaves as follows: we consider a non-linear parametric program R_{t_k, \bar{t}_k} , which considers a relaxation of the complementarity constraints that depends on the parameters t_k, \bar{t}_k . For each value $t_k > 0, \bar{t}_k > 0$ of decreasing sequences $\{t_k\}, \{\bar{t}_k\}$ we compute a stationary point of R_{t_k, \bar{t}_k} , which provides a sequence $\{x^{k+1}\}$. Such stationary points are computed using iterative methods that require an initial point. We use the previous stationary point as an initial point. The precision used to compute x^{k+1} may depend on the value of the parameters. For t_k, \bar{t}_k sufficiently small the sequence $\{x^{k+1}\}$ converge to an MPCC-stationary point of (MPCC). The relaxed

Data:

starting vector x^0 ; initial relaxation parameter (t_0, \bar{t}_0) ; update parameter $(\sigma_t, \sigma_{\bar{t}}) \in (0, 1)^2$; p_{min} minimal value of the parameter; ϵ precision ;

1 **Begin** ;

2 Set $k := 0$;

3 **while** $t_k > p_{min}$ **do**

4 x^{k+1} stationary point of R_{t_k, \bar{t}_k} with x^k initial point;

5 $(t_{k+1}, \bar{t}_{k+1}) := (t_k \sigma_t, \bar{t}_k \sigma_{\bar{t}})$;

Algorithm 1: Generic relaxation method for (MPCC), with a corresponding relaxed non-linear program R_{t_k, \bar{t}_k} .

sub-problems $R_{t,\bar{t}}$ with parameters t and \bar{t} are written in a general form as

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

The parameter t is not necessarily a real number, but can also be a vector of parameters. Let $\#t$ be the length of the vector t . $\bar{t} \in \mathbb{R}$ is chosen such that $\lim_{\|t\| \rightarrow 0} \bar{t} = 0$. The map $\Phi : \mathbb{R}^q \times \mathbb{R}^q \rightarrow \mathbb{R}^q$ is the relaxation map that is used to reformulate the complementarity constraints. Most of the methods proposed in the literature consider sub-problems like $(R_{t,\bar{t}})$. One exception is the relaxation of G.-H. Lin & M. Fukushima, [126].

Through the rest of this chapter, we review some of the existing methods proposed in the literature. We mainly focus on the convergence properties of such methods and the regularity properties of the corresponding relaxed sub-problems. In particular, we are interested on the stationary properties of the limit point attained by the sequence generated by Algorithm 1. According to Chapter 4 our aim is to get an M-stationary point. This state of the art is split in three parts:

- Regular relaxations [182, 126, 52, 93]: Section 5.1;
- A local relaxation [185]: Section 5.2;
- The paradigm of M-relaxations [111, 112]: Section 5.3

5.1 Regular relaxations

The first attempt in the literature to use a relaxation technique to deal with (MPCC) goes back to S. Scholtes in 2001 in [182]. They consider a method with $t \in \mathbb{R}_{++}$ and $\bar{t} = 0$ defined by the following relaxation map, $\forall i \in \{1, \dots, q\}$

$$\Phi_i^{SS}(G(x), H(x); t) := G_i(x)H_i(x) - t^2. \tag{SS}$$

According to the study in [182, 184], this method converges to a C-stationary point if MPCC-MFCQ holds at its limit point and the non-linear program $R_{t,0}^{SS}$, defined as $(R_{t,\bar{t}})$ with (SS), satisfies classical MFCQ at any point feasible for $R_{t,0}^{SS}$ in a neighbourhood of any point feasible for (MPCC) that satisfies MPCC-MFCQ.

This relaxation is clearly more regular than (MPCC) since MFCQ is violated at any feasible point of the original problem. The idea to additionally relax the positivity constraints, i.e. consider $\bar{t} > 0$, has been introduced in [52] as an extension to the relaxation (SS). Although, in [52], the motivation of the authors was not to decrease the two parameters simultaneously.

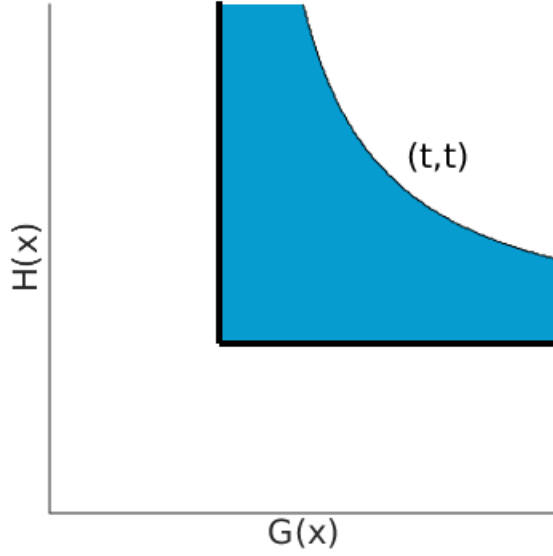


Figure 5.1: Feasible set of H. Scheel & S. Scholtes relaxation (SS).

The relaxation (SS) are very unlikely to satisfy LICQ at a point x^* with $\mathcal{I}^{00}(x^*) \neq \emptyset$, since this would mean that for $t = 0$ three constraints are active for only two gradients. G.-H.Lin & M.Fukushima propose in [126] a relaxation with fewer constraints in order to improve the regularity of the relaxed program by considering

$$\begin{aligned}
 & \min_{x \in \mathbb{R}^n} f(x) \\
 & \text{s.t. } h(x) = 0, \quad g(x) \leq 0, \\
 & \quad G_i(x)H_i(x) \leq t^2, \quad i = 1, \dots, q, \\
 & \quad (G_i(x) + t)(H_i(x) + t) \geq t^2, \quad i = 1, \dots, q.
 \end{aligned} \tag{R_t^{LF}}$$

The authors show in particular that if MPCC-LICQ holds at a point $x^* \in \mathcal{Z}$, this relaxation satisfies the classical LICQ in any feasible points of R_t^{LF} in a neighbourhood of x^* . Moreover, this relaxation converges to a C-stationary point as $t \downarrow 0$ as shown in [113].

In [93], M. Haddou extends the approach of [182] to a general family of relaxations, considering $t \in \mathbb{R}$, $\bar{t} = 0$ and $\forall i \in \{1, \dots, q\}$

$$\Phi_i^\theta(G(x), H(x); t) := \theta_t(G_i(x)) + \theta_t(H_i(x)) - 1, \tag{\theta}$$

where $\theta : \mathbb{R}_+ \rightarrow [0, 1]$ are the functions defined earlier in Chapter 1. The following lemma shows that (θ) is indeed a generalisation of (SS).

Lemma 5.1.1. *Let $(a, b) \in \mathbb{R}_+ \times \mathbb{R}_+$. It holds true that for all $t > 0$*

$$\theta_t^1(a) + \theta_t^1(b) - 1 \leq 0 \iff ab - t^2 \leq 0.$$

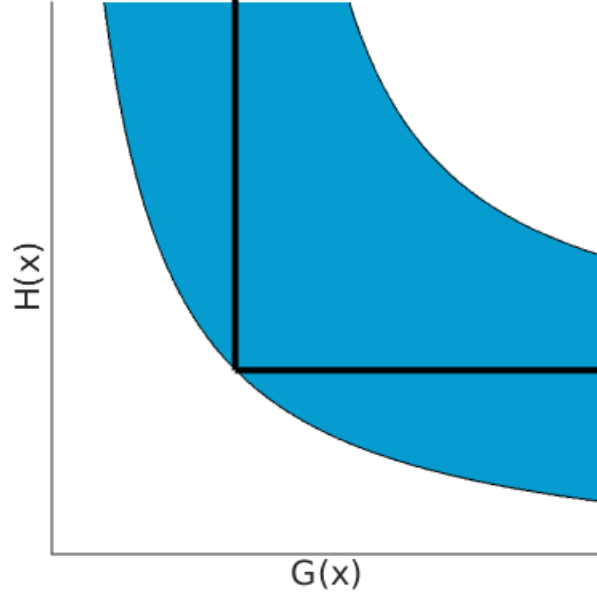


Figure 5.2: Lin & Fukushima relaxation (R_t^{LF})

Proof. By definition of θ^1 it holds that

$$\begin{aligned}
 \theta_t^1(a) + \theta_t^1(b) - 1 \leq 0 &\iff \frac{a}{a+t} + \frac{b}{b+t} - 1 \leq 0, \\
 &\iff \frac{a(b+t) + b(a+t) - (a+t)(b+t)}{(a+t)(b+t)} \leq 0, \\
 &\iff \frac{ab - t^2}{(a+t)(b+t)} \leq 0.
 \end{aligned}$$

The result follows since $(a+t)(b+t) > 0$. □

We now prove that a sequence of stationary point of $(R_{t,\bar{t}})$ with the relaxation map (θ) converges to a C-stationary point in a same way as the relaxation (SS).

Theorem 5.1.1. *Given $\{t_k\}, \{\bar{t}_k\}$ two sequences decreasing to zero. Let $\{x^k, \nu^k\}$ be a sequence of stationary points of R_{t_k, \bar{t}_k} with a relaxation map (θ) , with $x^k \rightarrow x^*$ such that MPCC-CRCQ holds at x^* . Then, x^* is a C-stationary point.*

Proof. First, we identify the expressions of the multipliers of the complementarity constraints in Definition 4.1.1 in function of the stationary points of R_{t_k, \bar{t}_k} . Let $\{x^k, \nu^{g,k}, \nu^{h,k}, \nu^{G,k}, \nu^{H,k}, \nu^{\Phi,k}\}$

be a sequence of KKT points of R_{t_k, \bar{t}_k} for all $k \in \mathbb{N}$, that by definition satisfies

$$\begin{aligned} 0 = & \nabla f(x^k) + \sum_{i=1}^p \nu_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \nu_i^{h,k} \nabla h_i(x^k) - \sum_{i=1}^q \nu_i^{G,k} \nabla G_i(x^k) \\ & - \sum_{i=1}^q \nu_i^{H,k} \nabla H_i(x^k) + \sum_{i=1}^q \nu_i^{\Phi,k} \nabla \Phi_i^\theta(G(x^k), H(x^k); t_k), \end{aligned}$$

with

$$\begin{aligned} \nu_i^{g,k} &= 0, \quad \forall i \notin \mathcal{I}_g(x^k) \text{ and } \lambda_i^{g,k} \geq 0, \quad \forall i \in \mathcal{I}_g(x^k), \\ \nu_i^{G,k} &= 0, \quad \forall i \notin \mathcal{I}_G(x^k; \bar{t}_k) \text{ and } \lambda_i^{G,k} \geq 0, \quad \forall i \in \mathcal{I}_G(x^k; \bar{t}_k), \\ \nu_i^{H,k} &= 0, \quad \forall i \notin \mathcal{I}_H(x^k; \bar{t}_k) \text{ and } \lambda_i^{H,k} \geq 0, \quad \forall i \in \mathcal{I}_H(x^k; \bar{t}_k), \\ \nu_i^{\Phi,k} &= 0, \quad \forall i \notin \mathcal{I}_\Phi(x^k; t_k) \text{ and } \lambda_i^{\Phi,k} \geq 0, \quad \forall i \in \mathcal{I}_\Phi(x^k; t_k). \end{aligned}$$

By the definition of Φ^θ it holds that for all $i \in \{1, \dots, q\}$

$$\nabla \Phi_i^\theta(x^k; t_k) = \theta'_{t_k}(G_i(x^k)) \nabla G_i(x^k) + \theta'_{t_k}(H_i(x^k)) \nabla H_i(x^k).$$

Thus, we can rewrite the equation above as

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

where

$$\begin{aligned} \eta_i^{G,k} &= \begin{cases} \nu_i^{G,k}, & \text{if } i \in \mathcal{I}_G(x^k; \bar{t}_k) \\ -\nu_i^{\Phi,k} \theta'_{t_k}(G_i(x^k)), & \text{if } i \in \mathcal{I}_\Phi(x^k; t_k) \\ 0, & \text{otherwise,} \end{cases} \\ \eta_i^{H,k} &= \begin{cases} \nu_i^{H,k}, & \text{if } i \in \mathcal{I}_H(x^k; \bar{t}_k) \\ -\nu_i^{\Phi,k} \theta'_{t_k}(H_i(x^k)), & \text{if } i \in \mathcal{I}_\Phi(x^k; t_k) \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

These multipliers are well-defined since $\mathcal{I}_G(x^k; \bar{t}_k) \cap \mathcal{I}_\Phi(x^k; t_k) = \emptyset$ and $\mathcal{I}_H(x^k; \bar{t}_k) \cap \mathcal{I}_\Phi(x^k; t_k) = \emptyset$. Furthermore, the following inclusions hold true

$$\mathcal{I}^{0+}(x^*) \cup \mathcal{I}^{00}(x^*) \subset \mathcal{I}_G(x^k; \bar{t}_k) \cup \mathcal{I}_\Phi(x^k; t_k),$$

$$\mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*) \subset \mathcal{I}_H(x^k; \bar{t}_k) \cup \mathcal{I}_\Phi(x^k; t_k).$$

Now, by Lemma 1.4.3 it holds that for $c^k = \|\nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\|_\infty$

$$\forall i \in \mathcal{I}^{+0}(x^*) \quad \lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|c^k\|_\infty} = 0 \text{ and } \forall i \in \mathcal{I}^{0+}(x^*) \quad \lim_{k \rightarrow \infty} \frac{\eta_i^{H,k}}{c^k} = 0.$$

Thus, applying Corollary 4.4.2 gives boundedness of the sequence $\{\nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\}$. Therefore, this sequence converges, up to some subsequence, to some limit

$$\{\nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\} \rightarrow (\nu^g, \nu^h, \eta^G, \eta^H).$$

We now prove that the point x^* with multipliers $(\nu^g, \nu^h, \eta^G, \eta^H)$ is a weak-stationary point and then that it is a C-stationary point.

Let $i \in \mathcal{I}^{+0}(x^*)$, weak-stationary conditions require that $\eta_i^G = 0$. It holds that for k sufficiently large $i \in \mathcal{I}_H(x^k; \bar{t}_k) \cup \mathcal{I}_\Phi(x^k; t_k)$. Consider two cases, since $\mathcal{I}_H(x^k; \bar{t}_k) \cap \mathcal{I}_\Phi(x^k; t_k) = \emptyset$:

- For $i \in \mathcal{I}^{+0}(x^*) \cap \mathcal{I}_H(x^k; \bar{t}_k)$, it holds that $\eta_i^{G,k} = 0$ and so $\eta_i^G = 0$.
- For $i \in \mathcal{I}^{+0}(x^*) \cap \mathcal{I}_\Phi(x^k; t_k)$, it holds that $\eta_i^{G,k} = -\nu_i^{\Phi,k} \theta'_{t_k}(G_i(x^k))$ and $\eta_i^{H,k} = -\nu_i^{\Phi,k} \theta'_{t_k}(H_i(x^k))$. Assume by contradiction that $\eta_i^G \neq 0$. Therefore, this assumption yields to $1/\nu_i^{\Phi,k} = O(\theta'_{t_k}(G_i(x^k)))$ since $\theta'_{t_k}(G_i(x^k)) \rightarrow 0$ by Lemma 1.4.3. However, this is a contradiction with boundedness of $\{\eta^{G,k}, \eta^{H,k}\}$ since by Lemma 1.4.4 it would follow that $|\eta_i^{H,k}| \rightarrow \infty$. So, $\eta_i^G = 0$.

By symmetry of the relaxation, it is similar to prove that $\eta_i^H = 0$ for all $i \in \mathcal{I}^{+0}(x^*)$.

Since $\eta_i^G = 0$ for $i \in \mathcal{I}^{+0}(x^*)$ and $\eta_i^H = 0$ for $i \in \mathcal{I}^{+0}(x^*)$, it is straightforward that $\nabla \mathcal{L}_{MPCC}(x^*, \nu^g, \nu^h, \eta^G, \eta^H) = 0$. Then, x^* is a weak-stationary point.

Let us verify that x^* is a C-stationary point by showing that $\eta_i^G \eta_i^H \geq 0$ for all $i \in \mathcal{I}^{00}(x^*)$. Let $i \in \mathcal{I}^{00}(x^*)$. Assume that for k sufficiently large $i \notin \mathcal{I}_{GH}(x^k; t_k)$, then $\eta_i^{G,k} \geq 0$, $\eta_i^{H,k} \geq 0$ and so $\eta_i^G \eta_i^H \geq 0$. Now, assuming that for k sufficiently large $i \in \mathcal{I}_{GH}(x^k; t_k)$ yields $\eta_i^{G,k} = -\nu_i^{\Phi,k} \theta'_{t_k}(G_i(x^k)) \leq 0$, $\eta_i^{H,k} = -\nu_i^{\Phi,k} \theta'_{t_k}(H_i(x^k)) \leq 0$ since θ is an increasing function and $\nu_i^{\Phi,k} \geq 0$. So, $\eta_i^G \eta_i^H \geq 0$. This completes the proof that x^* is a C-stationary point. \square

We already point out that similar result was proved in [184] for the relaxation (SS) under a slightly weaker condition that MPCC-CRCQ but stronger than MPCC-CPLD as defined in the previous chapter. For completeness, we now give a result on existence of stationary point in a neighbourhood of a regular point. This type of result is classical for relaxation techniques, but was missing in the study [93].

Theorem 5.1.2. *Let x^* be feasible for R_{t_k, \bar{t}_k} with (θ) , such that MPCC-LICQ holds at x^* . Then, there exists $\bar{t} > 0$ and a neighbourhood $U(x^*)$ of x^* such that for all $t \in (0, \bar{t}]$, LICQ for R_{t_k, \bar{t}_k} holds for all $x \in U(x^*)$ feasible for R_{t_k, \bar{t}_k} .*

Proof. In particular, we verify that the family of active gradients of R_{t_k, \bar{t}_k} with (θ) is linearly independent in a neighbourhood $U(x^*)$ of x^* . In other words, let us verify that the trivial solution is the only solution of the equation for all $x \in U(x^*)$:

$$\begin{aligned} 0 = & \sum_{i \in \mathcal{I}_g(x^*)} \lambda_i^g \nabla g_i(x) + \sum_{i=1}^m \lambda_i^h \nabla h_i(x) + \sum_{i \in \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{+0}(x^*)} (\lambda_i^\Phi \theta'_r(G_i(x)) - \lambda_i^G) \nabla G_i(x) \\ & + \sum_{i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*)} (\lambda_i^\Phi \theta'_r(H_i(x)) - \lambda_i^H) \nabla H_i(x). \end{aligned}$$

Now, since MPCC-LICQ holds at x^* and by continuity the gradients remain linearly independent in a neighbourhood $U(x^*)$ of x^* such that

$$\begin{aligned}\mathcal{I}^{00}(x) \cup \mathcal{I}^{0+}(x) &\subset \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{0+}(x^*), \\ \mathcal{I}^{00}(x) \cup \mathcal{I}^{+0}(x) &\subset \mathcal{I}^{00}(x^*) \cup \mathcal{I}^{+0}(x^*), \\ \mathcal{I}_g(x) &\subset \mathcal{I}_g(x^*).\end{aligned}$$

Thus, it follows that the solutions λ of the equation above satisfy

$$\begin{aligned}\lambda_i^g &= 0, \forall i \in \mathcal{I}_g(x), \quad \lambda_i^h = 0, \forall i = 1, \dots, m, \\ (\lambda_i^\Phi \theta'_r(G_i(x)) - \lambda_i^G) &= 0, \forall i \in \mathcal{I}^{00}(x) \cup \mathcal{I}^{0+}(x), \\ (\lambda_i^\Phi \theta'_r(H_i(x)) - \lambda_i^H) &= 0, \forall i \in \mathcal{I}^{00}(x) \cup \mathcal{I}^{+0}(x).\end{aligned}$$

Noticing that $\text{supp}(\lambda^G) \cap \text{supp}(\lambda^H) \cap \text{supp}(\lambda^\Phi) = \emptyset$ and that θ is an increasing function yields $\lambda^g = \lambda^h = \lambda^G = \lambda^H = \lambda^\Phi = 0$ and the result follows. \square

It is interesting to note that despite being an "extension" of the relaxation (SS). Splitting the two parts of the complementarity constraints gives a more regular formulation. The main reason of this phenomenon is that the gradient of relaxed complementarity constraint is now non-zero on the boundary of the domain.

All of these methods enjoy very nice regularity properties. However they may converge to some undesirable C-stationary points. This phenomenon is illustrated in the following example.

Example 5.1.1.

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

$x^* = (0, 0)$ is a C-stationary point with multipliers $\lambda^G = \lambda^H = -1$. The relaxations (SS), (R_t^{LF}) and (θ) may produce a sequence of stationary points (t_k, t_k) converging to the origin.

In [126], the authors study some second-order conditions and conditions on the smallest eigenvalue of the Hessian of the Lagrangian to ensure convergence to M-stationary points. However, it is not convenient to assume second-order conditions, when computing first order stationary points of the problem.

5.2 A local relaxation

In [185] and [190], S. Steffensen & M. Ulbrich study a family of local relaxations for all $i \in \{1, \dots, q\}$

$$\Phi_t^{SU}(G(x), H(x); t) := G_i(x) + H_i(x) - \begin{cases} |G_i(x) - H_i(x)| & \text{if } |G_i(x) - H_i(x)| \geq t \\ t\phi(\frac{G_i(x) - H_i(x)}{t}) & \text{otherwise.} \end{cases} \quad (\text{SU})$$

$\phi : \mathcal{D} \subset \mathbb{R} \rightarrow \mathbb{R}$, where \mathcal{D} is an open subset of \mathbb{R} with $[-1, 1] \subset \mathcal{D}$ and $\phi(z)$ satisfies the following conditions:

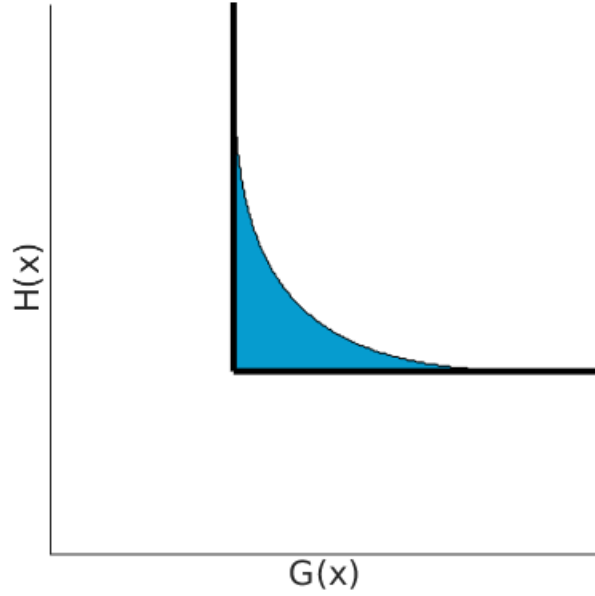


Figure 5.3: S. Steffensen & M. Ulbrich relaxation (SU)

- ϕ is twice continuously differentiable on $[-1, 1]$;
- $\phi(1) = \phi(-1) = 1$;
- $\phi'(-1) = -1$ and $\phi'(1) = 1$;
- $\phi''(-1) = \phi''(1) = 0$;
- ϕ is strictly convex on $[-1, 1]$.

As proposed in [185] a suitable choice of function ϕ can be $\phi(z) = \frac{2}{\pi} \sin\left(\frac{\pi}{2}z + \frac{3\pi}{2}\right) + 1$. The study in [184] reveals that this family of relaxations converges to a C-stationary point if MPCC-TCPLD holds at the limit point. The non-linear programs R_t^{SU} satisfy only ACQ in any point feasible in a neighbourhood of a point that verifies MPCC-LICQ. Therefore, the non-linear programs associated with (SU) are less regular than the relaxations (SS), (θ) , and (R_t^{LF}) . In the same way as the previous ones, these relaxations may converge to a C-stationary point in Example 5.1.1 with the sequence $\frac{2\pi}{\pi-2}(t, t)$.

5.3 The paradigm of M-relaxations

In 2009 A. Kadrani, J.-P. Dussault and A. Bechakroun introduce a method, which enjoys the desired goal to converge to an M-stationary point, see [111]. Their original method considers

an approximation of the complementarity constraints as a union of two boxes connected only on one point (t, t) , in the following way $\forall i \in \{1, \dots, q\}$:

$$\Phi_i^{KDB}(G(x), H(x); t) := (G_i(x) - t)(H_i(x) - t), \quad (\text{KDB})$$

with $\bar{t} = t$. This is not a relaxation but an approximation, since the feasible domain of

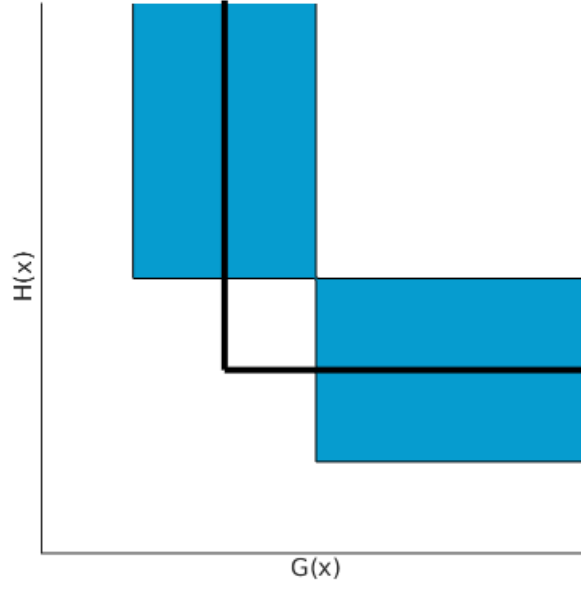


Figure 5.4: Boxes relaxation

R_t^{KDB} does not include the feasible domain of (MPCC). The gap between approximation and relaxation may be the source of computational problems as illustrated by the following example.

Example 5.3.1.

$$\begin{aligned} & \min_{x \in \mathbb{R}^2} x_1 - x_2 \\ & s.t. \quad x_1 \leq 0, \quad x_2 \leq 0, \\ & \quad \quad 0 \leq x_1 \perp x_2 \leq 0. \end{aligned}$$

The feasible set of relaxation KDB is always empty for $t > 0$.

This method has latter been extended to a relaxation in [184, 103] using a piecewise NCP function and considering $\forall i \in \{1, \dots, q\}$

$$\Phi_i^{KS}(G(x), H(x); t) := \phi(G_i(x) - t, H_i(x) - t), \quad (\text{KS})$$

where $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a continuously differentiable NCP-function with for instance

$$\phi(a, b) = \begin{cases} ab, & \text{if } a + b \geq 0 \\ -\frac{1}{2}(a^2 + b^2), & \text{if } a + b < 0. \end{cases}$$

Both methods enjoy a very nice convergence property as stated in the following theorem.

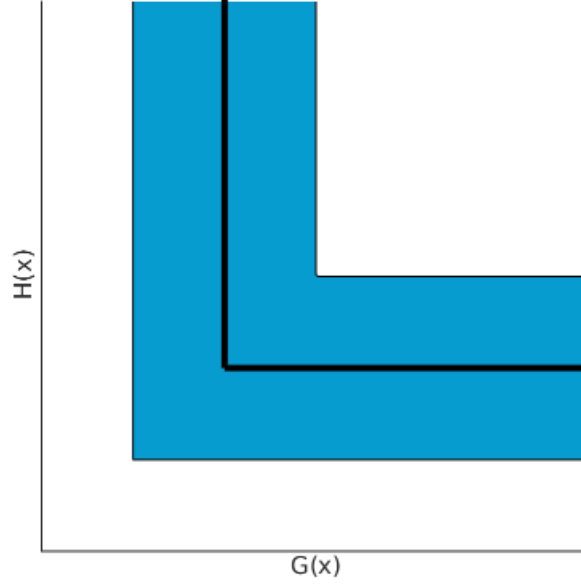


Figure 5.5: L-shape relaxation

Theorem 5.3.1. *Let $\{t_k\}$ be a decreasing sequence of positive parameters with $t_k \downarrow 0$. Let $\{x^k\}$ such that $x^k \rightarrow x^*$ and let $\{x^k, \lambda^k\} \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+^p \times \mathbb{R}_-^q \times \mathbb{R}_-^q \times \mathbb{R}_+^q$ be a sequence of first-order stationary point of $R_{t_k}^{KS}$ (resp. $R_{t_k}^{KDB}$), then*

1. *the sequence $\{x^k, \lambda^k\}$ converges to an M-stationary point of (MPCC) if MPCC-CCP holds at x^* ;*
2. *$R_{t_k}^{KS}$ (resp. $R_{t_k}^{KDB}$) satisfies GCQ in any feasible point of $R_{t_k}^{KS}$ in a neighbourhood of x^* , whenever MPCC-LICQ holds at x^* .*

This theorem is a sum up of two results: 1- has been proved recently in [165] and 2- is the original result from [111]. The former improves a previous result of [184], which considers convergence under MPCC-TCPLD. These methods attain one of the main goals, that is to converge to an M-stationary point. However, they do not handle the case where MPCC-LICQ holds at the limit point and then M-stationary points can be undesirable. One typical example is the following

Example 5.3.2.

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

This problem is unbounded as the global minimum is attained at $(0, \infty)$. Numerical methods can compute a sequence of points $\{(x_k, 0)^T\}_{k=1}^\infty$ converging to $(0, 0)$, which is an M-stationary point but obviously not a local minimum.

As pointed out in Section 4.1, M- and S-stationary points coincide if strict complementarity holds. That is why in the literature convergence to an S-stationary point is usually guaranteed under a hypothesis relative of the indices that do not satisfy strict complementarity. In the Theorem 4.1 of [111], the authors proved convergence of the approximation (KDB) to an S-stationary point if the sequence $\{x^k\}$ satisfies the asymptotically weakly non-degenerate assumption.

Definition 5.3.1. *A sequence $\{x^k\}$ is asymptotically weakly non-degenerate, if $\{x^k\} \rightarrow x^*$ as $\{t_k\} \rightarrow 0$, and there is a $t^* > 0$ such that for $t \in (0, t^*)$ one has*

$$-1 \leq \frac{G_i(x^k)}{H_i(x^k)} \leq 1, \quad i \in \mathcal{I}^{00}(x^*) \cap \{i | G_i(x^k) < t, H_i(x^k) > t\},$$

and

$$-1 \leq \frac{H_i(x^k)}{G_i(x^k)} \leq 1, \quad i \in \mathcal{I}^{00}(x^*) \cap \{i | G_i(x^k) > t, H_i(x^k) < t\}.$$

The following example illustrates this result.

Example 5.3.3.

$$\min x^2 - xy + \frac{1}{3}y^2 - 2x \text{ s.t. } x \geq 0, y \geq 0, xy \leq 0.$$

There is one S-stationary point $(1, 0)$. Using (KDB) we get two M-stationary point $(1, 0) \forall t$ and $(t, \frac{3}{2}t)$ which doesn't satisfy asymptotically weakly non-degenerate assumption.

It should be noted that this theorem does not provide the equivalence as shown by the following example.

Example 5.3.4.

$$\min_{(x_1, x_2)^T \in \mathbb{R}^2} (2x_1 - x_2)^4 + \frac{1}{x_1 x_2 + 1} \text{ s.t. } 0 \leq x_1 \perp x_2 \leq 0.$$

It is clear that $(0, 0)^T$ is the global minimum and that it is an S-stationary point. Picture proof shows that the relaxation KDB and KS possess a global minimum in $x(t) = (\frac{t}{2}, t)$, which converges to $(0, 0)^T$ as t goes to 0 and does not satisfy the asymptotically weakly non-degenerate assumption.

We conclude this chapter with Table 5.1 that sum up the methods presented here and their essential properties.

Name	Relaxation ?	Convergence	Regularity
SS	yes	C-stat.	MFCQ
θ	yes	C-stat.	LICQ
LF	yes	C-stat.	LICQ
SU	yes	C-stat.	ACQ
KDB	no	M-stat.	GCQ
KS	yes	M-stat.	GCQ

Table 5.1: Relaxations methods and their properties.

Chapter 6

The new butterfly relaxation method

We define in this chapter the new relaxation scheme so-called butterfly relaxation introduced in [64]. This new scheme extends the methods (KDB) and (KS) by handling two relaxing parameters instead of one.

In a first step, Section 6.1 introduces the new relaxation and some useful properties. Then, we study the theoretical properties in Section 6.2, 6.3 and 6.4. In particular, we are interested in convergence of a sequence of stationary points, existence of stationary points in a neighbourhood of a solution and discuss briefly convergence of a sequence of approximate stationary points.

6.1 The butterfly relaxation method

We propose a new family of relaxations with two positive parameters (t, r) defined such that for all $i \in \{1, \dots, q\}$

$$\Phi_i^B(G(x), H(x); t, r) = 0 \implies \Phi_i^B(G(x), H(x); t, r) := F_{1i}(x; r, t)F_{2i}(x; t, r), \quad (6.1)$$

and $\Phi_i^B(G(x), H(x); t, r)$ is extended in a continuously differentiable as a function with negative values for $\min(F_{1i}(x; r, t), F_{2i}(x; t, r)) < 0$ and as a function with positive values otherwise. Let us denote

$$F_{1i}(x; r, t) := H_i(x) - t\theta_r(G_i(x)) \text{ and } F_{2i}(x; t, r) := G_i(x) - t\theta_r(H_i(x)),$$

where $\theta_r : \mathbb{R} \rightarrow]-\infty, 1]$ are the functions presented in Section 1.4. This new relaxation handles two parameters r and t that satisfy

$$t\theta'(0) \leq r. \quad (6.2)$$

This condition insures that the intersection point between the sets $\{x \in \mathbb{R}^n \mid F_1(x; r, t) = 0\}$ and $\{x \in \mathbb{R}^n \mid F_2(x; t, r) = 0\}$ is reduced to the origin. In other words, the two branches of the relaxation does not cross each other. A typical choice will be to take $t = o(r)$ motivated by strong convergence properties as discussed in Section 6.2.

One way to write the relaxation for $t < \theta'(0)r$ uses the NCP function from [184] by considering

$$\Phi_i^B(G(x), H(x); t, r) := \begin{cases} F_{1i}(x; r, t)F_{2i}(x; t, r), & \text{if } F_{1i}(x; r, t) + F_{2i}(x; t, r) \geq 0, \\ -\frac{1}{2}(F_{1i}(x; r, t)^2 + F_{2i}(x; t, r)^2), & \text{if } F_{1i}(x; r, t) + F_{2i}(x; t, r) < 0. \end{cases} \quad (6.3)$$

This formulation will be used in the numerics and in the study of convergence of approximate points in Section 6.4. Most of the results presented in this article is only sensitive to the description of the constraint and its boundary.

Since these relaxations are an union of two convex sets connected on a single point we may also consider a relaxation of the positivity constraints parametrized by $\bar{t} \in \mathbb{R}_+$.

This method is an extension of the work of [111, 110] and [102, 112] since handling two parameters allows the two "wings" of the relaxation to be brought closer.

We now introduce some notations that will be extensively used in the sequel. Since the butterfly relaxation handles two parameters we denote $\hat{t} := (t, r)$ to simplify the notation and by extension $\hat{t}_k := (t_k, r_k)$. Let $\mathcal{X}_{\hat{t}, \bar{t}}^B$ be the feasible set of $R_{\hat{t}, \bar{t}}^B$, which corresponds to the non-linear program related to the butterfly relaxation of the complementarity constraints defined in (6.3), that is

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

and

$$\mathcal{X}_{\hat{t}, \bar{t}}^B := \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0, G(x) \geq -\bar{t}e, H(x) \geq -\bar{t}e, \Phi^B(G(x), H(x); \hat{t}) \leq 0\}.$$

Figure 6.1 shows the feasible set of the relaxed complementarity constraints for some relations between t and r . The sets of indices used in the sequel are defined in the following way

$$\begin{aligned} \mathcal{I}_G(x; \hat{t}) &:= \{i = 1, \dots, q \mid G_i(x) + \bar{t} = 0\}, \\ \mathcal{I}_H(x; \hat{t}) &:= \{i = 1, \dots, q \mid H_i(x) + \bar{t} = 0\}, \\ \mathcal{I}_{GH}(x; \hat{t}) &:= \{i = 1, \dots, q \mid \Phi_i^B(G(x), H(x); \hat{t}) = 0\}, \\ \mathcal{I}_{GH}^{0+}(x; \hat{t}) &:= \{i \in \mathcal{I}_{GH}(x; \hat{t}) \mid F_{1i}(x; \hat{t}) = 0, F_{2i}(x; \hat{t}) > 0\}, \\ \mathcal{I}_{GH}^{+0}(x; \hat{t}) &:= \{i \in \mathcal{I}_{GH}(x; \hat{t}) \mid F_{1i}(x; \hat{t}) > 0, F_{2i}(x; \hat{t}) = 0\}, \\ \mathcal{I}_{GH}^{++}(x; \hat{t}) &:= \{i \in \mathcal{I}_{GH}(x; \hat{t}) \mid F_{1i}(x; \hat{t}) > 0, F_{2i}(x; \hat{t}) > 0\}, \\ \mathcal{I}_{GH}^{00}(x; \hat{t}) &:= \{i \in \mathcal{I}_{GH}(x; \hat{t}) \mid F_{1i}(x; \hat{t}) = F_{2i}(x; \hat{t}) = 0\}. \end{aligned}$$

The following two lemmas give more insights on the relaxation.

Lemma 6.1.1. *Let $x \in \mathcal{X}_{\hat{t}, \bar{t}}^B$, then it is true for the relaxation (6.3) that:*

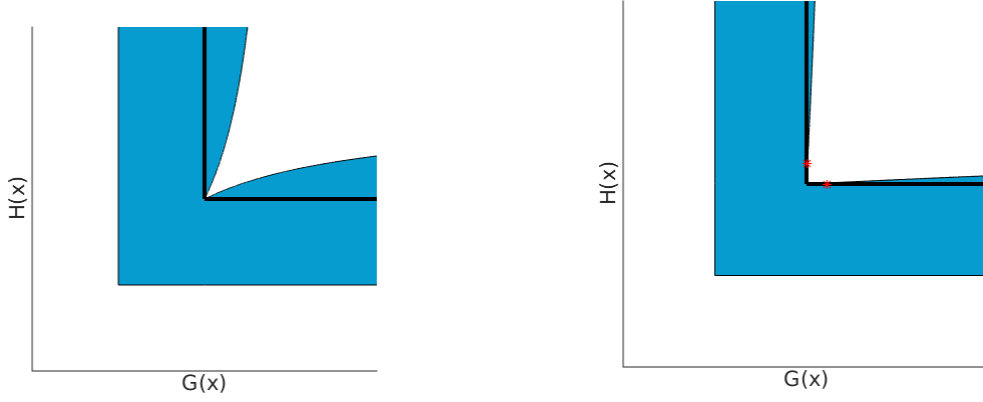


Figure 6.1: Feasible set of the butterfly relaxation for $\theta_r(z) = \frac{z}{z+r}$ with from the left to the right : $t = 2r$ and $t = r^{3/2}$.

$$(a) \{i \in \mathcal{I}_{GH}(x; \hat{t}) \mid F_{1i}(x; \hat{t}) = 0, F_{2i}(x; \hat{t}) < 0\} = \{i \in \mathcal{I}_{GH}(x; \hat{t}) \mid F_{1i}(x; \hat{t}) < 0, F_{2i}(x; \hat{t}) = 0\} = \emptyset;$$

$$(b) i \in \mathcal{I}_{GH}(x; \hat{t}) \implies G_i(x) \geq 0, H_i(x) \geq 0.$$

Proof. Case (a) is direct considering that $\Phi_i^B(G(x), H(x); t, r) \neq 0$ for $F_{1i}(x; \hat{t}) + F_{2i}(x; \hat{t}) < 0$.

By symmetry of the relaxation it is sufficient to assume that $F_{1i}(x; \hat{t}) = H_i(x) - t\theta_r(G_i(x)) = 0$ for some $i = 1, \dots, q$. Then, by definition of $F_{2i}(x; \hat{t})$ it holds that

$$F_{2i}(x; \hat{t}) = G_i(x) - t\theta_r(H_i(x)) = G_i(x) - t\theta_r(t\theta_r(G_i(x))),$$

so $G_i(x) \geq 0$ since in the other case by definition of the function θ it would follow that

$$F_{2i}(x; \hat{t}) = G_i(x)(1 - (\theta'(0)t/r)^2),$$

which would be negative if $G_i(x) < 0$. Finally, $G_i(x) \geq 0$ implies that $H_i(x) \geq 0$ since $F_{1i}(x; \hat{t}) = 0$. \square

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

Lemma 6.1.2. $\mathcal{X}_{t, \bar{t}}^B$ satisfy the following properties:

1. $\mathcal{X}_{0,0}^B = \mathcal{Z}$;
2. $\mathcal{X}_{t_a, \bar{t}_b}^B \subset \mathcal{X}_{t_b, \bar{t}_b}^B$ for all $0 < \frac{t_a}{r_a} < \frac{t_b}{r_b}$ and $0 < \bar{t}_a < \bar{t}_b$;
3. $\cap_{t, \bar{t} \geq 0} \mathcal{X}_{t, \bar{t}}^B = \mathcal{Z}$.

If the feasible set of the (MPCC) 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 $\mathcal{X}_{t,\bar{t}}^B$ is empty, then it implies that \mathcal{Z} is empty. Finally, if a local minimum of $R_{t,\bar{t}}^B$ already belongs to \mathcal{Z} , then it is a local minimum of the (MPCC).

We focus in the sequel on the properties of these new relaxation schemes. The asymptotic behaviour of these methods are highlighted by Lemma 1.4.5 and Lemma 6.1.4. Then, we move to convergence properties of the methods considering a sequence of stationary points, which is proved to converge to an A-stationary point in Theorem 6.2.1 and to an M-stationary point, Corollary 6.2.1, with some relation between the sequences $\{t_k\}$ and $\{r_k\}$. In both results we discuss weak constraint qualifications necessary to ensure convergence of these methods. The main motivation to consider relaxation methods for (MPCC) is to solve a sequence of regular problems. Under classical assumptions the butterfly relaxed non-linear programs satisfy the Guignard CQ, Theorem 6.3.1. A specific kind of butterfly methods, where $t_k \theta'(0) = r_k$, has improved properties since they satisfy Abadie CQ, Theorem 6.3.2. It is more realistic to consider a sequence of ϵ -stationary points instead of classical stationary points. Also there is a price to pay here since convergence properties of most of the relaxation methods are damaged. This is discussed in Theorem 6.4.1 for the formulation with $t_k = o(r_k)$.

Before moving to our main results regarding convergence and regularity properties of the butterfly relaxation, we provide some useful results. Direct computation gives the gradient of $\Phi^B(G(x), H(x); \hat{t})$ in the following lemma.

Lemma 6.1.3. *For all $i \in \{1, \dots, q\}$, the gradient of $\Phi_i^B(G(x), H(x); \hat{t})$ w.r.t. x for the relaxation (6.3) is given by*

$$\nabla_x \Phi_i^B(G(x), H(x); \hat{t}) = \begin{cases} (F_{1i}(x; \hat{t}) - t\theta'_r(G_i(x))F_{2i}(x; \hat{t})) \nabla G_i(x) \\ + (F_{2i}(x; \hat{t}) - t\theta'_r(H_i(x))F_{1i}(x; \hat{t})) \nabla H_i(x), & \text{if } F_{1i}(x; \hat{t}) \geq -F_{2i}(x; \hat{t}), \\ (t\theta'_r(G_i(x))F_{1i}(x; \hat{t}) - F_{2i}(x; \hat{t})) \nabla G_i(x) \\ + (t\theta'_r(H_i(x))F_{2i}(x; \hat{t}) - F_{1i}(x; \hat{t})) \nabla H_i(x), & \text{if } F_{1i}(x; \hat{t}) < -F_{2i}(x; \hat{t}). \end{cases}$$

The following lemma is a direct application of Lemma 1.4.5 on the convergence of the butterfly relaxation when t and r go to zero.

Lemma 6.1.4. *Assume that (MPCC) has a non-empty feasible set, i.e. $\mathcal{Z} \neq \emptyset$. Given two sequences $\{r_k\}$ and $\{t_k\}$, which converge to 0 as k goes to infinity and $\forall k \in \mathbb{N}$, $(r_k, t_k) \in \mathbb{R}_{++}^2$. Let $\{x^k\}$ be a sequence of points such that $\lim_{k \rightarrow \infty} x^k = x^*$ and satisfying for all $i \in \{1, \dots, q\}$ and for all $k \in \mathbb{N}$*

$$G_i(x^k) \geq -\bar{t}, \quad H_i(x^k) \geq -\bar{t}, \quad \Phi_i^B(G(x^k), H(x^k); \hat{t}_k) \leq 0.$$

Then, x^ is a feasible point for (MPCC) as long as $g(x^*) \leq 0$ and $h(x^*) = 0$.*

We conclude this introductory section by an example that shows that the butterfly relaxation may improve relaxations from [111] and [112]. Indeed, it illustrates an example where there are no sequences of stationary point that converge to some undesirable point.

Example 6.1.1.

