Convergence Rates of First-Order Operator Splitting Methods

Jingwei Liang

▶ **To cite this version:**

Jingwei Liang. Convergence Rates of First-Order Operator Splitting Methods. Optimization and Control [math.OC]. Normandie Université; GREYC CNRS UMR 6072, 2016. English. NNT : . tel-01388978v1

HAL Id: tel-01388978

https://hal.science/tel-01388978v1

Submitted on 27 Oct 2016 (v1), last revised 3 Nov 2016 (v2)

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.

Copyright

Normandie Université

THESE

Pour obtenir le diplôme de doctorat

Spécialité (Mathématiques Appliquées)

Préparée au sein de l'Université de Caen Normandie

Convergence Rates of First-Order Operator Splitting Methods

Présentée et soutenue par

Jingwei LIANG

Thèse soutenue publiquement le (17 Octobre 2016) devant le jury composé de			
Hedy	ATTOUCH	Université de Montpellier	Examinateur
Antonin	CHAMBOLLE	CMAP, École Polytechnique	Examinateur
Jalal	FADILI	GREYC, ENSICAEN	Directeur de thèse
Adrian	LEWIS	ORIE, Cornell University	Rapporteur
Russell	LUKE	University of Göttingen	Examinateur
Jérôme	MALICK	CNRS, Université Grenoble Alpes	Examinateur
Gabriel	PEYRE	CNRS, Université Paris-Dauphine	Codirecteur de thèse
Marc	TEBOULLE	Tel-Aviv University	Rapporteur

Thèse dirigée par Jalal FADILI (et Gabriel PEYRE), **laboratoire GREYC**

ENSICAEN
ÉCOLE NATIONALE SUPÉRIEURE D'INGÉNIEURS DE CAEN
& CENTRE DE RECHERCHE

ed 181 SIMEM

GREYC

To my beloved sister and parents.

To my beloved grand parents.

To my cats, especially the first one who murdered my sparrow.

To my dogs, especially the first one who accompanied me for almost ten years.

献给我最亲爱的家人!

Acknowledgements

About four years and a half ago, 6 March 2012, I had the interview with Jalal and Gabriel which I considered not a successful one. However, three days after the interview, they provided me the offer. Of course, I couldn't help sharing this exciting news with my family, and my parents came with a conclusion that "well, probably because your English is not that awful". I love them so much. Looking back at that moment, this definitely is one of the few best things happened in my life. The only minor thing is that, it took me nine moths to apply the visa, and after a certain point when my friends saw me in the campus, they would say "hey, why are you still here? I thought you have already left, so tired of seeing you again". Well, I love them too. I also wished that this journey could have started earlier, so that by now I would have been working with Jalal and Gabriel for more than four years.

I would like to express my deepest gratitude to Jalal, for his support, guidance, trust and encouragement over the past years. His creative ideas and critical thinking have driven me throughout my PhD study. I really miss the days when we were in the same office and could discuss almost whenever I wanted. From him, I not only learned the way of thinking and solving problems, and more importantly the taste of finding problems. In the meantime, besides being a great mentor, he is also a great friend to me. I really appreciate the conversations we had, the suggestions and understandings he provided me. Thank you very much, Jalal.

I am also deeply grateful to Gabriel. It is a shame that I did not visit him a lot during the past years, which I hope could be made up in the future. Nevertheless, each time we meet, the brain-storm is just unbelievable and enjoyable. He can always come up with very keen observations and inspiring ideas, this is what I have been looking up to and also has motivated me all along the PhD study. Moreover, he is very enthusiastic, passionate and has a great personality. With him and Jalal being my PhD advisors, I feel so blessed.

I feel very privileged to have Adrian Lewis and Marc Teboulle as the reviewers of this thesis, and Hedy Attouch, Antonin Chambolle, Russell Luke and Jérôme Malick as the members of the PhD committee. Their work has been inspiring me since graduate study, it is really a great honour to present the thesis in front of them.

I would like to thank the wonderful friends I made here at the lovely Caen, Maxime Daisy, Youssef Elrhabi, Pierre-Anthony Grenier, Felix Limmeroth, François Lozes and Matthieu Toutain. We had so much fun together, and I wish them the best.

I would like to thank the secretaries Nicole Delamotte and Gaëlle Lenogue (GREYC, ENSICAEN), and Isabelle Bellier (CEREMADE, Paris-Dauphine), for all the care they have taken of me. Especially for Nicole, I wish her a happy life after retirement.

I would like to thank Xiaoqun Zhang at Shanghai Jiao Tong University, and David Cai at Courant Institute of Mathematical Science, for all the help, guidance and encouragement they gave me. The fantastic two years I spent in the Institute of Natural Sciences, is where all these become possible.

Most importantly, I would like to thank my beloved sister Jingjing Liang, parents Fuxiang Qian and Ye Liang, for their endless love and support. This family means everything to me, and I dedicate this thesis to them.

Abstract

This manuscript is concerned with convergence analysis of first-order operator splitting methods that are ubiquitous in modern non-smooth optimization. It consists of three main theoretical advances on this class of methods, namely global convergence rates, novel operator splitting schemes and local linear convergence. First, we propose global (sub-linear) and local (linear) convergence rates for the inexact Krasnosel'skiĭ-Mann iteration built from non-expansive operators, and its application to a variety of monotone operator splitting schemes. Then we design two novel multi-step inertial operator splitting algorithms, both in the convex and non-convex settings, and prove their global convergence. Finally, building on the key concept of partial smoothness, we present a unified and sharp local linear convergence analysis for the class of first-order proximal splitting methods for optimization. We show that for all these algorithms, under appropriate non-degeneracy conditions, the iterates generated by each of these methods will (i) identify the involved partial smooth manifolds in finite time, and then (ii) enter a local linear convergence regime. The linear convergence rates are characterized precisely based on the structure of the optimization problems, that of the proximal splitting scheme, and the geometry of the identified active manifolds. Our theoretical findings are systematically illustrated on applications arising from inverse problems, signal/image processing and machine learning.

Keywords: Monotone inclusion, non-smooth optimization, first-order operator splitting, partial smoothness, (inertial) Forward–Backward/FISTA, Douglas–Rachford/ADMM, Primal–Dual splitting.