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

In this example, there are two stationary points: an S -stationary point $(1, 0)$ that is the global minimum and an M -stationary point $(0, 0)$, which is not a local minimum. Unlike the relaxations (KDB) and (KS) where for $t_k = \frac{1}{k}$ a sequence $x^k = (t_k \ 2t_k)^T$, with $\lambda^{\Phi, k} = k$, may converge to $(0, 0)$ as k goes to infinity, there is no sequences of stationary point that converges to this undesirable point with the butterfly relaxation.

6.2 Convergence

In this section, we focus on the convergence properties of the butterfly relaxation method and the constraint qualifications guaranteeing convergence of the sequence of stationary points generated by the method. Our aim is to compute an M -stationary point or at least to provide a certificate if we converge to an undesirable point.

We prove in Theorem 6.2.1 that the butterfly relaxation converges to an A -stationary point and provide a certificate independent of the multipliers in the case it converges to undesirable points. This result is improved to a convergence to M -stationary points for some choices on the parameters t and r in Corollary 6.2.1.

Another main concern in the literature is to find the weakest constraint qualification, which ensures convergence of the method. This has been extensively studied in the thesis [184] and related papers mentioned herein, where they prove convergence of most of the existing relaxation methods in the literature under a hypothesis close to MPCC-CRCQ. More recently in [165], Ramos proves convergence of the relaxation from [111] and [112] under MPCC-CCP.

Convergence of the butterfly relaxation under MPCC-CRCQ is proved in Proposition 6.2.1. An improved result for some choices of the parameter t and r is given in Proposition 6.2.2 that uses MPCC-CRSC. Example 6.2.3 shows that our methods may not converge under MPCC-CCP, since it requires boundedness of some multipliers.

Theorem 6.2.1. *Given two sequences $\{\hat{t}_k\}$ and $\{t_k\}$ of positive parameters satisfying (6.2) and decreasing to zero as k grows to infinity. Let $\{x^k, \nu^k\}$ be a sequence of points from $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^{3q}$ that are stationary points of $R_{\hat{t}, \bar{t}}^B$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$. Assume that the sequence*

$$\{\nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\}, \quad (6.4)$$

is bounded, where for all $i \in \{1, \dots, q\}$

$$\begin{aligned} \eta_i^{G,k} &:= \nu_i^{G,k} - \nu_i^{\Phi,k} \left(F_{1i}(x^k; \hat{t}_k) - t_k \theta'_{r_k} (G_i(x^k)) F_{2i}(x^k; \hat{t}_k) \right), \\ \eta_i^{H,k} &:= \nu_i^{H,k} - \nu_i^{\Phi,k} \left(F_{2i}(x^k; \hat{t}_k) - t_k \theta'_{r_k} (H_i(x^k)) F_{1i}(x^k; \hat{t}_k) \right). \end{aligned}$$

Then, one of the three following cases holds:

(i) x^* is an S -stationary point, if for all $i \in \mathcal{I}^{00}(x^*)$ one of the following holds:

- (i.a) there exists $\bar{k} \in \mathbb{N}$ such that $G_i(x^k) = H_i(x^k) = -\bar{t}_k$, $\forall k \geq \bar{k}$;
- (i.b) the sequence of multiplier $\{\nu_i^{\Phi,k}\}$ is bounded;
- (i.c) $\{\nu_i^{\Phi,k}\}$ is unbounded with $\lim_{k \rightarrow \infty} \nu_i^{\Phi,k} F_{2i}(x^k; \hat{t}_k) = 0$ for $i \in \mathcal{I}_{GH}^{0+}(x^k; \hat{t}_k)$ and $\lim_{k \rightarrow \infty} \nu_i^{\Phi,k} F_{1i}(x^k; \hat{t}_k) = 0$ for $i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k)$.

(ii) x^* is an M -stationary point, if for all $i \in \mathcal{I}^{00}(x^*)$ that does not satisfy conditions

(i) the sequence of multiplier $\{\nu_i^{\Phi,k}\}$ is unbounded and either for $i \in \mathcal{I}_{GH}^{0+}(x^k; \hat{t}_k)$, $\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(G_i(x^k)) = 0$ either for $i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k)$, $\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(H_i(x^k)) = 0$.

(iii) x^* is an A -stationary point, if the sequence of multipliers diverges or for all $i \in \mathcal{I}^{00}(x^*)$ that does not satisfy conditions (ii) the sequence of multiplier $\{\nu_i^{\Phi,k}\}$ is unbounded and either for all $i \in \mathcal{I}_{GH}^{0+}(x^k; \hat{t}_k)$, $\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(G_i(x^k)) > 0$ either for all $i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k)$, $\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(H_i(x^k)) > 0$.

The boundedness assumption on the sequence (6.4) is a classical assumption and is guaranteed under some constraint qualification as shown in the next Proposition 6.2.1.

Proof. First, we identify the expressions of the multipliers of the complementarity constraint in Definition 4.1.1 through the stationary points of $R_{\hat{t}, \hat{t}}^B$. Let $\{x^k, \nu^{g,k}, \nu^{h,k}, \nu^{G,k}, \nu^{H,k}, \nu^{\Phi,k}\}$ be a sequence of KKT points of $R_{\hat{t}, \hat{t}}^B$ for all $k \in \mathbb{N}$, that by definition satisfies

$$\begin{aligned} 0 = & \nabla f(x^k) + \sum_{i=1}^p \nu_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \nu_i^{h,k} \nabla h_i(x^k) - \sum_{i=1}^q \nu_i^{G,k} \nabla G_i(x^k) \\ & - \sum_{i=1}^q \nu_i^{H,k} \nabla H_i(x^k) + \sum_{i=1}^q \nu_i^{\Phi,k} \nabla \Phi_i^B(G(x), H(x); \hat{t}), \end{aligned}$$

with

$$\begin{aligned} \nu_i^{g,k} &= 0, \forall i \notin \mathcal{I}_g(x^k) \text{ and } \nu_i^{g,k} \geq 0, \forall i \in \mathcal{I}_g(x^k), \\ \nu_i^{G,k} &= 0, \forall i \notin \mathcal{I}_G(x^k) \text{ and } \nu_i^{G,k} \geq 0, \forall i \in \mathcal{I}_G(x^k), \\ \nu_i^{H,k} &= 0, \forall i \notin \mathcal{I}_H(x^k) \text{ and } \nu_i^{H,k} \geq 0, \forall i \in \mathcal{I}_H(x^k), \\ \nu_i^{\Phi,k} &= 0, \forall i \notin \mathcal{I}_{GH}(x^k; \hat{t}_k) \text{ and } \nu_i^{\Phi,k} \geq 0, \forall i \in \mathcal{I}_{GH}(x^k; \hat{t}_k). \end{aligned}$$

The representation of $\nabla \Phi^B$ immediately gives $\nabla \Phi_i^B(G(x^k), H(x^k); \hat{t}_k) = 0$, $\forall i \in \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)$ for all $k \in \mathbb{N}$. Thus, we can rewrite the equation above as

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

where

$$\eta_i^{G,k} = \begin{cases} \nu_i^{G,k}, & \text{if } i \in \mathcal{I}_G(x^k; \hat{t}_k) \\ -\nu_i^{\Phi,k} t_k \theta'_{r_k}(G_i(x^k)) F_{2i}(x^k; \hat{t}_k), & \text{if } i \in \mathcal{I}_{GH}^{0+}(x^k; \hat{t}_k) \\ \nu_i^{\Phi,k} F_{1i}(x^k; \hat{t}_k), & \text{if } i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k) \\ 0, & \text{otherwise,} \end{cases}$$

$$\eta_i^{H,k} = \begin{cases} \nu_i^{H,k}, & \text{if } i \in \mathcal{I}_H(x^k; \hat{t}_k) \\ -\nu_i^{\Phi,k} t_k \theta'_{r_k}(H_i(x^k)) F_{1i}(x^k; \hat{t}_k), & \text{if } i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k) \\ \nu_i^{\Phi,k} F_{2i}(x^k; \hat{t}_k), & \text{if } i \in \mathcal{I}_{GH}^{0+}(x^k; \hat{t}_k) \\ 0, & \text{otherwise.} \end{cases}$$

Noticing that whenever $i \in \{i = 1, \dots, q \mid F_{1i}(x^k; \hat{t}_k) = 0\}$ implies that $i \in \mathcal{I}_{GH}^{0+}(x^k; \hat{t}_k) \cup \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)$ or symmetrically $i \in \{i \mid F_{2i}(x^k; \hat{t}_k) = 0\}$ implies that $i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k) \cup \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)$ by concavity and $t_k \theta'(0) \leq r_k$ for all $k \in \mathbb{N}$.

We observe that the sequence $\{\nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\}$ is bounded, then it converges, up to a subsequence, to some limit denoted by $\{\nu^{g,*}, \nu^{h,*}, \eta^{G,*}, \eta^{H,*}\}$.

These multipliers are well-defined since

$$\begin{aligned} \mathcal{I}_G(x^k; \hat{t}_k) \cap \mathcal{I}_{GH}(x^k; \hat{t}_k) \cap (\{1, \dots, q\} \setminus \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)) &= \emptyset, \\ \mathcal{I}_H(x^k; \hat{t}_k) \cap \mathcal{I}_{GH}(x^k; \hat{t}_k) \cap (\{1, \dots, q\} \setminus \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)) &= \emptyset, \end{aligned}$$

and for k sufficiently large

$$\begin{aligned} \text{supp}(\nu^{G,k}) &\subseteq \mathcal{I}_G(x^k; \hat{t}_k), \\ \text{supp}(\nu^{H,k}) &\subseteq \mathcal{I}_H(x^k; \hat{t}_k), \\ \text{supp}(\nu^{G,k}) &\subseteq \mathcal{I}_{GH}(x^k; \hat{t}_k), \\ \text{supp}(\eta^{G,k}) &\subseteq \mathcal{I}_{GH}(x^k; \hat{t}_k) \cap (\{1, \dots, q\} \setminus \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)), \\ \text{supp}(\eta^{H,k}) &\subseteq \mathcal{I}_{GH}(x^k; \hat{t}_k) \cap (\{1, \dots, q\} \setminus \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)). \end{aligned}$$

Moreover, for k sufficiently large it holds

$$\begin{aligned} \text{supp}(\nu^{G,*}) &\subseteq \text{supp}(\nu^{G,k}), \\ \text{supp}(\nu^{H,*}) &\subseteq \text{supp}(\nu^{H,k}), \\ \text{supp}(\eta^{G,*}) &\subseteq \text{supp}(\eta^{G,k}), \\ \text{supp}(\eta^{H,*}) &\subseteq \text{supp}(\eta^{H,k}). \end{aligned}$$

The proof that shows convergence of the sequence and weak-stationary of x^* will be given in Section 6.4 by Lemma 6.4.1 on page 115 for $\epsilon_k = 0$.

Let us now verify that x^* is an A-stationary point by computing the multipliers for indices $i \in \mathcal{I}^{00}(x^*)$. Consider the various possible cases, where we denote

$$\nu_0^G := \{i = 1, \dots, q \mid \nu_i^{G,*} = \eta_i^{G,*}\} \text{ and } \nu_0^H := \{i = 1, \dots, q \mid \nu_i^{H,*} = \eta_i^{H,*}\} :$$

1. If $i \in \text{supp}(\nu^{G,*}) \cap \text{supp}(\nu^{H,*})$, then for k sufficiently large $i \in \text{supp}(\nu^{G,k}) \cap \text{supp}(\nu^{H,k})$. One has $\nu_i^{G,k} \geq 0$, $\nu_i^{H,k} \geq 0$ and

$$G_i(x^k) = H_i(x^k) = -\bar{t}_k.$$

2. If $i \in \text{supp}(\nu^{G,*}) \cap \text{supp}(\eta^{H,*})$, then for k sufficiently large $i \in \text{supp}(\nu^{G,k}) \cap \text{supp}(\eta^{H,k})$. One has $\nu_i^{G,k} \geq 0$, $G_i(x^k) = -\bar{t}_k$ and necessarily $i \in \mathcal{I}_{GH}(x^k; \hat{t}_k)$, which is not possible.
3. The case $i \in \text{supp}(\eta^{G,*}) \cap \text{supp}(\nu^{H,*})$ is completely similar.
4. If $i \in \text{supp}(\nu^{G,*}) \cap \nu_0^H$, then $\eta_i^{G,*} \geq 0$ and $\eta_i^{H,*} = 0$.
5. If $i \in \nu_0^G \cap \text{supp}(\nu^{H,*})$, then $\eta_i^{G,*} \geq 0$ and $\eta_i^{H,*} = 0$.
6. If $i \in \nu_0^G \cap \nu_0^H$, then $\eta_i^{G,*} = \eta_i^{H,*} = 0$.
7. If $i \in \nu_0^G \cap \text{supp}(\eta^{H,*})$, then $i \in \nu_0^G \cap \text{supp}(\eta^{H,k})$. Since $\eta^{G,k} = 0$ and $\eta^{H,k}$ free, one has $\nu_i^{\Phi,k} \geq 0$ and then $i \in \mathcal{I}_{GH}(x^k; \hat{t}_k)$.

$$\eta_i^{G,k} = 0 \iff F_{1i}(x^k; \hat{t}_k) = t_k \theta'_{r_k}(G_i(x^k)) F_{2i}(x^k; \hat{t}_k) \text{ or } \nu_i^{\Phi,k} = 0.$$

Moreover $t_k \theta'_{r_k}(G_i(x^k)) > 0$, so either $\nu_i^{\Phi,k} = 0$ either $F_{1i}(x^k; \hat{t}_k) = F_{2i}(x^k; \hat{t}_k) = 0$. It follows that $\eta_i^{G,*} = \eta_i^{H,*} = 0$.

8. The case $i \in \text{supp}(\eta^{G,*}) \cap \nu_0^H$ is completely similar to the previous case and leads to $\eta_i^{G,*} = \eta_i^{H,*} = 0$.
9. If $i \in \text{supp}(\eta^{G,*}) \cap \text{supp}(\eta^{H,*})$, then $i \in \text{supp}(\eta^{G,k}) \cap \text{supp}(\eta^{H,k})$ for k sufficiently large and $i \in \mathcal{I}_{GH}(x^k; \hat{t}_k)$.
 - (a). $i \in \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)$ implies that $F_{1i}(x^k; \hat{t}_k) = F_{2i}(x^k; \hat{t}_k)$, therefore $G(x^k) = H(x^k) = 0$ and $\eta_i^{G,*} = \eta_i^{H,*} = 0$.
 - (b). If $i \in \mathcal{I}_{GH}^{0+}(x^k)$, then $F_{1i}(x^k; \hat{t}_k) = 0$

$$0 < H_i(x^k) = t_k \theta_{r_k}(G_i(x^k)) < \frac{t_k \theta'(0)}{r_k} G_i(x^k),$$

therefore $F_{2i}(x^k; \hat{t}_k) > 0$. Assume $\nu_i^{\Phi,k}$ is not bounded, then going through the limit there is a non-negative constant C such that

$$\lim_{k \rightarrow \infty} \nu_i^{\Phi,k} F_{2i}(x^k; \hat{t}_k) = C \geq 0,$$

and so $\eta_i^{H,*} = -C$. If $\nu_i^{\Phi,k}$ is bounded, it corresponds to the case $C = 0$. Furthermore either one has

$$\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(G_i(x^k)) \geq 0,$$

and so $\eta_i^{G,*} \geq 0$ and $\eta_i^{H,*} \leq 0$. Either one has

$$\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(G_i(x^k)) = 0,$$

and so $\eta_i^{G,*} = 0$ and $\eta_i^{H,*} < 0$.

(c). The case $i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k)$ is completely similar to the previous case.

Indices that correspond to the first eight cases and 9.a) are indices that satisfy S-stationary condition. Furthermore, the indices in cases 9.b) and 9.c) when the constant $C = 0$ correspond to (i.c) and also have the sign of S-stationary indices. M- and A-stationary indices may appear only in the case 9.b) when $C \neq 0$ and either $t_k \theta'_{r_k}(G_i(x^k)) = 0$ (ii) either $t_k \theta'_{r_k}(G_i(x^k)) > 0$ (iii) for $i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k)$ and symmetrically in case 9.c). \square

The following proposition proves convergence of the sequence of multipliers under MPCC-CRSC by a direct application of Theorem 4.4.2. Here, we focus on the sequence of multipliers $\{\nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\}$ defined in (6.4), where we assume that the gradients associated with the non-vanishing multipliers in this sequence are linearly independent. Following the discussion before Theorem 4.4.2, this can be done without loss of generality.

Proposition 6.2.1. *Given two sequences $\{\hat{t}_k\}$ and $\{\bar{t}_k\}$ of positive parameters satisfying (6.2) and decreasing to zero as k goes to infinity. Let $\{x^k, \nu^k\}$ be a sequence of points that are stationary points of $R_{\hat{t}, \bar{t}}^B$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$ such that MPCC-CRCQ holds at x^* . Then, the sequence (6.4) is bounded.*

Proof. In order to apply Theorem 4.4.2, we prove that *Assumption 1* for $\{x^k, \nu^k\}$ is verified here. Denote $\|\eta^k\|_\infty := \|\nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\|_\infty$.

Obviously, (4.4) is satisfied $\forall k \in \mathbb{N}$, since $\{x^k, \nu^k\}$ are stationary points of $R_{\hat{t}, \bar{t}}^B$.

Let us now verify condition (4.5). By definition of $\{\nu^{g,k}\}$ it holds that $\mathcal{I}_g(x^k) \subset \mathcal{I}_g(x^*)$ and so $\forall i \notin \mathcal{I}_g(x^*) \lim_{k \rightarrow \infty} \nu_i^{g,k} = 0$. Let $i \in \mathcal{I}^{+0}$, we verify that $\lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} = 0$. The case $\forall i \in \mathcal{I}^{0+} \lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} = 0$ will follow symmetrically.

We can already notice that $\nu_i^{G,k} = 0$, since by the stationary condition it holds that $\nu_i^{G,k}(G_i(x^k) + \bar{t}_k) = 0$ and $G_i(x^k) \rightarrow G_i(x^*) > 0$. Assume by contradiction that $\lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} \neq 0$. By definition of $\eta^{G,k}$ and since $\nu_i^{G,k} = 0$, this implies that $\nu^{\Phi,k} > 0$. As a consequence $i \in \mathcal{I}_{GH}$ and in particular $i \in \mathcal{I}_{GH}^{0+}$, since $i \in \mathcal{I}_{GH}^{+0}$ would be a contradiction with $G_i(x^k) \rightarrow G_i(x^*)$. Besides, it also holds that $\nu^{H,k} = 0$, since $\text{supp}(\nu^{\Phi,k}) \cap \text{supp}(\nu^{H,k}) = \emptyset$. These simplifications yields

$$\eta_i^{G,k} = \nu_i^{\Phi,k} t_k \theta'_{r_k}(G_i(x^k)) F_{2i}(x^k; \hat{t}_k) \text{ and } \eta_i^{H,k} = -\nu_i^{\Phi,k} F_{2i}(x^k; \hat{t}_k).$$

However, by hypothesis on the sequences $\{r_k\}$ and $\{t_k\}$, this gives that

$$0 \leq \lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} \leq \lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{|\eta_i^{H,k}|} \leq \lim_{k \rightarrow \infty} t_k \theta'_{r_k}(G_i(x^k)) = 0,$$

leading to a contradiction, so $\forall i \in \mathcal{I}^{+0} \lim_{k \rightarrow \infty} \frac{\eta_i^{G,k}}{\|\eta^k\|_\infty} = 0$.

In conclusion, $\{x^k, \nu^{g,k}, \nu^{h,k}, \eta^{G,k}, \eta^{H,k}\}$ satisfies the hypothesis of Theorem 4.4.2. The result follows by its straightforward application. \square

In [165], the author proves similar convergence results for the relaxations [111] and [112] using the very weak constraint qualification MPCC-CCP, obtained by deriving the sequential optimality conditions from [14] in non-linear programming to (MPCC). However, this constraint qualification does not ensure boundedness of the sequence of multipliers (6.4), which is necessary for our previous theorem.

The following example shows that the result of Proposition 6.2.1 is sharp since convergence cannot be ensured assuming only that MPCC-MFCQ holds at the limit point. The convergence cannot be ensured under MPCC-CRSC, which is weaker than MPCC-MFCQ.

Example 6.2.1. *Consider the following two dimensional example*

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

MPCC-MFCQ holds at $(0, 0)^T$. However, MPCC-CRCQ obviously fails to hold at this point. In this case, the point $(0, 0)^T$ is even not a weak-stationary point.

In this case, there exists a sequence of stationary points of the relaxation that converges to the origin. Given a sequence $\{x^k\}$, with $\{1\} \in \mathcal{I}_{GH}(x^k; t_k)$, such that $x^k \rightarrow (0, 0)^T$ then $\nu^{G,k} = \nu^{H,k} = 0$ and we can choose $\nu^{\Phi,k}$ that satisfies

$$\eta^{G,k} = -\eta^{H,k} = \frac{1}{2x_2^k}.$$

The sequence of stationary points $\{x^k\}, \{\lambda^k\}$ converges to an undesirable point.

The main reason for this behaviour is that MPCC-GMFCQ does not give strong enough conditions in the neighbourhood of a non-M-stationary point.

Theorem 6.2.1 describes the various sequences that can arise from these relaxation methods in a constructive way. Indeed, it shows that in general the butterfly relaxation may converge to some undesirable A-stationary points. This theorem also provides a certificate independent of the computed multiplier that detects during the iterations, whether the method converges to this kind of undesirable point. According to condition (iii) from Theorem 6.2.1 if we detect for k sufficiently large that there exists an index $i \in \mathcal{I}^{00}(x^*)$ such that

$$\text{either } i \in \mathcal{I}_{GH}^{0+}(x^k; \hat{t}_k), \lim_{k \rightarrow \infty} t_k \theta'_{r_k}(G_i(x^k)) > 0 \text{ or } i \in \mathcal{I}_{GH}^{+0}(x^k; \hat{t}_k), \lim_{k \rightarrow \infty} t_k \theta'_{r_k}(H_i(x^k)) > 0$$

then x^k converges to an A-stationary point and not more. This is a priori not a trivial task, since the set of multipliers at those points is not bounded and an M-stationary point may be defined for only a subset of multipliers among the unbounded set of multipliers at this point. The following examples illustrate this phenomenon.

Example 6.2.2.

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

There exist two stationary points : an A-stationary point $(0, 0)^T$ and the global minimum $(\frac{1}{4}, 0)^T$. Similar computation gives two stationary points for the Butterfly relaxation (6.3) with $t_k = r_k$, $\forall k \in \mathbb{N} : (1/2, 0)^T$ and a point $(x_1, x_2)^T$ such that

$$x_1 = t\theta_r(x_2), \quad 0 = \frac{1}{2} + t\theta_r(x_2) + 2x_2 - t\theta'_r(x_2)(1 - 2t\theta_r(x_2) - x_2).$$

The following corollary of Theorem 6.2.1 shows that for some choice of parameters we can get rid of the undesirable A-stationary points. It is an essential result, since it shows that a subfamily of the butterfly relaxation has the desired convergence property to converge to an M-stationary point.

Corollary 6.2.1. *Given two sequences $\{\hat{t}_k\}$ and $\{\bar{t}_k\}$ of positive parameters satisfying (6.2), $t_k = o(r_k)$, and decreasing to zero as k goes to infinity. Let $\{x^k, \nu^k\}$ be a sequence of points that are stationary points of $R_{\hat{t}, \bar{t}}^B$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$. Assume that the sequence (6.4) is bounded. Then, x^* is an M-stationary point.*

Proof. The proof that x^* is an A-stationary follows the same path than in the proof of Theorem 6.2.1. Thus, it remains to verify that for $i \in \mathcal{I}^{00}$ there is no indices such that $\eta_i^{G,*} \eta_i^{H,*} < 0$.

In the proof of the Theorem 6.2.1 we consider all the possible cases, and it follows the case $\eta_i^{G,*} \eta_i^{H,*} < 0$ may only occur in the case (9).(b) (and by symmetry (9).(c)). In particular, if the sequence $\{t_k\}$, $\{r_k\}$ and $\{x_k\}$ satisfy

$$\lim_{k \rightarrow \infty} t_k \theta'_{r_k}(G_i(x^k)) > 0.$$

However, by Lemma 1.4.5 this is a contradiction with $t_k = o(r_k)$. So, x^* is a M-stationary point. This concludes the proof. \square

For this relation between the parameters t and r , we can improve the result of Proposition 6.2.1 by a straightforward application of Theorem 4.4.3.

Proposition 6.2.2. *Given two sequences $\{\hat{t}_k\}$ and $\{\bar{t}_k\}$ of positive parameters satisfying (6.2), $t_k = o(r_k)$, and decreasing to zero as k goes to infinity. Let $\{x^k, \nu^k\}$ be a sequence of points that are stationary points of $R_{\hat{t}, \bar{t}}^B$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$ such that MPCC-CRSC holds at x^* . Then, the sequence (6.4) is bounded.*

In conclusion, a sequence of stationary points of the butterfly relaxation with $t_k = o(r_k)$ that satisfies MPCC-CRSC at its limit point converges to an M-stationary point. The following example shows that this result is sharp, since it illustrates an example where MPCC-CRSC does not hold and the method converges to an undesirable weak-stationary point. This phenomenon only happens if the sequence of multipliers (6.4) is unbounded.

Example 6.2.3 (Continuation of Example 4.5.1).

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

The feasible set of this example is the set $\mathcal{Z} = \{(x_1, x_2)^T \in \mathbb{R}^2 \mid x_1 = 0\} \cup \{(x_1, x_2)^T \in \mathbb{R}^2 \mid x_1 = -x_2^2\}$. There is a unique stationary point $(0, 0)$, which is M -stationary with $(\lambda^G, \lambda^H = 0)$.

It is easy to verify that MPCC-CCP holds at this point. However, MPCC-CRSC fails to hold at any point $(0, a \in \mathbb{R})^T$ since the gradient of x_1^2 is non-zero for $x \neq 0$.

In this example the butterfly relaxation method may fail to converge to a weak-stationary point. Indeed, for $x^* = (0, a \neq 0)^T$ we can find a sequence x^k such that for t_k, r_k sufficiently small $F_2(x^k; \hat{t}_k) = 0$ and

$$x_1^k = t_k \theta'_{r_k}(x_1^k + a^2), \quad x_2^k = a, \quad \nu^{\Phi, k} F_1(x^k; \hat{t}_k) = \frac{1}{-t_k \theta'_{r_k}(x_1^k + a^2)}.$$

In this case, we have

$$\eta^{G, k} = \frac{1}{t_k \theta'_{r_k}(x_1^k + a^2)} \rightarrow \infty \text{ and } \eta^{H, k} = -1,$$

which is not a weak stationary point, since $\eta^{H, k} \neq 0$.

6.3 Existence of Lagrange multipliers of the relaxed sub-problems

In this part, we consider regularity properties of the relaxed non-linear programs. Indeed, in order to guarantee the existence of a sequence of stationary points the relaxed non-linear programs must satisfy some constraint qualifications in the neighbourhood of the limit point. The butterfly relaxation satisfies Guignard CQ as stated in Theorem 6.3.1, which is equivalent in terms of regularity to the relaxation (KS). The butterfly relaxation with $t\theta'(0) = r$ are more regular as they satisfy Abadie CQ, see Theorem 6.3.2. In our proofs we use the following results from [184] which allows the tangent cone of $\mathcal{X}_{t, \hat{t}}^B$ and its polar to be computed.

Lemma 6.3.1. [184, Lemma 8.10] For all $t > 0$ and all x feasible for $R_{\hat{t}, \bar{t}}^B$,

$$\mathcal{T}_{\mathcal{X}_{\hat{t}, \bar{t}}^B}(x) = \cup_{I \subseteq \mathcal{I}_{GH}^{00}(x; \hat{t})} \mathcal{T}_{\mathcal{X}(\hat{t}, I)}(x),$$

where $\mathcal{X}(\hat{t}, I)$ is the feasible set of the non-linear program $NLP_{t, I}(x)$ with $I \subseteq \mathcal{I}_{GH}^{00}(x; \hat{t})$

defined as

$$\begin{aligned}
& \min_{x \in \mathbb{R}^n} f(x) \\
& \text{s.t. } g(x) \leq 0, \quad h(x) = 0, \\
& \quad G(x) \geq -\bar{t}e, \quad H(x) \geq -\bar{t}e, \\
& \quad \Phi_i^B(G(x), H(x); \hat{t}) \leq 0, \quad i \notin \mathcal{I}_{GH}^{00}(x; \hat{t}), \\
& \quad F_{1i}(x; \hat{t}) \leq 0, \quad F_{2i}(x; \hat{t}) \geq 0, \quad i \in I, \\
& \quad F_{1i}(x; \hat{t}) \geq 0, \quad F_{2i}(x; \hat{t}) \leq 0, \quad i \in I^c,
\end{aligned} \tag{NLP}_{t,I}(x)$$

where $I \cup I^c = \mathcal{I}_{GH}^{00}(x; \hat{t})$ and $I \cap I^c = \emptyset$.

We also need the following lemma that links the gradients of G and H with the gradients of $F_1(x; \hat{t})$ and $F_2(x; \hat{t})$.

Lemma 6.3.2. *Let $I \in \mathcal{P}(\mathcal{I}_{GH}^{00}(x; \hat{t}))$. Assume that the gradients*

$$\begin{aligned}
& \{\nabla g_i(x) \ (i \in \mathcal{I}_g(x)), \ \nabla h_i(x) \ (i = 1, \dots, m), \\
& \nabla G_i(x) \ (i \in \mathcal{I}_{GH}^{00}(x; \hat{t}) \cup \mathcal{I}_{GH}^{+0}(x; \hat{t})), \ \nabla H_i(x^*) \ (i \in \mathcal{I}_{GH}^{00}(x; \hat{t}) \cup \mathcal{I}_{GH}^{+0}(x; \hat{t}))\}
\end{aligned}$$

are linearly independent. Then, MFCQ holds at x for $(\text{NLP}_{t,I}(x))$.

Proof. We show that the gradients of the constraints of $(\text{NLP}_{t,I}(x))$ are positively linearly independent. For this purpose, we prove that the trivial solution is the only solution to the equation

$$\begin{aligned}
0 = & \sum_{i \in \mathcal{I}_g(x)} \nu_i^g \nabla g_i(x) + \sum_{i=1}^m \nu_i^h \nabla h_i(x) + \sum_{i \in \mathcal{I}_G(x; \hat{t})} \nu_i^G \nabla G_i(x) + \sum_{i \in \mathcal{I}_H(x; \hat{t})} \nu_i^H \nabla H_i(x) \\
& + \sum_{i \in \mathcal{I}_{GH}^{+0}(x; \hat{t}) \cup \mathcal{I}_{GH}^{0+}(x; \hat{t})} \nu_i^\Phi \nabla \Phi_i^B(G(x), H(x); \hat{t}) + \sum_{i \in I} \nu_i^{F_1} \nabla F_{1i}(x; \hat{t}) - \sum_{i \in I} \nu_i^{F_2} \nabla F_{2i}(x; \hat{t}) \\
& - \sum_{i \in I^c} \mu_i^{F_1} \nabla F_{1i}(x; \hat{t}) + \sum_{i \in I^c} \mu_i^{F_2} \nabla F_{2i}(x; \hat{t}),
\end{aligned}$$

with $\nu^g, \nu^G, \nu^H, \nu^\Phi, \nu^{F_1}, \nu^{F_2}, \mu^{F_1}, \mu^{F_2} \geq 0$. By definition of F_1 and F_2 it holds that

$$\begin{aligned}
\nabla_x F_{1i}(x; \hat{t}) &= \nabla H_i(x) - t\theta'_r(G_i(x)) \nabla G_i(x), \\
\nabla_x F_{2i}(x; \hat{t}) &= \nabla G_i(x) - t\theta'_r(H_i(x)) \nabla H_i(x).
\end{aligned}$$

The gradient of Φ^B is given by Lemma 6.1.3. We now replace those gradients in the equation

above

$$\begin{aligned}
0 = & \sum_{i \in \mathcal{I}_g(x)} \nu_i^g \nabla g_i(x) + \sum_{i=1}^m \nu_i^h \nabla h_i(x) + \sum_{i=1}^q \nabla G_i(x) (\nu_i^G + \nu_i^\Phi (F_{1i}(x; \hat{t}) - F_{2i}(x; \hat{t}) t \theta'_r(G_i(x)))) \\
& + \sum_{i=1}^q \nabla G_i(x) (-\nu_i^{F_1} t \theta'_r(G_i(x)) - \nu_i^{F_2} + \mu_i^{F_1} t \theta'_r(G_i(x)) + \mu_i^{F_2}) \\
& + \sum_{i=1}^q \nabla H_i(x) (\nu_i^H + \nu_i^\Phi (F_{2i}(x; \hat{t}) - F_{1i}(x; \hat{t}) t \theta'_r(H_i(x)))) , \\
& + \sum_{i=1}^q \nabla H_i(x) (-\nu_i^{F_2} t \theta'_r(H_i(x)) - \nu_i^{F_1} + \mu_i^{F_2} t \theta'_r(H_i(x)) + \mu_i^{F_1}) ,
\end{aligned}$$

with $\text{supp}(\nu^g) \subset \mathcal{I}_g(x)$, $\text{supp}(\nu^G) \subset \mathcal{I}_G(x; \hat{t})$, $\text{supp}(\nu^H) \subset \mathcal{I}_H(x; \hat{t})$, $\text{supp}(\nu^\Phi) \subset \mathcal{I}_{GH}^{+0}(x; \hat{t}) \cup \mathcal{I}_{GH}^{0+}(x; \hat{t})$, $\text{supp}(\nu^{F_1}) \subset I$, $\text{supp}(\nu^{F_2}) \subset I$ and $\text{supp}(\mu^{F_1}) \subset I^c$, $\text{supp}(\mu^{F_2}) \subset I^c$ where $I \cup I^c = \mathcal{I}_{GH}^{00}(x; \hat{t})$ and $I \cap I^c = \emptyset$. Using the assumption of linear independence of the gradients gives that the solution of the equation above satisfy the following system of equations

$$\begin{aligned}
& \nu^g = 0, \nu^h = 0, \nu^G = 0, \nu^H = 0, \\
& -\nu_i^\Phi F_{2i}(x; \hat{t}) t \theta'_r(G_i(x)) = 0 \text{ and } \nu_i^\Phi F_{2i}(x; \hat{t}) = 0 \quad \forall i \in \mathcal{I}_{GH}^{0+}(x; \hat{t}), \\
& \nu_i^\Phi F_{1i}(x; \hat{t}) = 0 \text{ and } -\nu_i^\Phi F_{1i}(x; \hat{t}) t \theta'_r(H_i(x)) = 0 \quad \forall i \in \mathcal{I}_{GH}^{+0}(x; \hat{t}), \\
& -\nu_i^{F_1} t \theta'_r(G_i(x)) - \nu_i^{F_2} = 0 \text{ and } -\nu_i^{F_2} t \theta'_r(H_i(x)) - \nu_i^{F_1} = 0 \quad \forall i \in I, \\
& \mu_i^{F_1} t \theta'_r(G_i(x)) + \mu_i^{F_2} = 0 \text{ and } \mu_i^{F_2} t \theta'_r(H_i(x)) + \mu_i^{F_1} = 0 \quad \forall i \in I^c.
\end{aligned}$$

From the second and third equations it follows that $\nu_i^\Phi = 0$. The second last equation for $i \in I$ gives

$$\nu_i^{F_2} = -\nu_i^{F_1} t \theta'_r(G_i(x)).$$

This implies that $\nu_i^{F_1} = \nu_i^{F_2} = 0$ by non-decreasing hypothesis on θ and non-negativity of $\nu_i^{F_1}$ and $\nu_i^{F_2}$. We proceed in the exact same way with the last equation to get $\mu_i^{F_1} = \mu_i^{F_2} = 0$. This completes the proof that the trivial solution is the only solution to our first equation and so the result follows. \square

Now we move to the theorem stating the constraint qualifications satisfied by the butterfly relaxation.

Theorem 6.3.1. *Let $x^* \in \mathcal{Z}$ such that MPCC-LICQ holds at x^* . Then, there exists $t^* > 0$ and a neighbourhood $U(x^*)$ of x^* such that for all $t \in (0, t^*]$, if $x \in U(x^*) \cap \mathcal{X}_{t, \bar{t}}^B$, then standard GCQ for $R_{t, \bar{t}}^B$ holds at x .*

Proof. Let $x \in \mathbb{R}^n$ be such that $x \in U(x^*) \cap \mathcal{X}_{t, \bar{t}}^B$.

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

The linearised cone of $R_{\hat{t},\hat{t}}^B$ is given by

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

using that $\nabla \Phi_i^B(G(x), H(x); \hat{t}) = 0$ for all $i \in \mathcal{I}_{GH}^{00}(x, \hat{t})$. Let us compute the polar of the tangent cone. Consider the non-linear program $(\text{NLP}_{t,I}(x))$ with $I \subset \mathcal{I}_{GH}^{00}(x; \hat{t})$.

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

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

Therefore, we can apply Lemma 6.3.2 that gives that MFCQ holds for $(\text{NLP}_{t,I}(x))$ at x . Furthermore, by Lemma 6.3.1 and since MFCQ in particular implies Abadie CQ it follows

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

By Proposition 1.1.6, passing to the polar yields

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

and by Lemma 1.1.2

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

Taking $v \in \mathcal{T}_{\mathcal{X}_{\hat{t},\hat{t}}^B}(x)^\circ$ implies $v \in \mathcal{L}_{\text{NLP}(\hat{t}, I)}(x)^\circ$ for all $I \subseteq \mathcal{I}_{GH}^{00}(x; \hat{t})$. If we fix such I , then

there exists some multipliers ν^h and $\nu^g, \nu^G, \nu^H, \nu^\Phi \geq 0$ so that

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

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

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

By construction of t^* and $U(x^*)$ the gradients involved here are linearly independent and so the multipliers in both previous equations must be equal. Thus, the multipliers ν_i^G and ν_i^H with indices i in $I \cup I^c$ vanish.

Therefore, $v \in \mathcal{L}_{\mathcal{X}_{\hat{t}, \hat{t}}^B}(x)^\circ$ and as v has been chosen arbitrarily then $\mathcal{T}_{\mathcal{X}_{\hat{t}, \hat{t}}^B}(x)^\circ \subseteq \mathcal{L}_{\mathcal{X}_{\hat{t}, \hat{t}}^B}(x)^\circ$.

The result follows since it always holds that $\mathcal{L}_{\mathcal{X}_{\hat{t}, \hat{t}}^B}(x)^\circ \subseteq \mathcal{T}_{\mathcal{X}_{\hat{t}, \hat{t}}^B}(x)^\circ$. \square

This result is sharp as shown by the following example, since Abadie CQ does not hold except for the special case where $t\theta'(0) = r$.

Example 6.3.1.

$$\min_{(x_1, x_2) \in \mathbb{R}^2} f(x) \text{ s.t. } 0 \leq x_1 \perp x_2 \geq 0.$$

At $x^* = (0, 0)^T$ it holds that $\nabla \Phi^B(G(x^*), H(x^*); \hat{t}) = (0, 0)^T$ and so $\mathcal{L}_{\mathcal{X}_{\hat{t}, \hat{t}}^B}(x^*) = \mathbb{R}^2$, which is obviously different from the tangent cone at x^* for $t < \theta'(0)r$.

However, for $t\theta'(0) = r$ the tangent cone is the whole space and thus Abadie CQ holds at x^* in this case.

Regarding the butterfly relaxation with $t\theta'(0) = r$ an improved regularity result holds.

Theorem 6.3.2. *Let $x^* \in \mathcal{Z}$ such that MPCC-LICQ holds at x^* . Then, there exists $t^* > 0$ and a neighbourhood $U(x^*)$ of x^* such that the following holds for all $t \in (0, t^*]$. If $x \in U(x^*) \cap \mathcal{X}_{t\theta'(0)=r}^B$, then standard ACQ for $R_{t\theta'(0)=r}^B$ holds in x .*

Proof. Let $x \in \mathbb{R}^n$ be such that $x \in U(x^*) \cap \mathcal{X}_{t\theta'(0)=r}^B$.

The proof follows exactly the ones of Theorem 6.3.1, until we compute the tangent cone of $\mathcal{X}_{t\theta'(0)=r}^B$ as

$$\mathcal{T}_{\mathcal{X}_{t\theta'(0)=r}^B}(x) = \cup_{I \subseteq \mathcal{I}_{GH}^{00}(x;\hat{t})} \mathcal{T}_{NLP(t,I)}(x) = \cup_{I \subseteq \mathcal{I}_{GH}^{00}(x;\hat{t})} \mathcal{L}_{NLP(t,I)}(x),$$

since in particular ACQ holds in x for $(NLP_{t,I}(x))$. A simple computation gives the linearised tangent cone for $(NLP_{t,I}(x))$

$$\begin{aligned} \mathcal{L}_{NLP_{t,I}}(x) = \{d \in \mathbb{R}^n \mid & \nabla g_i(x)^T d \leq 0, \ i \in \mathcal{I}_g(x), \ \nabla h_i(x)^T d = 0, \ i = 1, \dots, m \\ & \nabla G_i(x)^T d \geq 0, \ i \in \mathcal{I}_G(x), \ \nabla H_i(x)^T d \geq 0, \ i \in \mathcal{I}_H(x) \\ & \nabla F_{1i}(x;\hat{t})^T d \leq 0, \ \nabla F_{2i}(x;\hat{t}) \geq 0, \ i \in I \\ & \nabla F_{1i}(x;\hat{t})^T d \geq 0, \ \nabla F_{2i}(x;\hat{t}) \leq 0, \ i \in I^c \\ & \nabla \Phi_i^B(G(x), H(x);\hat{t})^T d \leq 0, \ i \in \mathcal{I}_{GH}^{0+}(x;\hat{t}) \cup \mathcal{I}_{GH}^{+0}(x;\hat{t})\}. \end{aligned}$$

Moreover one has for all $I \subseteq \mathcal{I}_{GH}^{00}(x;\hat{t})$,

$$\begin{aligned} \nabla F_{1i}(x;\hat{t}) &= \nabla H_i(x) - t\theta'_r(G_i(x))\nabla G_i(x) \\ &= \nabla H_i(x) - \nabla G_i(x) = -\nabla F_{2i}(x;\hat{t}), \end{aligned}$$

since $t\theta'(0) = r$ and for all $i \in \mathcal{I}_{GH}^{00}(x;\hat{t})$ it follows that $G_i(x) = H_i(x) = 0$.

It is to be noted that

$$\cup_{I \subseteq \mathcal{I}_{GH}^{00}(x;\hat{t})} \{d \in \mathbb{R}^n \mid \nabla H_i(x)^T d \leq \nabla G_i(x), \ i \in I ; \ \nabla H_i(x)^T d \geq \nabla G_i(x), \ i \in I^c\} = \mathbb{R}^n.$$

Therefore, it holds that

$$\begin{aligned} \cup_{I \subseteq \mathcal{I}_{GH}^{00}(x;\hat{t})} \mathcal{L}_{NLP(t,I)}(x) &= \{d \in \mathbb{R}^n \mid \nabla g_i(x)^T d \leq 0, \ i \in \mathcal{I}_g(x), \ \nabla h_i(x)^T d = 0, \ i = 1, \dots, m \\ & \nabla G_i(x)^T d \geq 0, \ i \in \mathcal{I}_G(x), \ \nabla H_i(x)^T d \geq 0, \ i \in \mathcal{I}_H(x) \\ & \nabla \Phi_i^B(G(x), H(x);\hat{t})^T d \leq 0, \ i \in \mathcal{I}_{GH}^{0+}(x;\hat{t}) \cup \mathcal{I}_{GH}^{+0}(x;\hat{t})\}. \end{aligned}$$

Thus, $\cup_{I \subseteq \mathcal{I}_{GH}^{00}(x;\hat{t})} \mathcal{L}_{NLP(t,I)}(x) = \mathcal{L}_{\mathcal{X}_{t\theta'(0)=r}^B}(x)$ and the result follows. \square

This result is sharp since very weak constraint qualification CCP does not hold in the example 6.3.1.

Example 6.3.2. We show that CRSC and CCP do not hold in $x^* = (0,0)^T$ for the relaxation $R_{t\theta'(0)=r}^B$.

It holds that $\nabla \Phi_{t\theta'(0)=r}^B(x^*) = (0,0)^T$, therefore $\nabla \Phi_{t\theta'(0)=r}^B(x^*) \in -\mathcal{L}(x^*)^\circ$. However, for any x in a small neighbourhood around x^* the gradient $\nabla \Phi_{t\theta'(0)=r}^B(x^*) \neq (0,0)^T$. So, the rank is not constant in this neighbourhood and CRSC does not hold in x^* .