Résumé

Ce manuscrit traite de l'analyse de convergence des méthodes du premier ordre d'éclatement d'opérateurs qui sont omniprésents en optimisation non-lisse moderne. Il consiste en trois avancées théoriques principales sur la caractérisation de cette classe de méthodes, à savoir: leur taux de convergence globale, de nouveaux schémas d'éclatement et une analyse de leur convergence linéaire locale. Dans un premier temps, nous proposons des taux de convergence globaux (sous-linéaires) et locaux (linéaires) pour l'itération de Krasnosel'skiĭ-Mann inexacte, et ses applications à un large éventail de schémas d'éclatement d'opérateurs monotones. Ensuite, nous mettons au point deux algorithmes inertiels multi-pas d'éclatement d'opérateurs, pour le cas convexe et non-convexe, et établissons leurs garanties de convergence sur les itérées. Finalement, en s'appuyant sur le concept clé de la régularité partielle, nous présentons une analyse unifiée et précise de la convergence linéaire locale pour les méthodes d'optimisation proximales du premier ordre. Nous montrons que pour tous ces algorithmes, sous des conditions de non-dégénérescence appropriées, les itérées qu'ils génèrent (i) identifient les variétés actives de régularité partielle en temps finis, et ensuite (ii) entre dans un régime de convergence linéaire locale. Les taux de convergence linéaire sont caractérisés précisément, notamment en mettant en jeu la structure du problème d'optimisation, celle du schéma proximal, et la géométrie des variétés actives identifiées. Ces résultats théoriques sont systématiquement illustrés sur des applications issues des problèmes inverses, du traitement du signal et des images et de l'apprentissage.

Mots-clés: Inclusion monotone, optimisation non-lisse, éclatement d'oprateurs, régularité partielle, Implicite-explicite (inertiel), FISTA, Douglas–Rachford/ADMM, algorithmes primaux-duaux.

Table of contents

Chapter 1

Introduction

Contents

1.1 Context, motivations and objectives

1.1.1 Context

In various fields through science and engineering, such as inverse problems, signal/image processing and machine learning, many problems to handle end up solving a structured optimization problem. A classical example of these problems takes the form

$$\min_{x\in\mathcal{H}} \{\Phi(x) \stackrel{\text{def}}{=} F(x) + R(x)\}, \tag{$\mathcal{P}_{\text{opt}}$}$$

where $\mathcal{H}$ is a real Hilbert space, and

(**H.1**) $R : \mathcal{H} \to]-\infty, +\infty]$ is proper convex and lower semi-continuous (lsc).

(**H.2**) $F : \mathcal{H} \to]-\infty, +\infty[$ is convex and differentiable, and ∇F is $(1/\beta)$-Lipschitz continuous.

(**H.3**) $\operatorname{Argmin}(\Phi) \neq \emptyset$, *i.e.* the set of minimizers is non-empty.

More complex forms of problems ($\mathcal{P}_{\text{opt}}$), such as composition with linear operator and/or involving infimal convolutions will be considered in this manuscript.

Now let $x^\star \in \operatorname{Argmin}(\Phi)$ be a global minimizer of ($\mathcal{P}_{\text{opt}}$), then the first-order optimality condition reads

$$0 \in \partial R(x^\star) + \nabla F(x^\star), \tag{1.1.1}$$

where ∂R is the sub-differential of R which is maximal monotone (Definition 2.3.1), and ∇F is β-cocoercive (Definition 2.3.3). In plain words, solving the optimization problem ($\mathcal{P}_{\text{opt}}$) is equivalent to finding the zeros of $\partial R + \nabla F$. This equivalence can be abstracted into the following more general monotone inclusion problem, that is finding the zeros of the sum of two maximal monotone operators,

$$\text{find } \; x \in \mathcal{H} \; \text{ such that } \; 0 \in A(x) + B(x), \tag{$\mathcal{P}_{\text{inc}}$}$$

where

(**H.4**) $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a set-valued maximal monotone operator.

(**H.5**) $B : \mathcal{H} \to \mathcal{H}$ is maximal monotone and β-cocoercive.

(**H.6**) $\operatorname{zer}(A + B) \neq \emptyset$, *i.e.* the set of zeros of $A + B$ is non-empty.

More complex forms of problems ($\mathcal{P}_{\text{inc}}$), such as composition with linear operator and/or involving parallel sums (corresponding to infimal convolutions in the monotone operators language) will also be considered in this manuscript.

First-order operator splitting methods, for solving monotone inclusion problems as ($\mathcal{P}_{\text{inc}}$), are iterative schemes which evaluate each individual operator separately at various points in the course of iteration, rather than evaluating the whole summand in the same time. Taking ($\mathcal{P}_{\text{inc}}$) for example, let $B = 0$, then a classical approach to solve the problem is the Proximal Point Algorithm (PPA) [124, 151], which takes the form

$$x_{k+1} = \mathcal{J}_{\gamma_k A}(x_k),$$

where $\gamma_k > 0$ and $\mathcal{J}_{\gamma_k A} \stackrel{\text{def}}{=} (\mathrm{Id} + \gamma_k A)^{-1}$ is called the resolvent of $\gamma_k A$. This algorithm is very easy to implement if $\mathcal{J}_{\gamma_k A}$ can be computed in closed form, or approximated up to high precision. Now suppose B is non-trivial. In principle, PPA can still be applied. However, even if the resolvents of A and B can be computed in closed form separately, the resolvent of $A+B$ in general is very difficult to compute. In addition, the cocoercivity of B is not exploited. Therefore, a proper numerical scheme should take into account the structure of the problem and the properties of the operators, *i.e.* operator splitting. By doing so, one of the best-known algorithms to solve ($\mathcal{P}_{\text{inc}}$) is the Forward–Backward splitting method (FB) [119, 140], whose non-relaxed iteration takes the form

$$x_{k+1} = \mathcal{J}_{\gamma_k A}\big(x_k - B(x_k)\big),$$

where $\gamma_k \in]0, 2\beta[$ is the step-size. As shown by the formulation, the calculations of the two operators are split into two steps: a forward explicit step on B (gradient descent), and a backward implicit step on A (proximal point).

In the literature, based on the structure and properties of the problems at hand, numerous operator splitting algorithms have been proposed. Popular examples of them include, but not limited to,

- Proximal Point Algorithm [124, 151], and its inertial variants [4, 5].
- Forward–Backward splitting [119, 140], and its variants (*e.g.* inertial FB [131, 120, 115], FISTA [24, 50, 12]).
- Douglas–Rachford splitting (DR) [76, 119], alternating direction method of multipliers (ADMM) [80, 82, 83, 85], Peaceman–Rachford splitting [141].
- The class of Primal–Dual splitting methods (PD) [7, 51, 92, 173, 66, 54, 62].
- Generalized Forward–Backward splitting (GFB) [148], and Forward–Douglas–Rachford splitting (FDR) [40].

Most of the above methods will be studied in this manuscript.

1.1.2 Motivations

Before delving into the details, let us present two numerical experiments that motivated this work in the first place. Consider the following problem

$$\min_{x \in \mathbb{R}^n} R(x) + \frac{1}{2}\|\mathcal{K}x - f\|^2, \tag{1.1.2}$$

where $\mathcal{K} : \mathbb{R}^n \to \mathbb{R}^m$ is a linear operator with $m < n$, $f = \mathcal{K}x_{\text{ob}} + w$ is the observation of an object $x_{\text{ob}} \in \mathbb{R}^n$ contaminated by additive noise w. (1.1.2) is a standard formulation for regularized linear inverse problems, where R is the regularization function which is designed/chosen to promote objects resembling x_{ob}. Under the given setting, the Forward–Backward splitting is a natural candidate to solve (1.1.2).

Let us consider two examples where R is either the ℓ_1-norm (that promotes sparse vectors [122, 159]), or the nuclear norm (which promotes low-rank matrices [45]). For these two examples, the iteration profiles of the Forward–Backward splitting are shown below in Figure 1.1. For both plots of Figure 1.1, the red solid lines stand for $\|x_k - x^\star\|$. For the ℓ_1-norm, the black line shows the cardinality of the support of x_k (denoted as $|\text{supp}(x_k)|$), while for the nuclear norm, the black line denotes the rank of x_k. We can draw the following observations from the plots (keep in mind that the objective of (1.1.2) not strongly convex since $\mathcal{K}$ is underdetermined):

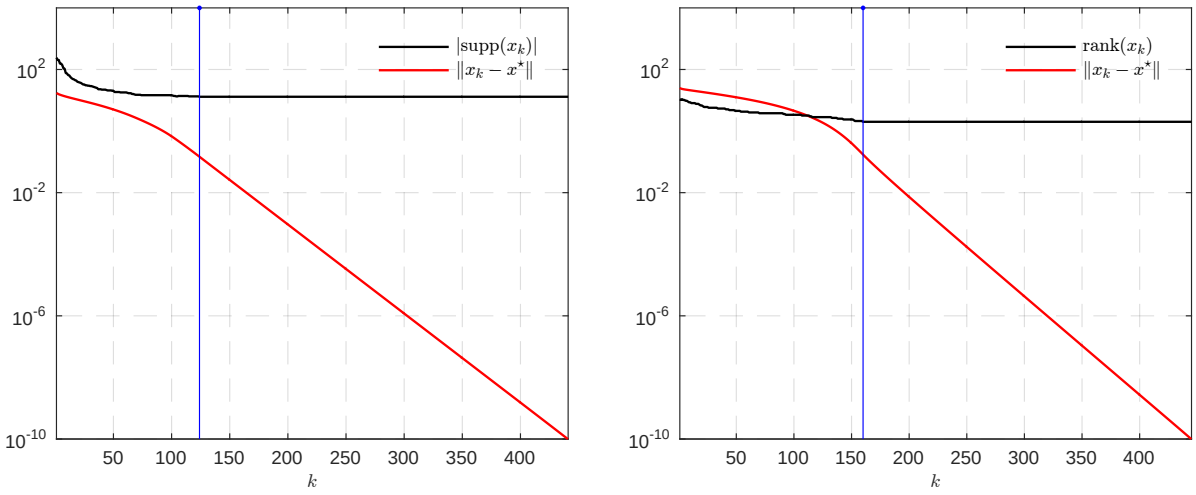
(i) $\|x_k - x^\star\|$ shows two distinct convergence regimes: sub-linear regime from the beginning (left side of the blue line of each figure), and then a local linear regime after certain point (right side of the blue line of each figure).

(ii) For the ℓ_1-norm, the change of $|\text{supp}(x_k)|$ also shows two distinct regimes: a regime where it varies (left side of the blue line) and then stabilizes after finite number of iterations (right side of the blue line). Moreover, the locations of $\text{supp}(x_k)$ also become stabilized which was verified for many runs of the experiments.

(iii) For the nuclear norm, a similar behaviour is observed on the rank of x_k: a regime where the rank varies (left side of the blue line) and then stabilizes after finitely many iterations (right side of the blue line).

(iv) This phase transition phenomenon between the sub-linear to the local linear convergence rates coincides with the points at which the support and the rank become constant.

It should be pointed out that these observations (global sub-linear $\to$ local linear convergence, sup-



(a) ℓ_1-norm (b) Nuclear norm

Figure 1.1: Convergence profiles of FB solving (1.1.2) with R being (a) ℓ_1-norm, and (b) nuclear norm.

port/rank stabilization) are not limited to the presented two functions and the FB method, and turn out to hold true for a large class of functions (*e.g.* $\ell_{1,2}, \ell_\infty$-norms, total variation, and even the non-convex ℓ_0 pseudo-norm, rank function) and a wide range of operator splitting methods (*e.g.* FB/FISTA, Douglas–Rachford/ADMM, Primal–Dual splitting).

Besides first-order operator splitting methods, in the past few years, an increasing attention has been drawn to the Riemannian manifold based optimization methods. The main advantage of Riemannian geometry based methods, is that they may converge much faster than the operator splitting ones. However, the key premise of these Riemannian manifold based methods is that the underlying manifold should be known a prior. For instance, for the fixed rank manifold, Riemannian geometry based optimization methods have been proposed for low-rank matrix recovery [34, 127, 172, 35], with excellent numerical results. However, in order to apply these methods, the rank of x_{ob} should be known, which is barely possible in general.

However, our observations from Figure 1.1 yield a hybrid practical way to overcome this difficulty: apply operator splitting methods first, and turn to the Riemannian manifold based methods whence the rank of the matrix becomes constant. However, another problem raises: how can we make sure that the applied operator splitting scheme will correctly identify the rank, and moreover in finite time? Otherwise one cannot turn to the Riemannian based optimization methods.