In order to verify that CCP does not hold in x^* , we check that there exists an AKKT sequence that does not converge to a KKT point. Given $\{x^k\}$, $\{\nu^{G,k}\}$, $\{\nu^{H,k}\}$, $\{\nu^{\Phi,k}\}$ such that

$$\begin{aligned} \lim_{k \rightarrow \infty} \nabla f(x^k) + \nu^{\Phi,k} \nabla \Phi_{(t\theta'(0)=r)}^B(x^k) - \nu^{G,k} \nabla G(x^k) - \nu^{H,k} \nabla H(x^k) &= 0, \\ \lim_{k \rightarrow \infty} \min(\nu^{\Phi,k}, -\Phi_{(t\theta'(0)=r)}^B(x^k)) &= 0, \\ \lim_{k \rightarrow \infty} \min(\nu^{G,k}, G(x^k)) &= 0, \\ \lim_{k \rightarrow \infty} \min(\nu^{H,k}, H(x^k)) &= 0. \end{aligned}$$

We can find a sequence x^k such that $t_k \theta'_{r_k}(H(x^k)) \rightarrow 0$, $F_{1i}(x^k; \hat{t}_k) = 0$, $F_{2i}(x^k; \hat{t}_k) \geq 0$ and $\nu^k = \frac{1}{k F_{2i}(x^k; \hat{t}_k)}$, $\nu^{G,k} \rightarrow \frac{1}{k}$, $\nu^{H,k} \rightarrow 0$ so that CCP does not hold at x^* .

The following example shows that we cannot have a similar result with MPCC-MFCQ instead of MPCC-LICQ for Theorem 6.3.2.

Example 6.3.3. Consider the two-dimensional set

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 + x_2^2 \perp x_1 \geq 0\}.$$

MPCC-MFCQ holds at $x^* = (0, 0)^T$, since the gradients are linearly dependent but only with coefficients $\lambda^G = -\lambda^H$ and thus the gradients are positively linearly independent.

Now, taking a sequence of stationary point such that $x^k \rightarrow x^* = (0, 0)^T$ and

$$F_2(x^k; \hat{t}_k) = 0, -t_k \theta'_{r_k}(H(x^k)) \rightarrow -1.$$

Since $\nabla G(x^*) = \nabla H(x^*)$ it holds that $\nabla F_2(x^*; 0) = (0 \ 0)^T$ and so MFCQ does not hold for $(\text{NLP}_{t,I}(x))$.

Both Theorem 6.3.1 and Theorem 6.3.2 are slightly disappointing since MPCC-LICQ is quite a strong assumption. Fortunately, the following result guarantees that the difficulties are only located in indices i of x^* that belongs to $\mathcal{I}^{00}(x^*)$.

Theorem 6.3.3. Let $x^* \in \mathcal{Z}$ be such that MPCC-LICQ holds at x^* . Then, there exists $t^* > 0$ and a neighbourhood $U(x^*)$ of x^* such that the following holds for all $t \in (0, t^*]$. If $x \in U(x^*) \cap \mathcal{X}_{\hat{t}, \bar{t}}^B$ and $\mathcal{I}_{GH}^{00}(x; \hat{t}) = \emptyset$, then standard LICQ for $R_{\hat{t}, \bar{t}}^B$ holds in x .

Proof. Following the same path as Lemma 6.3.2, the gradient of the Lagrangian for $R_{\hat{t}, \bar{t}}^B$ at $x \in U(x^*)$ for \hat{t} sufficiently small gives

$$\begin{aligned} 0 &= \sum_{i \in \mathcal{I}_g(x)} \nu_i^g \nabla g_i(x) + \sum_{i=1}^m \nu_i^h \nabla h_i(x) + \sum_{i \in \mathcal{I}_G(x)} \nabla G_i(x) \nu_i^G + \sum_{i \in \mathcal{I}_H(x)} \nabla H_i(x) \nu_i^H \\ &\quad + \sum_{i \in \mathcal{I}_{GH}^{+0}(x; \hat{t})} \nabla G_i(x) (\nu_i^{\Phi} (F_{1i}(x; r, t)_i(x; \hat{t}) - F_{2i}(x; \hat{t}) t \theta'_r(G_i(x)))) \\ &\quad + \sum_{i \in \mathcal{I}_{GH}^{0+}(x; \hat{t})} \nabla H_i(x) (\nu_i^{\Phi} (F_{2i}(x; \hat{t}) - F_{1i}(x; \hat{t}) t \theta'_r(H_i(x)))) . \end{aligned}$$

Using the assumption of linear independence of the gradients and

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

gives that the solution of the equation above satisfy the following system of equations

$$\begin{aligned}\nu^g &= 0, \nu^h = 0, \nu^G = 0, \nu^H = 0, \\ -\nu_i^\Phi F_{2i}(x; \hat{t}) t \theta'_r(G_i(x)) &= 0 \text{ and } \nu_i^\Phi F_{2i}(x; \hat{t}) = 0 \quad \forall i \in \mathcal{I}_{GH}^{0+}(x; \hat{t}), \\ \nu_i^\Phi F_{1i}(x; \hat{t}) &= 0 \text{ and } -\nu_i^\Phi F_{1i}(x; \hat{t}) t \theta'_r(H_i(x)) = 0 \quad \forall i \in \mathcal{I}_{GH}^{+0}(x; \hat{t}).\end{aligned}$$

From the second and third equations it follows that $\nu_i^\Phi = 0$. So, the only solution is the trivial solution. Thus, the result follows. \square

6.4 Convergence of the epsilon-stationary points

Non-linear programming algorithms usually compute sequences of approximate stationary points or epsilon-stationary points. This approach that has become an active subject recently alter significantly the convergence analysis of relaxation methods as stated in [111, 113, 115] and [165]. Previous results in the literature in [115] provide convergence to C-stationary point for the relaxation (SS) and the one from Lin and Fukushima, [126], at the limit point and under the hypothesis on the sequence ϵ_k , respectively $\epsilon_k = O(t_k)$ and $\epsilon_k = o(t_k^2)$. Furthermore, they provide a counter-example with sequences converging to a weak-stationary point if this conditions does not hold. Although in [115], the authors prove that relaxation (SU), (KDB) and (KS) converge only to a weak stationary point and require more hypothesis on the sequences ϵ_k and x_k to improve to a C- or an M-stationary limit point.

In the following theorem we prove that the situation is similar with the new butterfly relaxation method. For this study, we need more than just the description of the boundary of the constraint, we consider the butterfly relaxation defined in equation (6.3).

Lemma 6.4.1. *Given $\{\hat{t}_k\}$ a sequence of parameters satisfying (6.2) and $\{\epsilon_k\}$ a sequence of non-negative parameters such that both sequences decrease to zero as $k \in \mathbb{N}$ goes to infinity and for k sufficiently large $\epsilon_k = o(t_k)$. Let $\{x^k, \nu^k\}$ be a sequence of epsilon-stationary points of $(R_{\hat{t}, \bar{t}}^B)$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$. Let $\{\eta^{G,k}\}, \{\eta^{H,k}\}$ be two sequences such that*

$$\begin{aligned}\eta_i^{G,k} &:= \begin{cases} \nu_i^{G,k} + \nu_i^{\Phi,k} (t_k \theta'_{r_k}(G_i(x^k)) F_{2i}(x^k; \hat{t}_k) - F_{1i}(x^k; \hat{t}_k)) & \text{if } F_{1i}(x^k; \hat{t}_k) \geq -F_{2i}(x^k; \hat{t}_k) \\ \nu_i^{G,k} + \nu_i^{\Phi,k} (F_{2i}(x^k; \hat{t}_k) - t_k \theta'_{r_k}(G_i(x^k)) F_{1i}(x^k; \hat{t}_k)) & \text{if } F_{1i}(x^k; \hat{t}_k) < -F_{2i}(x^k; \hat{t}_k) \end{cases} \\ \eta_i^{H,k} &:= \begin{cases} \nu_i^{H,k} + \nu_i^{\Phi,k} (t_k \theta'_{r_k}(H_i(x^k)) F_{1i}(x^k; \hat{t}_k) - F_{2i}(x^k; \hat{t}_k)) & \text{if } F_{1i}(x^k; \hat{t}_k) \geq -F_{2i}(x^k; \hat{t}_k) \\ \nu_i^{H,k} + \nu_i^{\Phi,k} (F_{1i}(x^k; \hat{t}_k) - t_k \theta'_{r_k}(H_i(x^k)) F_{2i}(x^k; \hat{t}_k)) & \text{if } F_{1i}(x^k; \hat{t}_k) < -F_{2i}(x^k; \hat{t}_k) \end{cases}\end{aligned} \tag{6.5}$$

for $i \in \{1, \dots, q\}$. Assume that the sequence of multipliers $\{\nu^{h,k}, \nu^{g,k}, \eta^{G,k}, \eta^{H,k}\}$ is bounded. Then, x^* is a weak-stationary point of (MPCC).

Proof. By definition, since x^k is an epsilon-stationary point for $R_{\hat{t}, \bar{t}}^B$ it holds for all $k \in \mathbb{N}$

$$\left\| \nabla f(x^k) + \sum_{i=1}^p \nu_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \nu_i^{h,k} \nabla h_i(x^k) - \sum_{i=1}^q \nu_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^q \nu_i^{H,k} \nabla H_i(x^k) + \sum_{i=1}^q \nu_i^{\Phi,k} \nabla \Phi_i^B(G(x^k), H(x^k); \hat{t}_k) \right\|_{\infty} \leq \epsilon_k,$$

with

$$\begin{aligned} |h_i(x^k)| &\leq \epsilon_k, \quad \forall i \in \{1, \dots, m\}, \\ g_i(x^k) &\leq \epsilon_k, \quad \nu_i^{g,k} \geq 0, \quad \left| \nu_i^{g,k} g_i(x^k) \right| \leq \epsilon_k \quad \forall i \in \{1, \dots, p\}, \\ G_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu_i^{G,k} \geq 0, \quad \left| \nu_i^{G,k} (G_i(x^k) + \bar{t}_k) \right| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\ H_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu_i^{H,k} \geq 0, \quad \left| \nu_i^{H,k} (H_i(x^k) + \bar{t}_k) \right| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\ \Phi_i^B(G(x^k), H(x^k); \hat{t}_k) &\leq \epsilon_k, \quad \nu_i^{\Phi,k} \geq 0, \quad \left| \nu_i^{\Phi,k} \Phi_i^B(G(x^k), H(x^k); \hat{t}_k) \right| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}. \end{aligned}$$

The representation of $\Phi_i^B(G(x^k), H(x^k); \hat{t}_k)$ immediately gives $\nabla \Phi_i^B(G(x^k), H(x^k); \hat{t}_k) = 0$, $\forall i \in \mathcal{I}_{GH}^{00}(x^k; \hat{t}_k)$ for all $k \in \mathbb{N}$. Thus, we can rewrite the equation above as

$$\left\| \nabla f(x^k) + \sum_{i=1}^p \nu_i^{g,k} \nabla g_i(x^k) + \sum_{i=1}^m \nu_i^{h,k} \nabla h_i(x^k) - \sum_{i=1}^q \eta_i^{G,k} \nabla G_i(x^k) - \sum_{i=1}^q \eta_i^{H,k} \nabla H_i(x^k) \right\|_{\infty} \leq \epsilon_k.$$

Besides, the sequence of multipliers $\{\nu^{h,k}, \nu^{g,k}, \eta^{G,k}, \eta^{H,k}\}$ is assumed bounded. Therefore, it follows that the sequence converges up to some subsequence to some limit point so that

$$\{\nu^{h,k}, \nu^{g,k}, \eta^{G,k}, \eta^{H,k}\} \rightarrow (\nu^h, \nu^g, \eta^G, \eta^H).$$

It is to be noted that for k sufficiently large it holds

$$\begin{aligned} \text{supp}(\nu^{g,k}) &\subset \text{supp}(\nu^g), \\ \text{supp}(\eta^{G,k}) &\subset \text{supp}(\eta^G), \\ \text{supp}(\eta^{H,k}) &\subset \text{supp}(\eta^H). \end{aligned}$$

We prove that $(x^*, \nu^h, \nu^g, \eta^G, \eta^H)$ is a weak-stationary point. Obviously, since $\epsilon_k \downarrow 0$ it follows that $x^* \in \mathcal{Z}$, $\nabla_x \mathcal{L}_{MPCC}^1(x^*, \nu^h, \nu^g, \eta^G, \eta^H) = 0$ by previous inequality and that $\nu_i^g = 0$ for

$i \notin \mathcal{I}_g(x^*)$. It remains to show that for indices $i \in \mathcal{I}^{+0}(x^*)$, $\eta_i^G = 0$. The opposite case for indices $i \in \mathcal{I}^{0+}(x^*)$ would follow in a completely similar way. So, let i be in $\mathcal{I}^{+0}(x^*)$.

By definition of ϵ_k -stationarity it holds for all k that

$$|\nu_i^{G,k}(G_i(x^k) + \bar{t}_k)| \leq \epsilon_k.$$

Therefore, $\nu_i^{G,k} \rightarrow_{k \rightarrow \infty} 0$ since $\epsilon_k \downarrow 0$ and $G_i(x^k) \rightarrow G_i(x^*) > 0$.

Now, there is two possible cases either $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) \geq 0$ either $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) < 0$. Consider the case $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) \geq 0$ and denote

$$\begin{aligned} \alpha^H(x_i^k; \hat{t}_k) &:= -t_k \theta'_{r_k}(G_i(x^k)) F_{2i}(x^k; \hat{t}_k) + F_{1i}(x^k; \hat{t}_k), \\ \alpha^G(x_i^k; \hat{t}_k) &:= -t_k \theta'_{r_k}(H_i(x^k)) F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k). \end{aligned}$$

It remains to prove that $\alpha^H(x_i^k; \hat{t}_k) \nu_i^{\Phi,k} \rightarrow_{k \rightarrow \infty} 0$. Assume by contradiction that

$$\lim_{k \rightarrow \infty} \alpha^H(x_i^k; \hat{t}_k) \nu_i^{\Phi,k} = C < 0, \quad (6.6)$$

which is necessary a finite value by boundedness hypothesis of the sequence of multipliers. Obviously the sequence $\{\nu^{\Phi,k}\}$ must be unbounded otherwise (6.6) does not hold.

Additionally, $\lim_{k \rightarrow \infty} F_{1i}(x^k; \hat{t}_k) \nu_i^{\Phi,k} = 0$ since $|\nu_i^{\Phi,k} \Phi_i^B(G(x^k), H(x^k); \hat{t}_k)| \leq \epsilon_k$. So, by (6.3) we have that $\lim_{k \rightarrow \infty} \alpha_i^G(x^k; t_k) = G_i(x^*) > 0$ and therefore $\lim_{k \rightarrow \infty} \alpha^G(x^k; t_k) \nu_i^{\Phi,k} = \lim_{k \rightarrow \infty} G_i(x^k) \nu_i^{\Phi,k} = \infty$. Boundedness assumption in the statement of the theorem implies that η_i^H is bounded and so

$$\lim_{k \rightarrow \infty} |\nu_i^{H,k} - \alpha^G(x^k; t_k) \nu_i^{\Phi,k}| < \infty.$$

The complementarity conditions on $H_i(x^k) \geq -\bar{t}_k$ give that necessary $H_i(x^k) \sim -\bar{t}_k$ otherwise $\nu_i^{H,k}$ would be unbounded. However, this leads to a contradiction with $\nu^{\Phi,k} \rightarrow \infty$, since $\nu_i^{\Phi,k} F_{1i}(x^k; \hat{t}_k) \rightarrow 0$ gives that $\nu_i^{\Phi,k} \bar{t}_k \leq \epsilon_k$ and we assume in the statement of the theorem that $\epsilon_k = o(\bar{t}_k)$. So in the case $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) \geq 0$ it holds that $\eta_i^{G,*} = 0$.

Let us consider the case $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) < 0$. As pointed out above it is true by (6.3) that $F_{1i}(x^k; \hat{t}_k) \rightarrow H_i(x^*)$ and $F_{2i}(x^k; \hat{t}_k) \rightarrow G_i(x^*)$. Therefore, for k sufficiently large this case never happen since we choose $i \in \mathcal{I}^{0+}(x^*)$. This concludes the proof that $\eta_i^{G,*} = 0$. The case $i \in \mathcal{I}^{0+}(x^*)$ is completely similar by symmetry and gives that $\eta_i^{H,*} = 0$ for $i \in \mathcal{I}^{+0}(x^*)$. So, x^* is a weak-stationary point. \square

In order to attain the goal of computing a M-stationary, additional assumptions are required as illustrated by the following result.

Lemma 6.4.2. *Given $\{\hat{t}_k\}$ a sequence of parameters satisfying (6.2) and $\{\epsilon_k\}$ a sequence of non-negative parameters such that both sequences decrease to zero as $k \in \mathbb{N}$ goes to infinity. Assume that $\epsilon_k = o(\max(G(x^k), H(x^k), \bar{t}_k))$ and $t_k = o(r_k)$. Let $\{x^k, \nu^k\}_k$ be a sequence of epsilon-stationary points of $(R_{\hat{t}, \bar{t}}^B)$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$. Let $\{\eta^{G,k}\}, \{\eta^{H,k}\}$ be two sequences defined in (7.3). Assume that the sequence of multipliers $\{\nu^{h,k}, \nu^{g,k}, \eta^{G,k}, \eta^{H,k}\}$ is bounded. Then, x^* is an M-stationary point of (MPCC).*

The notation $\epsilon_k = o(\max(|G(x^k)|, |H(x^k)|))$ means here that for all $i = 1, \dots, q$, $\epsilon_k = o(\max(|G_i(x^k)|, |H_i(x^k)|))$.

Proof. By Lemma 6.4.1, we already known that x^* is a W-stationary point.

We now consider indices $i \in \mathcal{I}^{00}$. Our aim here is to prove that x^* is a M-stationary point, i.e. whenever $\eta_i^{G,k} \eta_i^{H,k} \rightarrow \eta_i^{G,*} \eta_i^{H,*}$ it holds that either $\eta_i^{G,*} \eta_i^{H,*} = 0$ or $\eta_i^{G,*} > 0, \eta_i^{H,*} > 0$.

Without loss of generality suppose that $\max(|G_i(x^k)|, |H_i(x^k)|) = |G_i(x^k)| \neq 0$, and so $\lim_{k \rightarrow \infty} \frac{\epsilon_k}{|G_i(x^k)|} = 0$. If $G_i(x^k) = 0$, then it follows that $H_i(x^k) = 0$ and we are done. Let α be such that

$$\lim_{k \rightarrow \infty} \frac{|G_i(x^k)|}{|t_k \theta_{r_k}(H_i(x^k))|} = \alpha.$$

It should be noticed that $\alpha > 1$, otherwise $|G_i(x^k)| \sim |t_k \theta_{r_k}(H_i(x^k))|$, which is a contradiction with $|G_i(x^k)| \geq |H_i(x^k)|$ and $t_k = o(r_k)$.

If the sequence $\{\lambda^{\Phi,k}\}$ is bounded, then $\eta_i^{G,*} \geq 0, \eta_i^{H,*} \geq 0$ and we are done by non-negativity of $\lambda^{G,k}$ and $\lambda^{H,k}$. So, we focus on an unbounded sequence $\{\lambda^{\Phi,k}\}$.

We consider separately the two cases $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) \geq 0$ and $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) < 0$.

a) When $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) \geq 0$, we have

$$\left| \lambda_i^{\Phi,k} F_{1i}(x^k; \hat{t}_k) \frac{F_{2i}(x^k; \hat{t}_k)}{G_i(x^k)} \right| = \left| \lambda_i^{\Phi,k} F_{1i}(x^k; \hat{t}_k) \left(1 - \frac{t_k \theta_{r_k}(H_i(x^k))}{G_i(x^k)} \right) \right| \leq \frac{\epsilon_k}{|G_i(x^k)|},$$

so $\lambda_i^{\Phi,k} F_{1i}(x^k; \hat{t}_k) \rightarrow 0$ and $F_{2i}(x^k; \hat{t}_k) > 0$, since $\alpha > 1$.

By the complementarity condition $|\lambda^{G,k}(G_i(x^k) + \bar{t}_k)| \leq \epsilon_k$, we obtain

$$\left| \lambda_i^{G,k} \left(1 + \frac{\bar{t}_k}{G_i(x^k)} \right) \right| \leq \frac{\epsilon_k}{|G_i(x^k)|}.$$

If $\lim_{k \rightarrow \infty} \lambda_i^{G,k} = 0$, then by boundedness assumption $\lim_{k \rightarrow \infty} \eta_i^{G,k} = 0$ and we are done. So, we consider $\lim_{k \rightarrow \infty} \lambda_i^{G,k} \neq 0$, which implies that $\lim_{k \rightarrow \infty} G_i(x^k)/\bar{t}_k = -1$.

In a similar way, if $\lim_{k \rightarrow \infty} \lambda_i^{H,k} = 0$, then $\lim_{k \rightarrow \infty} \eta_i^{H,k} = 0$ by boundedness assumption. So, we consider $\lim_{k \rightarrow \infty} \lambda_i^{H,k} \neq 0$, which implies that $\lim_{k \rightarrow \infty} H_i(x^k)/\bar{t}_k = -1$.

Using that $G_i(x^k) < 0$ and $F_{2i}(x^k; \hat{t}_k) > 0$, we have

$$0 \geq \frac{G_i(x^k)}{\bar{t}_k} \geq \frac{t_k \theta_{r_k}(H_i(x^k))}{\bar{t}_k} \sim \frac{t_k \theta'(0) H_i(x^k)}{r_k \bar{t}_k} \sim \frac{t_k \theta'(0)}{r_k},$$

where the first equivalence comes from Taylor formula of order 1 of functions θ s at 0. So, $\lim_{k \rightarrow \infty} G_i(x^k)/\bar{t}_k = 0$. However, this contradicts $\lim_{k \rightarrow \infty} G_i(x^k)/\bar{t}_k = -1$, which completes the proof in this case.

b) When $F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) < 0$, since $\left| \lambda_i^{\Phi,k} \Phi_i^B(G(x^k), H(x^k); \hat{t}_k) \right| \leq \epsilon_k$ we have

$$\left| \lambda_i^{\Phi,k} F_{2i}(x^k; \hat{t}_k)^2 \right| \leq 2\epsilon_k \iff \left| \lambda_i^{\Phi,k} F_{2i}(x^k; \hat{t}_k) \left(1 - \frac{t_k \theta_{r_k}(H_i(x^k))}{G_i(x^k)} \right) \right| \leq \frac{2\epsilon_k}{|G_i(x^k)|}.$$

This implies that $\lim_{k \rightarrow \infty} \lambda_i^{\Phi,k} F_{2i}(x^k; \hat{t}_k) = 0$, by assumption on ϵ_k and $\alpha > 1$. Now, by definition of functions θ s and their first order Taylor formula at 0 we obtain

$$\begin{aligned} F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k) &= G_i(x^k) + H_i(x^k) - t_k (\theta_{r_k}(G_i(x^k)) + \theta_{r_k}(H_i(x^k))), \\ &\sim (G_i(x^k) + H_i(x^k)) \left(1 - \frac{t_k \theta'(0)}{r_k} \right), \\ &\leq 2|G_i(x^k)| \left(1 - \frac{t_k \theta'(0)}{r_k} \right), \end{aligned}$$

and so $\lim_{k \rightarrow \infty} \lambda_i^{\Phi,k} (F_{1i}(x^k; \hat{t}_k) + F_{2i}(x^k; \hat{t}_k)) = 0$. As a consequence, it holds that $\eta_i^{G,*} = \lim_{k \rightarrow \infty} \lambda^{G,k} \geq 0$, $\eta_i^{H,*} = \lim_{k \rightarrow \infty} \lambda^{H,k} \geq 0$.

So, x^* is a M-stationary point. □

Theorem 6.4.1. *Given two sequences $\{t_k\}$ and $\{r_k\}$ satisfying (6.2) such that $\forall k \in \mathbb{N}$, $(t_k, r_k) \in \mathbb{R}_+^2$, both sequence decrease to zero as k goes to infinity. Let $\{x^k, \nu^k\}$ be a sequence of points that are epsilon-stationary points of $R_{i,\bar{t}}^B$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$ such that MPCC-CRSC holds at x^* . Furthermore assume that $t_k = o(r_k) \forall k \in \mathbb{N}$ sufficiently large and that the sequence $\{\epsilon_k\}$ is such that $\epsilon_k = o(\max(|G(x^k)|, |H(x^k)|))$ and $\epsilon_k = o(r_k)$. Then, x^* is an M-stationary point.*

Proof. The proof is direct by Lemma 7.4.1 and Corollary 4.4.3 that ensures boundedness of the sequence (7.3) under MPCC-CRSC. □

In the weaker conditions of Lemma 6.4.1 boundedness of the sequence should be expected under MPCC-CRCQ in similar ways as Proposition 6.2.1.

The following example from [112] shows that the butterfly relaxation with $t_k = o(r_k)$ may converge to an undesirable A-stationary point without the additional hypothesis that $\epsilon_k = o(\max(|G(x^k)|, |H(x^k)|))$.

Example 6.4.1.

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

Let $t_k = r_k^2$ and choose any positive sequences $\{r_k\}$ and $\{\epsilon_k\}$ such that $r_k, \epsilon_k \rightarrow 0$. Consider the following epsilon-stationary sequence

$$x^k = (\epsilon_k, \frac{\epsilon_k}{2})^T, \quad \nu^{G,k} = 0, \quad \nu^{H,k} = 1 - \nu^{\Phi,k} (r_k^2 \theta_{r_k} \left(\frac{\epsilon_k}{2} \right) F_1(x^k; \hat{t}_k) - F_2(x^k; \hat{t}_k))$$

and

$$\nu^{\Phi,k} = \frac{1}{r_k^2 \theta_{r_k} (\epsilon_k) F_2(x^k; \hat{t}_k) - F_1(x^k; \hat{t}_k)}.$$

This sequence converges to $x^* = (0, 0)$, which is an A-stationary point.

Based on this study it appears that the approximate resolution of the sub-problems must be handled with care. This problematic is the central concern of the following chapter.

Chapter 7

Theoretical study of the relaxed sub-problems

We discuss here the convergence of relaxation methods for MPCC with approximate sequence of stationary points by presenting a general framework to study these methods. It has been pointed out in the literature, [115], and in Chapter 6.4 that relaxation methods with approximate stationary points fail to give guarantees of convergence.

We show that by defining a new strong approximate stationary, we can attain the desired goal of computing an M-stationary point. We also provide an algorithmic strategy to compute such point. Existence of strong approximate stationary point in the neighbourhood of an M-stationary point is proved.

7.1 A unified framework for relaxation/approximation methods

In the past decade, several methods have been proposed to compute an M-stationary point of (MPCC). The first was the approximation scheme proposed by [111], which was later improved as a relaxation by [112]. This relaxation scheme has been generalised recently in [64] to a more general family of relaxation schemes. We propose in this section a unified framework that embraces those methods and may be used to derive new methods.

Consider the following non-linear parametric program $R_t(x)$ parametrised by the vector t :

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

with $\bar{t} : \mathbb{R}_+^l \rightarrow \mathbb{R}_+$ such that $\lim_{\|t\| \rightarrow 0} \bar{t}(t) \rightarrow 0$ and the relaxation map $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^q$. In the sequel we skip the dependency in t and denote \bar{t} to simplify the notation. It is to be noted here that t is a vector of an arbitrary size denoted l as for instance in the

previous chapter where $l = 2$. The generalised Lagrangian function of $(R_t(x))$ is defined for $\nu \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$ as

$$\mathcal{L}_{R_t}^r(x, \nu) := rf(x) + g(x)^T \nu^g + h(x)^T \nu^h - G(x)^T \nu^G - H(x)^T \nu^H + \Phi(G(x), H(x); t)^T \nu^\Phi.$$

Let \mathcal{I}_Φ be the set of active indices for the constraint $\Phi(G(x), H(x); t) \leq 0$ so that

$$\mathcal{I}_\Phi(x; t) := \{i \in \{1, \dots, q\} \mid \Phi_i(G(x), H(x); t) = 0\}.$$

The definition of a generic relaxation scheme is completed by the following hypothesis :

- $\Phi(G(x), H(x); t)$ is a continuously differentiable real valued map extended component by component, so that $\Phi_i(G(x), H(x); t) := \Phi(G_i(x), H_i(x); t)$.
- Direct computations give that the gradient with respect to x for $i \in \{1, \dots, q\}$ of $\Phi_i(G(x), H(x); t)$ for all $x \in \mathbb{R}^n$ is given by

$$\nabla_x \Phi_i(G(x), H(x); t) = \nabla G_i(x) \alpha_i^H(x; t) + \nabla H_i(x) \alpha_i^G(x; t),$$

where $\alpha^H(x; t)$ and $\alpha^G(x; t)$ are continuous maps by smoothness assumptions on Φ , which we assume satisfy $\forall x \in \mathcal{Z}$

$$\lim_{\|t\| \rightarrow 0} \alpha^H(x; t) = H(x) \text{ and } \lim_{\|t\| \rightarrow 0} \alpha^G(x; t) = G(x). \quad (\text{H2})$$

- At the limit when $\|t\|$ goes to 0, the feasible set of the non-linear parametric program $(R_t(x))$ must converge to the feasible set of (MPCC). In other words, given $\mathcal{F}(t)$ the feasible set of $(R_t(x))$ it holds that

$$\lim_{\|t\| \rightarrow 0} \mathcal{F}(t) = \mathcal{Z}, \quad (\text{H3})$$

where the limit is assumed pointwise.

- At the boundary of the feasible set of the relaxation of the complementarity constraint it holds that for all $i \in \{1, \dots, q\}$

$$\Phi_i(G(x^k), H(x^k); t) = 0 \iff F_{G_i}(x; t) = 0 \text{ or } F_{H_i}(x; t) = 0, \quad (\text{H4})$$

where

$$\begin{aligned} F_G(x; t) &:= G(x) - \psi(H(x); t), \\ F_H(x; t) &:= H(x) - \psi(G(x); t), \end{aligned} \quad (7.1)$$

and ψ is a continuously differentiable real valued function extended component by component. Note that the function ψ may be two different functions in (7.1) as long as they satisfy the assumptions below. Those functions $\psi(H(x); t)$, $\psi(G(x); t)$ are non-negative for all $x \in \{x \in \mathbb{R}^n \mid \Phi(G(x), H(x); t) = 0\}$ and satisfy $\forall z \in \mathbb{R}^q$

$$\lim_{\|t\| \rightarrow 0} \psi(z; t) = 0. \quad (\text{H5})$$

We prove in Lemma 7.4.1 that this generic relaxation scheme for (MPCC) converges to an M-stationary point requiring the following essential assumption on the functions ψ . As t goes to 0 the derivative with respect to the first variable of ψ satisfies $\forall z \in \mathbb{R}^q$

$$\lim_{\|t\| \rightarrow 0} \left. \frac{\partial \psi(x; t)}{\partial x} \right|_{x=z} = 0. \quad (\text{H6})$$

We conclude this section by giving an explicit formula for the relaxation map at the boundary of the feasible set.

Lemma 7.1.1. *Given $\Phi(G(x), H(x); t)$ be such that for all $i \in \mathcal{I}_\Phi(x; t)$*

$$\Phi_i(G(x), H(x); t) = F_{G_i}(x; t) F_{H_i}(x; t).$$

The gradient with respect to x of $\Phi_i(G(x), H(x); t)$ for $i \in \mathcal{I}_\Phi(x; t)$ is given by

$$\nabla_x \Phi_i(G(x), H(x); t) := \nabla G_i(x) \alpha_i^H(x; t) + \nabla H_i(x) \alpha_i^G(x; t),$$

with

$$\begin{aligned} \alpha_i^G(x; t) &= F_{G_i}(x; t) - \left. \frac{\partial \psi(x; t)}{\partial x} \right|_{x=H_i(x)} F_{H_i}(x; t), \\ \alpha_i^H(x; t) &= F_{H_i}(x; t) - \left. \frac{\partial \psi(x; t)}{\partial x} \right|_{x=G_i(x)} F_{G_i}(x; t). \end{aligned}$$

7.2 Existing methods under the unified framework

In this section, we illustrate the fact that the existing methods in the literature fall under this unified framework. Indeed, the approximation method (KDB) as well as the two relaxation methods (KS) and (6.3) satisfy those hypotheses. We conclude this section by presenting a new asymmetric relaxation method that also belongs to our framework. An optimisation method that satisfies all of the 6 hypotheses defined in the previous section is called an UF-method.

7.2.1 The boxes approximation

The approximation method (KDB) illustrated on Figure 7.1 belongs to the framework defined in previous section, despite the fact that the feasible domain of the relaxed problem does not include the feasible domain of (MPCC).

Proposition 7.2.1. *The approximation scheme $(R_t(x))$ with (KDB) is an UF-method.*

Proof. Continuity of the map Φ as well as (H3) has been proved in [111].

(H4) is satisfied by construction considering $\psi(z; t) = t$. In this case (H6) and (H5) are obviously satisfied.

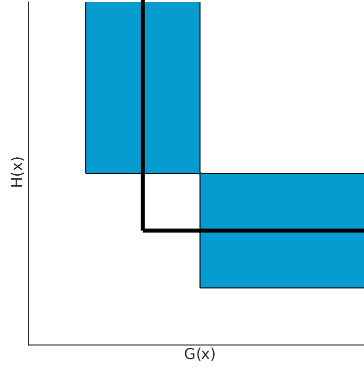


Figure 7.1: Feasible set of the approximation (KDB).

Now, we consider (H2). Direct computations give that the gradient of Φ for all $i \in \{1, \dots, q\}$ is given by

$$\nabla_x \Phi_i^{KDB}(G(x), H(x); t) = \nabla G_i(x)(H_i(x) - t) + \nabla H_i(x)(G_i(x) - t).$$

Therefore, α_i^G and α_i^H are given by

$$\begin{aligned}\alpha_i^H(x; t) &= H_i(x) - t, \\ \alpha_i^G(x; t) &= G_i(x) - t.\end{aligned}$$

It clearly holds that $\alpha_i^G(x; t) \rightarrow_{\|t\| \rightarrow 0} G_i(x)$ and $\alpha_i^H(x; t) \rightarrow_{\|t\| \rightarrow 0} H_i(x)$. So, in this case (H2) is satisfied.

This completes the proof that all of the 6 hypotheses are satisfied and so the approximation (KDB) is an UF-method. \square

7.2.2 The L-shape relaxation

The previous method has later been extended to a relaxation (KS) as illustrated on Figure 7.2. This extended method also satisfy our hypotheses.

Proposition 7.2.2. *The relaxation scheme $(R_t(x))$ with (KS) is an UF-method.*

Proof. Continuity of the map Φ as well as (H3) has been proved in [112].

(H4) is satisfied by construction considering $\psi(z; t) = t$. In this case (H6) and (H5) are obviously satisfied.

Now, we consider (H2). Direct computations give that the gradient of Φ for all $i \in \{1, \dots, q\}$ is given by

$$\nabla_x \Phi_i^{KS}(G(x), H(x); t) = \begin{cases} \nabla G_i(x)(H_i(x) - t) + \nabla H_i(x)(G_i(x) - t), & \text{if } H_i(x) - t + G_i(x) - t \geq 0, \\ -\nabla G_i(x)(G_i(x) - t) - \nabla H_i(x)(H_i(x) - t), & \text{otherwise.} \end{cases}$$

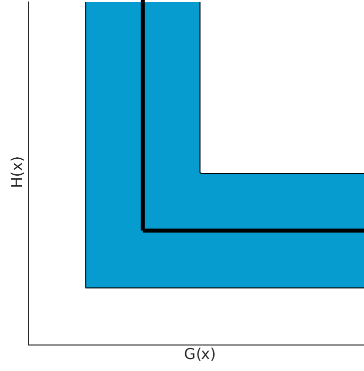


Figure 7.2: Feasible set of the relaxation (KS).

Therefore, α_i^G and α_i^H are given by

$$\alpha_i^H(x; t) = \begin{cases} H_i(x) - t, & \text{if } H_i(x) - t + G_i(x) - t \geq 0, \\ -(G_i(x) - t), & \text{otherwise,} \end{cases}$$

$$\alpha_i^G(x; t) = \begin{cases} G_i(x) - t, & \text{if } H_i(x) - t + G_i(x) - t \geq 0, \\ -(H_i(x) - t), & \text{otherwise.} \end{cases}$$

In the case $H_i(x) - t + G_i(x) - t \geq 0$ it clearly holds that $\alpha_i^G(x; t) \rightarrow G_i(x)$ and $\alpha_i^H(x; t) \rightarrow H_i(x)$. So, in this case (H2) is satisfied.

In the case $H_i(x) - t + G_i(x) - t < 0$ the opposite holds that is $\alpha_i^G(x; t) \rightarrow -H_i(x)$ and $\alpha_i^H(x; t) \rightarrow -G_i(x)$. However, it is to be noted that sequences x^t with $x^t \rightarrow_{\|t\| \rightarrow 0} x^*$ that belongs to this case satisfy $i \in \mathcal{I}^{00}(x^*)$. To sum up, in this case for $x \in \mathcal{Z}$ then $\alpha_i^G(x; t) \rightarrow H_i(x) = G_i(x) = 0$ and $\alpha_i^H(x; t) \rightarrow G_i(x) = H_i(x) = 0$. This proves that (H2) holds in this case too and so this hypothesis holds for this relaxation.

This completes the proof that all of the 6 hypotheses are satisfied and so the relaxation (KS) is an UF-method. \square

7.2.3 The butterfly relaxation

The butterfly relaxation introduced in Chapter 6 deals with two positive parameters (t_1, t_2) .

Proposition 7.2.3. *The relaxation scheme $(R_t(x))$ with (6.3) is an UF-method.*

Proof. Continuity of the map Φ as well as (H3) has been proved in [64].

(H4) is satisfied by construction considering $\psi(z; t) = t_1 \theta_{t_2}(z)$. In this case (H5) and (H6) are obviously satisfied. The latter being insured by $t_1 = o(t_2)$.

Now, we consider (H2). The gradient of Φ^B given in Lemma 6.1.3 yields α_i^G and α_i^H given by

$$\begin{aligned}\alpha_i^H(x; t) &= \begin{cases} F_{1i}(x; t) - t_1 \theta'_{t_2}(G_i(x)) F_{2i}(x; t), & \text{if } F_{1i}(x; t) + F_{2i}(x; t) \geq 0, \\ t_1 \theta'_{t_2}(G_i(x)) F_{1i}(x; t) - F_{2i}(x; t), & \text{otherwise,} \end{cases} \\ \alpha_i^G(x; t) &= \begin{cases} F_{2i}(x; t) - t_1 \theta'_{t_2}(H_i(x)) F_{1i}(x; t), & \text{if } F_{1i}(x; t) + F_{2i}(x; t) \geq 0, \\ t_1 \theta'_{t_2}(H_i(x)) F_{2i}(x; t) - F_{1i}(x; t), & \text{otherwise.} \end{cases}\end{aligned}$$

In the case $F_{1i}(x; t) + F_{2i}(x; t) \geq 0$ it clearly holds that $\alpha_i^G(x; t) \rightarrow G_i(x)$ and $\alpha_i^H(x; t) \rightarrow H_i(x)$. So, in this case (H2) is satisfied.

In the case $F_{1i}(x; t) + F_{2i}(x; t) < 0$ the opposite holds that is $\alpha_i^G(x; t) \rightarrow -H_i(x)$ and $\alpha_i^H(x; t) \rightarrow -G_i(x)$. However, it is to be noted that sequences x^t with $x^t \rightarrow_{\|t\| \rightarrow 0} x^*$ that belongs to this case satisfy $i \in \mathcal{I}^{00}(x^*)$. Therefore, in this case for $x \in \mathcal{Z}$ then $\alpha_i^G(x; t) \rightarrow H_i(x) = G_i(x) = 0$ and $\alpha_i^H(x; t) \rightarrow G_i(x) = H_i(x) = 0$. This proves that (H2) holds in this case too and so this hypothesis holds for this relaxation.

This completes the proof that all of the 6 hypotheses are satisfied and so the butterfly relaxation is an UF-method. \square

7.2.4 An asymmetric relaxation

Up till now we only consider relaxation methods that are symmetric. We can also define asymmetric relaxation methods illustrated on Figure 7.3 that respect the hypothesis of our unified framework. Let I_G and I_H be two sets of indices such that $I_G \cup I_H = \{1, \dots, q\}$ and $I_G \cap I_H = \emptyset$. Then, the relaxation constraint is defined with

$$\Phi_i(G(x), H(x); t) = \begin{cases} (G_i(x) - t)H_i(x), & \text{for } i \in I_G, \\ G_i(x)(H_i(x) - t), & \text{for } i \in I_H. \end{cases} \quad (7.2)$$

Proposition 7.2.4. *The relaxation scheme $(R_t(x))$ with (7.2) is an UF-method.*

Proof. Continuity of the map $\Phi(G(x), H(x); t)$ as well as (H3) can be easily deduced from the definition of (7.2).

(H4) is satisfied by construction considering $\psi(z; t) = t$ or 0. In this case (H5) and (H6) are obviously satisfied.

Now, we consider (H2). Direct computations give that the gradient of Φ for all $i \in \{1, \dots, q\}$ is given by

$$\nabla_x \Phi_i(G(x), H(x); t)(x) = \begin{cases} \nabla G_i(x)H_i(x) + \nabla H_i(x)(G_i(x) - t), & \text{for } i \in I_G, \\ \nabla G_i(x)(H_i(x) - t) + \nabla H_i(x)G_i(x), & \text{for } i \in I_H. \end{cases}$$

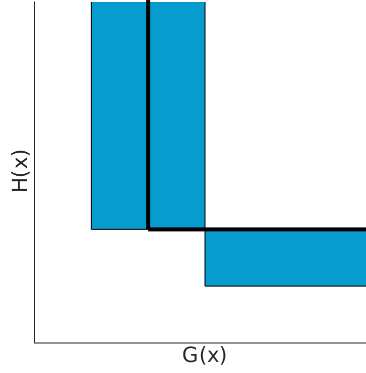


Figure 7.3: Feasible set of the relaxation (7.2) with $I_H = \{1\}$ and $I_G = \emptyset$.

Therefore, α_i^G and α_i^H are given by

$$\begin{aligned}\alpha_i^H(x; t) &= \begin{cases} H_i(x), & \text{for } i \in I_G, \\ H_i(x) - t, & \text{for } i \in I_H, \end{cases} \\ \alpha_i^G(x; t) &= \begin{cases} G_i(x) - t, & \text{for } i \in I_G, \\ G_i(x), & \text{for } i \in I_H. \end{cases}\end{aligned}$$

Clearly in both cases (H2) is satisfied.

This completes the proof that all of the 6 hypotheses are satisfied and so the relaxation (KS) is an UF-method. \square

7.3 Motivations on epsilon-solution to the regularised sub-problems

We have seen in the previous sections a general framework to define relaxations of (MPCC). From an algorithmic point of view, the main idea of relaxation methods to solve (MPCC) is to compute a sequence of stationary points, or more precisely approximate stationary points, for each value of a sequence of parameters $\{t_k\}$. The following definition is a specialisation of Definition 7.3.1 for $(R_t(x))$. It consists in replacing most “0” in (KKT) by small quantities ϵ .

Definition 7.3.1. x^k is an epsilon-stationary point for $(R_t(x))$ with $\epsilon_k \geq 0$ if there exists $\nu^k \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q$ such that

$$\|\nabla \mathcal{L}_{R_t}^1(x^k, \nu^k, t_k)\|_\infty \leq \epsilon_k$$

and

$$\begin{aligned}
|h(x^k)| &\leq \epsilon_k \quad \forall i \in \{1, \dots, m\}, \\
g_i(x^k) &\leq \epsilon_k, \quad \nu_i^{g,k} \geq 0, \quad |g_i(x^k) \nu_i^{g,k}| \leq \epsilon_k \quad \forall i \in \{1, \dots, p\}, \\
G_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu^{G,k}_i \geq 0, \quad |\nu^{G,k}_i (G_i(x^k) + \bar{t}_k)| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\
H_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu^{H,k}_i \geq 0, \quad |\nu^{H,k}_i (H_i(x^k) + \bar{t}_k)| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\
\Phi_i(G(x^k), H(x^k); t_k) &\leq \epsilon_k, \quad \nu_i^{\Phi,k} \geq 0, \quad \left| \nu_i^{\Phi,k} \Phi_i(G(x^k), H(x^k); t_k) \right| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}.
\end{aligned}$$

Unfortunately, it has been shown in [115] (Theorems 9 and 12) or Theorem 6.4 (115) for the KDB, L-shape and butterfly relaxations that under this definition, sequences of epsilon-stationary points only converge to weak-stationary point without additional hypothesis.

Our goal of computing an M-stationary point with a realistic method is far from obvious. Indeed, epsilon-stationary points have two main drawbacks considering our goal. The difficulties may come from the approximation of the complementarity condition and the approximate feasibility as shown in Example 7.3.1 or from the approximation of the feasibility of the relaxed constraint as illustrated in Example 7.3.2. In those examples, we consider the scheme (KDB) in order to simplify the presentation, but these observations can be easily generalised to the other methods.

Kanzow and Schwartz provide the following example exhibiting convergence to a W-stationary point.

Example 7.3.1.

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

If we perturb the relation $\nu^\Phi \Phi(x_1, x_2; t) \leq \epsilon$ (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. For the case KDB $\Phi(x_1, x_2; t) = (x_1 - t)(x_2 - t) = -\epsilon^2$ with $\epsilon = t^2$, the point $x(t, \epsilon) = (t - \epsilon, t + \epsilon)^T \geq (0, 0)^T$ is epsilon-stationary for small enough ϵ : $\Phi(x_1, x_2; t) \leq \epsilon$ and the choice $\nu^\Phi = \frac{1}{\epsilon}$ makes the Lagrangian $(-1, 1)^T + \nu^\Phi \Phi(x_1, x_2; t)$ vanish. $x(t)$ converges to the origin when $t, \epsilon \rightarrow 0$ but the origin is only weakly stationary.