In summary, from our discussion above, the following legitimate questions have to be answered:

(**Q1**) How fast is the global sub-linear convergence rates (*i.e.* $O(1/k)$ or $O(1/k^2)$, and in which sense)?

(**Q2**) Why $\mathrm{supp}(x_k)$ (or more generally the activity of x_k) becomes stable after a finite number of iterations? What are the possible mechanisms underlying this transition, and is this a general phenomenon that can be guaranteed for a large class of functions?

(**Q3**) Why $\|x_k - x^\star\|$ shows two convergence regimes, and what is the relation between local linear convergence of $\|x_k - x^\star\|$ and the stabilization of $\mathrm{supp}(x_k)$ (or $\mathrm{rank}(x_k)$)?

(**Q4**) How can we detect such an activity stabilization/identification? Can we exploit this to get faster identification and/or accelerate the local convergence behaviour?

In the literature, numerous works have been carried out in the recent years attempting to answer some of these questions. However most of them focus only on certain specific problems. As a consequent, their results are rather limited and cannot be extended to complicated general cases.

1.1.3 Objectives

The main objectives of this work is to answer all the questions above. In Chapter 3, we present global convergence rates for those operator splitting methods which can be cast into the form of the Krasnosel'skiĭ-Mann iteration and answer question **Q1**. A local linear convergence result under metric sub-regularity is also proposed which partly answers **Q3**. In Chapter 4, we also design two new and provably convergent operator splitting algorithms which numerically shows faster stabilization of activity, and also better local linear convergence speed. From Chapter 6 to 8, based on the notion of "partial smoothness", we build a unified analysis framework, under which we are able to answer questions **Q2**-**Q4** theoretically for a wide spectrum of proximal splitting algorithms.

1.2 Main contributions

1.2.1 Convergence rates of inexact Krasnosel'skiĭ-Mann iteration

Our first main result, which is at the heart of Chapter 3, establishes global sub-linear and local linear convergence of the inexact Krasnosel'skiĭ-Mann iteration [101, 123]. This allows to cover many operator splitting methods as special cases and to answer **Q1** and part of **Q3**.

A single-valued operator is called non-expansive operator if it is 1-Lipschitz continuous.

Definition 1.2.1 (Inexact Krasnosel'skiĭ-Mann iteration). Let $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator whose set of fixed points $\operatorname{fix}(\mathcal{F}) \stackrel{\text{def}}{=} \{z \in \mathcal{H} : z = \mathcal{F}(z)\}$ is non-empty. Let $\lambda_k \in]0, 1]$, and denote $\mathcal{F}_k = \lambda_k \mathcal{F} + (1 - \lambda_k)\mathrm{Id}$. Then the inexact Krasnosel'skiĭ-Mann iteration of $\mathcal{F}$ is defined by

$$z_{k+1} = z_k + \lambda_k\big(\mathcal{F}(z_k) + \varepsilon_k - z_k\big) = \mathcal{F}_k z_k + \lambda_k \varepsilon_k, \tag{1.2.1}$$

where ε_k is the error of approximating $\mathcal{F}(z_k)$. The residual of the iteration is defined as

$$e_k = (\mathrm{Id} - \mathcal{F})(z_k) = \frac{z_k - z_{k+1}}{\lambda_k} + \varepsilon_k.$$

Various first-order operator splitting methods in the literature can be cast as (1.2.1), see those reviewed in page 3 for examples.

1.2.1.1 Main results

For the inexact Krasnosel'skiĭ-Mann iteration (1.2.1), we prove the following results:

(i) Global pointwise and ergodic convergence rates.

(ii) Local linear convergence under metric sub-regularity of $\mathrm{Id} - \mathcal{F}$ (Definition 1.2.3).

(iii) Convergence rates of the non-stationary version of the Krasnosel'skiĭ-Mann iteration.

I - Global convergence rates The global convergence rate of the inexact Krasnosel'skiĭ-Mann iteration (1.2.1) consists of two aspects: pointwise and ergodic.

Define ℓ_+^1 as the set of non-negative summable sequences, and $\Lambda_k = \sum_{j=0}^k \lambda_j$ and $\bar{e}_k = \frac{1}{\Lambda_k}\sum_{j=0}^k \lambda_j e_j$. We prove the following theorem.

Theorem 1.2.2 (Global convergence rate). *For the inexact Krasnosel'skiĭ-Mann iteration (1.2.1),*

Pointwise *if* $0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < 1$ *and* $\{(k+1)\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell_+^1$, *then* $\|e_k\| = O(1/\sqrt{k})$.

Ergodic *Suppose that* $\{\lambda_k\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell_+^1$, *then* $\|\bar{e}_k\| = O(1/\Lambda_k)$.

In Chapter 3, we apply the above result to analyze the convergence rates of several operator splitting methods, including GFB, DR/ADMM and several Primal–Dual splitting methods.

II - Local linear convergence under metric sub-regularity Define $\mathcal{F}' \overset{\text{def}}{=} \mathrm{Id} - \mathcal{F}$.

Definition 1.2.3 (Metric sub-regularity [75]). A set-valued mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is called metrically sub-regular at $\tilde{z}$ for $\tilde{u} \in A(\tilde{z})$ if there exists $\kappa \geq 0$ along with neighbourhood $\mathcal{Z}$ of $\tilde{z}$ such that

$$\mathrm{dist}(z, A^{-1}(\tilde{u})) \leq \kappa \, \mathrm{dist}(\tilde{u}, A(z)), \ \forall z \in \mathcal{Z}. \tag{1.2.2}$$

The infimum of κ such that (1.2.2) holds is called the modulus of metric sub-regularity, and denoted by $\mathrm{subreg}(A; \tilde{z}|\tilde{u})$. The absence of metric regularity is signalled by $\mathrm{subreg}(A; \tilde{z}|\tilde{u}) = +\infty$.

Denote $d_k = \mathrm{dist}(z_k, \mathrm{fix}(\mathcal{F}))$, we prove the following theorem in Section 3.3.

Theorem 1.2.4 (Local linear convergence). *Let $z^\star \in \mathrm{fix}(\mathcal{F})$, suppose that $\mathcal{F}'$ is locally metric sub-regular at $z^\star$, let $\kappa > \mathrm{subreg}(\mathcal{F}'; z^\star|0)$ and $\lambda_k \in]0, 1]$. Assume moreover that $\sum_{k\in\mathbb{N}} \lambda_k \|\varepsilon_k\|$ is sufficiently small and $\lambda_k\|\varepsilon_k\|$ decays fast enough. Then there exists a neighbourhood $\mathcal{Z}$ of $z^\star$, such that for any starting point $z_0 \in \mathcal{Z}$, there exists $\rho \in]0, 1[$ such that for all $k \in \mathbb{N}$,*

$$d_{k+1} = O(\rho^k).$$

Two concrete illustrations of this theorem are provided in Chapter 3.

III - The non-stationary case The Krasnosel'skiĭ-Mann iteration considered in Definition 1.2.1 is stationary, namely the operator $\mathcal{F}$ is fixed along the iterations. In fact, $\mathcal{F}$ usually depends on certain parameters, and when these parameters change along iterations, we obtain the non-stationary Krasnosel'skiĭ-Mann iteration.

Definition 1.2.5 (Non-stationary Krasnosel'skiĭ-Mann iteration). Let $\mathcal{F}_\Gamma : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator depending on a parameter Γ. Let $\lambda_k \in]0, 1]$. Then the non-stationary Krasnosel'skiĭ-Mann iteration is defined by

$$z_{k+1} = z_k + \lambda_k\big(\mathcal{F}_{\Gamma_k}(z_k) + \varepsilon_k - z_k\big) = \mathcal{F}_{\Gamma_k,\lambda_k}(z_k) + \lambda_k\varepsilon_k, \tag{1.2.3}$$

with $\mathcal{F}_{\Gamma_k,\lambda_k} = \lambda_k\mathcal{F}_{\Gamma_k} + (1 - \lambda_k)\mathrm{Id}$.

We prove the following theorem.

Theorem 1.2.6 (Convergence of (1.2.3)). *Assume the following holds:*

(i) $\mathrm{fix}(\mathcal{F}_\Gamma) \neq \emptyset$.

(ii) $\forall k \in \mathbb{N}$, $\mathcal{F}_{\Gamma_k,\lambda_k}$ *is* $(1 + \beta_k)$*-Lipschitz with* $\beta_k \geq 0$, *and* $\{\beta_k\}_{k\in\mathbb{N}} \in \ell^1_+$.

(iii) $\lambda_k \in]0, 1[$ *such that* $0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < 1$.

(iv) $\{\lambda_k\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$.

(v) $\forall \rho \in [0, +\infty[$, *the sequence* $\{\lambda_k\Delta_{k,\rho}\}_{k\in\mathbb{N}} \in \ell^1_+$ *where* $\Delta_{k,\rho} = \sup_{\|z\|\leq\rho} \|\mathcal{F}_{\Gamma_k}(z) - \mathcal{F}_\Gamma(z)\|$.

Then $\{z_k\}_{k\in\mathbb{N}}$ converges weakly to a point $z^\star \in \mathrm{fix}(\mathcal{F}_\Gamma)$. Moreover, the convergence rates of Theorem 1.2.2 and Theorem 1.2.4 remain valid for (1.2.3) if $\{(k+1)\|\varepsilon_k\|\}_{k\in\mathbb{N}}$ and $\{(k+1)\Delta_{k,\rho}\}_{k\in\mathbb{N}} \in \ell^1_+$.

1.2.1.2 Relation to previous work

Global convergence rates After the appearance of our work, the $O(1/\sqrt{k})$ pointwise convergence rate was improved to $o(1/\sqrt{k})$ by the authors of [72].

The convergence rate of the exact Douglas–Rachford is studied in [91], where the authors show that $\|e_k\|$ converges to 0 at the rate of $O(1/\sqrt{k})$. Their result relies heavily on the structure of the fixed-point of DR iteration (*e.g.* the fixed-point operator of DR is firmly-non-expansive, see Definition

2.3.6). Our result goes much beyond this work by considering a more general iterative scheme with an operator that is only non-expansive and that can be evaluated approximately.

In [65], the authors consider the exact Krasnosel'skiĭ-Mann iteration and showed that $\|z^k - z^{k-1}\| = O(1/\sqrt{k})$. Our work differs from [65] in three main aspects: (i) we consider the inexact scheme without any compactness assumption on the domain of $\mathcal{F}$; (ii) we provide a shaper monotonicity property compared to [65, Proposition 11]; (iii) we establish global pointwise and ergodic convergence rates as well as local linear convergence.

Based on the enlargement of maximal monotone operators, in [157], a hybrid proximal extra-gradient method (HPE) is introduced to solve ($\mathcal{P}_{\text{inc}}$) with $B = 0$. The HPE framework encompasses some splitting algorithms of the literature, see [128]. Convergence of HPE is established in [157] and in [42] for its inexact version. The pointwise and ergodic convergence rate of the exact HPE on a similar error criterion as in our work were established in [128].

Local linear convergence In [107], local linear convergence of the distance to the set of zeros of a maximal monotone operator using the exact PPA is established by assuming metric sub-regularity to the maximal operator A of ($\mathcal{P}_{\text{inc}}$). Local convergence rate analysis of PPA under a higher-order extension of metric sub-regularity, namely metric q-sub-regularity $q \in]0, 1]$, is conducted in [113]. For our result, metric sub-regularity is assumed on $\mathrm{Id} - \mathcal{F}$, where $\mathcal{F}$ is the fixed point operator, *i.e.* the resolvent, rather than the maximal monotone operator in the case of PPA. Note also that the work of [107, 113] considers PPA only in its classical form, *i.e.* without errors nor relaxation.

Based on strong regularity, [109] proved local linear convergence of the Method of Alternating Projections (MAP) in the non-convex setting, where the sets are closed and one of which is suitably regular. The linear rate is associated with a modulus of regularity. This is refined later in [20]. In [95], the authors develop local linear convergence results for the MAP and DR to solve non-convex feasibility problems. Their analysis relies on a local version of firm non-expansiveness together with a coercivity condition. It turns out that this coercivity condition holds for mapping $\mathcal{F}$ for which the fixed points are isolated and $\mathrm{Id} - \mathcal{F}$ is metrically regular [94, Lemma 25]. The linear rate they establish, however, imposes a bound on the metric regularity modulus.

1.2.2 Multi-step inertial operator splitting algorithms

In Chapter 4, we partly answer question **Q4** by proposing two novel operator splitting methods, and proving their global convergence. We numerically show that they have faster activity identification and better local linear convergence rates. However, the tools applied in this chapter are not enough to theoretically justify this local behaviour, and this work is the core of the chapters in the second part of this manuscript.