Now, if the complementarity constraint is relaxed, but the complementarity condition is guaranteed convergence may occur to C-stationary points as shown in the following example.

Example 7.3.2.

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

We specialise the relations in Definition 7.3.1 as the following, for t and ϵ close to 0.

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

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

On this example, the relaxed regularised complementarity constraint is active for any small enough $t, \epsilon > 0$; moreover, the relaxed regularised stationary point is a local maximum for $t + 2\sqrt{\epsilon} < 1$. The origin is a local maximum for the original (MPCC). Another example might help understanding the phenomenon.

Example 7.3.3.

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

We again specialise the relations in Definition 7.3.1 as the following, for t and ϵ close to 0.

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

This time, the points $(t + \sqrt{\epsilon}, t + \sqrt{\epsilon})^T$ are no more epsilon-stationary but the points $x = (1, -t)^T$, $\nu^H = 1 + t$ and $x = (-t, 1)^T$, $\nu^G = 1 + t$ are. Their limits are $(1, 0)^T$ and $(0, 1)^T$ which are KKT points for the original MPCC with $\nu^H = 1, \nu^G = 0$ or $\nu^H = 0, \nu^G = 1$. The point $(-t, -t)^T$ with $\nu^H = 1 + t, \nu^G = 1 + t$ is also stationary, and, of course, converges to the origin, a local minimiser of the original MPCC.

In this example, the limit points are not minimisers for the original MPCC, but satisfy the first order KKT conditions for a minimiser. The second order conditions fail for those limit points. The two examples show limiting solutions of regularised sub-problems which are not local minimisers of the original MPCC. The first one fails to satisfy a first order condition while the second one satisfies such a first order condition but not the second order one (it is a maximum on the active set).

The Figure 7.4 gives an intuition that explains the weak convergence in Example 7.3.2 by showing the ϵ -feasible set of the butterfly relaxed complementarity constraint. It can be noticed that this feasible set is very similar to the relaxation from Scheel and Scholtes, [181]. Therefore, it is no surprise that we cannot expect more than convergence to a C-stationary point in these conditions.

7.4 Convergence of strong epsilon-stationary sequences

We now address the convergence of sequences of epsilon-stationary points. Motivated by the several issues pointed out in the previous section, we introduce the definition of a new kind of epsilon-stationary point called strong epsilon-stationary point, which is more stringent regarding the complementarity constraint.

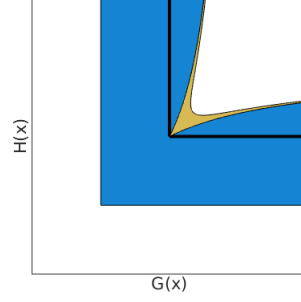


Figure 7.4: Butterfly relaxation with a constraint $\Phi^B(x; t) \leq \epsilon$.

Definition 7.4.1. x^k is a strong epsilon-stationary point for $(R_t(x))$ with $\epsilon_k \geq 0$ if there exists $\nu^k \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{3q}$ such that

$$\|\nabla \mathcal{L}_{R_t}^1(x^k, \nu^k; t_k)\|_\infty \leq \epsilon_k$$

and

$$\begin{aligned} |h(x^k)| &\leq \bar{t}_k + O(\epsilon_k) \quad \forall i \in \{1, \dots, m\}, \\ g_i(x^k) &\leq \epsilon_k, \quad \nu_i^{g,k} \geq 0, \quad |g_i(x^k) \nu_i^{g,k}| \leq \epsilon_k \quad \forall i \in \{1, \dots, p\}, \\ G_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu_i^{G,k} \geq 0, \quad \left| \nu_i^{G,k} (G_i(x^k) + \bar{t}_k) \right| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\ H_i(x^k) + \bar{t}_k &\geq -\epsilon_k, \quad \nu_i^{H,k} \geq 0, \quad \left| \nu_i^{H,k} (H_i(x^k) + \bar{t}_k) \right| \leq \epsilon_k \quad \forall i \in \{1, \dots, q\}, \\ \Phi_i(G(x^k), H(x^k); t_k) &\leq 0, \quad \nu_i^{\Phi,k} \geq 0, \quad \left| \nu_i^{\Phi,k} \Phi_i(G(x^k), H(x^k); t_k) \right| = 0 \quad \forall i \in \{1, \dots, q\}. \end{aligned}$$

In this case following a similar proof to the one of [115] and [64], we get an improved result that keep the nice properties of the relaxations without strong assumption on the sequence of $\{\epsilon_k\}$. The following lemma shows that a sequence of strong epsilon-stationary points converges to a weak-stationary point. This is not a new result since the same has been proved in [115] and [64] for a sequence of epsilon-stationary points. However, the new definition can be used to go further by showing convergence to an M-stationary point.

Lemma 7.4.1. Given $\{t_k\}$ a sequence of parameters and $\{\epsilon_k\}$ a sequence of non-negative parameters such that both sequences decrease to zero as $k \in \mathbb{N}$ goes to infinity. Assume that $\epsilon_k = o(\bar{t}_k)$. Let $\{x^k, \nu^k\}$ be a sequence of strong epsilon-stationary points of $(R_t(x))$ according to definition 7.4.1 for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$. Let $\{\eta^{G,k}\}, \{\eta^{H,k}\}$ be two sequences such that

$$\begin{aligned} \eta^{G,k} &:= \nu^{G,k} - \nu^{\Phi,k} \alpha^H(x^k; t_k), \\ \eta^{H,k} &:= \nu^{H,k} - \nu^{\Phi,k} \alpha^G(x^k; t_k). \end{aligned} \tag{7.3}$$

Assume that the sequence of multipliers $\{\nu^{h,k}, \nu^{g,k}, \eta^{G,k}, \eta^{H,k}\}$ is bounded. Then, x^* is an M-stationary point of (MPCC).

Proof. The proof is divided in two parts. We first show that x^* is a weak-stationary point and then we prove that it is an M-stationary point.

Let us prove the first part of the lemma. By definition $\{x^k, \nu^k\}$ is a sequence of strong epsilon-stationary points of $(R_t(x))$. We make the condition on the Lagrangian

$$\|\nabla \mathcal{L}_{R_t}^1(x^k, \nu^k; t_k)\|_\infty \leq \epsilon_k$$

more explicit. By construction of $\Phi(G(x), H(x); t)$ this condition becomes

$$\|\nabla f(x^k) + \nabla g(x^k)^T \nu^{g,k} + \nabla h(x^k)^T \nu^{h,k} - \nabla G(x^k)^T \eta^{G,k} - \nabla H(x^k)^T \eta^{H,k}\|_\infty \leq \epsilon_k. \quad (7.4)$$

Besides, the sequence of multipliers $\{\nu^{h,k}, \nu^{g,k}, \eta^{G,k}, \eta^{H,k}\}$ is assumed bounded. Therefore, it follows that the sequence converges to some limit point

$$\{\nu^{h,k}, \nu^{g,k}, \eta^{G,k}, \eta^{H,k}\} \rightarrow (\nu^h, \nu^g, \eta^G, \eta^H).$$

It is to be noted that for k sufficiently large it holds

$$\begin{aligned} \text{supp}(\nu^g) &\subset \text{supp}(\nu^{g,k}), \\ \text{supp}(\eta^G) &\subset \text{supp}(\eta^{G,k}), \\ \text{supp}(\eta^H) &\subset \text{supp}(\eta^{H,k}). \end{aligned}$$

We prove that $(x^*, \nu^h, \nu^g, \eta^G, \eta^H)$ is a weak-stationary point. Obviously, since $\epsilon_k \downarrow 0$ it follows that $x^* \in \mathcal{Z}$, $\nabla_x \mathcal{L}_{MPC}^1(x^*, \nu^h, \nu^g, \eta^G, \eta^H) = 0$ by (7.4) and that $\nu_i^g = 0$ for $i \notin \mathcal{I}_g(x^*)$. It remains to show that for indices $i \in \mathcal{I}^{+0}(x^*)$, $\eta_i^G = 0$. The opposite case for indices $i \in \mathcal{I}^{0+}(x^*)$ would follow in a completely similar way. So, let i be in $\mathcal{I}^{+0}(x^*)$.

By definition of strong ϵ_k -stationarity it holds for all k that

$$|\nu_i^{G,k}(G_i(x^k) + \bar{t}_k)| \leq \epsilon_k.$$

Therefore, $\nu_i^{G,k} \rightarrow_{k \rightarrow \infty} 0$ since $\epsilon_k \downarrow 0$ and $G_i(x^k) \rightarrow G_i(x^*) > 0$.

Without loss of generality we may assume that for k sufficiently large $\nu_i^{\Phi,k} \neq 0$ otherwise $\eta_i^G = 0$ and the proof is complete. By strong ϵ -stationarity $\nu_i^{\Phi,k} \neq 0$ implies that $F_{H,i}(x^k; t_k) = 0$ by (H4). (H2) yields $\alpha_i^H(x^k; t_k) \rightarrow H_i(x^*)$ and so $\eta_i^G = 0$ unless $\nu^{\Phi,k}$ diverges as k grows. We now prove that the latter case leads to a contradiction. Assume that $\nu^{\Phi,k} \rightarrow \infty$, boundedness hypothesis on η_i^G gives that there exists a finite non-vanishing constant C such that

$$\nu_i^{\Phi,k} \alpha_i^H(x^k; t_k) \rightarrow C.$$

Moreover, since η_i^H is finite and $\nu_i^{\Phi,k} \alpha_i^G(x^k; t_k) \rightarrow \infty$ as $G_i(x^k) > 0$ then necessarily $\nu_i^{H,k} \rightarrow \infty$. Furthermore, noticing that $F_H(x^k; t_k) = 0$ gives $H_i(x^k) \geq 0$, leads to a contradiction with $\nu_i^{H,k} \rightarrow \infty$ since by ϵ -stationarity we get

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

and $\epsilon_k = o(\bar{t}_k)$. We can conclude that for $i \in \mathcal{I}^{+0}(x^*)$, $\eta_i^G = 0$ and therefore x^* is a weak-stationary point.

Now, let us prove that x^* is even stronger than weak-stationary point since it is an M-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 $\eta_i^G \eta_i^H = 0$. It is clear that if $\nu_i^{\Phi,k} = 0$, then η_i^G and η_i^H are non-negative values and the result holds true. So, without loss of generality we may assume that $\nu_i^{\Phi,k} \geq 0$ and then $\Phi_i(G(x^k), H(x^k); t_k) = 0$ by Definition 7.4.1.

By construction of $\Phi_i(G(x^k), H(x^k); t_k)$ given in hypothesis (H4) it follows that

$$\Phi_i(G(x^k), H(x^k); t_k) = 0 \iff F_{G,i}(x^k; t_k) = 0 \text{ or } F_{H,i}(x^k; t_k) = 0$$

where we remind that

$$\begin{aligned} F_{G,i}(x^k; t_k) &= G_i(x^k) - \psi(H_i(x^k); t_k), \\ F_{H,i}(x^k; t_k) &= H_i(x^k) - \psi(G_i(x^k); t_k), \end{aligned}$$

Without loss of generality we assume that $F_{G,i}(x^k; t_k) = 0$ since the other case is completely similar. Furthermore by construction of $\Phi_i(G(x^k), H(x^k); t_k)$ it holds that $G_i(x^k)$ is non-negative in this case.

Considering one of the complementarity conditions of the strong ϵ -stationarity gives

$$\epsilon_k \geq |\nu_i^{G,k}(G_i(x^k) + \bar{t}_k)| = |\nu_i^{G,k}G_i(x^k)| + |\nu_i^{G,k}\bar{t}_k|,$$

since $G_i(x^k)$ is non-negative and it follows that

$$|\nu_i^{G,k}\bar{t}_k| \leq \epsilon_k.$$

Necessarily $\nu_i^{G,k} \rightarrow_{k \rightarrow \infty} 0$ as we assume in our statement that $\epsilon_k = o(\bar{t}_k)$.

Now at x^k we can use Lemma 7.1.1 that for $F_{G,i}(x^k; t_k) = 0$ gives

$$\begin{aligned} \alpha_i^G(x^k; t_k) &= - \left. \frac{\partial \psi(x; t_k)}{\partial x} \right|_{x=H_i(x^k)} F_{H,i}(x^k; t_k), \\ \alpha_i^H(x^k; t_k) &= F_{H,i}(x^k; t_k). \end{aligned}$$

Obviously, if $F_{H,i}(x^k; t_k) = 0$ we are done and so assume that $F_{H,i}(x^k; t_k) \neq 0$. By hypothesis (H6), it holds that $\left. \frac{\partial \psi(x; t_k)}{\partial x} \right|_{x=H_i(x^k)} \rightarrow_{k \rightarrow \infty} 0$. Therefore, $\alpha_i^G(x^k; t_k) \nu_i^{\Phi,k}$ going to a non-zero limits would imply that $\alpha_i^H(x^k; t_k) \nu_i^{\Phi,k}$ goes to infinity. However, this is a contradiction with η_i^G being finite. We can conclude that necessarily $\alpha_i^G(x^k; t_k) \nu_i^{\Phi,k}$ converges to zero.

Finally, we examine two cases regarding the sign of $F_{H,i}(x^k; t_k)$. For $F_{H,i}(x^k; t_k) \leq 0$, we get η_i^G, η_i^H non-negative, which satisfy the desired condition. For $F_{H,i}(x^k; t_k) \geq 0$ we get $\nu_i^{H,k} \rightarrow_{k \rightarrow \infty} 0$ using the same argument as for $\nu_i^{G,k}$. Thus, it follows that $\eta_i^H = 0$. This concludes the proof that x^* is an M-stationary point, since additionally to the proof of weak-stationary of x^* we proved for every $i \in \mathcal{I}^{00}(x^*)$ that either $\eta_i^H > 0$, $\eta_i^G > 0$ or $\eta_i^H \eta_i^G = 0$. \square

The following theorem is a direct consequence of both previous lemmas and is our main statement.

Theorem 7.4.1. *Given $\{t_k\}$ a sequence of parameters and $\{\epsilon_k\}$ a sequence of non-negative parameters such that both sequences decrease to zero as $k \in \mathbb{N}$ goes to infinity. Assume that $\epsilon_k = o(\bar{t}_k)$. Let $\{x^k, \nu^k\}$ be a sequence of epsilon-stationary points of $(R_t(x))$ according to definition 7.4.1 for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$ such that MPCC-CRSC holds at x^* . Then, x^* is an M-stationary point of (MPCC).*

Proof. The proof is direct by Lemma 7.4.1 and Corollary 2.3 of [64] that ensures boundedness of the sequence (7.3) under MPCC-CRSC. \square

Theorem (7.4.1) attains the ultimate goal. However it is not a trivial task to compute such a sequence of epsilon-stationary points. This is discussed later. Another important question is the existence of strong epsilon-stationary points in the neighbourhood of an M-stationary point. This problem is tackled in the following sections.

7.5 On Lagrange multipliers of the regularisation

The following example develops on Example 4.1.1 due to Kanzow and Schwartz exhibits a situation where the regularised sub-problems have no KKT point.

Example 7.5.1. *The KDB regularised problem is*

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

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 regularised 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 [112], but MPCC-LICQ is not satisfied by four constraints in \mathbb{R}^3 . It has been pointed out earlier that a practical algorithm may not be able to compute stationary point of the regularised sub-problem, but only some

approximate epsilon-stationary point. An intuitive idea would be that weaker constraint qualification may guarantee existence of such points. However, the following one-dimensional example shows that things are not that simple taking for instance the approximate method KDB, (KDB).

Example 7.5.2.

$$\min_{x \in \mathbb{R}} -x, \text{ s.t. } 0 \leq G(x) := x \perp H(x) := 0 \geq 0.$$

G and H are linear functions, so MPCC-CRCQ holds at each feasible point. Clearly, $x = 0$ is an M-stationary point with $\lambda^G = -1$ and $\lambda^H = 0$. Indeed, the gradient of the Lagrangian is given by

$$0 = -1 - \lambda^G.$$

We now verify that there is no epsilon-stationary point of the approximation KDB for this example :

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

Now, considering that the gradient of the Lagrangian function for this problem must be lower or equal as ϵ gives

$$\| -1 - \nu^G - t\nu^\Phi \| \leq \epsilon.$$

Noticing that $\nu^G, \nu^H, \nu^\Phi \geq 0$ leads to

$$1 + \nu^G + t\nu^\Phi \leq \epsilon,$$

which leads to a contradiction for $\epsilon < 1$. So, there is no sequence of approximate stationary points that goes to the origin.

The previous example illustrates the fact that even so strong constraint qualification holds for the problem existence of epsilon-stationary point are not ensured at an M-stationary point. Even so, this problem seems intractable by reformulating the (MPCC) with slack variables things could be slightly different.

Example 7.5.3. (Example 7.5.2 continued.) We now verify that there is a strong epsilon-stationary point of the approximation KDB written with slack variables for this example :

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

Given $\delta > 0$, consider the point $x = 0$, $(s_G, s_H) = (t, t + \delta)$ and Lagrange multiplier $(\nu^{s_G}, \nu^{s_H}, \nu^G, \nu^H, \nu^\Phi) = (-1, 0, 0, 0, \frac{1}{\delta})$.

- Condition on the gradient of the Lagrangian

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

- *Condition on the feasibility*

$$\begin{aligned}
|x - s_G| &= t \leq \epsilon, \\
|0 - s_H| &= t + \delta \leq \epsilon, \\
s_G + t &= 2t \geq -\epsilon, \quad s_H + t = 2t + \delta \geq -\epsilon, \\
(s_G - t)(s_H - t) &= 0.
\end{aligned}$$

- *Condition on the complementarity*

$$\begin{aligned}
|(s_G + t)\nu^G| &= 0, \quad |(s_H + t)\nu^H| = 0, \\
|(s_G - t)(s_H - t)\nu^\Phi| &= 0.
\end{aligned}$$

This completes the proof that there is a strong epsilon-stationary point for the formulation with slack variables.

This example motivates the use of slack variables to define the (MPCC) and in this case study the existence of strong epsilon-stationary point in a neighbourhood of an M-stationary point.

7.6 The MPCC with slack variables

Consider the following non-linear parametric program $R_t(x, s)$ parametrised by t :

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

with $\lim_{\|t\| \rightarrow 0} \bar{t} = 0^+$ and the relaxation map $\Phi(s_G, s_H; t) : \mathbb{R}^q \times \mathbb{R}^q \rightarrow \mathbb{R}^q$ is defined by replacing $G(x)$ and $H(x)$ by s_G and s_H in the map $\Phi(G(x), H(x); t)$. The generalised Lagrangian function of $(R_t^s(x, s))$ is defined as

$$\begin{aligned}
\mathcal{L}_s^r(x, s, \nu; t) &:= rf(x) + g(x)^T \nu^g + h(x)^T \nu^h - (G(x) - s_G)^T \nu^{s_G} - (H(x) - s_H)^T \nu^{s_H} \\
&\quad - s_G^T \nu^G - s_H^T \nu^H + \Phi(s_G, s_H; t)^T \nu^\Phi.
\end{aligned}$$

Let \mathcal{F}_s be the feasible set of $(R_t^s(x, s))$. The following result is a direct corollary of Theorem 7.4.1 stating that the reformulation with slack variables does not alter the convergence result.

Corollary 7.6.1. *Given $\{t_k\}$ a sequence of parameters and $\{\epsilon_k\}$ a sequence of non-negative parameters such that both sequences decrease to zero as $k \in \mathbb{N}$ goes to infinity. Assume that $\epsilon_k = o(\bar{t}_k)$. Let $\{x^k, \nu^k\}$ be a sequence of strong epsilon-stationary points of $(R_t^s(x, s))$ for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$ such that MPCC-CRSC holds at x^* . Then, x^* is an M-stationary point of (MPCC).*

Proof. Let $\tilde{h}(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^q$ 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 $(R_t^s(x, s))$ fall under the formulation $(R_t(x))$. Therefore, we can apply 7.4.1 to conclude this proof. \square

The following lemma giving an explicit form of the gradient of the Lagrangian function of $(R_t^s(x, s))$ can be deduced through direct computations.

Lemma 7.6.1. *The gradient of $\mathcal{L}_s^r(x, s, \nu; t)$ is given by*

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

There is two direct consequences of this result. First, it is easy to see from this lemma that computing a stationary point of $\mathcal{L}_s^r(x, s, \nu; t)$ is equivalent to computing a stationary point of $\mathcal{L}^r(x, \nu; t)$. Secondly, a stationary point of $\mathcal{L}_s^r(x, s, \nu; t)$ with $r = 1$ satisfies one of the conditions of weak-stationary point of (MPCC) that is $\nabla \mathcal{L}_{MPCC}^1(x, \nu) = 0$. In the next section, we now consider the existence of strong epsilon-stationary point for the relaxation with slack variables $(R_t^s(x, s))$.

7.7 Existence of strong epsilon-stationary points for the regularisation with slack variables

Before stating our main result, we give a series of additional hypothesis on the relaxation and the function ψ . It is essential to note once again that all these hypotheses are not restrictive, since they are satisfied by the existing methods in the literature.

7.7.1 Assumptions

The assumptions made in this section are divided in two parts. The first part concerns assumptions on the domain of the relaxation to be precised in the following subsection. The second part is assumptions on the relaxation function ψ that as been used to define the relaxation map on the boundary of the feasible set in Section 7.1.

Assumptions on the relaxations

We denote $\mathbb{B}_{c\epsilon}((-\bar{t}e, -\bar{t}e)^T)$ the ball of radius $c\epsilon$ and centred in $(-\bar{t}e, -\bar{t}e)^T$. We assume that the schemes considered in the sequel satisfy belong to one of two following cases for positive constants c, ϵ and \bar{t} :

Case 1

$$\mathbb{B}_{c\epsilon}((-\bar{t}e, -\bar{t}e)^T) \cap \{(s_G, s_H)^T \mid s_G \geq -\bar{t}e, s_H \geq -\bar{t}e\} \subset \mathcal{F}_s; \quad (\text{F1})$$

Case 2

$$\begin{aligned} & (\mathbb{B}_{ce}((- \bar{t}e, \psi(- \bar{t}e; t))^T) \cup \mathbb{B}_{ce}((\psi(- \bar{t}e; t), - \bar{t})^T)) \\ & \cap \{(s_G, s_H)^T \mid s_G \geq - \bar{t}e, s_H \geq - \bar{t}e, \Phi(s_G, s_H; t) = 0\} \subset \mathcal{F}_s. \end{aligned} \quad (\text{F2})$$

The first case includes the butterfly relaxation and the KS relaxation, while the second case includes the approximation KDB.

Assumptions on the relaxation function

For all $t \in \mathbb{R}_{++}^l$, we make the following supplementary assumptions on the function ψ for all $x \in \mathbb{R}^q$. We remind here that the functions ψ are separable with respect to x .

- $$\frac{\partial \psi(x; t)}{\partial t} > 0; \quad (\text{A1})$$

- $$\frac{\partial \psi(x; t)}{\partial x} \geq 0; \quad (\text{A2})$$

- $$\psi(\psi(\|t\|_\infty; t); t) \leq \|t\|_\infty; \quad (\text{A3})$$

- $$\psi(-\|t\|_\infty; t) \leq \|t\|_\infty. \quad (\text{A4})$$

Hypothesis (A1) in particular implies some monotonicity on the feasible set of the relaxed problems. Assumption (A4) is used for the second kind of relaxations only. It is to be noted here that the assumptions (A1),(A2),(A3) and (A4) are not the weakest for obtaining the following results. However, those assumptions are satisfied by all the relaxations defined in the literature.

Lemma 7.7.1. *Assume that (A1), (A2) and (A3) hold true. Then, giving constants $c > 0$, $\bar{t} > 0$ and $\epsilon > 0$ the following holds true for all $\|t\|_\infty \in (0, \bar{t} + c\epsilon)$*

$$\bar{t} + c\epsilon - \psi(\psi(\bar{t} + c\epsilon; t); t) > 0.$$

Proof. Using (A1), (A2) and that $\|t\|_\infty \in (0, \bar{t} + c\epsilon)$ yields

$$\bar{t} + c\epsilon - \psi(\psi(\bar{t} + c\epsilon; t); t) > \bar{t} + c\epsilon - \psi(\psi(\bar{t} + c\epsilon; e(\bar{t} + c\epsilon)); e(\bar{t} + c\epsilon)) \geq 0.$$

The conclusion comes from the assumption (A3). □

Lemma 7.7.2. *Assume that (A1) and (A3) holds true. Then, giving constants $c > 0$, $\bar{t} > 0$ and $\epsilon > 0$ the following holds true for all $\|t\|_\infty \in (0, \bar{t} + c\epsilon]$*

$$\psi(\bar{t} + c\epsilon; t) \leq \bar{t} + c\epsilon.$$

Proof. Using assumption (A1) and then (A3) gives

$$\psi(\bar{t} + c\epsilon; t) < \psi(\bar{t} + c\epsilon; e(\bar{t} + c\epsilon)) \leq \bar{t} + c\epsilon,$$

which concludes the proof. \square

Lemma 7.7.3. *Given positive constants \bar{t}, c, ϵ, K . There exists a $t^* > 0$ such that for all $t \in (0, t^*]$ it holds that*

$$\left. \frac{\partial \psi(x; t)}{\partial x} \right|_{x=\bar{t}+c\epsilon} \leq K\epsilon,$$

and

$$0 \leq \bar{t} - \left. \frac{\partial \psi(x; t)}{\partial x} \right|_{x=c\epsilon} \leq K\epsilon.$$

Proof. The proof is clear from Assumption (H6) on the relaxation. \square

Lemma 7.7.4. *Assume that (A1) and (A4) holds true. Then, giving constants $c > 0$, $\bar{t} > 0$ and $\epsilon > 0$ with $\bar{t} > c\epsilon$ the following holds true for all $\|t\|_\infty \in (0, \bar{t} - c\epsilon]$*

$$\psi(\bar{t} - c\epsilon; t) \leq \bar{t} + c\epsilon.$$

Proof. Using assumption (A1) and then (A4) gives

$$\psi(\bar{t} - c\epsilon; t) < \psi(\bar{t} - c\epsilon; e(\bar{t} - c\epsilon)) \leq \bar{t} - c\epsilon \leq \bar{t} + c\epsilon,$$

which concludes the proof. \square

7.7.2 Main theorem on existence of Lagrange multiplier

All of the supplementary assumptions made above are now used to derive the following result.

Theorem 7.7.1. *Let $x^* \in \mathcal{Z}$ be an M-stationary point and $\epsilon > 0$ be arbitrarily small. Furthermore, assume that the hypothesis (A1),(A2),(A3),(A4) on ψ and the hypothesis (F1) or (F2) on the relaxation introduced above hold true. Then, there exists positive constants c , t^* and \bar{t}^* with $\bar{t}^* > c\epsilon$ and a neighbourhood $U(x^*)$ of $(x^*, G(x^*), H(x^*))^T$ such that for all $t \in (0, t^*)$ and $\bar{t} \in (0, \bar{t}^*)$ there exists $(x, s)^T \in U(x^*)$, which is a strong epsilon-stationary point of the relaxation $(R_t^s(x, s))$.*

Regarding the value of t^* we need at least that $\|t\|_\infty \leq \bar{t} - c\epsilon$. The constant c is given in the proof and depends on the multipliers of the M-stationary point.

Proof. The proof is conducted in two steps. First, we construct a point based on the solution that is a candidate to be a strong epsilon-stationary point. Then, we verify that this candidate is actually a strong epsilon-stationary point.

x^* is assumed to be an M-stationary point. Therefore, there exists $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ such that

$$\begin{aligned} \nabla \mathcal{L}_{MPCC}^1(x^*, \lambda) &= 0, \\ \min(\lambda^g, -g_i(x^*)) &= 0, \quad \lambda_{\mathcal{I}^{+0}(x^*)}^G = 0, \quad \lambda_{\mathcal{I}^{0+}(x^*)}^H = 0, \\ \text{either } \lambda_i^G > 0, \lambda_i^H > 0 &\text{ either } \lambda_i^G \lambda_i^H = 0 \text{ for } i \in \mathcal{I}^{00}(x^*). \end{aligned}$$

Let c be the positive constant bounding the value of the Lagrange multipliers so that

$$c := \max_{i \in \text{supp}(\lambda^G), j \in \text{supp}(\lambda^H)} \frac{1}{|\lambda_i^G|} + \frac{1}{|\lambda_j^H|}.$$

Construction of the point $(\hat{x}, \hat{s}, \hat{\nu})$ Let us construct a point $(\hat{x}, \hat{s}, \hat{\nu})$ that satisfies the strong epsilon-stationary conditions (7.4.1).

$$\hat{x} := x^*, \quad \hat{\nu}^g := \lambda^g, \quad \hat{\nu}^h := \lambda^h, \quad \hat{\nu}^{sG} := \lambda^G, \quad \hat{\nu}^{sH} := \lambda^H.$$

We now split into two cases **(A)** and **(B)** corresponding to the two different kinds of relaxations. Denote the following set of indices

$$\begin{aligned} \mathcal{I}_{-0}^{00}(x^*, \lambda) &:= \{i = 1, \dots, q \mid \lambda_i^G < 0, \lambda_i^H = 0\}, \\ \mathcal{I}_{0-}^{00}(x^*, \lambda) &:= \{i = 1, \dots, q \mid \lambda_i^G = 0, \lambda_i^H < 0\}, \\ \mathcal{I}_{\nu^G} &:= \text{supp}(\hat{\nu}^{sG}) \setminus (\mathcal{I}_{-0}^{00}(x^*, \lambda) \cup \mathcal{I}_{0-}^{00}(x^*, \lambda)), \\ \mathcal{I}_{\nu^H} &:= \text{supp}(\hat{\nu}^{sH}) \setminus (\mathcal{I}_{-0}^{00}(x^*, \lambda) \cup \mathcal{I}_{0-}^{00}(x^*, \lambda)), \end{aligned}$$

A) Consider the Case 1, we choose $\hat{s}_G, \hat{s}_H, \hat{\nu}^G, \hat{\nu}^H$ and $\hat{\nu}^\Phi$ such that :

$$\begin{aligned} \hat{s}_G &:= \begin{cases} \psi(\bar{t} + c\epsilon; t), & i \in \mathcal{I}_{-0}^{00}(x^*, \lambda) \\ \bar{t} + c\epsilon, & i \in \mathcal{I}_{0-}^{00}(x^*, \lambda) \\ \frac{\epsilon - \bar{t}\hat{\nu}^G}{\hat{\nu}^G}, & i \in \mathcal{I}_{\nu^G} \\ \in \mathcal{F} & \text{otherwise,} \end{cases} \\ \hat{s}_H &:= \begin{cases} \bar{t} + c\epsilon, & i \in \mathcal{I}_{0-}^{00}(x^*, \lambda) \\ \psi(\bar{t} + c\epsilon; t), & i \in \mathcal{I}_{-0}^{00}(x^*, \lambda) \\ \frac{\epsilon - \bar{t}\hat{\nu}^H}{\hat{\nu}^H}, & i \in \mathcal{I}_{\nu^H} \\ \in \mathcal{F} & \text{otherwise,} \end{cases} \\ \hat{\nu}^G &:= \begin{cases} \hat{\nu}_i^{sG} & \text{for } i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*) \setminus (\mathcal{I}_{-0}^{00}(x^*, \lambda) \cup \mathcal{I}_{0-}^{00}(x^*, \lambda)) \\ 0 & \text{otherwise,} \end{cases} \\ \hat{\nu}^H &:= \begin{cases} \hat{\nu}_i^{sH} & \text{for } i \in \mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*) \setminus (\mathcal{I}_{-0}^{00}(x^*, \lambda) \cup \mathcal{I}_{0-}^{00}(x^*, \lambda)) \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

and finally

$$\hat{\nu}^\Phi := \begin{cases} \frac{-\hat{\nu}_i^{sG}}{\alpha_i^H(s;t)} & \text{for } i \in \mathcal{I}_{-0}^{00}(x^*, \lambda) \\ \frac{-\hat{\nu}_i^{sH}}{\alpha_i^G(s;t)} & \text{for } i \in \mathcal{I}_{0-}^{00}(x^*, \lambda) \\ 0 & \text{otherwise.} \end{cases}$$

$(\hat{x}, \hat{s}, \hat{\nu})$ **satisfy the stationary** : Finally, we verify that in both cases we satisfy the strong epsilon-stationary conditions, that is

$$\|\nabla \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t)\|_\infty \leq \epsilon,$$

and

$$\begin{aligned} |h_i(\hat{x})| &\leq \bar{t} + c\epsilon \quad \forall i \in \{1, \dots, m\}, \\ g_i(\hat{x}) &\leq \epsilon, \quad \hat{\nu}_i^g \geq 0, \quad |g_i(\hat{x})\hat{\nu}_i^g| \leq \epsilon \quad \forall i \in \{1, \dots, p\}, \\ |G_i(\hat{x}) - \hat{s}_{G,i}| &\leq \bar{t} + c\epsilon \quad \forall i \in \{1, \dots, q\}, \\ |H_i(\hat{x}) - \hat{s}_{H,i}| &\leq \bar{t} + c\epsilon \quad \forall i \in \{1, \dots, q\}, \\ \hat{s}_{G,i} + \bar{t} &\geq -\epsilon, \quad \hat{\nu}_i^G \geq 0, \quad |\hat{\nu}_i^G(\hat{s}_{G,i} + \bar{t})| \leq \epsilon \quad \forall i \in \{1, \dots, q\}, \\ \hat{s}_{H,i} + \bar{t} &\geq -\epsilon, \quad \hat{\nu}_i^H \geq 0, \quad |\hat{\nu}_i^H(\hat{s}_{H,i} + \bar{t})| \leq \epsilon \quad \forall i \in \{1, \dots, q\}, \\ \Phi_i(\hat{s}_G, \hat{s}_H; t) &\leq 0, \quad \hat{\nu}_i^\Phi \geq 0, \quad |\hat{\nu}_i^\Phi \Phi_i(\hat{s}_G, \hat{s}_H; t)| \leq 0 \quad \forall i \in \{1, \dots, q\}. \end{aligned}$$

We split the rest of the proof of **A)** in 6 parts :

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

Let us now run through these 6 conditions.

I. Since, $\hat{x} = x^*$ and $(\hat{\nu}^g, \hat{\nu}^h, \nu^{sG}, \nu^{sH}) = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$ it holds that

$$\|\nabla_x \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t)\|_\infty = 0$$

and

$$\begin{aligned} |h_i(\hat{x})| &= 0 \quad \forall i \in \{1, \dots, m\}, \\ g_i(\hat{x}) &\leq 0, \quad \hat{\nu}_i^g \geq 0, \quad |g_i(\hat{x})\hat{\nu}_i^g| \leq 0 \quad \forall i \in \{1, \dots, p\}. \end{aligned}$$

II. The gradient of the Lagrangian with respect to s is given by

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

In the case $\mathcal{I}_{-0}^{00}(x^*, \lambda)$ (the case $\mathcal{I}_{-0}^{00}(x^*, \lambda)$ is similar by symmetry) it is true that $\hat{\nu}_i^G = \hat{\nu}_i^H = 0$ and $\hat{\nu}_i^\Phi = \frac{-\hat{\nu}_i^{s_G}}{\alpha_i^H(\hat{s}; t)}$. Therefore

$$\begin{aligned}\nabla_{s_{G,i}} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) &= 0, \\ \nabla_{s_{H,i}} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) &= \frac{-\hat{\nu}_i^{s_G} \alpha_i^G(\hat{s}; t)}{\alpha_i^H(\hat{s}; t)} = \hat{\nu}_i^{s_G} \frac{\partial \psi(x; t)}{\partial x} \Big|_{x=\bar{t}+c\epsilon},\end{aligned}$$

since for $i \in \mathcal{I}_{-0}^{00}(x^*, \lambda)$ by construction of \hat{s}_G, \hat{s}_H it holds that $F_{G,i} = 0$ and $\hat{s}_{H,i} = \bar{t} + c\epsilon$. The conclusion follows by Lemma 7.7.3, which gives

$$\frac{\partial \psi(x; t)}{\partial x} \Big|_{x=\bar{t}+c\epsilon} \leq \epsilon.$$

Now, in the cases $\mathcal{I}^{0+}(x^*) \cup \mathcal{I}^{00}(x^*) \setminus (\mathcal{I}_{-0}^{00}(x^*, \lambda) \cup \mathcal{I}_{0-}^{00}(x^*, \lambda))$ and $\mathcal{I}^{+0}(x^*) \cup \mathcal{I}^{00}(x^*) \setminus (\mathcal{I}_{-0}^{00}(x^*, \lambda) \cup \mathcal{I}_{0-}^{00}(x^*, \lambda))$ the construction of the multipliers gives directly that $\nabla_{(s_G, s_H)} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) = 0$.

This concludes the proof of II.

III. \hat{x} feasible for the MPCC yields to

$$\begin{aligned}|G_i(\hat{x}) - \hat{s}_{G,i}| &= |\hat{s}_{G,i}| \text{ and } |H_i(\hat{x}) - \hat{s}_{H,i}| = 0, \text{ for } i \in \mathcal{I}^{0+}(x^*), \\ |G_i(\hat{x}) - \hat{s}_{G,i}| &= 0 \text{ and } |H_i(\hat{x}) - \hat{s}_{H,i}| = |\hat{s}_{H,i}|, \text{ for } i \in \mathcal{I}^{+0}(x^*), \\ |G_i(\hat{x}) - \hat{s}_{G,i}| &= |\hat{s}_{G,i}| \text{ and } |H_i(\hat{x}) - \hat{s}_{H,i}| = |\hat{s}_{H,i}|, \text{ for } i \in \mathcal{I}^{00}(x^*).\end{aligned}$$

By symmetry it is sufficient to consider the variables s_G . We analyse the cases where $i \in \mathcal{I}_{0-}^{00}(x^*, \lambda), \mathcal{I}_{-0}^{00}(x^*, \lambda)$ and \mathcal{I}_{ν^G} .

- Let $i \in \mathcal{I}_{0-}^{00}(x^*, \lambda)$, then $\hat{s}_{G,i} = \bar{t} + c\epsilon$;
- Let $i \in \mathcal{I}_{-0}^{00}(x^*, \lambda)$, then $\hat{s}_{G,i} = \psi(\bar{t} + c\epsilon; t) \leq \bar{t} + c\epsilon$ by Lemma 7.7.2;
- Let $i \in \mathcal{I}_{\nu^G}$, then $\hat{s}_{G,i} = \left| \frac{\epsilon}{\hat{\nu}_i^G} - \bar{t} \right| \leq \left| \frac{\epsilon}{\hat{\nu}_i^G} \right| + \bar{t} \leq \bar{t} + c\epsilon$.

In every case it holds that $|\hat{s}_{G,i}| \leq \bar{t} + c\epsilon$ and so this III is verified.

IV. By construction $\hat{s}_{G,i}$ and $\hat{s}_{H,i}$ are both non-negative as $\psi(\cdot; t)$ is assumed non-negative for indices i such that $\Phi_i(\hat{s}_G, \hat{s}_H; t) = 0$.

It remains to verify the condition in the case where $i \in \mathcal{I}_{\nu^G}$ and $i \in \mathcal{I}_{\nu^H}$. However, in both cases it holds that

$$\begin{aligned}\forall i \in \mathcal{I}_{\nu^G}, \quad \hat{s}_{G,i} + \bar{t} &= \frac{\epsilon - \bar{t}\hat{\nu}_i^G}{\hat{\nu}_i^G} + \bar{t} = \frac{\epsilon}{\hat{\nu}_i^G} > 0 \geq -\epsilon, \\ \forall i \in \mathcal{I}_{\nu^H}, \quad \hat{s}_{H,i} + \bar{t} &= \frac{\epsilon - \bar{t}\hat{\nu}_i^H}{\hat{\nu}_i^H} + \bar{t} = \frac{\epsilon}{\hat{\nu}_i^H} > 0 \geq -\epsilon.\end{aligned}$$

So, the feasibility in condition IV is satisfied. Now, regarding the complementarity condition it holds that

$$\begin{aligned}\forall i \in \mathcal{I}_{\nu^G}, \quad |(\hat{s}_{G,i} + \bar{t})\hat{\nu}_i^G| &= \left| \left(\frac{\epsilon - \bar{t}\hat{\nu}_i^G}{\hat{\nu}_i^G} + \bar{t} \right) \hat{\nu}_i^G \right| = \epsilon, \\ \forall i \in \mathcal{I}_{\nu^H}, \quad |(\hat{s}_{H,i} + \bar{t})\hat{\nu}_i^H| &= \left| \left(\frac{\epsilon - \bar{t}\hat{\nu}_i^H}{\hat{\nu}_i^H} + \bar{t} \right) \hat{\nu}_i^H \right| = \epsilon.\end{aligned}$$

This proves that the complementarity condition holds true for the relaxed positivity constraints and so condition IV is verified.

V. The feasibility $\Phi_i(\hat{s}_G, \hat{s}_H; t) \leq 0$ and the complementarity condition $|\hat{\nu}_i^\Phi \Phi_i(\hat{s}_G, \hat{s}_H; t)| \leq 0$ are satisfied by construction and by hypothesis on the relaxation.

VI. The multiplier $\hat{\nu}^\Phi$ is non-negative since for $i \in \mathcal{I}_{-0}^{00}(x^*, \lambda)$ it holds that

$$\alpha_i^H(\hat{s}; t) = F_{Hi}(\hat{s}; t) - \frac{\partial \psi(x; t)}{\partial x} \Big|_{x=s_{G,i}} F_{Gi}(\hat{s}; t) = F_{Hi}(\hat{s}; t) > 0,$$

since $F_{Gi}(\hat{s}; t) = 0$ by construction of $\hat{s}_{G,i}$ and $F_H(\bar{t} + c\epsilon; t) > 0$ by Lemma 7.7.1. The case $i \in \mathcal{I}_{0-}^{00}(x^*, \lambda)$ follows by symmetry.

The other multipliers are obviously non-negative by construction. This concludes the case VI.

The verification of all 6 cases proves that the point constructed above is strong epsilon-stationary, which concludes the proof of the relaxations **(A)**.

B) Consider the Case 2. Let $\hat{s}_G, \hat{s}_H, \hat{\nu}^G, \hat{\nu}^H$ and $\hat{\nu}^\Phi$ be such that :

$$\begin{aligned}\hat{s}_G &:= \begin{cases} \psi(\bar{t} + c\epsilon; t), & i \in \mathcal{I}_{-0}^{00}(x^*, \lambda) \\ \bar{t} + c\epsilon, & i \in \mathcal{I}_{0-}^{00}(x^*, \lambda) \\ \frac{\epsilon - \bar{t}\hat{\nu}_i^G}{\hat{\nu}_i^G}, & i \in \mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H} \\ \psi\left(\frac{\epsilon - \bar{t}\hat{\nu}_i^H}{\hat{\nu}_i^H}; t\right), & i \in \mathcal{I}_{\nu^H} \setminus \mathcal{I}_{\nu^G} \\ \in \mathcal{F} & \text{otherwise,} \end{cases} \\ \hat{s}_H &:= \begin{cases} \bar{t} + c\epsilon, & i \in \mathcal{I}_{0-}^{00}(x^*, \lambda) \\ \psi(\bar{t} + c\epsilon), & i \in \mathcal{I}_{-0}^{00}(x^*, \lambda) \\ \psi\left(\frac{\epsilon - \bar{t}\hat{\nu}_i^G}{\hat{\nu}_i^G}; t\right), & i \in \mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H} \\ \frac{\epsilon - \bar{t}\hat{\nu}_i^H}{\hat{\nu}_i^H}, & i \in I_2 \\ \in \mathcal{F} & \text{otherwise,} \end{cases} \\ \hat{\nu}^G &:= \begin{cases} \hat{\nu}_i^{s_G} & \text{for } i \in \mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H} \\ 0 & \text{otherwise,} \end{cases}\end{aligned}$$

$$\hat{\nu}^H := \begin{cases} \hat{\nu}_i^{sH} & \text{for } i \in \mathcal{I}_{\nu^H} \setminus \mathcal{I}_{\nu^G} \\ 0 & \text{otherwise,} \end{cases}$$

$$\hat{\nu}^\Phi := \begin{cases} \frac{-\hat{\nu}_i^{sG}}{\alpha_i^H(\hat{s};t)} & \text{for } i \in \mathcal{I}_{-0}^{00}(x^*, \lambda) \\ \frac{-\hat{\nu}_i^{sH}}{\alpha_i^G(\hat{s};t)} & \text{for } i \in \mathcal{I}_{0-}^{00}(x^*, \lambda) \cup \mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H} \\ 0 & \text{otherwise.} \end{cases}$$

Once again we run through the 6 conditions. It is to be noted that the variables involved in I. have not been changed so this condition stands true.