1.2.2.1 Main results

In Chapter 4, we propose two multi-step inertial operator splitting methods:

(i) For the monotone inclusion problem ($\mathcal{P}_{\text{inc}}$), we propose a *variable metric multi-step inertial operator splitting method*, coined "MUSTARD";

(ii) In the context of *non-convex* optimization, we propose a *multi-step inertial Forward–Backward splitting method*, dubbed "ncvx-MiFB".

I - Variable metric multi-step inertial operator splitting Let $\nu \in]0, +\infty[$ be a positive constant, define $\mathcal{M}_\nu \overset{\text{def}}{=} \{\mathcal{V} : \mathcal{H} \to \mathcal{H} \,|\, \mathcal{V} = \mathcal{V}^*, \mathcal{V} \succcurlyeq \nu \mathrm{Id}\}$, *i.e.* the set of *self-adjoint positive definite operators* whose spectrum is bounded from below by ν. The details of MUSTARD for solving ($\mathcal{P}_{\text{inc}}$) are described in Algorithm 1.

Algorithm 1: The MUSTARD algorithm

Initial: let $s \geq 0$ be an positive integer and $\mathcal{S} \overset{\text{def}}{=} \{0, ..., s-1\}$. $x_0 \in \mathcal{H}$, $x_{-i} = x_0, i \in \mathcal{S}$.

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}}, \{b_{i,k}\}_{i\in\mathcal{S}} \in]-1, 2]^s$, $\mathcal{V}_k \in \mathcal{M}_\nu$ and $\gamma_k \in]0, 2\beta\nu[$:

$$\left\lfloor \begin{aligned} y_{a,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}), \\ y_{b,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}), \\ x_{k+1} &= \mathcal{J}_{\gamma_k \mathcal{V}_k^{-1} A}\big(y_{a,k} - \gamma_k \mathcal{V}_k^{-1} B(y_{b,k})\big). \end{aligned} \right. \tag{1.2.4}$$

$k = k + 1$;

until *convergence*;

We prove the global convergence results for the MUSTARD algorithm, that we describe here informally as follows:

(i) $s \geq 2$: We prove in Theorem 4.2.3 conditional convergence of $\{x_k\}_{k\in\mathbb{N}}$ for a fixed metric $\mathcal{V}$. The terminology "conditional convergence" means that the inertial parameters $\{a_{i,k}\}_{i\in\mathcal{S}}$, $\{b_{i,k}\}_{i\in\mathcal{S}}$ need to be chosen depending on $\{x_k\}_{k\in\mathbb{N}}$.

(ii) For the inertial steps $s = 1$:

(a) In Theorem 4.2.5, we establish conditional convergence of $\{x_k\}_{k\in\mathbb{N}}$ for a variable metric $\mathcal{V}_k$.

(b) Moreover, in Theorem 4.2.9, we provide proper choices of the inertial parameters $\{a_{i,k}\}_{i\in\mathcal{S}}$, $\{b_{i,k}\}_{i\in\mathcal{S}}$ and operators $\{\mathcal{V}_k\}_{k\in\mathbb{N}}$ for which $\{x_k\}_{k\in\mathbb{N}}$ is guaranteed to be convergent.

The formal statements and their proofs are in Section 4.2.

Since the MUSTARD algorithm solves ($\mathcal{P}_{\text{inc}}$), it can be readily applied to schemes that solve optimization problems and which can also be cast as an equivalent monotone inclusion problem. One can think for instance to DR, GFB and several Primal–Dual splitting methods. The corresponding results are detailed in Section 4.3.

II - Multi-step inertial FB for non-convex optimization Now consider problem ($\mathcal{P}_{\text{opt}}$), and remove the convexity assumption, *i.e.*

(**H.7**) $R : \mathcal{H} \to]-\infty, +\infty]$ is proper lsc.

(**H.8**) $F : \mathcal{H} \to]-\infty, +\infty[$ is finite-valued, differentiable, and ∇F is $(1/\beta)$-Lipschitz continuous.

We propose the following multi-step inertial FB splitting method for solving this problem in Section 4.4. Given $\gamma > 0$ is a descent step-size, and

$$\operatorname{prox}_{\gamma R}(\cdot) \overset{\text{def}}{=} \operatorname{Argmin}_{x\in\mathcal{H}} R(x) + \tfrac{1}{2\gamma}\|x - \cdot\|^2,$$

denotes the proximity operator of R, which is assumed to be non-empty.

Algorithm 2: The ncvx-MiFB algorithm

Initial: let $s \geq 0$ be an positive integer and $\mathcal{S} \overset{\text{def}}{=} \{0, ..., s-1\}$. $x_0 \in \mathcal{H}$, $x_{-i} = x_0, i \in \mathcal{S}$.

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}}, \{b_{i,k}\}_{i\in\mathcal{S}} \in]-1, 2]^s$ and $\gamma_k \in]0, \beta[$:

$$\left\lfloor \begin{aligned} y_{a,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}), \\ y_{b,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}), \\ x_{k+1} &\in \operatorname{prox}_{\gamma_k R}\big(y_{a,k} - \gamma_k \nabla F(y_{b,k})\big). \end{aligned} \right.$$

$k = k + 1$;

until *convergence*;

With the help of Kurdyka-Łojasiewicz property (KL, see Definition 4.4.2), we prove the following global convergence result for ncvx-MiFB.

Theorem 1.2.7 (Convergence of ncvx-MiFB). *For problem* ($\mathcal{P}_{\text{opt}}$)*, suppose that* (**H.7**)-(**H.8**) *hold,* Φ *is proper lsc and satisfies the KL property, and* R *is bounded from below. For Algorithm 2, if we choose properly the values of* $\gamma_k, a_{i,k}, b_{i,k}$*, then each bounded sequence* $\{x_k\}_{k\in\mathbb{N}}$ *satisfies*

(i) $\{x_k\}_{k\in\mathbb{N}}$ *has finite length, i.e.* $\sum_{k\in\mathbb{N}} \|x_k - x_{k-1}\| < +\infty$.

(ii) *There exists a critical point* $x^\star \in \operatorname{crit}(\Phi)$ *such that* $\lim_{k\to\infty} x_k = x^\star$.

1.2.2.2 Relation to previous work

The MUSTARD algorithm By form, the proposed MUSTARD Algorithm 1 is the most general Forward–Backward splitting method we are aware of. When $s \geq 2$, our proposed Algorithm 1 is new to the literature.

When $s = 0$, Algorithm 1 recovers the variable metric Forward–Backward splitting method proposed in [62]. If further $\mathcal{V}_k \equiv \mathrm{Id}$, then it becomes the classical Forward–Backward splitting method [119, 140]. Moreover for optimization problem ($\mathcal{P}_{\text{opt}}$), Algorithm 1 recovers the gradient descent method when $R = 0$, and the PPA [151] when $F = 0$.

When $s = 1$ and $\mathcal{V}_k \equiv \mathrm{Id}$, based on the choice of the inertial parameters a_k and b_k, Algorithm 1 recovers the following special cases:

- $a_k \in [0, \bar{a}], b_k = 0$: this is the case studied in [131] for ($\mathcal{P}_{\text{inc}}$). In the context of optimization with $R = 0$, Algorithm 1 recovers the heavy ball method by Polyak [146].
- $a_k \in [0, \bar{a}]$, $b_k = a_k$: this corresponds to the work of [120] for solving ($\mathcal{P}_{\text{inc}}$). If moreover restrict $\gamma_k \in]0, \beta]$ and let $a_k \to 1$, then Algorithm 1 recovers the FISTA schemes [24, 50, 13, 12] developed for optimization.
- When $a_k \in [0, \bar{a}]$, $b_k \in]0, \bar{b}]$, $a_k \neq b_k$, Algorithm 1 recovers the general inertial Forward–Backward splitting method we proposed in [115].

The ncvx-MiFB algorithm In the context of non-convex optimization, the convergence property of Forward–Backward for solving ($\mathcal{P}_{\text{opt}}$) was first established in [10] under the assumption that the objective Φ satisfies the Kurdyka-Łojasiewicz property. Following their footprints, [33, 136] established convergence of the special inertial schemes of [131] in the non-convex setting.

1.2.3 Local linear convergence under partial smoothness

From the local linear convergence analysis in Chapter 3, we observe the following drawbacks of metric sub-regularity:

(i) In general, the modulus of metric sub-regularity is difficult to compute;

(ii) It requires that the fixed-point operator of the operators splitting methods should be non-expansive. For instance, this is not valid for the MUSTARD algorithm or FISTA;

(iii) It does not take advantage of the structure of the underlying composite problem (*e.g.* ($\mathcal{P}_{\text{opt}}$)), and thus does not allow to explain and exploit the activity identification, *e.g.* support/rank stabilization of ℓ_1-norm and nuclear norm.

As a result, a more powerful tool is needed to answer questions **Q2**-**Q4**. This tool relies on the key concept of "partial smoothness" that has been developed for optimization in finite-dimensional Euclidean spaces $\mathbb{R}^n$.

1.2.3.1 Partial smoothness

The notion of partial smoothness was first introduced in [108]. This concept, as well as that of identifiable surfaces [174], captures the essential features of the geometry of non-smoothness which are along the so-called active/identifiable manifold. For convex functions, a closely related idea is developed in [106]. Loosely speaking, a partly smooth function behaves smoothly if we move along the identifiable submanifold, and sharply if we move normal to the manifold. In fact, the behaviour of the function and of its minimizers depend essentially on its restrictions to this manifold, hence offering a powerful framework for algorithmic and sensitivity analysis theory.

Given a set $\mathcal{S} \subset \mathbb{R}^n$, define $\mathrm{par}(\mathcal{S})$ is smallest subspace in $\mathbb{R}^n$ which is parallel to $\mathcal{S}$. For simplicity, the definition below is adapted from [108] to the case of convex functions.

Definition 1.2.8 (Partly smooth function). Let $R : \mathbb{R}^n \to]-\infty, +\infty]$ be proper convex lsc. R is said to be *partly smooth at x relative to a set $\mathcal{M}$* containing x if $\partial R(x) \neq \emptyset$, and moreover

Smoothness $\mathcal{M}$ is a C^2-manifold around x, R restricted to $\mathcal{M}$ is C^2 around x.
Sharpness The tangent space $\mathcal{T}_{\mathcal{M}}(x) = \mathrm{par}(\partial R(x))^{\perp}$.
Continuity The set-valued mapping ∂R is continuous at x relative to $\mathcal{M}$.

Many popular regularization functionals, including $\ell_1, \ell_{1,2}, \ell_\infty$-norms, nuclear norm and total variation, are partly smooth.

1.2.3.2 Main results

From Chapter 6 to 8, we present finite identification and local linear convergence analysis for the class of Forward–Backward-type splitting methods, Douglas–Rachford/ADMM, and the class of Primal–Dual splitting methods respectively.

The outcomes of our proposed analysis framework in these three chapters can be summarized in the following abstract result. Let the composite objective function Φ such that each of its individual component functions is partly smooth relative to an active manifold. Let $\{x_k\}_{k\in\mathbb{N}}$ be the sequence generated by a proximal splitting method applied to minimize Φ. Suppose that $\{x_k\}_{k\in\mathbb{N}}$ is convergent to $x^\star \in \mathrm{Argmin}(\Phi)$, and an appropriate non-degeneracy condition holds. Then,

(i) finite activity identification (answers **Q2**): there exists K large enough such that for all $k \geq K$, x_k belongs to an active manifold $\mathcal{M}_{x^\star}$.

(ii) The identification iteration can be bounded and increases as the non-degeneracy condition becomes more demanding (answers **Q4**);

(iii) Local linearization: the proximal splitting method locally linearizes, and the matrix M in this linearization depends on the structure of the active manifolds;

(iv) Local linear convergence (answers **Q3**): the iterates $\{x_k\}_{k\in\mathbb{N}}$ enter a local linear convergence regime whose rate is precisely determined by the spectral properties of M;

(v) Locally, the parameters of the proximal splitting method can be optimized to attain the fastest convergence rate (answers **Q4**).

Specializing the above result to each class of proximal splitting methods, we obtain additional key results:

Forward–Backward-type methods We explain why FISTA (with convergent sequences) locally oscillates and is eventually locally slower than FB (Section 6.4). In addition, we provide clues on how to avoid these oscillations.

Douglas–Rachford/ADMM We provide conditions under which finite convergence of the sequence generated by DR (or ADMM) can by obtained. This is a striking new result (Section 7.5).