II. As pointed out earlier, the gradient of the Lagrangian with respect to s is given by

$$\begin{aligned} \nabla_{s_G} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) &= \hat{\nu}^{sG} - \hat{\nu}^G + \hat{\nu}^\Phi \alpha^H(\hat{s}; t), \\ \nabla_{s_H} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) &= \hat{\nu}^{sH} - \hat{\nu}^H + \hat{\nu}^\Phi \alpha^G(\hat{s}; t). \end{aligned}$$

For indices i in $\mathcal{I}_{-0}^{00}(x^*, \lambda)$ and $\mathcal{I}_{0-}^{00}(x^*, \lambda)$ we refer to case **(A)**. Let us consider indices i in $\mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H}$ and $\mathcal{I}_{\nu^H} \setminus \mathcal{I}_{\nu^G}$. For $i \in \mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H}$, then $\hat{\nu}_i^G > 0$, $\hat{\nu}_i^H = 0$ and $\hat{\nu}_i^\Phi = \frac{-\hat{\nu}_i^{sH}}{\alpha_i^G(\hat{s};t)}$ and so the gradient of the Lagrangian with respect to s becomes

$$\begin{aligned} \nabla_{s_{G,i}} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) &= \hat{\nu}_i^{sG} - \hat{\nu}_i^G + \hat{\nu}_i^\Phi \alpha_i^H(\hat{s}; t) = \frac{-\hat{\nu}_i^{sH} \alpha_i^H(\hat{s}; t)}{\alpha_i^G(\hat{s}; t)}, \\ \nabla_{s_{H,i}} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) &= \hat{\nu}_i^{sH} - \hat{\nu}_i^H + \hat{\nu}_i^\Phi \alpha_i^G(\hat{s}; t) = 0. \end{aligned}$$

By construction of $\hat{s}_{G,i}$ and $\hat{s}_{H,i}$, it holds that $F_{H,i}(\hat{s}; t) = 0$ and so

$$\nabla_{s_G} \mathcal{L}_s^1(\hat{x}, \hat{s}, \hat{\nu}; t) = \frac{-\hat{\nu}_i^{sH} \alpha_i^H(\hat{s}; t)}{\alpha_i^G(\hat{s}; t)} = \hat{\nu}_i^{sH} \frac{\partial \psi(x; t)}{\partial x} \Big|_{x=s_{G,i}} = \hat{\nu}_i^{sH} \frac{\partial \psi(x; t)}{\partial x} \Big|_{x=\frac{\epsilon}{\hat{\nu}_i^{sG}} - \bar{t}} \leq \epsilon,$$

for some $t \in (0, t^*)$ according to Lemma 7.7.3.

Now, for indices $i \in \mathcal{I}_{\nu^H} \setminus \mathcal{I}_{\nu^G}$ it holds that $\hat{\nu}_i^{sG} = 0$ so by the choice of multipliers ν^G, ν^H and ν^Φ the gradient of the Lagrangian with respect to s vanishes.

It follows that the condition II holds true.

III. Since \hat{x} is feasible for the MPCC therefore

$$\begin{aligned} |G_i(\hat{x}) - \hat{s}_{G,i}| &= |\hat{s}_{G,i}| \text{ and } |H_i(\hat{x}) - \hat{s}_{H,i}| = 0, \text{ for } i \in \mathcal{I}^{0+}(x^*), \\ |G_i(\hat{x}) - \hat{s}_{G,i}| &= 0 \text{ and } |H_i(\hat{x}) - \hat{s}_{H,i}| = |\hat{s}_{H,i}|, \text{ for } i \in \mathcal{I}^{+0}(x^*), \\ |G_i(\hat{x}) - \hat{s}_{G,i}| &= |\hat{s}_{G,i}| \text{ and } |H_i(\hat{x}) - \hat{s}_{H,i}| = |\hat{s}_{H,i}|, \text{ for } i \in \mathcal{I}^{00}(x^*). \end{aligned}$$

Let us consider the case where $i \in \mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H}$ noticing that $i \in \mathcal{I}_{\nu^H} \setminus \mathcal{I}_{\nu^G}$ is similar by symmetry. The other cases have been checked in case **(A)** of this proof. It follows that

$$\begin{aligned} |\hat{s}_{G,i}| &= \left| \frac{\epsilon - \bar{t} \hat{\nu}_i^G}{\hat{\nu}_i^G} \right| \leq \left| \frac{\epsilon}{\hat{\nu}_i^G} \right| + \bar{t} \leq \bar{t} + c\epsilon, \\ |\hat{s}_{H,i}| &= \left| \psi \left(\frac{\epsilon - \bar{t} \hat{\nu}_i^G}{\hat{\nu}_i^G}; t \right) \right|. \end{aligned}$$

The second condition is ensured by Lemma 7.7.4.

IV. and **V.** These conditions are straightforward and follow the same path that condition IV. in the case **A**.

VI. The proof for indices is very similar to the condition VI in case **(A)** except in the case where i is in $\mathcal{I}_{\nu^G} \cap \mathcal{I}_{\nu^H}$. In this case

$$\hat{\nu}_i^\Phi = \frac{-\hat{\nu}_i^{sH}}{\alpha_i^G(\hat{s}; t)} = \frac{-\hat{\nu}_i^{sH}}{F_{Gi}(\hat{s}; t)},$$

since by construction of \hat{s}_G and \hat{s}_H , $F_{Hi}(\hat{s}; t) = 0$ here. Now, by definition of $F_G(\hat{s}; t)$ it follows

$$\hat{\nu}_i^\Phi = \frac{-\hat{\nu}_i^{sH}}{\hat{s}_{G,i} - \psi(\hat{s}_{H,i}; t)} = \frac{-\hat{\nu}_i^{sH}}{\epsilon/\hat{\nu}_i^{sG} - \bar{t} - \psi(\hat{s}_{H,i}; t)} \geq 0,$$

for $c\epsilon \leq \bar{t}$ since ψ is non-negative whenever s_G and s_H are chosen such that $F_{Hi}(s; t) = 0$ (Assumption (H5)).

The verification of all 6 cases proves that the point constructed above is strong epsilon-stationary, which concludes the proof of the relaxations **(B)** and complete the whole proof. \square

It is to be noted that no constraint qualification is required for this result. This is a clear improvement over what was obtained in the literature in the ideal case of sequence of stationary points. For instance, Theorem 5.1 in [111] requires some second-order information to get a result on existence of stationary points for the regularised sub-problems.

Theorem 7.7.2. *For any M -stationary point of (MPCC) that satisfies MPCC-CRSC, there exists a sequence of strong epsilon-stationary points of the relaxation $(R_t^s(x, s))$ that converges to that point.*

Proof. Theorem 7.7.1 gives more relations between the parameters that are compatible with Corollary 7.6.1. Indeed for a chosen sequence of arbitrarily small parameters $\{\epsilon^k\}$, Corollary 7.6.1 requires that $\epsilon^k = o(\bar{t}_k)$ and Theorem 7.7.1 requires that $\bar{t}_k > c\epsilon$ and t_k must be sufficiently small, in particular smaller than $\bar{t}_k - c\epsilon_k$.

Thus, a straightforward application of both of these results provides the result. \square

Previous section points out that such result cannot be obtained without a formulation with slack variables. The various results obtained through this chapter leads to an algorithmic strategy with strong theoretical properties to solve the sub-problems of the relaxation method. A discussion of the numerical implementation will be discussed in the last part, Part III.

Notes and perspectives of Part II

Through the Part II, we have discussed stationary conditions for MPCC as well as relaxation methods to attain some optimality conditions. We raise the issue of solving approximately the sub-problems of the relaxation scheme and suggest a specific approach to deal with this issue. We briefly discuss here the perspectives from this work on relaxation methods. In particular, there are two different extensions of this work. A first approach is to use algorithms for MPCC to solve more difficult problems, while another approach is to apply the technique used to derive the butterfly relaxation to the Mathematical Programs with Vanishing Constraints (MPVCs) and Optimisation Problems with Cardinality Constraints (OPCCs). We give some more details about these perspectives.

Bilevel programming

The bilevel problem has been very popular for many applications, for instance in biology [36] and in network design [144] to cite just a few examples. A very used technique is to cast this problem as a (MPCC), see the monographs [53, 55] and some recent papers, for instance [54]. Bilevel optimisation problems are optimisation problems where the feasible set is determined (in part) using the graph of the solution set mapping of a second parametric optimisation problem. This problem is given as

$$\min_y f(x, y) \text{ s.t } g(x, y) \leq 0, \ y \in T, \quad (\text{Foll}(x))$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$, $T \subseteq \mathbb{R}^m$ is a (closed) set. Let $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ denote the feasible set mapping:

$$Y(x) := \{y : g(x, y) \leq 0\},$$
$$\varphi(x) := \min_y \{f(x, y) : g(x, y) \leq 0, \ y \in T\},$$

the optimal value function, and $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ the solution set mapping of the problem (Foll(x)) for a fixed value of x :

$$\Psi(x) := \{y \in Y(x) \cap T : f(x, y) \leq \varphi(x)\}.$$

Let

$$\mathbf{gph} \Psi := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in \Psi(x)\}$$

be the graph of the mapping Ψ . Then the bilevel optimisation problem is given as

$$\min_x \{F(x, y) : G(x) \leq 0, (x, y) \in \mathbf{gph} \Psi, x \in X\}, \quad (\text{BP})$$

where $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $G : \mathbb{R}^n \rightarrow \mathbb{R}^q$, $X \subseteq \mathbb{R}^n$ is a closed set. Problem (Foll(x)) is called the follower's problem and (BP) the leader's problem. The latter is the bilevel optimisation problem.

Example (Example 1.2, [55]).

$$\min_x x^2 + y : y \in \arg \min_y \{-xy : 0 \leq y \leq 1\}.$$

We see on these figures that for $x = 0$ if the solution $y = 0$ is taken, then there exists an

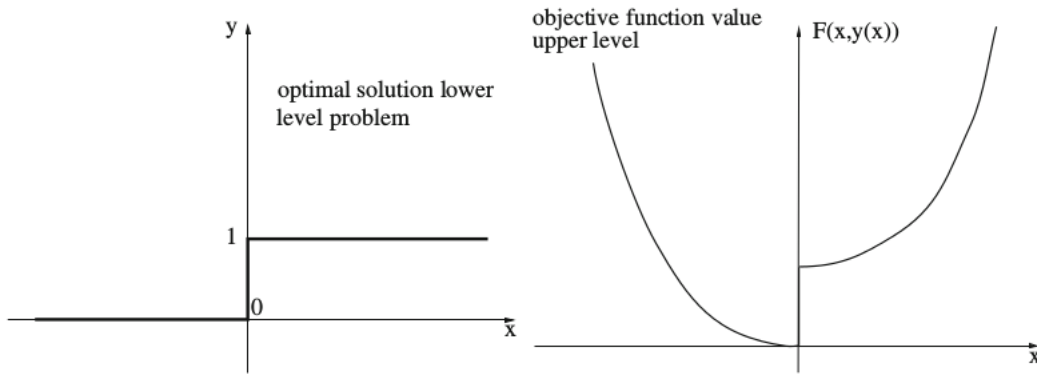


Figure 7.5: Mapping to be minimised in Example 7.7.2.

optimal solution (optimistic solution), otherwise the infimum exists but there is no solution. Also, note that the case $y = 1$ is called pessimistic case.

It follows from the example (7.7.2) that the problem (BP) is not well-posed if the follower's answer is not a singleton for some values of x . That is why we use quotation marks in (BP) to indicate this ambiguity. To overcome such an unpleasant situation, the leader has a number of possibilities, we briefly mention here two:

- The follower co-operate with the leader. In this case, we select the solution of the problem (Foll(x)) that minimises the function $F(x, \cdot)$. This approach is discussed in [55].
- The leader has no possibility to influence the follower's selection and neither he/she has an intuition about the follower's choice. In this case, the leader has to accept a worst case scenario. A penalisation technique is introduced in [27].

Among others, one approach to solve (BP) is to reformulate it as a one-level problem by replacing the follower problem by its KKT conditions. Since the KKT condition is a complementarity problem, this approach leads to the study of an (MPCC).

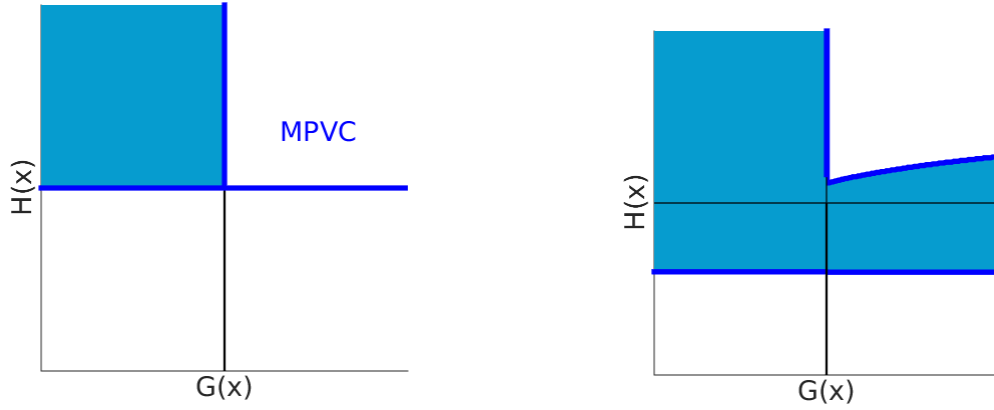


Figure 7.6: Feasible set of the MPVC and the butterfly relaxation for $\theta_r(z) = \frac{z}{z+r}$ with $t_1 = t_2^{3/2}$.

Extensions of the butterfly relaxation

The study of numerical methods has been the motivation of extensions of this technique to problems that are similar in the structure, but require an independent study. Among others two problems that corresponds to this description are the Mathematical Program with Vanishing Constraint (MPVC) and the Optimisation Program with Cardinality Constraints (OPCC).

Mathematical programs with vanishing constraints

The (MPVC) is defined as the following optimisation problem:

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

with $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ and $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^q$ that are assumed continuously differentiable. The feasible set of the "vanishing constraint" is given in Figure 7.6. This problem was first proposed by Achtziger in [9] motivated by applications in topology design and in mechanical structures problems. This problem can be reformulated as a (MPCC), however this runs at some constraint qualifications issues explaining why we need to propose specific numerical methods even so they are very close to methods for (MPCC). Relaxation methods from previous sections can be adapted to this case by considering the

non-linear parametric program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t. } & g(x) \leq 0, \quad h(x) = 0, \\ & H(x) \geq -t_3, \\ & \Phi(G(x), H(x); t) \leq 0. \end{aligned}$$

Recent literature extend relaxation methods for (MPCC) to (MPVC): [8, 106] deal with the smooth method by Scheel & Scholtes, [7] consider a non-smooth reformulation, [101] uses the local regularisation from Steffensen & Ulbrich and finally [102] adapt the new paradigm of relaxation method with convergence to M-stationary point in particular the method by Kanzow & Schwarz. Following the same path, we can extend the Butterfly relaxation method for MPVC:

$$\Phi(G(x), H(x); t) := G_i(x)(H_i(x) - t_1 - t_1\theta_{t_2}(G_i(x))), \quad i = 1, \dots, q$$

where θ is defined as before and with $t \in \mathbb{R}_{++}^3$. The feasible set of the vanishing constraint relaxed is given in Figure 7.6.

In order to validate this approach, we run the method on an application of MPVC to truss topology optimisation that was described in depth in chapter 9 of the monograph [100].

Example.

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

The feasible set of this example is given in Figure 7.7. 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, and then the unique optimal solution to the problem is the point $\bar{x} = (0, 5)^T$.

We run butterfly relaxation tailored to (MPVC) on Example 7.7.2 using an initial point inside the feasible domain $x^0 = (6, 6)^T$. Results are presented in Table 7.1 with solvers SNOPT, IPOPT and MINOS. In two cases the butterfly method manages to converge to the global optimum and in the third case it converges to the point $(0, 5)$ which is a local minimum.

Optimisation problems with cardinality constraints

The (OPCC) is defined as the following non-smooth optimisation problem:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t. } x \in X, \quad \|x\|_0 \leq \kappa, \quad (\text{OPCC})$$

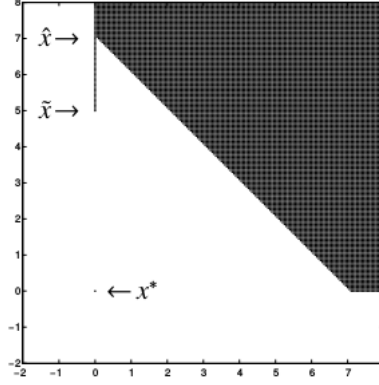


Figure 7.7: The feasible set of Example 7.7.2 from [100].

sovler	x^*	$f(x^*)$	last value of t
SNOPT	$(0, 5)^T$	10	0, 5
IPOPT	$(0, 0)^T$	0	$4, 67.10^{-4}$
MINOS	$(0, 0)^T$	0	0, 5

Table 7.1: Butterfly (MPVC) relaxation with $t = r^{3/2}$ on Example 7.7.2 with initial point $(6, 6)^T$.

with a set $X \subset \mathbb{R}^n$ described by some standard constraints

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

for $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ and $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$. f, g and h are assumed to be continuously differentiable and $\kappa < n$, otherwise there would be no cardinality constraints. This problem has many important applications, including portfolio optimisation [32], subset selection problem in regression [151], or the compressed sensing technique used in signal processing [42]. Due to the combinatorial behaviour of the ℓ_0 norm, (OPCC) is usually treated as a Mixed Integer Non-linear Program in the literature. However, recent development in the literature has started considering a formulation with continuous variables of (OPCC), see for instance [67, 35, 34, 43, 32]. In order to obtain a suitable reformulation of the cardinality constraints, we may consider the mixed-integer problem

$$\begin{aligned}
& \min_{x \in \mathbb{R}^n} f(x) \\
& \text{s.t } g(x) \leq 0, h(x) = 0, \\
& \quad e^T y \geq n - \kappa, \\
& \quad x_i y_i = 0, \forall i = 1, \dots, n, \\
& \quad y \in \{0, 1\}.
\end{aligned}$$

Next, we consider the standard relaxation of this mixed-integer problem

$$\begin{aligned}
& \min_{x \in \mathbb{R}^n} f(x) \\
& \text{s.t. } g(x) \leq 0, h(x) = 0, \\
& \quad e^T y \geq n - \kappa, \\
& \quad x_i y_i = 0, \quad \forall i = 1, \dots, n, \\
& \quad y \in [0, 1].
\end{aligned} \tag{7.5}$$

One of the main interests to study the relaxed problem (7.5) is the following two results from [35, 34]. The first result concerns the equivalence of a global solution.

Theorem (Theorem 2, [35]). *A vector $x^* \in \mathbb{R}^n$ is a solution of (OPCC) if and only if there exists a vector $y^* \in \mathbb{R}^n$ such that the pair (x^*, y^*) is a solution of the relaxed problem (7.5).*

As it could be expected, the same relation does not hold in general for local minimisers. Nevertheless, one can still obtain useful result using local minimisers of the relaxed problem (7.5).

Theorem (Theorem 5, [35]). *Let (x^*, y^*) be a local minimiser of problems (7.5) satisfying $\|x^*\|_0 = \kappa$. Then, x^* is a local minimiser of the cardinality-constrained problem (OPCC).*

We now briefly discuss the extension of regularisation techniques applied to the problem (7.5).

Relaxation methods

The relaxation (SS) and (KS) discussed earlier in Chapter 5 were extended to (7.5) respectively in [34] and [32] by considering

$$\begin{aligned}
& \min_{x \in \mathbb{R}^n} f(x) \\
& \text{s.t. } g(x) \leq 0, h(x) = 0, \\
& \quad e^T y \geq n - \kappa, \\
& \quad x_i y_i = 0, \quad \forall i = 1, \dots, n, \\
& \quad \Phi^+(x, y; t) \leq 0, \Phi^-(x, y; t) \leq 0, \\
& \quad y \in [0, 1],
\end{aligned}$$

where for (KS) relaxation $\Phi_i^+(x, y; t) = \varphi(x_i, y_i; t)$ and $\Phi_i^-(x, y; t) = \varphi(-x_i, y_i; t)$ with

$$\varphi(a, b; t) = \begin{cases} (a - t)(b - t), & \text{if } a + b \geq 2t \\ -\frac{1}{2}((a - t)^2 + (b - t)^2), & \text{if } a + b < 2t. \end{cases}$$

and for (SS) relaxation $\Phi_i^+(x, y; t) = x_i y_i - t$ and $\Phi_i^-(x, y; t) = -x_i y_i - t$. Thus, a natural application of the Butterfly technique to this formulation would consider for $t \in \mathbb{R}^2$

$$\Phi_i^+(x, y; t) = \begin{cases} (x_i - t_1 \theta_{t_2}(y_i))(y_i - t_1 \theta_{t_2}(x_i)), & \text{if } x_i - t_1 \theta_{t_2}(y_i) + y_i - t_1 \theta_{t_2}(x_i) \geq 0 \\ -\frac{1}{2}((x_i - t_1 \theta_{t_2}(y_i))^2 + (y_i - t_1 \theta_{t_2}(x_i))^2), & \text{otherwise.} \end{cases}$$

and $\Phi_i^-(x, y; t) = \Phi_i^+(-x, y; t)$.

Sparse theta-approximation

Another approach could use the relaxation of the ℓ_0 norm presented in Chapter 2 to tackle the problem (OPCC). Applying this technique would lead to an iterative method that solves the following sub-problem at each step

$$\min_{x, y \in \mathbb{R}^n \times \mathbb{R}^n} f(x) \text{ s.t. } (x, y) \in \mathcal{X}, \sum_{i=1}^n \theta_r(y_i) \leq \kappa,$$

with a set $\mathcal{X} \subset \mathbb{R}^n \times \mathbb{R}^n$ described by some standard constraints

$$\mathcal{X} := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0, y \geq 0, -y \leq x \leq y\}.$$

In both cases, a careful theoretical study supported by numerical test is needed to show the validity of these new methods.

Part III

Numerics

We discussed in the previous part relaxation methods to solve the (MPCC). Until now the discussion has been essentially focus on the theoretical properties of the various existing methods. We now focus on the numerical behaviour of these methods.

In Chapter 8, we present a numerical comparison between these methods. The non-linear regularised sub-problems have been solved using different solvers available via AMPL. The classical set of test problems used in the literature, and also used in this document, to compare the method is MacMPEC, [123]. The comparison has been made after defining criteria of success and also using a methodology introduced by Dolan and Moré in [57] to compare non-linear problems.

In Chapter 7, we pointed out that in order to keep the strong theoretical guarantees of the relaxations, the regularised sub-problems need to be solved with care. We present in Chapter 9 an implementation in JULIA of an algorithmic strategy that maintain the strong guarantees of the relaxation methods. Preliminary results regarding this implementation are given.

Chapter 8

Numerical comparison between the relaxation methods

In this chapter, we focus on the numerical implementation of the butterfly relaxation method. Our aim is to compare this new method with the existing ones in the literature and show some of their features. This comparison uses the collection of test problems MacMPEC [123], that has been widely used in the literature to compare relaxation methods as in [100, 111, 185]. The test problems included in MacMPEC are extracted from the literature and real-world applications.

Our comparison is based on two experiments. The first one, in Section 8.2, is a comparison of the relaxation methods for various criteria and for several strategies of parameter with a fixed precision. The second method, in Section 8.3, considers a methodology proposed in the literature to benchmark non-linear programs independently of the stopping criteria of the solvers. One of the benefit of this approach is to determine the precision that must be given to the solver.

Finally, we also present in Section 8.4 an example of an (MPCC) that illustrates some of the difficulties that may occur by dealing with epsilon-stationary points.

8.1 On the implementation of the butterfly relaxation

As pointed out in Chapter 6, the butterfly relaxation handles two parameters t and r . It can be practical to choose a relation between both parameters. Among the infinite possibilities of relationship between t and r , at least two are specific:

- (i) $t = r$, since as stated in Theorem 6.3.2 this relaxation is more regular, but may converge to undesirable A-stationary points according to Theorem 6.2.1;
- (ii) $t = o(r)$, for instance $t = r^{3/2}$, which ensures convergence to M-stationary points as stated in Corollary 6.2.1.

Practical implementation could also consider a slightly different model by adding a new parameter s in order to move the intersection of both wings in the point $(G(x), H(x)) = (s, s)$.

This can be done by redefining $F_1(x; r, t)$ and $F_2(x; t, r)$ such that for all $i \in \{1, \dots, q\}$

$$\begin{aligned} F_{1i}(x; r, s, t) &= (H_i(x) - s - t\theta_r(G_i(x) - s)), \\ F_{2i}(x; r, s, t) &= (G_i(x) - s - t\theta_r(H_i(x) - s)). \end{aligned}$$

Even so we did not give any theoretical proof regarding this modified system, this modification does not alter the behaviour of the butterfly relaxation. Besides, this formulation is clearly an extension of the method (KS).

The numerical comparison of the butterfly relaxation with other existing methods consider the three following schemes:

1. $B_{(t=r)}: (s = 0, t = r);$
2. $B_{(t=r^{3/2})}: (s = 0, t = r^{3/2});$
3. $B_{(s=t, 2t=r)}: (s = t, 2t = r).$

Each scheme can be duplicated whether we omit or not the relaxation of the positivity constraints. We do not present here the results with the relaxation of the positivity constraints, i.e. $\bar{t} = 0$, since after preliminary tests it does not seem to be helpful.

8.2 Comparison of the relaxation methods

We provide in this section and in Algorithm 2 some more details on the implementation and the comparison between relaxation methods. It is to be noted that our aim is to compare the methods and so no attempt to optimise any method has been carried out. We use 101 test problems from MacMPEC, where are omitted the problems that exceed the limit of 300 variables or constraints and some problems with the evaluation error of the objective function or the constraints. Algorithm 2 is coded in Matlab and uses the AMPL API.

$R_{\hat{t}_k}$ denotes the relaxed non-linear program associated with a generic relaxation, where except from the butterfly method the parameter r_k does not play any role. At each step we compute x^{k+1} as a solution of $R_{\hat{t}_k}$ starting from x^k . Therefore, at each step the initial point is more likely to be infeasible for $R_{\hat{t}_k}$. The iterative process stops when t_k and r_k are smaller than some tolerance, denoted p_{\min} which is set as 10^{-15} here, or when the solution x^{k+1} of $R_{\hat{t}_k}$ is considered an ϵ -solution of (MPCC). To consider x^{k+1} as an ϵ -solution with ϵ set as 10^{-7} we check three criteria:

- a) Feasibility of the last relaxed non-linear program: $\nu_f(x) := \max(-g(x), |h(x)|, -\Phi(x)),$
- b) Feasibility of the complementarity constraints: $\nu_{comp}(x) := \min(G(x), H(x))^2,$
- (c) The complementarity between the Lagrange multipliers and the constraints of the last relaxed non-linear program:

$$\begin{aligned} \nu_c(x) &:= \max(\|g(x) \circ \lambda^g\|_\infty, \|h(x) \circ \lambda^h\|_\infty, \|G(x) \circ \lambda^G\|_\infty, \\ &\quad \|H(x) \circ \lambda^H\|_\infty, \|\Phi^B(G(x), H(x); \hat{t}) \circ \lambda^\Phi\|_\infty). \end{aligned}$$

Obviously, it is hard to ask a tighter condition on the complementarity constraints, since the feasibility only guarantees that the product component-wise is less than ϵ . Using these criteria, we define a measure of optimality

$$\min_local(x) := \max(\nu_f(x), \nu_{comp}(x), \nu_c(x)).$$

A fourth criterion could be the dual feasibility, that is the norm of the gradient of the Lagrangian. However, solvers like SNOPT or MINOS do not use this criterion as the stopping criterion. One reason among others for the solvers to discard such a criterion is the numerical issues implied by the degeneracy in the KKT conditions.

In the case of an infeasible or unbounded sub-problem $R_{\hat{t}_k}$, the algorithm stops and return a certificate.

Data:

starting vector x^0 ; initial relaxation parameter t_0 ; update parameter $(\sigma_t, \sigma_r) \in (0, 1)^2$ and p_{min} the minimum parameter value, ϵ the precision tolerance ;

1 **Begin** ;

2 Set $k := 0$;

3 **while** $\max(t_k, r_k) > p_{min}$ **and** $\min_local(x) > \epsilon$ **do**

4 x^{k+1} solution of R_{t_k, r_k} with x^k initial point;

5 $(t_{k+1}, r_{k+1}) := (t_k \sigma_t, r_k \sigma_r)$;

6 **return:** f_{opt} the optimal value at the solution x_{opt} or a decision of infeasibility or unboundedness.

Algorithm 2: Basic Relaxation methods for (MPCC), with a relaxed non-linear program $R_{\hat{t}_k}$.

Step 4 in Algorithm 2 is performed using three different solvers accessible through AMPL, [79], that are SNOPT 7.2-8 [84], MINOS 5.51 [155] and IPOPT 3.12.4 [192] with their default parameters. Previous similar comparison in the literature only consider SNOPT to solve the sub-problems.

We compare the butterfly schemes with the relaxations (SS) and (KS) that we respectively denote SS and KS. Moreover, we also take into account results of the non-linear programming solver without specific MPCC tuning and denote it NL.

In order to compare the various relaxation methods, we try to have a coherent use of the parameters. In a similar way as in [184] we consider the value of the "intersection between G and H", which is (t, t) for KDB, KS and Butterfly, (\sqrt{t}, \sqrt{t}) for SS and $\frac{2\pi}{\pi-2}(t, t)$ for SU. Then, we run a sensitivity analysis on several values of the parameters $T \in \{100, 25, 10, 5, 1, 0.5, 0.05\}$ and $S \in \{0.1, 0.075, 0.05, 0.025, 0.01\}$, which corresponds to t_0 and σ_t as described in Table 8.1.

In [103], the authors consider as a stopping criterion the feasibility of the last non-linear parametric program in particular by considering the complementarity constraint by the minimum component-wise. Table 8.3 provides our results with this criterion. We provide elementary statistics by considering the percentage of success for each set of parameters. A problem is considered solved in this case if criteria a) and b) are satisfied.

First, we see that the method NL is giving decent results. It is not a surprise as was pointed out in [76]. Practical implementation of relaxation methods would select the best

Relaxation	NL	SS	KS	Butterfly
t_0	none	T^2	T	T
σ_t	none	S^2	S	S

Table 8.1: Parameter links among the methods

SS	KS	Butterfly
(\sqrt{t}, \sqrt{t})	(t, t)	(t, t)

Table 8.2: Link between the parameters of relaxation methods.

choice of parameters, so that we focus most of our attention to the line 'best'. In all cases, the relaxations manage to improve or at least equal the number of problems solved by NL. By using SNOPT, KS and the butterfly with $t = r^{3/2}$ get 1% of improvement and with IPOPT the butterfly with $t = r^{3/2}$ is the only one that attains 100%. The relaxation methods seem to make a significant improvement over NL with MINOS. In this case, it is clear that the butterfly methods benefit from the introduction of the parameter s and the method with $s = t, 2t = r$ is very competitive.

Our goal by solving (MPCC) is to compute a local minimum. The results using the local minimum criterion defined above as a measure of success are given in Table 8.4. Once again we provide a percentage of success. In comparison with Table 8.3, this new criterion is more selective. Independently of the solver, the relaxation methods with some correct choices of parameters provide improved results. Using SNOPT as a solver, the methods KS and butterfly give the highest number of success. The butterfly method with $t = r^{3/2}$ even improved the number of problems solved by SNOPT alone in average. In a similar way as in the previous experiment the butterfly method benefits of the introduction of the parameter s with the solver MINOS.

The methods (KDB) and (SU) have been discarded after preliminary results. We already point out in Example 5.3.1 one reason for the method (KDB). It is to be noted that both methods have received a special attention in [110] and [185] to solve the sub-problems that handle potential issues.

8.3 Benchmarking MPCCs solvers

Recent works have tried to propose methodologies to benchmark solvers and algorithms on some set of problems independently of the differences of technique and stopping criteria used. In [57], the authors proposed a methodology to compare non-linear programming solvers that uses the performance profile introduced in [56].

After giving some detailed about this methodology, we provide computational results of an adaptation of this methodology to the (MPCC).

Solver SNOPT	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
best	97.03	97.03	98.02	97.03	97.03	98.02
average	97.03	95.02	94.71	95.39	93.89	94.88
worst	97.03	91.09	91.09	92.08	91.09	91.09
std	0	1.64	2.09	1.50	1.97	2.42
Solver MINOS	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
best	89.11	94.06	93.07	90.10	95.05	89.11
average	89.11	91.20	90.89	83.54	91.06	81.92
worst	89.11	87.13	87.13	77.23	86.14	76.24
std	0	1.50	1.44	2.81	2.15	2.89
Solver IPOPT	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
best	98.02	99.01	98.02	99.01	98.02	100
average	98.02	98.16	96.38	94.03	93.89	94.79
worst	98.02	95.05	93.07	89.11	88.12	88.12
std	0	0.97	1.99	2.62	2.80	3.60

Table 8.3: Sensitivity analysis for MacMPEC test problems considering the feasibility of (MPCC). Results are a percentage of success. best : percentage of success with the best set of parameters, worst : percentage of success with the worst set of parameters, average : average percentage of success among the distribution of (T, s) , std : standard deviation

8.3.1 NLP benchmarking

The aim of this section is to define a test of convergence (up to some precision τ) for the problem

$$\min f(x) \text{ s.t } l \leq c(x) \leq u. \quad (8.1)$$

This problem is an equivalent formulation to (NLP) used earlier in this manuscript. This test of convergence verify the following essential properties:

- Does not need access inside the code of the solvers;
- Takes into account the absolute and relative error;
- Uses only the primal information given by the solvers;
- Is invariant of any scaling.

We use the following proximity measure

$$d(x, y) := \min \left(|x - y|, \frac{|x - y|}{|x| + |y|} \right).$$

The set of indices of active constraints is given here as

$$\mathcal{A}_\tau(x) := \{k : \min(d(c_k(x), l_k), d(c_k(x), u_k)) \leq \tau\}.$$

Solver SNOPT	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
best	92.08	94.06	96.04	96.04	97.03	96.04
average	92.08	90.78	91.17	92.08	90.04	92.33
worst	92.08	83.17	86.14	87.13	82.18	87.13
std	0	3.15	2.59	2.45	2.86	2.77
Solver MINOS	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
best	85.15	94.06	93.07	88.11	94.06	87.13
average	85.15	90.94	90.18	81.92	90.04	80.11
worst	85.15	87.13	86.14	76.23	85.15	74.26
std	0	1.50	1.62	2.65	2.31	2.95
Solver IPOPT	NL	SS	KS	$B_{(t=r)}$	$B_{(s=t, 2t=r)}$	$B_{(t=r^{3/2})}$
best	91.09	93.07	93.07	94.06	93.07	94.06
average	91.09	91.82	89.84	89.05	88.80	89.02
worst	91.09	90.10	86.14	84.16	84.16	81.19
std	0	1.14	2.19	3.09	2.72	3.86

Table 8.4: Sensitivity analysis for MacMPEC test problems considering the optimality of (MPCC). Results are a percentage of success. best : percentage of success with the best set of parameters, worst : percentage of success with the worst set of parameters, average : average percentage of success among the distribution of (T, s) , std : standard deviation

Since we want to use only primal information given by the solver, we need to compute ourselves the Lagrange multipliers by solving the following linear problem

$$\min_v \|\nabla f(x) - \nabla c(x)v\|_\infty \text{ s.t } v \in \mathcal{S}_\tau(x),$$

where $\mathcal{S}_\tau(x)$ is the cone of multipliers, defined such that whenever x is a stationary point then (x, ν) is a KKT-pair. Besides, we consider three measures:

1. feasibility:

$$\nu_f := \|\min(d(c(x), l), d(c(x), u)) \text{ if } c(x) \notin [c, u]\|_\infty \leq \|\text{median}(c(x) - l, 0, c(x) - u)\|_\infty;$$

2. complementarity: $\nu_c := \|\min(d(c(x), l), d(c(x), u)) \text{ pour } k \in \mathcal{A}_\tau(x)\|_\infty \leq \nu_f;$

3. stationarity: $\nu_s := \|d(\nabla f(x), \nabla c(x)\lambda(x, \tau))\|_\infty \leq \|\nabla f(x) - \nabla c(x)\lambda(x, \tau)\|_\infty.$

A point is said acceptable whenever $\nu_f \leq \tau$ and $\nu_s \leq \tau$.

Now, that we have defined what is a success, we consider Algorithm 3. In theory, the more precise the resolution is required the more time is needed to solve the problem. Thus, we may compare the resolution of a problem by comparing the time required to solve the last iteration.

Data:

$\tau := 10^{-6}$; $\epsilon_0 :=$ default parameter ;

1 **Begin** ;

2 Set $k := 0$;

3 **while** $(\nu_f^k > \tau \text{ or } \nu_s^k > \tau)$ and $\epsilon^k > 10^{-16}$ **do**

4 $(\nu_f^k, \nu_s^k) :=$ resolution of (8.1) ;

5 reduce ϵ^k ;

6 **return:** f_{opt} the optimal value at the solution x_{opt} or a decision of infeasibility or unboundedness.

Algorithm 3: Dolan & More methodology to solve (8.1).

8.3.2 Application to MacMPEC

We now apply the methodology of Dolan & More to the (MPCC). In a similar way as in Section 8.2 we add as a supplementary criterion the complementarity constraint ν_{comp} . The computation of the Lagrange multipliers is done in AMPL by using the solver gjh that computes gradient, hessian matrix of the objective function and the jacobian matrix of the constraints of an optimisation problem. At this point we notice that the result will most probably be different from those presented in the previous sections, since we consider the dual feasibility as a new criterion.

We first run Algorithm 3 to compare the three solvers SNOPT, MINOS and IPOPT on the test problems without any regularisation. Results are given in the performance profile in Figure 8.1.

We see from this figure that MINOS is the fastest solver, since it solves approximately 80% of the problems faster than the others. SNOPT manages to solve the largest number of problems that is slightly up to 90 %. We also notice that MINOS does not improve with more time.

We now apply the same approach with the relaxation methods KS, SS, $B_{(t=r^{2/3})}$ and $B_{(s=t, 2t=r)}$ for $T = 0.5$ and $s = 0.01$. The selected methods are the ones with the best results according to the study in Section 8.2, while this choice of parameters rather median. The tests are done with the three solvers and results are presented in Figure 8.2.

As a general remark we see on these results that despite its weaker theoretical properties the relaxation method (SS) solves most of the problems faster. Moreover, it also solves more problems than the other methods in most of the cases. The butterfly method with $t = r^{2/3}$ manages to solve more problems with the solver MINOS and the same number of problems with solver SNOPT. This method that has stronger convergence guarantees than (SS) is clear second in the number of problems solved in a short time compared to the butterfly relaxation with $(s=t, 2t=r)$ and the relaxation (KS). According to this study it seems that all the methods benefit from using SNOPT as a solver for the regularised sub-problems.

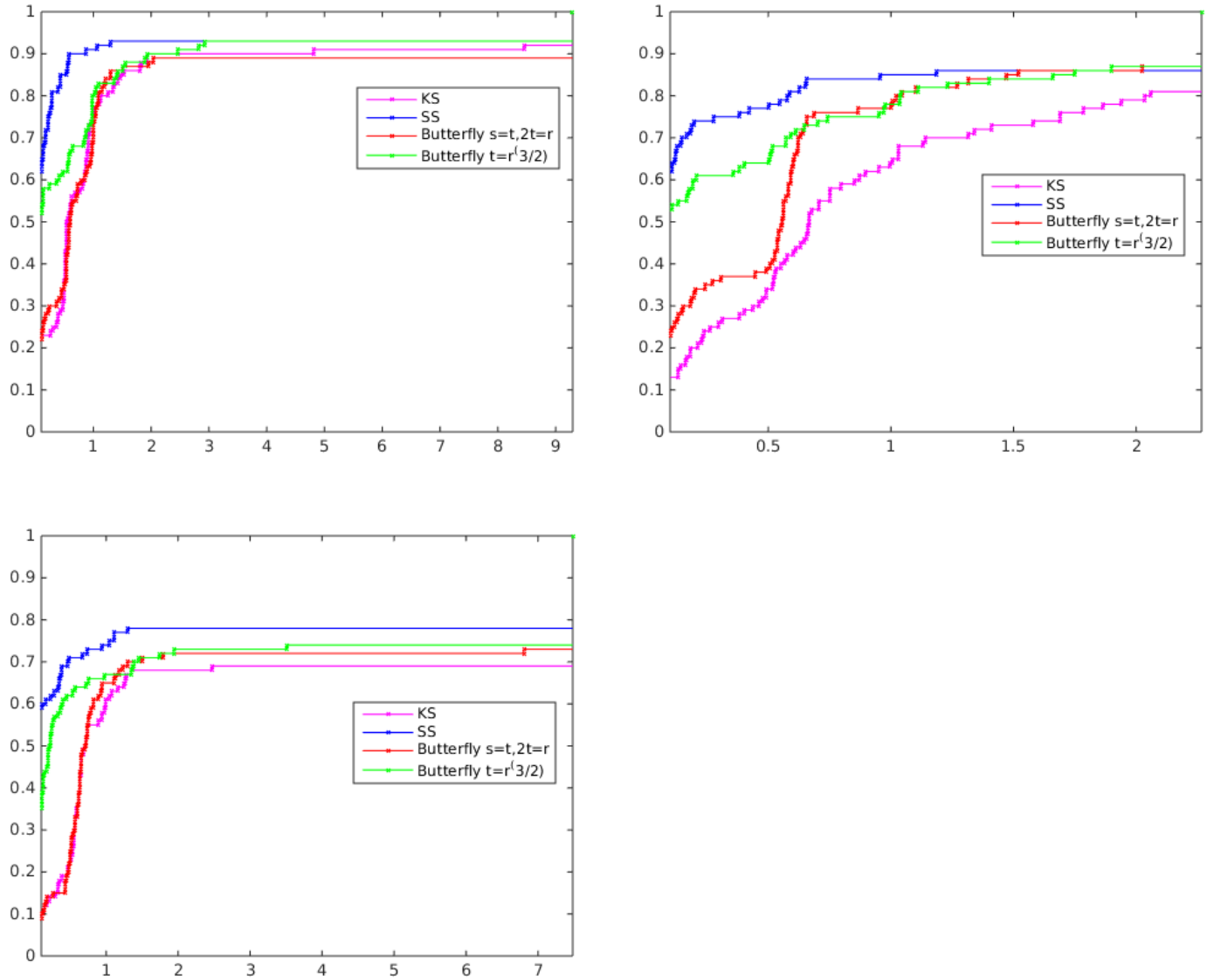


Figure 8.2: Performance profile of relaxations KS, SS, $B_{(t=r^{2/3})}$ and $B_{(s=t, 2t=r)}$ for $T = 0.5$ and $s = 0.01$. From the left to the right, up to above: relaxations with SNOPT, relaxations with MINOS and relaxations with IPOPT.

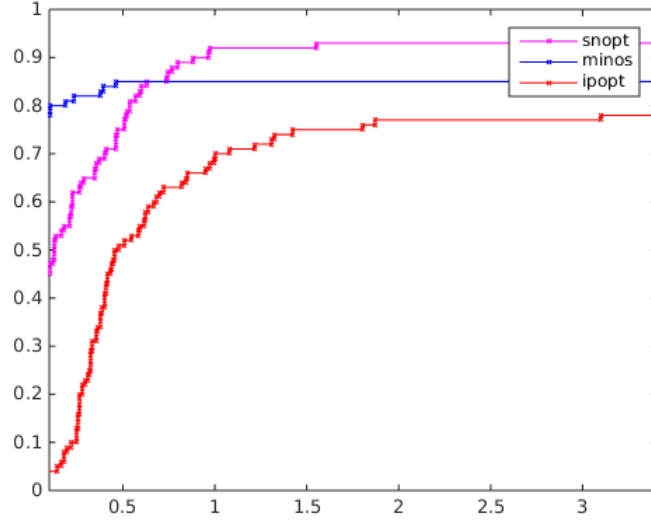


Figure 8.1: Performance profile that compares the time to solve the last iteration of Algorithm 3 for the test problems of the library MacMPEC.

8.4 An example of numerical difficulties

In this section, we illustrate the possible numerical difficulties that can arise by solving a (MPCC) with relaxation methods.

Example 8.4.1. *Consider the problem*

$$\begin{aligned}
 & \min_{x \in \mathbb{R}^4} \exp(-x_1^2 - x_2^2) + \exp(-x_3) \\
 & s.t. \quad x_3^2 \leq (x_1^2 + x_2^2 - 1)(x_1^2 + x_2^2 - 10) + x_4, \\
 & \quad \quad x_1^2 + x_2^2 - 10 \leq 0, x_4 \leq 0, \\
 & \quad \quad 0 \leq x_1^2 + x_2^2 - 1 \perp x_3(-x_1^2 - x_2^2 + 10) \geq 0.
 \end{aligned}$$

The feasible set is the union of two circles, $\{x \in \mathbb{R}^4 \mid x_3 = x_4 = 0, x_1^2 + x_2^2 = 1\}$ and $\{x \in \mathbb{R}^4 \mid x_3 = x_4 = 0, x_1^2 + x_2^2 = 10\}$. In this example, all the feasible points are local minima.

Let us now compute the stationary points of the problem. The gradient of MPCC-Lagrangian function equal to zero yields

$$\begin{aligned}
 -2 \exp(-x_1^2 - x_2^2)x_1 - 2\lambda_1^g x_1((x_1^2 + x_2^2 - 10) + (x_1^2 + x_2^2 - 1)) + 2\lambda_2^g x_1 - 2\lambda^G x_1 + 2\lambda^H x_1 x_3 &= 0, \\
 -2 \exp(-x_1^2 - x_2^2)x_2 - 2\lambda_1^g x_2((x_1^2 + x_2^2 - 10) + (x_1^2 + x_2^2 - 1)) + 2\lambda_2^g x_2 - 2\lambda^G x_2 + 2\lambda^H x_2 x_3 &= 0, \\
 -\exp(-x_3) + 2\lambda_1^g x_3 - \lambda^H(-x_1^2 - x_2^2 + 10) &= 0, \\
 -\lambda_1^g + 2\lambda_3^g x_4 &= 0.
 \end{aligned}$$

$x_1^0 \backslash x_2^0$	0	1	2	3
0	o	C	C	M
1	C	C	C	M
2	C	C	M	M
3	M	M	M	M

Table 8.5: Sensitivity analysis depending on the initial point $(x_1^0, x_2^0, 0, 0)$ on Example 8.4.1 by using the butterfly relaxation method $t = r^{3/2}$ with $T = 0.5, s = 0.01$ and SNOPT as a non-linear solver. **Legend:** o: error, C: circle $x^2 + y^2 = 1$, M: circle $x^2 + y^2 = 10$.

It is clear that necessarily $x_3 = x_4 = 0$, thus $\lambda_1^g = 0$ and

$$\begin{aligned} -\exp(-x_1^2 - x_2^2)x_1 + \lambda_2^g x_1 - \lambda^G x_1 &= 0, \\ -\exp(-x_1^2 - x_2^2)x_2 + \lambda_2^g x_2 - \lambda^G x_2 &= 0, \\ -1 &= \lambda^H(-x_1^2 - x_2^2 + 10). \end{aligned}$$

The third equality gives that $x_1^2 + x_2^2 \neq 10$, thus $\lambda_2^g = 0$. Furthermore, by the inequality constraints it is necessary that $x_1^2 + x_2^2 = 1$ and so either x_1 or x_2 is non-zero. It follows that $\lambda^H < 0$ and

$$-\exp(-1) = \lambda^G < 0.$$

To sum up, any point that satisfies $x_1^2 + x_2^2 = 1$ is C-stationary and is a local minimum, while any point that satisfies $x_1^2 + x_2^2 = 10$ is not stationary, despite the fact that it is a global minimum.

Up to this point, we may notice that the points that belong to the circle of centre 0 and radius $\sqrt{10}$ that are the global minima of the problem are sequentially M-stationary. Indeed, let $(x_1^k, x_2^k, x_3^k, x_4^k) = (0, \sqrt{10} - \frac{1}{k}, 0, 1/k)$, $\lambda^{H,k} = -\frac{1}{10-x_2^{k,2}} < 0$, $\lambda^{G,k} = 0$, $\lambda_1^{g,k} = \frac{-\exp(-x_1^{k,2}-x_2^{k,2})}{2\lambda_1^{g,k}(-x_2^{k,2}+\frac{11}{2})}$, $\lambda_2^{g,k} = 0$ and $2\lambda_3^{g,k} = k\lambda_1^{g,k}$.

We run Algorithm 2 with $T = 0.5$ and $s = 0.01$. Table 8.5 shows that the butterfly relaxation with $t = r^{3/2}$ may converge to both circles depending on the initial point. Note that for $(x_1^0, x_2^0) = (0, 0)$ the algorithm declares the problem infeasible. We do not give the results for other methods and other solvers here, but it has a similar behavior.

Those results may be surprising since it is proved that this method should converge to an M-stationary point and not less. So, in theory the algorithm should have some difficulties to compute Lagrange multiplier at this point. We run Algorithm 1 with methods NL, SS, KS and butterfly $t = r^{3/2}$ on this example. Results are presented in Table 8.6.

We see that independently of the solver all of the methods converge to a C-stationary point. In the cases of IPOPT and MINOS, the solvers exit with a success output and even more, they satisfy our local minimum criterion.

Those disturbing results are explained by Theorem 6.4.1 and related results in the literature that illustrate the fact that computing ϵ -stationary point may perturb the convergence

relaxation	solver	output (last iter.)	last parameter	x_1	x_2	x_3	x_4
NL	SNOPT	401	.	0.1929	0.9812	0.0117	0.0001
SS		401	$2.5e - 13$	0.1929	0.9812	0.0112	0.0001
KS		401	$5.0e - 15$	0.1930	0.9811	0.0112	0.0001
$B_{(t=r^{3/2})}$		401	$5.0e - 15$	0.1927	0.9812	0.0116	0.0001
NL	MINOS	0	.	0.7266	0.6870	0.0005	2.8389e-7
SS		0	$2.5e - 9$	0.7266	0.6870	0.0007	5.3595e-7
KS		0	$5.0e - 5$	0.7265	0.6870	0.0005	3.1903e-7
$B_{(t=r^{3/2})}$		0	$5.0e - 5$	0.7266	0.6869	0.0005	3.5130e-7
NL	IPOPT	0	.	0.1819	0.9833	0.0100	9.9999e-5
SS		0	0.25	0.1961	0.9805	0.0100	9.9999e-5
KS		0	0.5	0.1961	0.9805	0.0100	9.9999e-5
$B_{(t=r^{3/2})}$		0	0.5	0.1961	0.9805	0.0100	9.9999e-5

Table 8.6: Example 8.4.1 with initial point $(0.1, 0.5)$. output 0 is a success and output 401 is iteration limit message.

properties of these methods. We also point out here that local minima of the problem are not M-stationary and so by Theorem 4.1.1 MPCC-GCQ does not hold at these points. Moreover, this example does not contradict the Theorem 6.4.1 since in particular MPCC-CRSC is not verified at any feasible point of the problem.

Chapter 9

How to compute strong epsilon-stationary points

Chapter 7 introduces the new concept of strong epsilon-stationary point of the relaxed subproblems. In this chapter, we answer the non-trivial question of how to compute such an approximate stationary point. We present here a generalisation of the penalisation with active set scheme proposed in [110] and illustrate the fact that it has the desired property.

9.1 A penalized formulation

The following minimisation problem aims at finding $(x, s) \in \mathbb{R}^n \times \mathbb{R}^{2q}$ so that for $\rho > 0$

$$\begin{aligned} \min_{x,s} \Psi_\rho(x, s) &:= f(x) + \frac{1}{2\rho} \phi(x, s) \\ \text{s.t. } s_G &\geq -\bar{t}e, \quad s_H \geq -\bar{t}e, \quad \Phi(s_G, s_H; t) \leq 0, \end{aligned} \quad (P(x, s))$$

where ϕ is the quadratic penalty function

$$\phi(x, s) := \|\max(g(x), 0), h(x), G(x) - s_G, H(x) - s_H\|^2.$$

An adaptation of Theorem 7.4.1 gives the following result that validates the penalisation approach.

Theorem 9.1.1. *Given a decreasing sequence $\{\rho_k\}$ of positive parameters and $\{\epsilon_k\}$ a sequence of non-negative parameters that decrease to zero as $k \in \mathbb{N}$ goes to infinity. Assume that $\epsilon_k = o(\bar{t}_k)$. Let $\{x^k, \nu^k\}$ be a sequence of strong epsilon-stationary points of $(P(x, s))$ according to definition 7.4.1 for all $k \in \mathbb{N}$ with $x^k \rightarrow x^*$ such that MPCC-CRSC holds at x^* . If x^* is feasible, then it is an M-stationary point of (MPCC).*

Proof. Assuming that x^* is feasible for (MPCC), the result is a straightforward adaptation of Theorem 7.4.1. \square

Unfortunately, the strong assumption on the previous theorem that x^* must be feasible is hard to avoid. Indeed, it is a classical pitfall of penalisation methods in optimisation to possibly compute a limit point that minimises the linear combination of the constraints. In other words, we compute a point x^* infeasible that satisfies.

$$\begin{aligned} \sum_{i=1}^m h_i(x^*) \nabla h_i(x^*) + \sum_{i=1}^p \max(-g_i(x^*), 0) \nabla g_i(x^*) - \sum_{i=1}^q \max(G_i(x^*), 0) \nabla G_i(x^*) \\ - \sum_{i=1}^q \max(H_i(x^*), 0) \nabla H_i(x^*) = 0. \end{aligned}$$

This phenomenon has been well-known in non-linear programming methods for instance with filter methods. Such a point is sometimes called infeasible stationary point.

It is interesting to note that the way the penalty parameter ρ behave may provide some information on the stationarity of the limit point. Indeed, if we find a stationary point of the initial problem without driving ρ to zero, then we get an S-stationary point. This observation was introduced in [48] in the context of elastic interior-point for (MPCC) and then adapted to the penalisation technique from [110].

Theorem 9.1.2. *Let (x, s) be a strong epsilon-stationary point of $(P(x, s))$ with $\rho > 0$. If x is feasible for (MPCC), then x is an S-stationary point of (MPCC).*

This fact was already observed in Theorem 2 of [110] in a slightly weaker but similar framework. We do not repeat the proof here, but gives an interpretation of this result. It has been made clear in the proof of the convergence theorem, Theorem 7.4.1, that the case where x^* is an M-stationary point only occur if the sequence of multipliers $\{\nu^{\Phi, k}\}$ diverges. Therefore, it is to be expected that the penalty parameter must be driven to its limit to observe such phenomenon.

9.2 Active-set method for the penalised problem

We discuss here an active set method to solve the penalised problem $(P(x, s))$. This method is an extension of the method proposed in [110] to the general class of methods presented in previous sections. Let $\beta_{t, \bar{t}}(x, s)$ denote the measure of feasibility of $(R_t^s(x, s))$

$$\begin{aligned} \beta_{t, \bar{t}}(x, s) := & \|\max(g(x), 0)\|^2 + \|h(x)\|^2 + \|G(x) - s_G\|^2 + \|H(x) - s_H\|^2 + \\ & \|\max(-s_G + \bar{t}e, 0)\|^2 + \|\max(-s_H + \bar{t}e, 0)\|^2 + \|\max(-\Phi(s_G, s_H; t), 0)\|^2. \end{aligned}$$

Let $\mathcal{W}(s; t, \bar{t})$ be the set of active constraints among the constraints

$$s_G \geq -\bar{t}e, \quad s_H \geq -\bar{t}e, \quad \Phi(s_G, s_H; t) \leq 0.$$

We can be even more specific when for some $i \in \{1, \dots, q\}$ the relaxed constraint is active since

$$\Phi_i(s_G, s_H; t) = 0 \iff s_{H,i} = \psi(s_{G,i}; t) \text{ or } s_{G,i} = \psi(s_{H,i}; t).$$

Remark 9.2.1. *It is essential to note here that active constraints act almost like bound constraints since an active constraint means that for some $i \in \{1, \dots, q\}$ one (possibly both) of the two cases holds*

$$\begin{aligned} s_{G,i} &= -\bar{t} \text{ or } \psi(s_{H,i}; t), \\ &\text{or} \\ s_{H,i} &= -\bar{t} \text{ or } \psi(s_{G,i}; t). \end{aligned}$$

Considering the relaxation from Kanzow and Schwartz it is obviously a bound constraint since $\psi(s_{G,i}; t) = \psi(s_{H,i}; t) = t$. The butterfly relaxation gives $\psi(s_{G,i}; t) = t_1\theta_{t_2}(s_{H,i})$ and $\psi(s_{H,i}; t) = t_1\theta_{t_2}(s_{G,i})$. This is not a bound constraint but we can easily use a substitution technique. This key observation is another motivation to use a formulation with slack variables.

Furthermore, a careful choice of the function ψ may enable to compute an analytical solution of the following equation in α for given values of $s_G, s_H, d_{s_G}, d_{s_H}$:

$$s_{G,i} + \alpha d_{s_{G,i}} - \psi(s_{H,i} + \alpha d_{s_{H,i}}, t) = 0.$$

Solving exactly this equation is very useful while computing the largest step so that the iterates remain feasible along a given direction. For the butterfly relaxation with $\theta_{t_2}(x) = \frac{x}{x+t_2}$, the equation above is reduced to the following second order polynomial equation if $s_{H,i} + \alpha d_{s_{H,i}} \geq 0$:

$$(s_{H,i} + \alpha d_{s_{H,i}} + r)(s_{G,i} + \alpha d_{s_{G,i}}) - t(s_{H,i} + \alpha d_{s_{H,i}}) = 0. \quad (9.1)$$

Algorithm 4 presents an active-set scheme to solve $(P(x, s))$, which is described in depth in the sequel.

Data:

Initial data x^{k-1}, s^{k-1} ; precision $\epsilon > 0$;

$\sigma_\rho \in (0, 1)$ update in ρ , $\rho_0 > 0$ initial value of ρ , $\tau_{vio} \in (0, 1)$, ρ_{max} ;

Initial estimate of the multiplier ν^0 ;

Initial working set \mathcal{W}_0 of active constraints, A_0 initial matrix of gradients of active constraints.;

sat:=true

1 **Begin** ;

2 Set $j := 0$, $\rho := \rho_0$;

3 $(x^{k-1,0}, s^{k-1,0}) = \text{Projection of } (x^{k-1}, s^{k-1}) \text{ if not feasible for } (P(x, s))$;

4 **while** sat **and**

$(\|\nabla \mathcal{L}^1(x^{k,j}, s_G^{k,j}, s_H^{k,j}, \nu^j; t_k)\|_\infty^2 > \epsilon \|\nu^j\|_\infty \text{ or } \min(\nu^j) < 0 \text{ or } \beta_{t_k, \bar{t}_k}(x^{k,j}, s^{k,j}) > \epsilon)$
do

5 Substitution of the variables that are fixed by the active constraints in \mathcal{W}_j ;

6 Compute a feasible direction d^j that lies in the subspace defined by the working set \mathcal{W}_j (see (9.2)) and satisfy the conditions (SDD);

7 Compute $\bar{\alpha}$ the maximum non-negative feasible step along d^j

$$\bar{\alpha} := \sup\{\alpha : z^j + \alpha d^j \in \mathcal{F}_{t, \bar{t}}\}$$

Compute a step length $\alpha_j \leq \bar{\alpha}$ (see (9.1)) such that Armijo condition (9.3) holds ;

8 **if** $\alpha_j = \bar{\alpha}$ **then**

9 Update the working set $\rightarrow \mathcal{W}_{j+1}$ and compute A_{j+1} the matrix of gradients of active constraints
10 $(x^{k,j+1}, s^{k,j+1}) = (x^{k,j}, s^{k,j}) + \alpha_j d^j$;

11 j:=j+1 ;

12 **if** $\beta_{t_k, \bar{t}_k}(x^{k,j+1}, s^{k,j+1}) \geq \max(\tau_{vio}\beta_{t_k, \bar{t}_k}(x^{k,j}, s^{k,j}), \epsilon)$ **and** $\rho < \rho_{max}$ **then**

13 $\rho = \sigma_\rho \rho$

14 **else**

15 Determine the approximate multipliers $\nu^{j+1} = (\nu^G, \nu^H, \nu^\Phi)$ by solving

$$\nu^{j+1} \in \arg \min_{\nu \in \mathbb{R}^{\mathcal{W}_j}} \|A_{j+1}^T \nu - \nabla \Psi_\rho(z^j, t)\|^2$$

Relaxing rule : **if** $\exists i, \nu_i^{j+1} < 0$ **and** (satisfy (9.4) **or** $\alpha_j = 0$) **then**

16 Update of the working set \mathcal{W}_{j+1} (with an anti-cycling rule) ;

17 sat:= $\|d^j\| > \epsilon$

18 **return:** x^k, s^k or a decision of unboundedness.

Algorithm 4: Outer loop iteration : active set method for relaxed non-linear program $(P(x, s))$.

At each step, the set \mathcal{W}_j denotes the set of active constraints of the current iterate $x^{k,j}$. As pointed out in Remark 9.2.1, these active constraints fix some of the variables. Therefore, by replacing these fixed variables we can rewrite the problem in a subspace of the initial

domain. Thus, we consider the following minimisation problem

$$\begin{aligned}
\min_{(x,s) \in \mathbb{R}^n \times \mathbb{R}^{2q}} \quad & \tilde{\Psi}_\rho(x, s_{\mathcal{S}_G \cup \mathcal{S}_H}) := f(x) + \frac{1}{2\rho} \tilde{\phi}(x, s_{\mathcal{S}_G \cup \mathcal{S}_H}) \\
\text{s.t.} \quad & s_{G,i} \geq -\bar{t} \text{ for } i \in \mathcal{S}_G, \quad s_{H,i} \geq -\bar{t} \text{ for } i \in \mathcal{S}_H, \\
& \Phi_i(s_G, s_H; t) \leq 0 \text{ for } i \in \mathcal{S}_G \cup \mathcal{S}_H,
\end{aligned} \tag{9.1}$$

with

$$\begin{aligned}
\mathcal{I}_G &:= \{i \in \{1, \dots, q\} \mid s_{Gi} = -\bar{t}\}, \\
\mathcal{I}_H &:= \{i \in \{1, \dots, q\} \mid s_{Hi} = -\bar{t}\}, \\
\mathcal{I}_{GH}^{0+} &:= \{i \in \{1, \dots, q\} \mid s_{Hi} = \psi(s_G; t)\}, \\
\mathcal{I}_{GH}^{+0} &:= \{i \in \{1, \dots, q\} \mid s_{Gi} = \psi(s_H; t)\}, \\
\mathcal{I}_{GH}^{00} &:= \{i \in \{1, \dots, q\} \mid s_{Gi} = s_{Hi} = \psi(0; t)\}, \\
\mathcal{S}_G &:= \{1, \dots, q\} \setminus (\mathcal{I}_G \cup \mathcal{I}_{GH}^{0+} \cup \mathcal{I}_{GH}^{00}), \\
\mathcal{S}_H &:= \{1, \dots, q\} \setminus (\mathcal{I}_H \cup \mathcal{I}_{GH}^{+0} \cup \mathcal{I}_{GH}^{00}).
\end{aligned}$$

\mathcal{S}_G and \mathcal{S}_H respectively denote the set of indices where the variables s_G and s_H are free. Some of the fixed variables are replaced by a constant and others are replaced by an expression that depends on the free variables. It is rather clear from this observation that the use of slack variables is an essential tool to handle the non-linear bounds.

A major consequence here is that the gradient of Ψ in this subspace must be done with care using the composition of the derivative formula :

$$\nabla \tilde{\Psi}_\rho(x, s_{\mathcal{S}_G \cup \mathcal{S}_H}) = J_{\mathcal{W}_j}^T \nabla \Psi_\rho(x, s), \tag{9.2}$$

where $J_{\mathcal{W}_j}$ is an $(n + 2q) \times (n + \#\mathcal{S}_G + \#\mathcal{S}_H)$ matrix defined such that

$$J_{\mathcal{W}_j} := \begin{pmatrix} J_{\mathcal{W}_j}^x \\ J_{\mathcal{W}_j}^{s_G} \\ J_{\mathcal{W}_j}^{s_H} \end{pmatrix}.$$

The three sub-matrices used to define $J_{\mathcal{W}_j}$ are computed in the following way

$$\begin{aligned}
J_{\mathcal{W}_j}^x &= Id_n, \\
J_{\mathcal{W}_j,i}^{s_G} &= \begin{cases} e_i^T, & \text{for } i \in \mathcal{S}_G \\ \frac{\partial \psi(x;t)}{\partial x} \Big|_{x=s_H} e_i^T, & \text{for } i \in \mathcal{I}_{GH}^{+0} \\ 0, & \text{for } i \in (\{1, \dots, q\} \setminus \mathcal{S}_G) \setminus \mathcal{I}_{GH}^{+0}, \end{cases}
\end{aligned}$$

$$J_{\mathcal{W}_j,i}^{s_H} = \begin{cases} e_i^T, & \text{for } i \in \mathcal{S}_H \\ \frac{\partial \psi(x;t)}{\partial x} \Big|_{x=s_G} e_i^T, & \text{for } i \in \mathcal{I}_{GH}^{0+} \\ 0, & \text{for } i \in (\{1, \dots, q\} \setminus \mathcal{S}_H) \setminus \mathcal{I}_{GH}^{0+}, \end{cases}$$

where $J_{\mathcal{W}_j,i}$ denotes the i -th line of a matrix and e_i is a vector of zero whose i -th component is one. We may proceed in a similar way to compute the hessian matrix of $\tilde{\Psi}_\rho(x, s_{\mathcal{S}_G \cup \mathcal{S}_H}, t)$.

The feasible direction d^j is constructed to lie in a subspace defined by the working set and satisfying the sufficient-descent direction conditions:

$$\begin{aligned} \nabla \Psi(z^j) d^j &\leq -\mu_0 \|\nabla \Psi(z^j)\|^2, \\ \|d^j\| &\leq \mu_1 \|\nabla \Psi(z^j)\|, \end{aligned} \tag{SDD}$$

where $\mu_0 > 0$, $\mu_1 > 0$.

The step length $\alpha_j \in (0, \bar{\alpha}]$ is respectively computed to satisfy the Armijo and Wolfe conditions:

$$\Psi(z^j + \alpha_j d^j) \leq \Psi(z^j) + \tau_0 \alpha_j \nabla \Psi(z^j)^T d^j, \quad \tau_0 \in (0, 1), \tag{9.3}$$

$$\nabla \Psi(z^j + \alpha_j d^j)^T d^j \geq \tau_1 \nabla \Psi(z^j)^T d^j, \quad \tau_1 \in (\tau_0, 1). \tag{9.4}$$

If $\bar{\alpha}$ satisfies the Armijo condition (9.3), the active set strategy adds a new active constraint and the Wolfe condition (9.4) is not enforced. Otherwise, the Armijo condition requires $\alpha < \bar{\alpha}$ and the Wolfe condition is enforced.

The relaxing rule is given by the following scheme. Relax some constraint i_0 if and only if the two following conditions are fulfilled:

1. $\nu_{i_0}^j < 0$;
2. No constraint was added at the arrival point z^j and no constraint was deleted at the previous iteration.

The convergence result will rely on the fact that at least one step satisfying Wolfe's condition will be performed before removing an active constraint.

Convergence of the original algorithm has been shown in [110] and its adaptation to Algorithm 4 is the subject of future research. It is clear by construction of the Algorithm that the multipliers of the relaxed constraints satisfy exactly the complementarity condition and the relaxed constraints are satisfied exactly.

9.3 Numerical results

In what follows, we present a small set of instances to show the behaviour of our algorithm. Beforehand, we give some supplementary information regarding the implementation of Algorithm 2 and Algorithm 4. An extended butterfly relaxation has been used, which consider

$$\psi(z; r, s, t) = s + t\theta_r^1(z - s),$$

where $\theta_r^1(z - s) = \frac{z-s}{z-s+r}$ for $z \geq s$ and $\theta_r^1(z - s) = \frac{z-s}{r} - \frac{(z-s)^2}{2r^2}$ for $z < s$. The list of parameters used in the process is detailed in Table 9.1. It is to be noted that no attempt has been made to optimise the performance of the algorithm and the results come from a straightforward implementation of the algorithm in the JULIA programming language.

The direction d^j used in Algorithm 4 is computed through a Newton method. The computation of the constrained step length along this direction is computed through a back-tracking line search technique. A comparison between some methods to compute the descent direction has been conducted in [110]. We now introduce three examples and give the result

parameter	function	default value
Parameters for Algorithm 2		
(r_0, s_0, t_0)	relaxation parameter	(0.1,0.1,0.01)
$(\sigma_r, \sigma_s, \sigma_t)$	update of relaxation parameters	(0.1,0.1,0.01)
\bar{t}		
ϵ	sequence of precision	$\max(r, s, t)$
ϵ_∞	precision of (MPCC)	10^{-4}
Parameters for Algorithm 4		
ρ_0		1
σ_ρ		2
ρ_{max}		10^8
τ_{vio}		0.5
τ_0	Armijo parameter	0.25
τ_1	Wolfe parameter	0.9

Table 9.1: List of parameters for Algorithm 2 and their default values.

of our method. The first example is the continuation of Example 7.3.1, which illustrates a case where $(0, 0)$ is a weak-stationary point.

Example 9.3.1.

$$\begin{aligned} \min_{(x_1, x_2) \in \mathbb{R}^2} \quad & x_1 - x_2 \\ \text{s.t.} \quad & 0 \leq x_1 \perp x_2 \geq 0, \quad x_2 \leq 1. \end{aligned}$$

By starting from the initial point $(x_1, x_2) = (1.0, 1.0)$, our algorithm finds the solution after 9 iterations. Table 9.2 summarises the results of the algorithm. The last example is a continuation of Example 4.1.1 and illustrates a case where the solution $(0, 0, 0)^t$ is an M-stationary point.

Example 9.3.2.

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

Iter	x^k	s^k	$f(x^k)$	ρ	# inner Iter
0	(1.0,1.0)	-	0.0	-	-
1	(-0.2,1.1)	(-0.1,1.1)	-1.3	10.0	5
2	(-0.06,1.05)	(-0.06,1.05)	-1.11	20.0	5
3	(-0.026,1.025)	(-0.001,1.025)	-1.051	40.0	4
4	(-0.0126,1.0125)	(-0.0001,1.0)	-1.0251	80.0	4
5	(-0.0031,1.0031)	(-1.0e-5,1.0031)	-1.0062	320.0	5
6	(-0.0031,1.0015)	(-1.0e-6,1.0015)	-1.0031	640.0	5
7	(-0.0007,1.0)	(-1.0e-7,1.0007)	-1.0015	1280.0	4
8	(-0.0001,1.0002)	(-1.0e-8,1.002)	-1.0003	5120.0	5
9	(-9.7657e-5,1.0001)	(-1.0e-9,1.0001)	-1.0001	10240.0	4

Table 9.2: Solutions and optimal values of Example 9.3.1.

Iter	x^k	s^k	$f(x^k)$	ρ	# inner Iter
0	(0.5,1.0,0.0)	-	1.5	-	-
1	(0.2078,0.2140,0.8687)	(0.1053,0.2140)	-0.4468	20.0	4
2	(0.0600,0.0632,0.2591)	(0.0100,0.0632)	-0.1358	40.0	4
3	(0.0260,0.0275,0.1133)	(0.0010,0.0275)	-0.0598	80.0	4
4	(0.0063,0.0067,0.0277)	(0.0001,0.0067)	-0.0146	320.0	5
5	(0.0015,0.0016,0.0068)	(1.0000e-5,0.0016)	-0.0036	1280.0	5
6	(0.0007,0.0008,0.0034)	(1.0e-6,0.0008)	-0.0018	2560.0	4
7	(0.0003,0.0004,0.0017)	(1.0e-7,0.0004)	-0.0009	5120.0	4
8	(0.0001,0.0002,0.0008)	(1.0e-8,0.0002)	-0.0004	10240.0	4

Table 9.3: Solutions and optimal values of Example 9.3.2.

By starting from the initial point $(x_1, x_2, x_3) = (0.5, 1.0, 1.0)$, our algorithm finds the solution after 8 iterations. Table 9.3 summarises the results of the algorithm.

In both cases, we see that the algorithm manages to converge to the optimal solution, while satisfying exactly the complementarity constraint. Further research may focus on improvement of the algorithm and its application to some test problems coming from non-linear programs and (MPCC).

Concluding remarks

This manuscript presents different regularisation/relaxation methods for complementarity problems and related formulations. In particular, in Section 1.4 we introduce a family of smoothing functions that are central in the regularisations proposed through this document. Considering here a general family of smoothing functions allows us to extend existing results/methods from the literature or derive new ones.

By construction, these functions have been used to approximate the ℓ_0 -norm of a vector and therefore we use them in Chapter 2 for the problem of finding the sparsest solution over a polyhedron. This approach was a generalisation of other approaches in the literature and in particular was used to obtain a new exact penalisation theorem that gives a strong stopping criterion for an algorithm. We also suggest an implementation of this method that uses a successive linearisation algorithm. This implementation may thus be interpreted as a weighted ℓ_1 method.

Then, we used the smoothing functions to propose a regularisation technique for complementarity problems. This technique has been applied in Chapter 3 to (AVE), which is an equivalent formulation of (LCP). Additionally, to convergence of the method we provide an error bound and numerical results that motivates the validity of our approach.

Thereafter, we considered the extension of this technique to solve the (MPCC). One of the difficulty here is that we no longer try to just find a feasible point, but we optimise an objective function on a thin domain. The difficulties and our goal has been clarified in Chapter 4.

The direct application of the regularisation technique used in Chapter 3 has first been used in [93] and we give in Chapter 5 additional results on this method. As pointed out in this section, this technique does not satisfy our goal since it may converge to spurious solutions. We also mention a series of other methods with the same drawbacks.

Inspired by a method from [111] latter extended in [112], we propose in Chapter 6 a new regularisation method so-called butterfly relaxation method. This method that uses the θ 's functions in a different way can be seen as an extension to the existing methods. Besides, we prove that this new method has similar strong theoretical guarantees than methods in [111] and [112].

In order to continue the study of the family of methods that satisfy our goal, we consider in Chapter 7 a general framework of relaxation methods. We prove that existing methods including the butterfly relaxation belong to this framework and some new methods can also be derived using this framework. A key result in this chapter concern the convergence of sequence of approximate stationary points of the relaxation. Indeed, it has been pointed out in [115] and in Chapter 6 that by solving approximatively the sub-problems of the relaxation the strong theoretical properties are not longer guaranteed. In order to overcome this difficulty, we provide a specific definition of approximate stationary point that would be sufficient to keep the strong convergence of the methods. Moreover, we prove existence of these strong approximate stationary point in the neighborhood of a solution.

The last part of this manuscript focus on the numerical behavior of the regularisation methods applied to the (MPCC). In Chapter 8, we compare the various methods discussed in this document on the test problems MacMPEC, [123]. The sub-problems of the methods are here solved using well-known non-linear programming solver that are available via AMPL¹.

We run two experiments. First, after defining a success criterion we compare the percentage of success of the methods for various strategies of reducing the relaxation parameters. This experiment shows the validity of the relaxation approach that manages to improve the results given by just applying the non-linear solver to the (MPCC). Furthermore, the butterfly method shows to be very competitive. A second experiment uses the methodology proposed in [57] to benchmark non-linear programs. This methodology has the benefit to handle the choice of the precision used to solve the sub-problems. This experiment confirms the conclusion from the previous one and once again shows that the butterfly relaxation is competitive.

The last chapter of this document is focus on the implementation of relaxation methods. We defined in Chapter 7 a specific kind of approximate stationary point that are needed to keep the strong theoretical properties. We propose in Chapter 9 a penalisation-active set strategy that attains this goal. We discuss here the algorithm as well as preliminary results in JULIA on a small set of examples. We let detailed numerical application of this implementation to further research.

¹www.ampl.com

Appendices

Appendix A

A generalized direction in interior point method for monotone linear complementarity problems

This chapter is a paper submitted to Optimization Letters entitled:

A Generalized Direction in Interior Point Method for Monotone Linear Complementarity Problems

Authors:

M. Haddou, IRMAR-INSA Rennes

T. Migot, IRMAR-INSA Rennes

J. Omer, IRMAR-INSA Rennes

Abstract: In this paper, we present a new interior point method with full Newton step for monotone linear complementarity problems. The specificity of our method is to compute the Newton step using a modified system similar to that introduced by Darvay in [49]. We prove that this new method possesses the best known upper bound complexity for these methods. Moreover, we extend results known in the literature since we consider a general family of smooth concave functions in the Newton system instead of the square root.

Mathematics Subject Classification. 49M05 and 90C33 and 90C51

Keywords : concave functions ; interior-point methods ; linear programming ; linear complementarity problems ; polynomial time complexity

A.1 Introduction

In this paper we focus on the simplest interior point methods (IPMs) : the full Newton step interior point method [179], this method is one among the IPMs with the best worst-case complexity. Therefore, naturally any new try in the IPM framework must be validated on this method. In particular, we focus on this method applied to the monotone complementarity problem [119]. This problem is a generalization of the linear optimization problem and has been very popular in the literature due to its numerous applications [66]. In the literature of IPM, this problem has also been an active subject [119].

Darvay [49, 50], introduces a modification in the interior point method for finding search directions for linear optimization problems, based on an algebraic transformation of the central path. In particular, he applied the square root function to both sides of the centering equation, and he used Newton's method to obtain the new direction. He proved that the corresponding full Newton step algorithm has $O(\sqrt{n}L)$ iteration complexity. This new direction using the square root has become an active subject in the past few years [5, 4, 6, 18, 17, 20, 19, 194, 195, 193, 196]. Several authors generalized this approach to a wide class of optimization problems, for example for linear complementarity problems [5, 6, 18, 17, 20, 19, 196], convex quadratic programming [4], second-order cone optimization [195], semidefinite optimization [194] and symmetric cone optimization [19, 193].

Inspired by Darvay's new approach we introduce here a new class of IPMs by considering a large family of smooth concave functions instead of the square root. This new class of methods generalizes the classical path-following IPMs, since we obtain them as a special case. The technique presented here does not include Darvay's algorithm, but we can consider a smoother version that belongs to our family of methods. Our main contribution is that we prove that the algorithm with the new directions converges to a solution with the best known complexity for this family of methods.

In Section 2, we introduce the problem and our new directions. In Section 3, we show the polynomial complexity of our new class of methods.

Notations

Through this paper we will use the following notations: $\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\}$, $\mathbb{R}_{++}^n = \{x \in \mathbb{R}^n \mid x > 0\}$ and \mathbf{e} denotes the vector with all entries equal to one and whose dimension can be deduced from the context. Given two vectors $z, s \in \mathbb{R}^n$, we denote by $z^T s$ the usual scalar product and by zs the Hadamard product of two vectors, that is $zs = (z_i s_i)_{1 \leq i \leq n}$. Moreover, we extend this component-wise operation to the division of two vectors and to the square root, that is $\forall z \in \mathbb{R}^n, s \in (\mathbb{R} \setminus \{0\})^n$, $z/s = (z_i/s_i)_{1 \leq i \leq n}$ and $\forall z \in \mathbb{R}_{++}^n$, $\sqrt{z} = (\sqrt{z_i})_{1 \leq i \leq n}$ with $z_i \geq 0$ for all $i \in 1 \leq i \leq n$.

A.2 Preliminaries and Problem Settings

An LCP consists in finding $z, s \in \mathbb{R}^n$ such that for a square matrix $M \in \mathbb{R}^{n \times n}$ and a vector $q \in \mathbb{R}^n$

$$s = Mz + q, \quad z, s \geq 0, \quad zs = 0. \quad (\text{LCP})$$

A couple (z, s) such that $s = Mz + q$ is said to be feasible for the LCP if we have $z, s \geq 0$ and strictly feasible if $z, s > 0$. From now on, we use standard notation $\mathcal{F}_+ := \{(z, s) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n \mid s = Mz + q\}$ for the set of strictly feasible points of (LCP). In this paper we consider a monotone linear complementarity problem, i.e. an LCP where M is positive semi-definite. In this case the set of solutions of (LCP) is a convex set.

The main strategy of IPMs is to follow the central path $(z(\mu), s(\mu))$ for $\mu \in \mathbb{R}_{++}^n$, defined by

$$s = Mz + q, \quad z, s \geq 0, \quad zs = \mu. \quad (\text{LCP}_\mu)$$

The couples $(z(\mu), s(\mu))$ are also called μ -centers and define an homotopic path. The limit when $\mu \rightarrow 0$ satisfies the complementarity condition, and hence yields optimal solutions whenever the limit exists.

IPMs follow the central path approximately by solving an approached version of the non-linear system in (LCP_μ) for several values of μ . The main tool to solve such a system is the Newton method. A Newton step $(\Delta z, \Delta s)$ is given as the solution of the following linear system

$$\begin{cases} M\Delta z = \Delta s \\ z\Delta s + s\Delta z = \mu - zs \end{cases} \quad (\text{A.1})$$

There exists a wide variety of different IPMs that are based on this principle. In this paper we focus on the simplest IPM (see Algorithm 5): the full Newton step interior point method (FN-IPM).

Input: an accuracy parameter $\epsilon > 0$;
a sequence of update parameters $\{\theta^k\}$, $0 < \theta^k < 1 \ \forall k \in \mathbb{N}$;
initial values $(z^0, s^0) \in \mathcal{F}_+$, $\mu^0 = z^0 s^0$;

- 1 $z := z^0, s := s^0, \mu := \mu^0, k := 0$;
- 2 **while** $z^T s \geq n\epsilon$ **do**
- 3 $\mu := (1 - \theta^k)\mu$;
- 4 solve system (A.1) to find $(\Delta z, \Delta s)$;
- 5 $(z, s) := (z, s) + (\Delta z, \Delta s)$;
- 6 $k := k + 1$;

Algorithm 5: Full Newton step IPM (FN-IPM)

In [50], Darvay introduces a modification in (LCP_μ) by considering

$$s = Mz + q, \quad z, s \geq 0, \quad \varphi(zs) = \mu, \quad (\text{A.2})$$

where $\varphi : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$ is assumed to be a smooth function such that $\varphi(0) = 0$ and φ is defined by a component-wise extension of a real-valued function φ to \mathbb{R}^n , i.e. for $\mathbf{t} \in \mathbb{R}^n$ $\varphi(\mathbf{t}) = (\varphi(t_i))_{1 \leq i \leq n}$.

Darvay's method modifies the Newton steps. More precisely, the Newton step $(\Delta z, \Delta s)$ is given by the linear system

$$\begin{cases} M\Delta z = \Delta s \\ \varphi'(zs)(z\Delta s + s\Delta z) = \boldsymbol{\mu} - \varphi(zs) \end{cases} \quad (\text{A.3})$$

In this paper we consider functions $\varphi \in C^3(\mathbb{R}_+)$ such that $\varphi(0) = 0$, φ is increasing and concave, and $\varphi'''(t) \geq 0 \ \forall t \in \mathbb{R}_+$. This class of functions allows for a generalization of the classical IPMs, since we obtain the classical central path system (LCP $_{\mu}$) for $\varphi(t) = t$. The square root function does not belong to this family since it is not differentiable at 0, but we can build a smooth version with $\varphi_{\epsilon>0} : t \mapsto \sqrt{t+\epsilon} - \sqrt{\epsilon}$. We modify Algorithm 5 to solve (A.3) instead of (A.1) at step 4, and call the resulting algorithm φ -FN-IPM. The main result of this article is that φ -FN-IPM, converges to an ϵ -solution in at most $\mathcal{O}(\sqrt{n} \log(\frac{n}{\epsilon}))$ iterations. This upper bound is the best one known for the FN-IPM.

A.3 Polynomial Complexity

In this section, we consider the worst-case complexity of the φ -FN-IPM described in Algorithm 6 with $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ verifying

- (i) $\varphi(0) = 0$;
- (ii) $\varphi \in C^3([0, +\infty))$;
- (iii) $\varphi'(t) > 0, \ \forall t \geq 0$, i.e. φ is increasing;
- (iv) $\varphi''(t) \leq 0, \ \forall t \geq 0$, i.e. φ is concave;
- (v) $\varphi'''(t) \geq 0, \ \forall t \geq 0$.

These functions are invertible and can be extended in a smooth way for negative t by considering : $\varphi(t) = t\varphi'(0) + \frac{t^2}{2}\varphi''(0) + \frac{t^3}{6}\varphi'''(0)$. Function $\boldsymbol{\varphi} : \mathbb{R}^n \mapsto \mathbb{R}^n$ is then defined component-wisely: $\forall \mathbf{t} \in \mathbb{R}^n, \boldsymbol{\varphi}(\mathbf{t}) = (\varphi(t_i))_{1 \leq i \leq n}$.

One important characteristic of φ is the existence of a constant T defined by

$$-\varphi''(0) = T(\varphi'(0))^2. \quad (\text{A.4})$$

By conditions (iii) and (iv), $T \geq 0$ and $T = 0$ for $\varphi(t) = t$. It should be noted that arbitrary values of $T > 0$ can be achieved by scaling φ .

Note that $t \mapsto \sqrt{t}$, in the same way as any function $t \mapsto t^q, 0 < q < 1$, does not satisfy these hypotheses since it is not differentiable in 0. However we can consider a smooth version for $\epsilon > 0$ with $t \mapsto (t + \epsilon)^q - (\epsilon)^q$. As said in the introduction the classical method is given

Input: an accuracy parameter $\epsilon > 0$;
a sequence of update parameters $\{\theta^k\}$, $0 < \theta^k < 1 \ \forall k \in \mathbb{N}$;
initial values $(z^0, s^0) \in \mathcal{F}_+$, $\boldsymbol{\mu}^0 = z^0 s^0$;
1 $z := z^0, s := s^0, \boldsymbol{\mu} := \boldsymbol{\mu}^0, k := 0$;
2 **while** $\sum_{i=1}^n \varphi(z_i s_i) \geq n\epsilon$ **do**
3 $\boldsymbol{\mu} := (1 - \theta^k)\boldsymbol{\mu}$;
4 solve system (A.3) to find $(\Delta z, \Delta s)$;
5 $(z, s) := (z, s) + (\Delta z, \Delta s)$;
6 $k := k + 1$;

Algorithm 6: φ -Full Newton step IPM (φ -FN-IPM)

by $\varphi(t) = t$. Other examples are $\varphi : t \mapsto \log(1 + t)$ and functions constructed as in [94, 150]: for instance $\varphi : t \mapsto \frac{t}{t+1}$ and $\varphi : t \mapsto 1 - \exp(-t)$. It is interesting to note two latter functions φ are bounded. Moreover, notice that if a function φ satisfies all these hypotheses, then $t \mapsto \alpha\varphi(Ct)$ with $\alpha, C \in \mathbb{R}_{++}$ also satisfies these hypotheses.

The main result of this section (Theorem A.3.5) states the polynomial worst-case complexity of the φ -FN-IPM described in Algorithm 6. In order to achieve this result we define a measure of the proximity to the central path in Section A.3.1. Then, in Section A.3.2, Theorem A.3.1 estimates the error made at each Newton step. Sections A.3.3 and A.3.4 present conditions to ensure the correct behaviour of the algorithm: strict feasibility of the iterates (Theorem A.3.2) and quadratic convergence of the Newton process (Theorem A.3.3). Section A.3.5 provides the sequence of update parameters (Theorem A.3.4).

A.3.1 Proximity Measure

At each iteration, after updating parameter $\boldsymbol{\mu}$, we compute the Newton direction $(\Delta z, \Delta s)$ as a solution of system (A.3). Then we update the iterates with

$$z^+ = z + \alpha\Delta z \text{ and } s^+ = s + \alpha\Delta s . \quad (\text{A.5})$$

Note that here we consider a damping factor $\alpha \in [0, 1]$ to be more general. In this case we denote by $(\alpha\Delta z, \alpha\Delta s)$ the Newton step with length α and call it the α -Newton step. Then the full Newton step is given for $\alpha = 1$.

In order to measure the distance to the target on the central path we consider a proximity measure $\delta_\varphi(z, s, \boldsymbol{\mu})$ defined by

$$\delta_\varphi(z, s, \boldsymbol{\mu}) := \frac{1}{2} \left\| \frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(zs)} \left((v_\varphi(z, s, \boldsymbol{\mu})^{-1} - v_\varphi(z, s, \boldsymbol{\mu})) \right) \right\|_2 ,$$

with

$$v_\varphi(z, s, \boldsymbol{\mu}) := \sqrt{\frac{\boldsymbol{\varphi}(zs)}{\boldsymbol{\mu}}} \text{ and } v_\varphi(z, s, \boldsymbol{\mu})^{-1} := \sqrt{\frac{\boldsymbol{\mu}}{\boldsymbol{\varphi}(zs)}} .$$

We may omit the arguments of $v_\varphi(z, s, \boldsymbol{\mu})$ and $\delta_\varphi(z, s, \boldsymbol{\mu})$, when it is clear from the context. Notice that this proximity measure is a generalization of the one presented in [179], where the authors consider

$$\delta_{t \mapsto t} := \frac{1}{2} \left\| \sqrt{\frac{\boldsymbol{\mu}}{zs}} - \sqrt{\frac{zs}{\boldsymbol{\mu}}} \right\|_2.$$

Both proximity measures are equal for $\varphi(t) = t$. Moreover for any function φ , the two proximity measures are asymptotically similar (for $zs \downarrow 0$).

The following two lemmas link the iterates and the proximity measure.

Lemma A.3.1. *If $(z, s) \in \mathcal{F}_+$ and $\delta_\varphi \leq 1$, then $\varphi(zs) \leq 6\boldsymbol{\mu}$.*

Proof. Assume by contradiction that there exists $i \in \{1, \dots, n\}$ such that $\varphi(z_i s_i) > 6\boldsymbol{\mu}_i$. Since φ' is decreasing and $zs > 0$:

$$\begin{aligned} 2\delta_\varphi &\geq \|v_\varphi^{-1} - v_\varphi\|_2 = \left\| \frac{\sqrt{\boldsymbol{\mu}}}{\sqrt{\varphi(zs)}} - \frac{\sqrt{\varphi(zs)}}{\sqrt{\boldsymbol{\mu}}} \right\|_2 \geq \left| \frac{\sqrt{\boldsymbol{\mu}_i}}{\sqrt{\varphi(z_i s_i)}} - \frac{\sqrt{\varphi(z_i s_i)}}{\sqrt{\boldsymbol{\mu}_i}} \right| \\ &= \frac{\sqrt{\varphi(z_i s_i)}}{\sqrt{\boldsymbol{\mu}_i}} - \frac{\sqrt{\boldsymbol{\mu}_i}}{\sqrt{\varphi(z_i s_i)}} \\ &> \sqrt{6} - \frac{1}{\sqrt{6}} \approx 2,04, \end{aligned}$$

where the penultimate step comes from the increasing property of the function $x \mapsto x - 1/x$ on \mathbb{R}_{++} . This is in contradiction with $\delta_\varphi \leq 1$. \square \square

The following lemma gives bounds on δ_φ that depend on some constant $\Gamma(\|\boldsymbol{\mu}\|_\infty)$ defined as

$$\Gamma(\|\boldsymbol{\mu}\|_\infty) := \left(1 - \frac{\varphi''(0)\varphi^{-1}(6\|\boldsymbol{\mu}\|_\infty)}{\varphi'(\varphi^{-1}(6\|\boldsymbol{\mu}\|_\infty))} \right).$$

Lemma A.3.2. *Let $\delta = \frac{1}{2} \|v_\varphi^{-1} - v_\varphi\|_2$, then*

$$\delta \leq \delta_\varphi \leq \left(1 - \frac{\varphi''(0)\|zs\|_\infty}{\varphi'(\|zs\|_\infty)} \right) \delta.$$

Furthermore in a close neighbourhood of the central path, i.e. $\delta_\varphi \leq 1$, we have

$$\delta \leq \delta_\varphi \leq \Gamma(\|\boldsymbol{\mu}\|_\infty) \delta. \quad (\text{A.6})$$

Proof. By convexity of function φ' for all $i \in \{1, \dots, n\}$

$$\varphi'(z_i s_i) \geq \varphi'(0) + \varphi''(0) z_i s_i.$$

Then, for all $i \in \{1, \dots, n\}$

$$1 \leq \frac{\varphi'(0)}{\varphi'(z_i s_i)} \leq 1 - \frac{\varphi''(0) z_i s_i}{\varphi'(z_i s_i)} \leq \max_i \left(1 - \frac{\varphi''(0) z_i s_i}{\varphi'(z_i s_i)} \right).$$

Hence, by definition of δ and δ_φ

$$\delta \leq \delta_\varphi \leq \left(1 - \frac{\varphi''(0)\|zs\|_\infty}{\varphi'(\|zs\|_\infty)}\right) \delta.$$

The sharpest result when $\delta_\varphi \leq 1$ is deduced from Lemma A.3.1. \square \square

In the previous lemma, equation (A.6), we define Γ as a function of $\|\boldsymbol{\mu}\|_\infty$ which depends on the choice of φ . For $\varphi(t) = t$, we get $\Gamma(\|\boldsymbol{\mu}\|_\infty) = 1$ for all $\boldsymbol{\mu}$. Moreover, for any function φ , Γ is increasing with respect to $\|\boldsymbol{\mu}\|_\infty$, and converges to 1 as $\|\boldsymbol{\mu}\|_\infty \downarrow 0$. Moreover, in the course of the proof we showed that if $\delta_\varphi \leq 1$

$$\frac{\varphi'(0)}{\varphi'(z_i s_i)} \leq \Gamma(\|\boldsymbol{\mu}\|_\infty), \forall i \in \{1, \dots, n\}. \quad (\text{A.7})$$

This result will be useful in a future proof.

A.3.2 Error Bound of the Newton Step

We use the first order Taylor-Lagrange formula applied to φ in zs . There exists $\xi \in [z^+ s^+, zs]$ (or $\xi \in [zs, z^+ s^+]$ if $zs < z^+ s^+$) such that

$$\begin{aligned} \varphi(z^+ s^+) &= \varphi(zs) + \alpha \varphi'(zs)(z\Delta s + s\Delta z) \\ &\quad + \alpha^2 \Delta z \Delta s \varphi'(zs) + \frac{\varphi''(\xi)}{2} (z^+ s^+ - zs)^2, \end{aligned} \quad (\text{A.8})$$

with $(\Delta z, \Delta s)$ solution of (A.3). The update of $\boldsymbol{\mu}$ will be chosen such that $\varphi(z^+ s^+) < \varphi(zs)$ and thus $0 < z^+ s^+ < zs$. Therefore the error we make when we say that $\varphi(z^+ s^+)$ is the μ -center is

$$\eta(\alpha) := \alpha^2 \Delta z \Delta s \varphi'(zs) + \frac{\varphi''(\xi)}{2} (z^+ s^+ - zs)^2. \quad (\text{A.9})$$

The following sequence of lemmas aims to bound this error in terms of the proximity measure. Before doing so, we recall a useful lemma from [105].

Lemma A.3.3 (Lemma 5.1, [105]). *Let $(z, s) \in \mathcal{F}_+$ and $a \in \mathbb{R}^n$. Assume that matrix M is a positive semidefinite matrix. Let $(\Delta z, \Delta s)$ be the solution of*

$$\begin{cases} -M\Delta z + \Delta s = 0 \\ s\Delta z + z\Delta s = a \end{cases}.$$

Then

$$\begin{aligned} \|\Delta z \Delta s\|_1 &\leq C_1 \left\| \frac{a}{\sqrt{zs}} \right\|_2^2, \\ \|\Delta z \Delta s\|_2 &\leq C_2 \left\| \frac{a}{\sqrt{zs}} \right\|_2^2, \end{aligned}$$

$$\|\Delta z \Delta s\|_\infty \leq C_\infty \left\| \frac{a}{\sqrt{zs}} \right\|_2^2,$$

with $C_1 = 1/2$, $C_2 = 1/(2\sqrt{2})$ and $C_\infty = 1/4$.

Through the rest of this article we denote by C_p for $p \in \{1, 2, \infty\}$ the constants defined as

$$C_1 = \frac{1}{2}, \quad C_2 = \frac{1}{2\sqrt{2}} \text{ and } C_\infty = \frac{1}{4}.$$

Straightforward application of this lemma for $a = \frac{\mu - \varphi(zs)}{\varphi'(zs)}$ and $v_\varphi = \sqrt{\frac{\varphi(zs)}{\mu}}$ gives the following lemma.

Lemma A.3.4. *Let $(z, s) \in \mathcal{F}_+$, $(\Delta z, \Delta s)$ be the solution of (A.3) and $p \in \{1, 2, \infty\}$, then*

$$\|\Delta z \Delta s\|_p \leq C_p \left\| \frac{\sqrt{\varphi(zs)} \sqrt{\mu}}{\sqrt{zs} \varphi'(zs)} (v_\varphi^{-1} - v_\varphi) \right\|_2^2.$$

The next lemma will bound $\|\Delta z \Delta s \varphi'(zs)\|_p$ for $p \in \{1, 2, \infty\}$.

Lemma A.3.5. *Let $(z, s) \in \mathcal{F}_+$, $(\Delta z, \Delta s)$ be the solution of (A.3) and $p \in \{1, 2, \infty\}$, then*

$$\|\Delta z \Delta s \varphi'(zs)\|_p \leq \|\Delta z \Delta s \varphi'(0)\|_p \leq 4C_p \|\mu\|_\infty \delta_\varphi^2.$$

Proof. By concavity of φ we have that $\varphi(z_i s_i) \leq \varphi'(0) z_i s_i$, $\forall i$, so

$$\frac{\varphi(zs) \varphi'(0)}{zs (\varphi'(zs))^2} \leq \frac{(\varphi'(0))^2}{(\varphi'(zs))^2}. \quad (\text{A.10})$$

Furthermore for $p \in \{1, 2, \infty\}$ and using Lemma A.3.4, followed by (A.10)

$$\begin{aligned} \|\Delta z \Delta s \varphi'(zs)\|_p &\leq \|\Delta z \Delta s\|_p \varphi'(0) \\ &\leq C_p \left\| \frac{\sqrt{\varphi(zs)} \sqrt{\mu}}{\sqrt{zs} \varphi'(zs)} (v_\varphi^{-1} - v_\varphi) \right\|_2^2 \varphi'(0) \\ &\leq \left(\sum_{i=1}^n C_p \frac{\varphi(z_i s_i) \mu_i}{z_i s_i (\varphi'(z_i s_i))^2} (v_{\varphi_i}^{-1} - v_{\varphi_i})^2 \right) \varphi'(0) \\ &\leq \sum_{i=1}^n C_p \mu_i \left(\frac{\varphi'(0)}{\varphi'(z_i s_i)} \right)^2 (v_{\varphi_i}^{-1} - v_{\varphi_i})^2 \\ &\leq 4C_p \|\mu\|_\infty \delta_\varphi^2. \end{aligned}$$

□

□

Now we move to the main result which gives a bound for the complete error.

Lemma A.3.6. *Let $(z, s) \in \mathcal{F}_+$ and $(\Delta z, \Delta s)$ be the solution of (A.3). For $p \in \{2, \infty\}$ we have*

$$\begin{aligned} \|\eta\|_p &\leq (4C_p + 2T\varphi(\|zs\|_\infty)) \alpha^2 \|\boldsymbol{\mu}\|_\infty \delta_\varphi^2 \\ &\quad + \left(C_\infty \sqrt{\varphi(\|zs\|_\infty)} + \alpha C_p^2 \sqrt{\|\boldsymbol{\mu}\|_\infty} \delta_\varphi \right) 8\alpha^3 T \|\boldsymbol{\mu}\|_\infty^{3/2} \delta_\varphi^3 \end{aligned}$$

Proof. By equation (A.9) we have

$$\begin{aligned} \eta &= \alpha^2 \Delta z \Delta s \boldsymbol{\varphi}'(zs) + \frac{\alpha^2 \boldsymbol{\varphi}''(\xi)}{2} (z \Delta s + s \Delta z)^2 + \frac{\alpha^4 \boldsymbol{\varphi}''(\xi)}{2} (\Delta s \Delta z)^2 \\ &\quad + \alpha^3 \boldsymbol{\varphi}''(\xi) (\Delta s \Delta z) (z \Delta s + s \Delta z) \end{aligned}$$

Taking the p-norm for $p \in \{2, \infty\}$ and using triangle inequalities

$$\begin{aligned} \|\eta\|_p &\leq \alpha^2 \|\Delta z \Delta s \boldsymbol{\varphi}'(zs)\|_p + \alpha^2 \left\| \frac{\boldsymbol{\varphi}''(\xi)}{2} (z \Delta s + s \Delta z)^2 \right\|_p \\ &\quad + \alpha^4 \left\| \frac{\boldsymbol{\varphi}''(\xi)}{2} (\Delta s \Delta z)^2 \right\|_p + \alpha^3 \|\boldsymbol{\varphi}''(\xi) (\Delta s \Delta z) (z \Delta s + s \Delta z)\|_p. \end{aligned}$$

We now bound each term of the above right-hand side. First, Lemma A.3.5 gives

$$\|\Delta z \Delta s \boldsymbol{\varphi}'(zs)\|_p \leq 4C_p \|\boldsymbol{\mu}\|_\infty \delta_\varphi^2.$$

Using successively (A.4), $(\boldsymbol{\mu} - \boldsymbol{\varphi}(zs))^2 = (\boldsymbol{\mu} v_\varphi (v_\varphi^{-1} - v_\varphi))^2 = \boldsymbol{\mu} \varphi(zs) (v_\varphi^{-1} - v_\varphi)^2$ and $\varphi(z_i s_i) \leq \varphi(\|zs\|_\infty)$, $\forall i \in \{1, \dots, n\}$, we obtain step by step

$$\begin{aligned} \left\| \frac{\boldsymbol{\varphi}''(\xi)}{2} (z \Delta s + s \Delta z)^2 \right\|_p &\leq \left\| \frac{\boldsymbol{\varphi}''(0)}{2} (z \Delta s + s \Delta z)^2 \right\|_p \\ &= \frac{T}{2} \left\| \left(\frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(zs)} \right)^2 (\boldsymbol{\mu} - \boldsymbol{\varphi}(zs))^2 \right\|_p \\ &\leq \frac{T\varphi(\|zs\|_\infty) \|\boldsymbol{\mu}\|_\infty}{2} \left\| \left(\frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(zs)} (v_\varphi^{-1} - v_\varphi) \right)^2 \right\|_p \\ &\leq 2T\varphi(\|zs\|_\infty) \|\boldsymbol{\mu}\|_\infty \delta_\varphi^2. \end{aligned}$$

To bound the third term, we use Lemma A.3.5, equality (A.4) and $\|u^2\|_p \leq \|u\|_p^2 \forall u \in \mathbb{R}^n$

$$\begin{aligned} \left\| \frac{\boldsymbol{\varphi}''(\xi)}{2} (\Delta s \Delta z)^2 \right\|_p &\leq \frac{T}{2} \|(\boldsymbol{\varphi}'(0) \Delta s \Delta z)^2\|_p \leq \frac{T}{2} \|\boldsymbol{\varphi}'(0) \Delta s \Delta z\|_p^2 \\ &\leq \frac{T}{2} (\|\boldsymbol{\mu}\|_\infty 4C_p \delta_\varphi^2)^2 = 8TC_p^2 \|\boldsymbol{\mu}\|_\infty^2 \delta_\varphi^4. \end{aligned}$$

Finally, the definition of v_φ implies that

$$\begin{aligned}\|\varphi'(0)(z\Delta s + s\Delta z)\|_p &= \left\| \frac{\varphi'(0)}{\varphi'(zs)} \boldsymbol{\mu} v_\varphi(v_\varphi^{-1} - v_\varphi) \right\|_p \\ &\leq \left\| \frac{\varphi'(0)}{\varphi'(zs)} \sqrt{\|\boldsymbol{\mu}\|_\infty} \sqrt{\varphi(\|zs\|_\infty)} (v_\varphi^{-1} - v_\varphi) \right\|_p\end{aligned}$$

Using the above inequality, as well as $\|uw\|_p \leq \|u\|_\infty \|w\|_p$, $\forall (u, w) \in \mathbb{R}^n \times \mathbb{R}^n$, and (A.4), we get

$$\begin{aligned}\|\varphi''(\xi)(\Delta s \Delta z)(z\Delta s + s\Delta z)\|_p &\leq T \|\varphi'(0)(\Delta s \Delta z)\|_\infty \|\varphi'(0)(z\Delta s + s\Delta z)\|_p \\ &\leq 8TC_\infty \sqrt{\varphi(\|zs\|_\infty)} \|\boldsymbol{\mu}\|_\infty^{3/2} \delta_\varphi^3,\end{aligned}$$

which completes the proof. \square \square

In the special case where we are in a close neighbourhood of the central path we get an improved version of the result:

Theorem A.3.1. *Let $(z, s) \in \mathcal{F}_+$, $\delta_\varphi \leq 1$ and $(\Delta z, \Delta s)$ be the solution of (A.3) and $p \in \{2, \infty\}$, we have*

$$\|\eta\|_p \leq (4C_p + 12T\|\boldsymbol{\mu}\|_\infty)\alpha^2 \|\boldsymbol{\mu}\|_\infty \delta_\varphi^2 + (\sqrt{6}C_\infty + \alpha C_p^2 \delta_\varphi) \alpha^3 8T \|\boldsymbol{\mu}\|_\infty^2 \delta_\varphi^3 \quad (\text{A.11})$$

Proof. The proof is similar to the proof of Lemma A.3.6, but we use $\varphi(zs) \leq 6\boldsymbol{\mu}$ from Lemma A.3.1 instead of $\varphi(z_i s_i) \leq \varphi(\|zs\|_\infty)$, $\forall i$. \square \square

For instance, with $\alpha = 1$, using $\delta_\varphi^q \leq \delta_\varphi$ for $q \geq 1$, (A.11) becomes

$$\|\eta\|_\infty \leq \left(\|\boldsymbol{\mu}\|_\infty + \left(\frac{25 + 4\sqrt{6}}{2} \right) T \|\boldsymbol{\mu}\|_\infty^2 \right) \delta_\varphi^2.$$

For $\varphi(t) = t$ we get the same result as in [177]: $\|\eta\|_\infty \leq \|\boldsymbol{\mu}\|_\infty \delta_\varphi^2$.

A.3.3 Feasibility of the Newton Step

A Newton step is feasible (strictly feasible) if the couple (z^+, s^+) defined by (A.5) is feasible (strictly feasible).

Theorem A.3.2. *Let α be in $[0, 1]$ and $\delta_\varphi \leq 1$. The α -Newton step is strictly feasible for (z, s) if*

$$\alpha \delta_\varphi^2 < \frac{1}{1 + \left(\frac{25 + 4\sqrt{6}}{2} \right) T \|\boldsymbol{\mu}\|_\infty}. \quad (\text{A.12})$$

Note that for $\alpha = 1$, the above condition makes the full Newton step be strictly feasible.

For $\alpha = 1$ condition (A.12) implies that the proximity measure, δ_φ , must be less than 1, which will not be a restrictive assumption.

Proof. For $\alpha = 0$ the result trivially holds. Therefore, we can assume without loss of generality that $\alpha \in]0, 1]$. Let $\beta \in]0, \alpha]$, and define $z^\beta := z + \beta\Delta z$ and $s^\beta := s + \beta\Delta s$, where $(\Delta z, \Delta s)$ is the solution of (A.3). The proof rests upon a continuity argument. Indeed, we assume $z, s > 0$ so $\varphi(zs) > 0$ and, using equations (A.3) and (A.8):

$$\begin{aligned}\varphi(z^\beta s^\beta) &= \varphi(zs) + \beta\varphi'(zs)(z\Delta s + s\Delta z) + \eta(\beta) \\ &= \varphi(zs)(1 - \beta) + \beta \left(\boldsymbol{\mu} + \frac{\eta(\beta)}{\beta} \right) \\ &\geq \varphi(zs)(1 - \beta) + \beta \left(\boldsymbol{\mu} - e \frac{\|\eta(\beta)\|_\infty}{\beta} \right)\end{aligned}$$

which is positive for all $z, s > 0$ if $\|\boldsymbol{\mu}\|_\infty > \|\eta(\beta)\|_\infty/\beta$. Using Lemma A.3.6 this condition holds if

$$\begin{aligned}\|\boldsymbol{\mu}\|_\infty &> (4C_\infty + 2T\varphi(\|zs\|_\infty))\beta\|\boldsymbol{\mu}\|_\infty\delta_\varphi^2 \\ &\quad + \left(C_\infty\sqrt{\varphi(\|zs\|_\infty)} + \beta C_\infty^2\sqrt{\|\boldsymbol{\mu}\|_\infty}\delta_\varphi \right) 8\beta^2 T\|\boldsymbol{\mu}\|_\infty^{3/2}\delta_\varphi^3.\end{aligned}$$

The right-hand side is increasing with respect to β , so it is sufficient to verify

$$\begin{aligned}\|\boldsymbol{\mu}\|_\infty &> (4C_\infty + 2T\varphi(\|zs\|_\infty))\alpha\|\boldsymbol{\mu}\|_\infty\delta_\varphi^2 \\ &\quad + \left(C_\infty\sqrt{\varphi(\|zs\|_\infty)} + \alpha C_\infty^2\sqrt{\|\boldsymbol{\mu}\|_\infty}\delta_\varphi \right) 8\alpha^2 T\|\boldsymbol{\mu}\|_\infty^{3/2}\delta_\varphi^3.\end{aligned}$$

Therefore, since $\delta_\varphi \leq 1$, using Lemma A.3.1, it suffices to have

$$1 - \delta_\varphi^2 \alpha \left(4C_\infty + \left(12 + 8\sqrt{6}C_\infty + 8C_\infty^2 \right) T\|\boldsymbol{\mu}\|_\infty \right) > 0,$$

which corresponds to our assumption. It follows that for all $\beta \in [0, \alpha]$, $\varphi(z^\beta s^\beta) > 0$. By continuity of φ , this implies that none of z^β or s^β vanish for $\beta \in [0, \alpha]$, so the result follows. \square \square

A.3.4 Quadratic Decrease of the Proximity Measure

The Newton method is known to behave well in a close neighbourhood of the solution. The following theorem states a condition on the proximity measure, $\delta_\varphi := \delta_\varphi(z, s, \boldsymbol{\mu})$, that ensures a quadratic convergence of the Newton step. We denote by $\overline{\delta}_\varphi := \delta_\varphi(z^+, s^+, \boldsymbol{\mu})$ and $\overline{v}_\varphi := v_\varphi(z^+, s^+, \boldsymbol{\mu})$ the proximity measure and the function v_φ after the Newton step.

Theorem A.3.3. *Let $(z, s) \in \mathcal{F}_+$ and (z^+, s^+) be defined as in (A.5) for $\alpha = 1$. Let $\Gamma(\|\boldsymbol{\mu}\|_\infty)$ defined in (A.6) and*

$$Q(\|\boldsymbol{\mu}\|_\infty) := \frac{1 - (\Gamma(\|\boldsymbol{\mu}\|_\infty)(\sqrt{2} + (13 + 2\sqrt{6})T\|\boldsymbol{\mu}\|_\infty))^2/4}{1 + \left(\frac{25+4\sqrt{6}}{2} \right) T\|\boldsymbol{\mu}\|_\infty}. \quad (\text{A.13})$$

If $\delta_\varphi^2 \leq Q(\|\boldsymbol{\mu}\|_\infty)$, then $\overline{\delta}_\varphi \leq \delta_\varphi^2$.

Proof. Let $\Gamma(\|\boldsymbol{\mu}\|_\infty)$ be defined as in Section A.3.1:

$$\overline{\delta_\varphi} \leq \frac{\Gamma(\|\boldsymbol{\mu}\|_\infty)}{2} \|(\overline{v_\varphi})^{-1} - (\overline{v_\varphi})\|_2 = \frac{\Gamma(\|\boldsymbol{\mu}\|_\infty)}{2} \|(\overline{v_\varphi})^{-1}(\mathbf{e} - (\overline{v_\varphi})^2)\|_2 \quad (\text{A.14})$$

$$\leq \frac{\Gamma(\|\boldsymbol{\mu}\|_\infty)}{2} \left\| \frac{\frac{\eta(\alpha)}{\boldsymbol{\mu}}}{\sqrt{\mathbf{e} + \frac{\eta(\alpha)}{\boldsymbol{\mu}}}} \right\|_2 \quad (\text{A.15})$$

$$\leq \frac{\Gamma(\|\boldsymbol{\mu}\|_\infty) \left\| \frac{\eta(\alpha)}{\boldsymbol{\mu}} \right\|_2}{2\sqrt{1 - \left\| \frac{\eta(\alpha)}{\boldsymbol{\mu}} \right\|_\infty}}. \quad (\text{A.16})$$

Let $\delta_\varphi^2 \leq Q(\|\boldsymbol{\mu}\|_\infty) \leq 1$, using Theorem A.3.1 with a full step, i.e. $\alpha = 1$, it follows

$$\overline{\delta_\varphi} \leq \delta_\varphi^2 \Gamma(\|\boldsymbol{\mu}\|_\infty) \frac{4C_2 + (12 + 8\sqrt{6}C_\infty + 8C_2^2)T\|\boldsymbol{\mu}\|_\infty}{2\sqrt{1 - \delta_\varphi^2(4C_\infty + (12 + 8\sqrt{6}C_\infty + 8C_\infty^2)T\|\boldsymbol{\mu}\|_\infty)}}.$$

So, $\overline{\delta_\varphi} \leq \delta_\varphi^2$ if $\delta_\varphi^2 \leq Q(\|\boldsymbol{\mu}\|_\infty)$. □ □

Considering $\varphi(t) = t$, the condition of Theorem A.3.3 becomes the same as in [177]: $\delta_\varphi^2 \leq 1/2$.

Remark A.3.1. *The condition in Theorem A.3.3 implies the condition in Theorem A.3.2. So, if the iterates locate in the neighbourhood of quadratic convergence, the full Newton step will provide strictly feasible iterates.*

Remark A.3.2. *Notice that since the proximity measure is always non-negative, the condition from Theorem A.3.3 can hold only when μ is sufficiently small, i.e. when*

$$\Gamma(\|\boldsymbol{\mu}\|_\infty) \left(\sqrt{2} + (13 + 2\sqrt{6})T\|\boldsymbol{\mu}\|_\infty \right) \leq 2.$$

This is not a restrictive assumption, because we can always scale a given initial point so that it satisfies this condition.

A.3.5 Updating Parameter Strategy

The sequence of parameters $\{\theta^k\}$ must be chosen such that the iterates remain strictly feasible and satisfy the condition of Theorem A.3.3. In this section, Proposition A.3.1 gives an upper bound on the proximity measure after an update on $\boldsymbol{\mu}$, that is $\boldsymbol{\mu}^+ = \boldsymbol{\mu}(1 - \theta^k)$, and then Theorem A.3.4 describes how to build the sequence $\{\theta^k\}$.

First, we provide an upper bound of the proximity measure after an update of the parameter, denoted $\delta_\varphi^+ := \delta_\varphi(z^+, s^+, \boldsymbol{\mu}(1 - \theta^k))$, in terms of the update θ^k , the proximity measure before this update, denoted $\overline{\delta_\varphi} := \delta_\varphi(z^+, s^+, \boldsymbol{\mu})$, and the proximity measure before the Newton step, denoted $\delta_\varphi := \delta_\varphi(z, s, \boldsymbol{\mu})$. The computation of this upper bound is based on the following lemma.

Lemma A.3.7. Let $(z, s) \in \mathcal{F}_+$ and $(\Delta z, \Delta s)$ be the solution of (A.3). Then,

$$\sum_{i=1}^n \frac{\varphi(z_i^+ s_i^+)}{\boldsymbol{\mu}_i} \leq n + 2\delta_\varphi.$$

Proof. We first show that

$$\left\| \frac{\Delta z \Delta s \boldsymbol{\varphi}'(zs)}{\boldsymbol{\mu}} \right\|_1 \leq 2\delta_\varphi. \quad (\text{A.17})$$

For this, we apply Lemma A.3.3 with $\Delta z/\sqrt{\boldsymbol{\mu}}$, $\Delta s/\sqrt{\boldsymbol{\mu}}$ and $\mathbf{a} = \frac{\boldsymbol{\mu} - \boldsymbol{\varphi}(zs)}{\sqrt{\boldsymbol{\mu} \boldsymbol{\varphi}'(zs)}}$ to obtain

$$\left\| \frac{\Delta z \Delta s}{\boldsymbol{\mu}} \right\|_1 \leq \frac{1}{2} \left\| \frac{\boldsymbol{\mu} - \boldsymbol{\varphi}(zs)}{\sqrt{\boldsymbol{\mu} \boldsymbol{\varphi}'(zs)} \sqrt{zs}} \right\|_2^2.$$

Then, we get (A.17) by following the same steps as in the proof of Lemma A.3.5.

From equation (A.8) and by concavity of φ , $\boldsymbol{\varphi}(z^+ s^+)/\boldsymbol{\mu} \leq \mathbf{e} + \Delta z \Delta s \boldsymbol{\varphi}'(zs)/\boldsymbol{\mu}$. So, $\sum_{i=1}^n \boldsymbol{\varphi}(z_i^+ s_i^+)/\boldsymbol{\mu}_i \leq \sum_{i=1}^n 1 + |\Delta z_i \Delta s_i \boldsymbol{\varphi}'(z_i s_i)/\boldsymbol{\mu}_i| \leq n + 2\delta_\varphi$. \square \square

Proposition A.3.1. Let $\overline{v}_\varphi := v_\varphi(z^+, s^+, \boldsymbol{\mu})$ and $v_\varphi^+ := (z^+, s^+, \boldsymbol{\mu}^+)$, where $\boldsymbol{\mu}^+ := (1 - \theta^k)\boldsymbol{\mu}$. Then,

$$(\delta_\varphi^+)^2 \leq (1 - \theta^k)(\overline{\delta}_\varphi)^2 + \frac{\Gamma(\|\boldsymbol{\mu}\|_\infty)^2}{4(1 - \theta^k)} (n(\theta^k)^2 + (4\theta^k - 2(\theta^k)^2)\delta_\varphi^2).$$

Furthermore, assuming that $\delta_\varphi^2 \leq Q(\|\boldsymbol{\mu}\|_\infty)$ yields

$$(\delta_\varphi^+)^2 \leq (1 - \theta^k)Q(\|\boldsymbol{\mu}\|_\infty)^2 + \frac{\Gamma(\|\boldsymbol{\mu}\|_\infty)^2}{4(1 - \theta^k)} (n(\theta^k)^2 + (4\theta^k - 2(\theta^k)^2)Q(\|\boldsymbol{\mu}\|_\infty)).$$

Proof. Noticing that $v_\varphi^+ = \overline{v}_\varphi/\sqrt{1 - \theta^k}$, it follows that

$$\begin{aligned} (\delta_\varphi^+)^2 &= \frac{1}{4} \sum_{i=1}^n \left(\frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(z_i^+ s_i^+)} \right)^2 ((v_{\varphi_i}^+)^{-2} + (v_{\varphi_i}^+)^2 - 2) \\ &= \frac{1}{4} \sum_{i=1}^n \left(\frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(z_i^+ s_i^+)} \right)^2 \left((1 - \theta^k)(\overline{v}_{\varphi_i})^{-2} + \frac{\overline{v}_{\varphi_i}^2}{(1 - \theta^k)} - 2 \right) \\ &= (1 - \theta^k) \left(\frac{1}{4} \sum_{i=1}^n \left(\frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(z_i^+ s_i^+)} \right)^2 ((\overline{v}_{\varphi_i})^{-2} + \overline{v}_{\varphi_i}^2 - 2) \right) \\ &\quad + \frac{1}{4} \sum_{i=1}^n \left(\frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(z_i^+ s_i^+)} \right)^2 \left(-2\theta^k + \frac{2\theta^k - (\theta^k)^2}{1 - \theta^k} \overline{v}_{\varphi_i}^2 \right) \\ &= (1 - \theta^k)(\overline{\delta}_\varphi)^2 + \frac{1}{4} \sum_{i=1}^n \left(\frac{\boldsymbol{\varphi}'(0)}{\boldsymbol{\varphi}'(z_i^+ s_i^+)} \right)^2 \left(-2\theta^k + \frac{2\theta^k - (\theta^k)^2}{1 - \theta^k} \overline{v}_{\varphi_i}^2 \right) \end{aligned}$$

Using successively equation (A.7), Lemma A.3.7 and $\Gamma(\|\boldsymbol{\mu}^+\|_\infty) \leq \Gamma(\|\boldsymbol{\mu}\|_\infty)$, we obtain

$$\begin{aligned} (\delta_\varphi^+)^2 &\leq (1 - \theta^k)(\overline{\delta_\varphi})^2 + \Gamma(\|\boldsymbol{\mu}^+\|_\infty)^2 \left(-\frac{2\theta^k}{4}n + \frac{2\theta^k - (\theta^k)^2}{4(1 - \theta^k)}(n + 2\delta_\varphi^2) \right) \\ &\leq (1 - \theta^k)(\overline{\delta_\varphi})^2 + \Gamma(\|\boldsymbol{\mu}\|_\infty)^2 \left(\frac{n(\theta^k)^2}{4(1 - \theta^k)} + \frac{2\theta^k - (\theta^k)^2}{2(1 - \theta^k)}\delta_\varphi^2 \right). \end{aligned}$$

This proves the first part of the proposition.

Now, assuming that $\delta_\varphi \leq Q(\|\boldsymbol{\mu}\|_\infty)$ allows us to use Theorem A.3.3 and so $\overline{\delta_\varphi} \leq \delta_\varphi^2 \leq Q(\|\boldsymbol{\mu}\|_\infty)^2$ gives the result. \square

We conclude this section by a description of the choice of the update parameters θ^k .

Theorem A.3.4. *Let $\boldsymbol{\mu}$ be such that $Q(\|\boldsymbol{\mu}\|_\infty) > 0$ and $(z, s) \in \mathcal{F}_+$ such that $\delta_\varphi^2 \leq Q(\|\boldsymbol{\mu}\|_\infty)$. Define θ^k as*

$$\theta^k = \frac{-b + \sqrt{b^2 - 4ac}}{2a},$$

with

$$\begin{aligned} a &= \Gamma(\|\boldsymbol{\mu}\|_\infty)^2 n - 2\Gamma(\|\boldsymbol{\mu}\|_\infty)^2 Q(\|\boldsymbol{\mu}\|_\infty) + 4Q(\|\boldsymbol{\mu}\|_\infty)^2, \\ b &= 4\Gamma(\|\boldsymbol{\mu}\|_\infty)^2 Q(\|\boldsymbol{\mu}\|_\infty) - 8Q(\|\boldsymbol{\mu}\|_\infty)^2 + 4Q(\|\boldsymbol{\mu}\|_\infty), \\ c &= 4Q(\|\boldsymbol{\mu}\|_\infty)^2 - 4Q(\|\boldsymbol{\mu}\|_\infty). \end{aligned}$$

The proximity measure $\delta_\varphi^+ := \delta_\varphi(z^+, s^+, \boldsymbol{\mu}(1 - \theta^k))$ satisfies the conditions of feasibility in Theorem A.3.2 and quadratic convergence of the Newton step in Theorem A.3.3.

By definition $Q(\|\boldsymbol{\mu}\|_\infty) < 1$, thus c is negative. Furthermore, for n sufficiently large a is positive and so $b^2 - 4ac$ and θ^k are positive.

We would also like to point out that the value of θ^k is of order $1/\sqrt{n}$ for n large. This observation is fundamental considering the complexity of the algorithm.

Proof. As pointed out earlier in Remark A.3.1 the condition of Theorem A.3.2 is weaker than the condition of Theorem A.3.3. Thus, it is sufficient to satisfy the latter condition to ensure strict feasibility of the iterates.

According to the condition of Theorem A.3.3, after an update of $\boldsymbol{\mu}$, i.e. $\boldsymbol{\mu}^+ = (1 - \theta^k)\boldsymbol{\mu}$, the proximity measure δ_φ^+ must satisfy

$$(\delta_\varphi^+)^2 \leq Q(\|\boldsymbol{\mu}^+\|_\infty).$$

As Q is decreasing with respect to $\|\boldsymbol{\mu}\|_\infty$, it is sufficient to ensure that

$$(\delta_\varphi^+)^2 \leq Q(\|\boldsymbol{\mu}\|_\infty).$$

By Proposition A.3.1 in the case $\delta_\varphi^2 \leq Q(\|\boldsymbol{\mu}\|_\infty)$, we can choose any θ^k satisfying

$$(1 - \theta^k)Q(\|\boldsymbol{\mu}\|_\infty)^2 + \frac{\Gamma(\|\boldsymbol{\mu}\|_\infty)^2}{4(1 - \theta^k)} (n(\theta^k)^2 + (4\theta^k - 2(\theta^k)^2)Q(\|\boldsymbol{\mu}\|_\infty)) \leq Q(\|\boldsymbol{\mu}\|_\infty).$$

Therefore, it is sufficient to choose $\theta^k > 0$ such that

$$\theta^k \leq \frac{-b + \sqrt{b^2 - 4ac}}{2a},$$

with a, b and c defined as in the statement of the theorem. □ □

Remark A.3.3. *For a more explicit characterization of θ^k , we can study its asymptotical behaviour. By definition of Γ and Q ,*

$$\lim_{\mu \rightarrow 0} Q(\|\mu\|_\infty) = 1/2 \text{ and } \lim_{\mu \rightarrow 0} \Gamma(\|\mu\|_\infty) = 1,$$

$$\text{so } \lim_{\mu \rightarrow 0} \theta^k = \frac{-1 + \sqrt{1 + 4n}}{2n} \leq \frac{1}{\sqrt{n}}.$$

A.3.6 Complexity Analysis of the Full Newton Step IPM

The complexity of this algorithm is obtained by the extension of a classical lemma, whose proof can be found for instance in [179].

Lemma A.3.8. *Let $\bar{\theta}$ be such that $0 < \bar{\theta} \leq \theta^k \forall k \in \mathbb{N}$. The φ -FN-IPM for monotone LCP described in Algorithm 6 provides an ϵ -solution (z, s) , which satisfies $\sum_{i=1}^n \varphi(z_i s_i) \leq n\epsilon$ after at most $\log(\sum_{i=1}^n \varphi(z_i^0 s_i^0)/\epsilon) / \bar{\theta}$ iterations.*

The sequence $\{\theta^k\}$ is given by Theorem A.3.4. As already stated, θ^k is of order $1/\sqrt{n}$ for n large, which justify the existence of $\bar{\theta}$. Moreover, we can choose $\bar{\theta}$ of order $1/\sqrt{n}$ without loss of generality. As a result, we can now state our main theorem.

Theorem A.3.5. *Let $\mu^0 = z^0 s^0$. Algorithm 6, with the sequence of update parameters θ^k described above, guarantees strict feasibility of the iterates and quadratic convergence of the proximity measure. Moreover, it provides an ϵ -solution (z, s) , which satisfies $\sum_{i=1}^n \varphi(z_i s_i) \leq n\epsilon$ after at most $\mathcal{O}(\sqrt{n} \log(\frac{n}{\epsilon}))$ iterations.*

Preliminary computational experiments give similar results to the classical method on a small selection of LPs from the NETLIB repository¹. This confirms the validity of our approach. Some informations regarding these results are presented on Appendix A.5. Further investigations on more sophisticated methods may get the best out of this new direction.

A.4 Conclusions

The method presented in this article shows a generalization of the FN-IPM with polynomial upper bound complexity for monotone LCPs considering a general family of smooth increasing concave functions. The main contributions of this article are that we extend the

¹<http://www.netlib.org/lp/>

classical path-following method and Darvay’s method with $\varphi(t) = \sqrt{t}$ and prove that these new methods have the best known worst-case complexity.

Further research may extend this result to a more general family of LCPs such that $P_*(\kappa)$ -LCPs or P_0 -LCPs with bounded level sets as in [18, 17, 196].

Recent developments by Roos in [178] consider an infeasible IPM with full Newton step using only one feasibility step. More investigations regarding the method presented in this paper could extend the results in [178].

Despite having the best worst case upper-bound for IPMs the full Newton step is not the most used approach for numerics, since this upper-bound is attained in general [179]. We believe that the philosophy applied in this paper can be generalized to other IPMs approaches. In particular, we are planning to study a predictor-corrector implementation of this approach.

A.5 Numerics

To validate the theoretical results, we first implemented the φ -FN-IPM described in Algorithm 6 with the sequence of update parameters given by Theorem A.3.4 and with the functions φ given in Table A.1. The datasets and the residuals after convergence are de-

Name	Id.	θ^1	log	$(\)^\alpha$
Fct.	t	$\frac{t}{t+1}$	$\log(1+t)$	$(t+10^{-3})^\alpha - (10^{-3})^\alpha$ with $\alpha \in (0, 1)$

Table A.1: φ -functions used in the computational tests

tailed in Table A.2. The stopping criterion is $z^T s \leq n\epsilon$, where $\epsilon = 10^{-6}$, and we compute the residuals as $\|zs\|_\infty \times 10^6$. For every φ function that appears in Table A.1, the method converged in the same number of iterations, so we only display the number of iterations once. This phenomenon is not surprising, since the φ -FN-IPM used here stays very close to the central path, as shown by Theorem A.3.3. Therefore, this implementation does not exploit fully the new directions. Nonetheless, some small differences remain in the residuals as illustrated in Table A.2.

To illustrate the possible differences between the functions φ , we run another experiment that considers a fixed value for $\theta^k = \frac{1}{\sqrt{n}}$ for all $k \in \mathbb{N}$. Table A.3 illustrates the different behaviours observed for different choices of φ . We notice that a smaller number of iterations seems to be required when the derivative in zero is larger for the methods with $\alpha = 0.25, 0.5$ and 0.75 . Thus, further research exploring interior-point methods in large neighbourhood of the central path may get the best out of these differences.

Instance	Iter.	Id.	θ^1	log	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$
ADLITTLE	268	0.995583	0.995853	0.995856	0.995851	0.995855	0.995854
AFIRO	170	0.943919	0.944204	0.944207	0.944200	0.944205	0.944205
BEACONFD	418	0.995479	0.995728	0.995730	0.995724	0.995727	0.995728
BOEING2	420	0.974064	0.974307	0.974309	0.974309	0.974309	0.974307
BLEND	264	0.984332	0.984599	0.984602	0.984600	0.984599	0.984599
GROW7	406	0.983538	0.983785	0.983788	0.983780	0.983787	0.983785
ISRAEL	428	0.988911	0.989157	0.989155	0.989148	0.989154	0.989156
KB2	203	0.944257	0.944530	0.944532	0.944530	0.944531	0.944530
RECIPELP	332	0.960634	0.960884	0.960887	0.960884	0.960885	0.960885
SC50A	218	0.951084	0.951353	0.951356	0.951350	0.951354	0.951354
SC50B	218	0.951084	0.951353	0.951356	0.951350	0.951354	0.951354
SC105	316	0.973217	0.973472	0.973475	0.973473	0.973474	0.973473
SCAGR7	342	0.985376	0.985631	0.985633	0.985632	0.985632	0.985632
SHARE1B	372	0.968387	0.968633	0.968635	0.968634	0.968632	0.968634
SHARE2B	310	0.972849	0.973105	0.973107	0.973102	0.973107	0.973106
STOCFOR1	324	0.982774	0.983031	0.983033	0.983033	0.983031	0.983031

Table A.2: Value of $\text{Res.} = \|zs\|_\infty \times 10^6$, after the algorithm reaches $z^T s \leq n\epsilon = 10^{-6}n$

Instance	Id.		θ^1		log		$\alpha = 0.25$		$\alpha = 0.5$		$\alpha = 0.75$	
	Res.	Iter.	Res.	Iter.	Res.	Iter.	Res.	Iter.	Res.	Iter.	Res.	Iter.
ADLITTLE	0.995	268	0.975	255	0.993	263	0.967	212	0.983	226	0.966	245
AFIRO	0.943	170	0.988	161	0.930	167	0.970	134	0.972	143	0.954	155
BEACONFD	0.995	418	0.998	397	1.000	410	0.984	330	0.996	352	0.968	382
BOEING2	0.974	420	0.974	399	0.977	412	0.986	331	0.967	354	0.975	383
BLEND	0.984	264	0.975	251	0.986	259	0.950	209	0.954	223	0.973	241
GROW7	0.983	406	0.973	386	0.996	398	0.993	320	0.982	342	0.992	370
ISRAEL	0.988	428	0.976	407	0.987	420	0.976	338	0.972	361	0.998	390
KB2	0.944	203	0.938	193	0.957	199	0.968	160	0.954	171	0.959	185
RECIPELP	0.960	332	0.978	315	0.992	325	0.963	262	0.991	279	0.995	302
SC50A	0.951	218	0.960	207	0.946	214	0.962	172	0.999	183	0.944	199
SC50B	0.951	218	0.960	207	0.946	214	0.962	172	0.999	183	0.944	199
SC105	0.973	316	0.982	300	0.976	310	0.989	249	0.983	266	0.982	288
SCAGR7	0.985	342	0.981	325	0.968	336	0.978	270	0.988	288	0.981	312
SHARE1B	0.968	372	0.984	353	0.969	365	0.989	293	0.983	313	0.979	339
SHARE2B	0.972	310	0.996	294	0.981	304	0.957	245	0.981	261	0.962	283
STOCFOR1	0.982	324	0.975	308	0.979	318	0.967	256	0.980	273	0.961	296

Table A.3: Value of $\text{Res.} = \|zs\|_\infty \times 10^6$ after the algorithm reaches $z^T s \leq n\epsilon = 10^{-6}n$ and number of iterations (Iter.).

Appendix B

Formations complémentaires et participations aux conférences

Au-delà des travaux présentés dans ce document, les trois années de la thèse ont été l'occasion de participer à des conférences et des formations complémentaires. Cette section présente quelques détails à propos de ces activités parallèles.

J'ai participé à des formations complémentaires en optimisation lors de mini-cours organisés par le GdR MOA à Limoges du 15 au 17 mai 2015 ("Numerical methods for complementarity problems and applications" par Mounir Haddou, "Bundle methods for nonsmooth optimization" par Dominikus Noll et "Nonsmooth, nonconvex optimization with applications to polynomial and eigenvalue optimization" par Michael Overton) et à Toulouse le 21 et 22 mars 2016 (Problèmes de contrôles par Frédéric Bonnans).

Par ailleurs, j'ai suivi des cours lors d'une école d'été à Rome du 6 au 10 juin 2016 intitulée COST/MINO PhD School on Advanced Optimization Methods ("Polyhedral Combinatorics" par Santanu Dey, "Interior Point methods" par Jordi Castro, "Structured Dantzig-Wolfe Decomposition" par Antonio Frangioni et "Semidefinite Programming" par Veronica Piccialli).

Le 25 et 26 janvier 2017, j'ai suivi un séminaire du programme Gaspard Monge donné par Marc Teboulle et intitulé "First Order Optimization Methods".

À la frontière entre formation et vulgarisation, j'ai participé à deux reprises à la Semaine d'Étude Math-Entreprise (SEME) organisée par AMIES. La première fois à Rouen en octobre 2014 sur le thème "Modelling Gas Regulators" proposé par le groupe GCE tandis que la deuxième participation était à Nice en janvier 2016 sur le thème "Recalibration de modèles pharmacocinétiques" proposé par l'entreprise ExactCure.

Au cours de ces compléments en optimisation, j'ai également bénéficié de la possibilité de réaliser trois stages de recherche à l'étranger: Université Libanaise à Beyrouth au Liban invité par Lina Abdallah, 1 semaine en novembre 2015 ; Université de Sherbrooke au Canada invité par Jean-Pierre Dussault, 4 mois entre septembre et décembre 2016 et 1 mois en juin/juillet 2017. Ces stages de recherche ont été soutenus financièrement par des bourses de mobilités :

- (2016) bourse de mobilité de l'Ecole des Docteurs de l'Université Bretagne-Loire et du Conseil Régional de Bretagne, montant 2000 euros ;
- (2017) bourse de mobilité du GDR Recherche Opérationnelle, montant 700 euros ;
- (2017) bourse de mobilité de l'Ecole des Docteurs de l'Université Bretagne-Loire et du Conseil Régional de Bretagne, montant 1000 euros.

Outre les compléments de formations en optimisation, j'ai également eu l'opportunité d'être moniteur et encadrant. J'ai donné des cours à l'INSA de Rennes sur l'année 2015-16 (64 h Analyse 1 en 1ère année de Licence) et 2016-17 (48 h Géométrie en 2ème année de Licence ; 16 h Optimisation numérique en 3ème année de Licence). Par ailleurs, j'ai participé à l'encadrement de deux stages de Master:

- (avril à juin 2017) "Relaxation methods for MPCC" par Cao Van Kien de l'Université Paris 13 ;
- (avril à juin 2017) "Optimization methods for complementarity problems" par Nguyen Dinh Duong de l'Université Paris Est.

J'ai participé aux tâches administratives de la vie du laboratoire à l'INSA de Rennes en étant d'une part membre du bureau de la composante IRMAR-INSA et d'autre part membre du comité d'organisation de la conférence de fin de projet ANR HJNET à Rennes du 30 mai au 3 juin 2016 et du Groupe de Travail en Programmation Mathématiques du GdR RO à Rennes le 13 et 14 juin 2016.

Tout au long de la durée de la thèse j'ai participé à des conférences en donnant des exposés:

- Conférence invité : *A Smoothing Method for Sparse Optimization over Polyhedral Sets*, Groupe de Travail Programmation Mathématiques, Dijon, 2015.
- Exposés lors de séminaires :
 - *Problèmes de complémentarité en optimisation non lisse*, Séminaire LANDAU des jeunes doctorants en analyse, Rennes, 2015.
 - *Méthodes numériques pour l'optimisation non linéaire*, Séminaire LANDAU des jeunes doctorants en analyse, Rennes 2016.
 - *Une méthode numérique pour les problèmes d'optimisation biniveaux*, Séminaire LANDAU des jeunes doctorants en analyse, Rennes 2017.
 - *Problèmes d'optimisation sous contraintes et parcimonie*, séminaire informatique de l'Université de Sherbrooke. Sherbrooke, 2017
- Exposés lors de conférences :

- *A Smoothing Method for Sparse Optimization over Polyhedral Sets*, MCO Metz, 2015.
- *A new direction in polynomial time interior-point methods for monotone linear complementarity problem*, Journées SMAI-MODE, Toulouse, 2016.
- *A new relaxation method for Mathematical Program with Complementarity Constraint*, poster in HJNET, Rennes, 2016.
- *A new relaxation method for Mathematical Program with Complementarity Constraint*, INFORMS Annual Meeting, Nashville, 2016.
- *How to Compute a Stationary Point of the MPCC?*, EUROPT, Montréal, 2017.
- *Computation of a Local Minimum of the MPCC*, PARAOPT XI, Prague, 2017.

Enfin, certaines collaborations extérieures au projet de thèse ont mené à des présentations lors de conférences :

- avec F. Monteiro et al. (Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg). *Robust Prediction of Minimal Medium Composition Using Sparse Optimization*, poster in 4th Conference on Constraint-Based Reconstruction and Analysis, Heidelberg 2015.
- avec J. Erhel (INRIA Rennes). *About Some Numerical Models for Geochemistry*. HPSC Hanoi, 2015.
- avec J. Erhel (INRIA Rennes). *About Some Numerical Models for Geochemistry*. Workshop MoMas on reactive transport, 2015.
- avec J. Erhel and S. Sabit (INRIA Rennes). *Reactive transport simulations using a global approach*. Computational Methods in Water Resources, Toronto, 2016.

Bibliography

- [1] J. Abadie. *On the Kuhn-Tucker theorem*. Nonlinear Programming, 1967.
- [2] L. Abdallah, M. Haddou, and T. Migot. “Solving absolute value equation using complementarity and smoothing functions”. In: *Journal of Computational and Applied Mathematics* 327 (2018), pp. 196–207.
- [3] Lina Abdallah and Mounir Haddou. “An exact penalty approach for mathematical programs with equilibrium constraints.” In: *Journal of Advances in Mathematics* 9.8 (2013).
- [4] M. Achache. “A new primal-dual path-following method for convex quadratic programming”. In: *Computational & Applied Mathematics* 25.1 (2006), pp. 97–110.
- [5] M. Achache. “A weighted path-following method for the linear complementarity problem”. In: *Universitatis Babes. Bolyai. Series Informatica* (49)(1) (2004), pp. 61–73.
- [6] M. Achache. “Complexity analysis and numerical implementation of a short-step primal-dual algorithm for linear complementarity problems”. In: *Applied Mathematics and computation* 216.7 (2010), pp. 1889–1895.
- [7] W. Achtziger, T. Hoheisel, and C. Kanzow. “A smoothing-regularization approach to mathematical programs with vanishing constraints”. In: *Computational Optimization and Applications* 55.3 (2013), pp. 733–767.
- [8] W. Achtziger, T. Hoheisel, and C. Kanzow. *On a relaxation method for mathematical programs with vanishing constraints*. Inst. of Math., 2012.
- [9] W. Achtziger and C. Kanzow. “Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications”. In: *Mathematical Programming* 114.1 (2008), pp. 69–99.
- [10] S. Albrecht and M. Ulbrich. “Mathematical programs with complementarity constraints in the context of inverse optimal control for locomotion”. In: *Optimization Methods and Software* 6788.October (Sept. 2016), pp. 1–29.
- [11] R. Andreani, G. Haeser, and J. M. Martínez. “On sequential optimality conditions for smooth constrained optimization”. In: *Optimization* 60.5 (2011), pp. 627–641.
- [12] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J.S. Silva. “A relaxed constant positive linear dependence constraint qualification and applications”. In: *Mathematical Programming* 135.1-2 (2012), pp. 255–273.

- [13] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J.S. Silva. “Two new weak constraint qualifications and applications”. In: *SIAM Journal on Optimization* 22.3 (2012), pp. 1109–1135.
- [14] R. Andreani, J. M. Martinez, A. Ramos, and P. J.S. Silva. “A cone-continuity constraint qualification and algorithmic consequences”. In: *SIAM Journal on Optimization* 26.1 (2016), pp. 96–110.
- [15] R. Andreani, J. M. Martinez, and M. L. Schuverdt. “The CPLD condition of Qi and Wei implies the quasinormality qualification”. In: *Journal of Optimization Theory and Applications* 125 (2005), pp. 473–485.
- [16] M. Anitescu. *On Using the Elastic Mode in Nonlinear Programming Approaches to Mathematical Programs with Complementarity Constraints*. Vol. 15. 4. 2005, pp. 1203–1236.
- [17] S. Asadi and H. Mansouri. “A New Full-Newton Step $O(n)$ Infeasible Interior-Point Algorithm for $P_*(\kappa)$ -horizontal Linear Complementarity Problems.” In: *The Computer Science Journal of Moldova* 22.1 (2014), pp. 37–61.
- [18] S. Asadi and H. Mansouri. “Polynomial interior-point algorithm for $P_*(\kappa)$ horizontal linear complementarity problems”. In: *Numerical Algorithms* 63.2 (2013), pp. 385–398.
- [19] S. Asadi, H. Mansouri, and Z. Darvay. “An infeasible full-NT step IPM for horizontal linear complementarity problem over Cartesian product of symmetric cones”. In: *Optimization* 66.2 (2017), pp. 225–250.
- [20] S. Asadi, H. Mansouri, and Z. Darvay. “An infeasible interior-point method with improved centering steps for monotone linear complementarity problems”. In: *Asian-European Journal of Mathematics* 8.03 (2015), p. 1550037.
- [21] A. Auslender and M. Teboulle. *Asymptotic cones and functions in optimization and variational inequalities*. Springer Science & Business Media, 2006.
- [22] M. Babaie-Zadeh and C. Jutten. “On the stable recovery of the sparsest overcomplete representations in presence of noise”. In: *Signal Processing, IEEE Transactions on* 58.10 (2010), pp. 5396–5400.
- [23] M. S. Bazaraa and C. M. Shetty. *Foundations of optimization*. Vol. 122. Springer Science & Business Media, 2012.
- [24] A. Beck and M. Teboulle. “A fast iterative shrinkage-thresholding algorithm for linear inverse problems”. In: *SIAM journal on imaging sciences* 2.1 (2009), pp. 183–202.
- [25] J. Y. Bello Cruz, O. P. Ferreira, and L. F. Prudente. “On the global convergence of the inexact semi-smooth Newton method for absolute value equation”. In: *Computational Optimization and Applications* 65.1 (Sept. 2016), pp. 93–108.
- [26] M. Bergounioux and M. Haddou. “A new relaxation method for a discrete image restoration problem”. In: *Journal of Convex Analysis* 17.3 (2010), pp. 421–435.

- [27] M. Bergounioux and M. Haddou. “A regularization method for ill-posed bilevel optimization problems”. In: *RAIRO-Operations Research* 40.1 (2006), pp. 19–35.
- [28] S. Bi, X. Liu, and S. Pan. “Exact penalty decomposition method for zero-norm minimization based on MPEC formulation”. In: *SIAM Journal on Scientific Computing* 36.4 (2014), A1451–A1477.
- [29] T. Blumensath and M. E. Davies. “Gradient pursuits”. In: *IEEE Transactions on Signal Processing* 56.6 (2008), pp. 2370–2382.
- [30] J.-F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2013.
- [31] P. S. Bradley and O. L. Mangasarian. “Feature selection via concave minimization and support vector machines.” In: *ICML*. Vol. 98. 1998, pp. 82–90.
- [32] M. Branda, M. Bucher, M. Červinka, and A. Schwartz. “Convergence of a Scholtes-type Regularization Method for Cardinality-Constrained Optimization Problems with an Application in Sparse Robust Portfolio Optimization”. In: *arxiv.org* (2017).
- [33] A. M. Bruckstein, D. L. Donoho, and M. Elad. “From sparse solutions of systems of equations to sparse modeling of signals and images”. In: *SIAM review* 51.1 (2009), pp. 34–81.
- [34] O. P. Burdakov, C. Kanzow, and A. Schwartz. “Mathematical programs with cardinality constraints: reformulation by complementarity-type conditions and a regularization method”. In: 26.1 (2016), pp. 397–425.
- [35] O. P. Burdakov, C. Kanzow, and A. Schwartz. “On a Reformulation of Mathematical Programs with Cardinality Constraints”. In: *Advances in Global Optimization*. Ed. by David Gao, Ning Ruan, and Wenxun Xing. Cham: Springer International Publishing, 2015, pp. 3–14.
- [36] A. P. Burgard, P. Pharkya, and C. D. Maranas. “Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization”. In: *Biotechnology and bioengineering* 84.6 (2003), pp. 647–657.
- [37] L. Caccetta, B. Qu, and G. Zhou. “A globally and quadratically convergent method for absolute value equations”. In: *Computational Optimization and Applications* 48.1 (2011), pp. 45–58.
- [38] E. J. Candes, J. K. Romberg, and T. Tao. “Stable signal recovery from incomplete and inaccurate measurements”. In: *Communications on pure and applied mathematics* 59.8 (2006), pp. 1207–1223.
- [39] E. J. Candès, J. K. Romberg, and T. Tao. “Stable signal recovery from incomplete and inaccurate measurements”. In: *Communications on Pure and Applied Mathematics* 59.8 (2006), pp. 1207–1223. ISSN: 1097-0312. DOI: 10.1002/cpa.20124. URL: <http://dx.doi.org/10.1002/cpa.20124>.

- [40] E. J. Candes, J. Romberg, and T. Tao. “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information”. In: *IEEE Transactions on Information Theory* 52.2 (Feb. 2006), pp. 489–509. ISSN: 0018-9448. DOI: 10.1109/TIT.2005.862083.
- [41] E. J. Candes and T. Tao. “Decoding by linear programming”. In: *IEEE Transactions on Information Theory* 51.12 (Dec. 2005), pp. 4203–4215. ISSN: 0018-9448. DOI: 10.1109/TIT.2005.858979.
- [42] E. J. Candès and M. B. Wakin. “An introduction to compressive sampling”. In: *IEEE signal processing magazine* 25.2 (2008), pp. 21–30.
- [43] M. Červinka, C. Kanzow, and A. Schwartz. “Constraint qualifications and optimality conditions for optimization problems with cardinality constraints”. In: *Mathematical Programming* 160.1-2 (Nov. 2016), pp. 353–377.
- [44] R. Chartrand. “Exact reconstruction of sparse signals via nonconvex minimization”. In: *Signal Processing Letters, IEEE* 14.10 (2007), pp. 707–710.
- [45] S. S. Chen, D. L. Donoho, and M. A. Saunders. “Atomic decomposition by basis pursuit”. In: *SIAM Journal on Scientific Computing* 20.1 (1998), pp. 33–61.
- [46] F. H. Clarke. *Optimization and nonsmooth analysis*. SIAM, 1990.
- [47] R. W. Cottle. “William Karush and the KKT theorem”. In: *Doc. Math., Extra volume: Optimization stories* (2012), pp. 255–269.
- [48] Z. Coulibaly and D. Orban. “An ℓ_1 Elastic Interior-Point Method for Mathematical Programs with Complementarity Constraints”. In: *SIAM Journal on Optimization* 22.1 (Jan. 2012), pp. 187–211.
- [49] Z. Darvay. “A new algorithm for solving self-dual linear optimization problems”. In: *Studia Universitatis Babe-Bolyai, Series Informatica* 47 (2002), pp. 15–26.
- [50] Z. Darvay. “New interior-point algorithms in linear programming”. In: *Advanced Modelling and Optimization* 5.1 (2003), pp. 51–92.
- [51] I. Daubechies, M. Defrise, and C. De Mol. “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint”. In: *Communications on pure and applied mathematics* 57.11 (2004), pp. 1413–1457.
- [52] V. DeMiguel, M. P. Friedlander, F. J. Nogales, and S. Scholtes. “A two-sided relaxation scheme for mathematical programs with equilibrium constraints”. In: *SIAM Journal on Optimization* 16.2 (2005), pp. 587–609.
- [53] S. Dempe. *Foundations of bilevel programming*. Springer Science & Business Media, 2002.
- [54] S. Dempe and J. Dutta. “Is bilevel programming a special case of a mathematical program with complementarity constraints?” In: *Mathematical programming* 131.1-2 (2012), pp. 37–48.

- [55] S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnykova. *Bilevel programming problems: theory, algorithms and applications to energy networks*. Springer, 2015.
- [56] E. D. Dolan and J. J. Moré. “Benchmarking optimization software with performance profiles”. In: *Mathematical programming* 91.2 (2002), pp. 201–213.
- [57] E. D. Dolan, J. J. Moré, and T. S. Munson. “Optimality Measures for Performance Profiles”. In: *SIAM Journal on Optimization* 16.3 (Jan. 2006), pp. 891–909.
- [58] D. L. Donoho. “Compressed sensing”. In: *IEEE Transactions on Information Theory* 52.4 (Apr. 2006), pp. 1289–1306. ISSN: 0018-9448. DOI: 10.1109/TIT.2006.871582.
- [59] D. L. Donoho. “De-noising by soft-thresholding”. In: *IEEE transactions on information theory* 41.3 (1995), pp. 613–627.
- [60] D. L. Donoho. “For most large underdetermined systems of linear equations the minimal ℓ_1 -norm solution is also the sparsest solution”. In: *Communications on pure and applied mathematics* 59.6 (2006), pp. 797–829.
- [61] D. L. Donoho. *Neighborly polytopes and sparse solutions of underdetermined linear equations*. Tech. rep. Department of Statistics, Stanford University, 2004.
- [62] D. L. Donoho, M. Elad, and V. N. Temlyakov. “Stable recovery of sparse overcomplete representations in the presence of noise”. In: *Information Theory, IEEE Transactions on* 52.1 (2006), pp. 6–18.
- [63] D. L. Donoho and X. Huo. “Uncertainty principles and ideal atomic decomposition”. In: *IEEE Transactions on Information Theory* 47.7 (2001), pp. 2845–2862.
- [64] J.-P. Dussault, M. Haddou, and T. Migot. “The New Butterfly Relaxation Methods for Mathematical Program with Complementarity Constraint”. In: *optimization-online.org* (2016).
- [65] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems*. Vol. 1. Springer, 2003.
- [66] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.
- [67] M. Feng, J. E. Mitchell, J.-S. Pang, X. Shen, and A. Wächter. “Complementarity Formulations of Optimization Problems”. In: *Cmmi* (2015), pp. 1–32.
- [68] Mingbin Feng, John E Mitchell, Jong-Shi Pang, Xin Shen, and Andreas Wächter. *Complementarity formulations of l_0 -norm optimization problems*. Tech. rep. Technical report, Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, 2013.
- [69] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming*. Society for Industrial and Applied Mathematics, Jan. 1990.
- [70] M. L. Flegel. “Constraint Qualifications and Stationarity Concepts for Mathematical Programs with Equilibrium Constraints”. PhD thesis. University of Würzburg, 2005.

- [71] M. L. Flegel and C. Kanzow. *A direct proof for M-stationarity under MPEC-GCQ for mathematical programs with equilibrium constraints*. Springer, 2006.
- [72] M. L. Flegel and C. Kanzow. “Abadie-type constraint qualification for mathematical programs with equilibrium constraints”. In: *Journal of Optimization Theory and Applications* 124.3 (2005), pp. 595–614.
- [73] M. L. Flegel and C. Kanzow. “On M-stationary points for mathematical programs with equilibrium constraints”. In: *Journal of Mathematical Analysis and Applications* 310.1 (2005), pp. 286–302.
- [74] M. L. Flegel and C. Kanzow. “On the Guignard constraint qualification for mathematical programs with equilibrium constraints”. In: *Optimization* 54.6 (2005), pp. 517–534.
- [75] R. Fletcher and S. Leyffer. “Numerical experience with solving MPECs as NLPs”. In: (2002).
- [76] R. Fletcher and S. Leyffer. “Solving mathematical programs with complementarity constraints as nonlinear programs”. In: *Optimization Methods and Software* 19.1 (2004), pp. 15–40.
- [77] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes. “Local Convergence of SQP Methods for Mathematical Programs with Equilibrium Constraints”. In: *SIAM Journal on Optimization* 17.1 (2006), pp. 259–286.
- [78] S. Foucart and M.-J. Lai. “Sparsest solutions of underdetermined linear systems via ℓ_q -minimization for $0 < q \leq 1$ ”. In: *Applied and Computational Harmonic Analysis* 26.3 (2009), pp. 395–407.
- [79] R. Fourer, D. Gay, and B. Kernighan. *Ampl*. Vol. 119. Boyd & Fraser, 1993.
- [80] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*. Vol. 1. Springer series in statistics New York, 2001.
- [81] J.-J. Fuchs. “On sparse representations in arbitrary redundant bases”. In: *IEEE transactions on Information theory* 50.6 (2004), pp. 1341–1344.
- [82] G. M. Fung and O. L. Mangasarian. “Equivalence of minimal ℓ_0 and ℓ_p norm solutions of linear equalities, inequalities and linear programs for sufficiently small p ”. In: *Journal of Optimization Theory and Applications* 151.1 (2011), pp. 1–10.
- [83] D. Ge, X. Jiang, and Y. Ye. “A note on the complexity of L_p minimization”. In: *Mathematical programming* 129.2 (2011), pp. 285–299.
- [84] P. E. Gill, W. Murray, and M. A. Saunders. “SNOPT: An SQP algorithm for large-scale constrained optimization”. In: *SIAM Rev.* 47 (2005), pp. 99–131.
- [85] I. F. Gorodnitsky, J. S. George, and B. D. Rao. “Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm”. In: *Electroencephalography and clinical Neurophysiology* 95.4 (1995), pp. 231–251.

- [86] F. J. Gould and J. W. Tolle. “A Necessary and Sufficient Qualification for Constrained Optimization”. In: *SIAM Journal on Applied Mathematics* 20.2 (Mar. 1971), pp. 164–172.
- [87] R. Gribonval and M. Nielsen. “Sparse decomposition in unions of bases”. In: *IEEE Trans. Inf. Theory* 49.12 (2003), pp. 3320–3325.
- [88] M. Guignard. “Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space”. In: *SIAM Journal on Control* 7.2 (1969), pp. 232–241.
- [89] L. Guo, G. H. Lin, and J. J. Ye. “Solving mathematical programs with equilibrium constraints”. In: *Journal of Optimization Theory and Applications* 166.1 (2015), pp. 234–256.
- [90] L. Guo and G.-H. Lin. “Notes on some constraint qualifications for mathematical programs with equilibrium constraints”. In: *Journal of Optimization Theory and Applications* 156.3 (2013), pp. 600–616.
- [91] L. Guo, G.-H. Lin, and J. J. Ye. “Second-order optimality conditions for mathematical programs with equilibrium constraints”. In: *Journal of Optimization Theory and Applications* 158.1 (2013), pp. 33–64.
- [92] Inc. Gurobi Optimization. *Gurobi Optimizer Reference Manual*. 2015. URL: <http://www.gurobi.com>.
- [93] M. Haddou. “A new class of smoothing methods for mathematical programs with equilibrium constraints”. In: *Pacific Journal of Optimization* 5.1 (2009), p. 86.
- [94] M. Haddou and P. Maheux. “Smoothing methods for nonlinear complementarity problems”. In: *Journal of Optimization Theory and Applications* 160.3 (2014), pp. 711–729.
- [95] M. Haddou and T. Migot. “A Smoothing Method for Sparse Optimization over Polyhedral Sets”. In: *hal.archives-ouvertes.fr* (2016).
- [96] M. Haddou, T. Migot, and J. Omer. “A Generalized Direction in Interior Point Method for Monotone Linear Complementarity Problems”. In: *hal.archives-ouvertes.fr* (2016).
- [97] F. K. Haghani. “On Generalized Traub’s Method for Absolute Value Equations”. In: *Journal of Optimization Theory and Applications* (2015), pp. 1–7.
- [98] M. Hintermüller and I. Kopacka. “Mathematical Programs with Complementarity Constraints in Function Space: C- and Strong Stationarity and a Path-Following Algorithm”. In: *SIAM Journal on Optimization* 20.2 (Jan. 2009), pp. 868–902.
- [99] A. J. Hoffman. “On approximate solutions of systems of linear inequalities”. In: *Journal of Research of the National Bureau of Standard* 49.2 (1952), pages.
- [100] T. Hoheisel. “Mathematical programs with vanishing constraints”. PhD thesis. 2009.

- [101] T. Hoheisel, C. Kanzow, and A. Schwartz. “Convergence of a local regularization approach for mathematical programmes with complementarity or vanishing constraints”. In: *Optimization Methods and Software* 27.3 (2012), pp. 483–512.
- [102] T. Hoheisel, C. Kanzow, and A. Schwartz. “Mathematical programs with vanishing constraints: a new regularization approach with strong convergence properties”. In: *Optimization* 61.6 (2012), pp. 619–636.
- [103] T. Hoheisel, C. Kanzow, and A. Schwartz. “Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints”. In: *Mathematical Programming* 137.1-2 (Feb. 2013), pp. 257–288.
- [104] X. M. Hu and D. Ralph. “Convergence of a Penalty Method for Mathematical Programming with Complementarity Constraints”. In: *Journal of Optimization Theory and Applications* 123.2 (Nov. 2004), pp. 365–390.
- [105] T. Illés, C. Roos, and T. Terlaky. “Polynomial Affine-Scaling Algorithms for $P_*(\kappa)$ Linear Complementary Problems”. In: *Recent Advances in Optimization*. Springer, 1997, pp. 119–137.
- [106] A. F. Izmailov and M. V. Solodov. “Mathematical programs with vanishing constraints: Optimality conditions, sensitivity, and a relaxation method”. In: *Journal of Optimization Theory and Applications* 142.3 (2009), pp. 501–532.
- [107] R. Janin. *Directional derivative of the marginal function in nonlinear programming*. Springer, 1984.
- [108] F. John. *Extremum problems with inequalities as subsidiary conditions*. Wiley-Interscience New York, 1948, pp. 187–204.
- [109] A. Kadrani. “Nouvelles Régularisations pour l’optimisation avec contraintes de complémentarité”. PhD thesis. Université de Sherbrooke, 2008.
- [110] A. Kadrani, J.-P. Dussault, and A. Benchakroun. “A globally convergent algorithm for MPCC”. In: *EURO Journal on Computational Optimization* 3.3 (2015), pp. 263–296.
- [111] A. Kadrani, J.-P. Dussault, and A. Benchakroun. “A new regularization scheme for mathematical programs with complementarity constraints”. In: *SIAM Journal on Optimization* 20.1 (2009), pp. 78–103.
- [112] C. Kanzow and A. Schwartz. “A New Regularization Method for Mathematical Programs with Complementarity Constraints with Strong Convergence Properties”. In: *SIAM Journal on Optimization* 23.2 (Apr. 2013), pp. 770–798.
- [113] C. Kanzow and A. Schwartz. “Convergence properties of the inexact Lin-Fukushima relaxation method for mathematical programs with complementarity constraints”. In: *Computational Optimization and Applications* 59.1-2 (Oct. 2014), pp. 249–262.
- [114] C. Kanzow and A. Schwartz. “Mathematical programs with equilibrium constraints: enhanced Fritz John-conditions, new constraint qualifications, and improved exact penalty results”. In: *SIAM Journal on Optimization* 20.5 (2010), pp. 2730–2753.

- [115] C. Kanzow and A. Schwartz. “The Price of Inexactness: Convergence Properties of Relaxation Methods for Mathematical Programs with Complementarity Constraints Revisited”. In: *Mathematics of Operations Research* 40.2 (May 2015), pp. 253–275.
- [116] W. Karush. “Minima of Functions of Several Variables with Inequalities as Side Constraints”. PhD thesis. M.Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago, Chicago, Illinois., 1939.
- [117] S. Ketabchi and H. Moosaei. “An efficient method for optimal correcting of absolute value equations by minimal changes in the right hand side”. In: *Computers & Mathematics with Applications* 64.6 (2012), pp. 1882–1885.
- [118] S. Ketabchi, H. Moosaei, and S. Fallahi. “Optimal error correction of the absolute value equation using a genetic algorithm”. In: *Mathematical and Computer Modelling* 57.9 (2013), pp. 2339–2342.
- [119] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A unified approach to interior point algorithms for linear complementarity problems*. Vol. 538. Springer Science & Business Media, 1991.
- [120] H. W. Kuhn and A. W. Tucker. “Nonlinear programming”. In: *Proceedings of 2nd Berkeley Symposium* (1951), pp. 481–492.
- [121] M.-J. Lai and J. Wang. “An Unconstrained ℓ_q Minimization with $0 \leq q \leq 1$ for Sparse Solution of Underdetermined Linear Systems”. In: *SIAM Journal on Optimization* 21.1 (2011), pp. 82–101.
- [122] H. A. Le Thi, T. P. Dinh, H. M. Le, and X. T. Vo. “DC approximation approaches for sparse optimization”. In: *European Journal of Operational Research* 244.1 (2015), pp. 26–46.
- [123] S. Leyffer. “MacMPEC: AMPL collection of MPECs”. In: *Argonne National Laboratory. Available at www.mcs.anl.gov/leyflier/MacMPEC* (2000).
- [124] S. Leyffer, G. López-Calva, and J. Nocedal. “Interior Methods for Mathematical Programs with Complementarity Constraints”. In: *SIAM Journal on Optimization* 17.1 (2006), pp. 52–77.
- [125] S. Leyffer and T. S. Munson. “A globally convergent filter method for MPECs”. In: April (2009).
- [126] G.-H. Lin and M. Fukushima. “A modified relaxation scheme for mathematical programs with complementarity constraints”. In: *Annals of Operations Research* 133.1-4 (2005), pp. 63–84.
- [127] X. Liu and J. Sun. *Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints*. Vol. 101. 1. 2004, pp. 231–261.
- [128] G. Liuzzi and F. Rinaldi. “Solving ℓ_0 -penalized problems with simple constraints via the Frank–Wolfe reduced dimension method”. In: *Optimization Letters* 9.1 (Jan. 1, 2015), pp. 57–74. ISSN: 1862-4480. DOI: 10.1007/s11590-014-0757-3. URL: <https://doi.org/10.1007/s11590-014-0757-3>.

- [129] D. Di Lorenzo, G. Liuzzi, F. Rinaldi, F. Schoen, and M. Sciandrone. “A concave optimization-based approach for sparse portfolio selection”. In: *Optimization Methods and Software* 27.6 (2012), pp. 983–1000. DOI: 10.1080/10556788.2011.577773. eprint: <http://dx.doi.org/10.1080/10556788.2011.577773>. URL: <http://dx.doi.org/10.1080/10556788.2011.577773>.
- [130] T. Lotfi and H. Veisheh. “A note on unique solvability of the absolute value equation”. In: (2013).
- [131] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints*. Cambridge University Press, 1996.
- [132] D. Malioutov and A. Aravkin. “Iterative log thresholding”. In: *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 7198–7202.
- [133] S. Mallat. *A wavelet tour of signal processing*. Academic press, 1999.
- [134] O. L. Mangasarian. “A generalized Newton method for absolute value equations”. In: *Optimization Letters* 3.1 (2009), pp. 101–108.
- [135] O. L. Mangasarian. “A hybrid algorithm for solving the absolute value equation”. In: *Optimization Letters* (2015), pp. 1–6.
- [136] O. L. Mangasarian. “Absolute value equation solution via concave minimization”. In: *Optimization Letters* 1.1 (2007), pp. 3–8.
- [137] O. L. Mangasarian. “Absolute value equation solution via dual complementarity”. In: *Optimization Letters* 7.4 (2013), pp. 625–630.
- [138] O. L. Mangasarian. “Absolute value equation solution via linear programming”. In: *Journal of Optimization Theory and Applications* 161.3 (2014), pp. 870–876.
- [139] O. L. Mangasarian. “Linear complementarity as absolute value equation solution”. In: *Optimization Letters* 8.4 (2014), pp. 1529–1534.
- [140] O. L. Mangasarian. “Machine learning via polyhedral concave minimization”. In: *Applied Mathematics and Parallel Computing*. Springer, 1996, pp. 175–188.
- [141] O. L. Mangasarian. “Machine learning via polyhedral concave minimization”. In: *Applied Mathematics and Parallel Computing*. Springer, 1996, pp. 175–188.
- [142] O. L. Mangasarian. “Minimum-support solutions of polyhedral concave programs”. In: *Optimization* 45.1-4 (1999), pp. 149–162.
- [143] O. L. Mangasarian and R. R. Meyer. “Absolute value equations”. In: *Linear Algebra and Its Applications* 419.2 (2006), pp. 359–367.
- [144] P. Marcotte. “Network design problem with congestion effects: A case of bilevel programming”. In: *Mathematical programming* 34.2 (1986), pp. 142–162.
- [145] M. Marechal and R. Correa. “A DC (Difference of Convex functions) approach of the MPECs”. In: *optimization-online.org* (2014).

- [146] MATLAB. *version 8.4.0.150421 (R2014b)*. Natick, Massachusetts: The MathWorks Inc., 2014.
- [147] MATLAB. *version R2013*. Natick, Massachusetts: The MathWorks Inc., 2014.
- [148] T. Migot, J.-P. Dussault, M. Haddou, and A. Kadrani. “How to Compute a Local Minimum of the MPCC”. In: *optimization-online.org* (2017).
- [149] T. Migot and J. Erhel. *Analyse mathématique de modèles géochimiques*. Tech. rep. INRIA Rennes, équipe SAGE, 2014.
- [150] T. Migot and M. Haddou. “A Smoothing Method for Sparse Optimization over Polyhedral Sets”. In: *Modelling, Computation and Optimization in Information Systems and Management Sciences*. Springer, 2015, pp. 369–379.
- [151] A. Miller. *Subset selection in regression*. CRC Press, 2002.
- [152] H. Mohimani, M. Babaie-Zadeh, and C. Jutten. “A fast approach for overcomplete sparse decomposition based on smoothed norm”. In: *Signal Processing, IEEE Transactions on* 57.1 (2009), pp. 289–301.
- [153] M. T. T. Monteiro and J. F. P. Meira. “A penalty method and a regularization strategy to solve MPCC”. In: *International Journal of Computer Mathematics* 88.1 (Jan. 2011), pp. 145–149.
- [154] H. Moosaei, S. Ketabchi, M. A. Noor, J. Iqbal, and V. Hooshyarbakhsh. “Some techniques for solving absolute value equations”. In: *Applied Mathematics and Computation* 268 (2015), pp. 696–705.
- [155] B. A. Murtagh and M. A. Saunders. “MINOS 5.51 USER’S GUIDE by”. In: (1983).
- [156] B. K. Natarajan. “Sparse approximate solutions to linear systems”. In: *SIAM Journal on Computing* 24.2 (1995), pp. 227–234.
- [157] D. Needell and J. A. Tropp. “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples”. In: *Applied and Computational Harmonic Analysis* 26.3 (2009), pp. 301–321.
- [158] J. V. Outrata. “Mathematical Programs with Equilibrium Constraints: Theory and Numerical Methods”. In: *Nonsmooth Mechanics of Solids*. 1. Vienna: Springer Vienna, 2002, pp. 221–274.
- [159] J.-S. Pang and M. Fukushima. “Complementarity Constraint Qualifications and Simplified B-Stationarity Conditions for Mathematical Programs with Equilibrium Constraints”. In: (1998).
- [160] W. B. Pennebaker and J. L. Mitchell. *JPEG: Still image data compression standard*. Springer Science & Business Media, 1992.
- [161] O. Prokopyev. “On equivalent reformulations for absolute value equations”. In: *Computational Optimization and Applications* 44.3 (2009), pp. 363–372.

- [162] L. Qi and Z. Wei. “On the constant positive linear dependence condition and its application to SQP methods”. In: *SIAM Journal on Optimization* 10.4 (2000), pp. 963–981.
- [163] Y.-E. Qi and M.-Z. Zhu. “The nonlinear HSS-like iteration method for absolute value equations”. In: *Applied Mathematics and Computation* 265.11271174 (Mar. 2014), pp. 266–274.
- [164] A. U. Raghunathan and L. T. Biegler. “An Interior Point Method for Mathematical Programs with Complementarity Constraints (MPCCs)”. In: *SIAM Journal on Optimization* 15.3 (2005), pp. 720–750.
- [165] A. Ramos. “Mathematical Programms with Equilibrium Constraints: A sequential optimality condition, new constraint qualifications and algorithmic consequences.” In: *Optimization Online* (2016).
- [166] F. Rinaldi. “Concave programming for finding sparse solutions to problems with convex constraints”. In: *Optimization Methods and Software* 26.6 (2011), pp. 971–992.
- [167] F. Rinaldi. “Mathematical programming methods for minimizing the zero-norm over polyhedral sets”. PhD thesis. PhD thesis, Sapienza, University of Rome, 2009.
- [168] Francesco Rinaldi, Fabio Schoen, and Marco Sciandrone. “Concave programming for minimizing the zero-norm over polyhedral sets”. In: *Computational Optimization and Applications* 46.3 (2010), pp. 467–486.
- [169] S. M. Robinson. “An application of error bounds for convex programming in a linear space”. In: *SIAM Journal on Control* 13.2 (1975), pp. 271–273.
- [170] R. T. Rockafellar. *Convex analysis*. Princeton university press, 1970.
- [171] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*. Vol. 317. Springer Science & Business Media, 2009.
- [172] J. Rohn. “A theorem of the alternatives for the equation $Ax + B|x| = b$ ”. In: *Linear and Multilinear Algebra* 52.6 (2004), pp. 421–426.
- [173] J. Rohn. “An algorithm for solving the absolute value equation”. In: *Electronic Journal of Linear Algebra* 18.589-599 (2009), p. 5.
- [174] J. Rohn. *An algorithm for solving the absolute value equation: An improvement*. Tech. rep. 2010.
- [175] J. Rohn. “On unique solvability of the absolute value equation”. In: *Optimization Letters* 3.4 (2009), pp. 603–606.
- [176] J. Rohn, V. Hooshyarbakhsh, and R. Farhadsefat. “An iterative method for solving absolute value equations and sufficient conditions for unique solvability”. In: *Optimization Letters* 8.1 (2014), pp. 35–44.
- [177] C. Roos. “A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization”. In: *SIAM Journal on Optimization* 16.4 (2006), pp. 1110–1136.

- [178] C. Roos. “An Improved and Simplified Full-Newton Step $O(n)$ Infeasible Interior-Point Method for Linear Optimization”. In: *SIAM Journal on Optimization* 25.1 (2015), pp. 102–114.
- [179] C. Roos, T. Terlaky, and J.-Ph. Vial. *Interior point methods for linear optimization*. Springer Science & Business Media, 2006.
- [180] D. K. Salkuyeh. “The Picard–HSS iteration method for absolute value equations”. In: *Optimization Letters* 8.8 (2014), pp. 2191–2202.
- [181] H. Scheel and S. Scholtes. “Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity”. In: *Mathematics of Operations Research* 25.1 (2000), pp. 1–22.
- [182] S. Scholtes. “Convergence Properties of a Regularization Scheme for Mathematical Programs with Complementarity Constraints”. In: *SIAM Journal on Optimization* 11.4 (2001), pp. 918–936.
- [183] S. Scholtes and M. Stöhr. “How stringent is the linear independence assumption for mathematical programs with complementarity constraints?” In: *Mathematics of Operations Research* 26.4 (2001), pp. 851–863.
- [184] A. 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.
- [185] S. Steffensen and M. Ulbrich. “A new relaxation scheme for mathematical programs with equilibrium constraints”. In: *SIAM Journal on Optimization* 20.5 (2010), pp. 2504–2539.
- [186] R. Tibshirani. “Regression shrinkage and selection via the lasso”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), pp. 267–288.
- [187] J. A. Tropp. “Just relax: Convex programming methods for identifying sparse signals in noise”. In: *IEEE transactions on information theory* 52.3 (2006), pp. 1030–1051.
- [188] J. Tropp, S. J. Wright, et al. “Computational methods for sparse solution of linear inverse problems”. In: *Proceedings of the IEEE* 98.6 (2010), pp. 948–958.
- [189] V. Vapnik. *The nature of statistical learning theory*. Springer science & business media, 2013.
- [190] S. Veelken. “A new relaxation scheme for mathematical programs with equilibrium constraints: Theory and numerical experience”. PhD thesis. Julius-Maximilians-Universität Würzburg, 2009.
- [191] S. Voronin and R. Chartrand. “A new generalized thresholding algorithm for inverse problems with sparsity constraints”. In: *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE. 2013, pp. 1636–1640.
- [192] A. Wächter and L. T. Biegler. “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming”. In: *Mathematical programming* 106.1 (2006), pp. 25–57.

- [193] G.-Q. Wang and Y.-Q. Bai. “A new full Nesterov–Todd step primal–dual path-following interior-point algorithm for symmetric optimization”. In: *Journal of Optimization Theory and Applications* 154.3 (2012), pp. 966–985.
- [194] G.-Q. Wang and Y.-Q. Bai. “A new primal-dual path-following interior-point algorithm for semidefinite optimization”. In: *Journal of Mathematical Analysis and Applications* 353.1 (2009), pp. 339–349.
- [195] G.-Q. Wang and Y.-Q. Bai. “A primal-dual interior-point algorithm for second-order cone optimization with full Nesterov–Todd step”. In: *Applied Mathematics and Computation* 215.3 (2009), pp. 1047–1061.
- [196] G.-Q. Wang, X.J. Fan, D.T. Zhu, and D.Z. Wang. “New complexity analysis of a full-Newton step feasible interior-point algorithm for $P_*(\kappa)$ -LCP”. In: *Optimization Letters* 9.6 (2015), pp. 1105–1119.
- [197] J. Weston, A. Elisseeff, B. Schölkopf, and M. Tipping. “Use of the zero-norm with linear models and kernel methods”. In: *Journal of machine learning research* 3.Mar (2003), pp. 1439–1461.
- [198] J. J. Ye. “Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints”. In: *Journal of Mathematical Analysis and Applications* 307.1 (2005), pp. 350–369.
- [199] J. J. Ye. “Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints”. In: *Journal of Mathematical Analysis and Applications* 307.1 (2005), pp. 350–369.
- [200] J. J. Ye and X. Y. Ye. “Necessary optimality conditions for optimization problems with variational inequality constraints”. In: *Mathematics of Operations Research* 22.4 (1997), pp. 977–997.
- [201] J. J. Ye and J. Zhang. “Enhanced Karush-Kuhn-Tucker condition and weaker constraint qualifications”. In: *Mathematical Programming* 139.1-2 (2013), pp. 353–381.
- [202] J. J. Ye and J. Zhang. “Enhanced Karush-Kuhn-Tucker Conditions for Mathematical Programs with Equilibrium Constraints”. In: *Journal of Optimization Theory and Applications* 163.3 (2014), pp. 777–794.
- [203] J. J. Ye and J. Zhang. “Enhanced Karush–Kuhn–Tucker conditions for mathematical programs with equilibrium constraints”. In: *Journal of Optimization Theory and Applications* 163.3 (2014), pp. 777–794.
- [204] L. Yong. “A Smoothing Newton Method for Absolute Value Equation”. In: *International Journal of Control and Automation* 9.2 (Feb. 2016), pp. 119–132.
- [205] L. Yong. “Particle swarm optimization for absolute value equations”. In: *Journal of Computational Information Systems* 6.7 (2010), pp. 2359–2366.
- [206] J.-J. Zhang. “The relaxed nonlinear PHSS-like iteration method for absolute value equations”. In: *Applied Mathematics and Computation* 265 (2015), pp. 266–274.

- [207] Y.-B. Zhao and M. Kočvara. “A New Computational Method for the Sparsest Solutions to Systems of Linear Equations”. In: *SIAM Journal on Optimization* 25.2 (2015), pp. 1110–1134.

AVIS DU JURY SUR LA REPRODUCTION DE LA THESE SOUTENUE

Titre de la thèse:

Contributions aux méthodes numériques pour les problèmes de complémentarité et problèmes d'optimisation sous contraintes de complémentarité

Nom Prénom de l'auteur : MIGOT TANGI

Membres du jury :

- Monsieur LEY Olivier
- Monsieur HADDOU Mounir
- Monsieur ADLY Samir
- Monsieur DUSSAULT Jean-Pierre
- Monsieur KANZOW Christian
- Madame ERHEL Jocelyne
- Monsieur GILBERT Jean-Charles
- Madame SAGASTIZABAL Claudia

Président du jury : *Jean Charles Gilbert*

Date de la soutenance : 06 Octobre 2017

Reproduction de la these soutenue

- ☒ Thèse pouvant être reproduite en l'état
☐ Thèse pouvant être reproduite après corrections suggérées

Fait à Rennes, le 06 Octobre 2017

Signature du président de jury
Jean Charles Gilbert
JCG

Le Directeur,

M'hamed DRISS



Résumé

Dans cette thèse, nous avons étudié les méthodes de régularisation pour la résolution numérique de problèmes avec équilibres.

Dans une première partie, nous nous sommes intéressés aux problèmes de complémentarité au travers de deux applications : les équations en valeur absolue et les problèmes de parcimonie.

Dans une seconde partie, nous avons étudié les problèmes d'optimisation sous contraintes de complémentarité. Après avoir défini des conditions d'optimalité pour ces problèmes nous avons proposé une nouvelle méthode de régularisation appelée méthode des papillons. A partir d'une étude de la résolution des sous-problèmes de la régularisation nous avons défini un algorithme avec des propriétés de convergence forte.

Tout au long de ce manuscrit nous nous sommes concentrés sur les propriétés théoriques des algorithmes ainsi que sur leurs applications numériques.

La dernière partie de ce document est consacrée aux résultats numériques des méthodes de régularisation.

Abstract

In this thesis, we studied the regularization methods for the numerical resolution of problems with equilibria.

In the first part, we focused on the complementarity problems through two applications that are the absolute value equation and the sparse optimization problem.

In the second part, we concentrated on optimization problems with complementarity constraints. After studying the optimality conditions of this problem, we proposed a new regularization method, so-called butterfly relaxation. Then, based on an analysis of the regularized sub-problems we defined an algorithm with strong convergence property.

Throughout the manuscript, we concentrated on the theoretical properties of the algorithms as well as their numerical applications.

In the last part of this document, we presented numerical results using the regularization methods for the mathematical programs with complementarity constraints.