1.2.3.3 Relation to previous work

In [89, 90, 88], the authors have shown finite identification of active manifolds associated to partly smooth functions for a few algorithms, namely the (sub)gradient projection method, Newton-like methods, the proximal point algorithm and the algorithm in [167]. Their work extends that of *e.g.* [174] on identifiable surfaces . The algorithmic framework we consider encompasses all the aforementioned methods as special cases. Moreover, in all these works, the local convergence behaviour was not studied.

Forward–Backward-type methods Finite support identification and local linear convergence of FB for solving a special instance of ($\mathcal{P}_{\text{opt}}$) where F is quadratic and R the ℓ_1-norm, though in infinite-dimensional setting, is established in [36]. A similar result is proved in [87], for F being a smooth convex and locally C^2 function and R the ℓ_1-norm, under restricted injectivity and non-degeneracy assumptions. The ℓ_1-norm is polyhedral, hence partly smooth function, and is therefore covered by our results. [3] proved local linear convergence of FB to solve ($\mathcal{P}_{\text{opt}}$) for F satisfying restricted smoothness and strong convexity assumptions, and R being a so-called convex decomposable regularizer which is a subclass of partly smooth functions. Local linear convergence rate of FB for nuclear norm regularization is studied in [98] under local strong convexity assumption. Local linear convergence of FISTA for the Lasso problem (*i.e.* ($\mathcal{P}_{\text{opt}}$) for F quadratic and R the ℓ_1 norm) has been recently addressed, for instance in [165], and also [99] under some additional constraints on the inertial parameters.

Douglas–Rachford/ADMM There are problem instances in the literature where the (stationary) DR and ADMM algorithms are proved to converge linearly either globally or locally. For instance, in [119, Proposition 4], it is assumed that the "internal" function is strongly convex with a Lipschitz continuous gradient. This local linear convergence result is further investigated in [72, 84] under smoothness and strong convexity assumptions. The special case of basis pursuit, *i.e.* ℓ_1 minimization with an affine constraint, is considered in [73] and an eventual local linear convergence is shown in the absence of strong convexity. In [26], the author analyses the local convergence behaviour of ADMM for quadratic or linear programs, and shows local linear convergence if the optimal solution is unique and strict complementarity holds. A similar result is reported recently in [8]. For the case of two subspaces (though in general real Hilbert space), linear convergence of DR with the optimal rate being the cosine of the Friedrichs angle between the subspaces is proved in [17]. It turns out that [73, 26, 17] are special cases of our framework, and our results generalize theirs to a larger class of problems.

Our finite convergence result complements and extends that of [19] who established finite convergence of (unrelaxed stationary) DR in the presence of Slater's condition, for solving convex feasibility problems where one set is an affine subspace and the other is a polyhedron.

Primal–Dual splitting Recently in [161], a local linear convergence result is presented for a Primal–Dual splitting method applied to the problem of minimizing the sum of a convex quadratic function and a convex gauge under a linear constraint. However, this structure is far too restrictive (for instance sublinearity is irrelevant). This result is a very special case of those we develop in Chapter 8.

1.3 Organisation of the manuscript

This manuscript consists of two parts and nine chapters.

Chapter 2: this chapter collects the necessary mathematical material used throughout the manuscript.

Chapter 3: in this chapter, we present the global convergence rates of the inexact Krasnosel'skiĭ-Mann iteration built from non-expansive operators. Our results include two main parts: global pointwise and ergodic convergence rates (Theorem 3.2.1 and 3.2.4), and then, under a metric sub-regularity assumption, we establish a local linear convergence result (Theorem 3.3.3).

Chapter 4: this chapter is devoted to the following two novel multi-step inertial operator splitting methods:

(i) A variable metric multi-step inertial operator splitting method, dubbed as "MUSTARD", for general monotone inclusion problems.

(ii) A multi-step inertial Forward–Backward splitting method, which is called "ncvx-MiFB", for non-convex optimization.

Detailed convergence properties of the two algorithms are discussed in Section 4.2 and 4.4 respectively.

Chapter 5: in this chapter we introduce the notion of partial smoothness which lays the ground for finite identification and local linear convergence of proximal splitting methods.

Chapter 6: in this chapter, we present our results on finite activity identification and local linear convergence for a wide class of Forward-Backward-type splitting methods (*e.g.* FB, (multi-step) inertial schemes, and sequence convergent FISTA), under the assumption that the non-smooth part of the optimization problem is partly smooth. Our local convergence analysis allows us to establish and explain why FISTA locally oscillates and is actually locally slower than FB (Section 6.4). Extension to the non-convex case is also presented.

Chapter 7: this chapter is dedicated to understanding the local linear convergence behaviour of DR/ADMM when the involved functions are partly smooth. When both functions are locally polyhedral, we show that the optimal convergence rate is given in terms of the cosine of the Friedrichs angle (Definition 7.4.2) between the tangent spaces of the identified manifolds. We also characterize situations where finite convergence of DR occurs.

Chapter 8: in this chapter, we consider the class of Primal-Dual splitting methods for minimizing composite structured non-smooth optimization problems involving smooth terms, partly smooth ones, infimal convolutions and composition by linear operators. We establish finite identification and characterize the linear convergence regime. Under special scenarios, we also draw connections between Primal–Dual splitting on the one hand, and the FB and DR/ADMM on the other hand.

Chapter 9: this last chapter summarizes our contributions, and draws important conclusions. It also discusses several interesting perspectives and open problems.

Chapter 2

Mathematical Background

Contents

In this chapter, we collect the necessary mathematical material used in the manuscript.

Denote $\mathcal{H}$ the real Hilbert space equipped with scalar inner product $\langle \cdot, \cdot \rangle$ and the associated norm $\| \cdot \|$. Id denotes the identity operator on $\mathcal{H}$. Denote $\mathbb{R}^n$ the n-dimensional real Euclidean space, $\mathbb{R}_+ =]0, +\infty[$ the set of positive values and $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ the extended real value, ℓ_+^1 denotes the set of summable sequences in $[0, +\infty[$. $\mathbb{N}$ denotes the set of non-negative integers, $\mathbb{N}_+$ denotes the set of positive integers. Let $\mathcal{S} \subset \mathcal{H}$ be a nonempty subset of $\mathcal{H}$, we denote $\mathrm{ri}(\mathcal{S})$ and $\mathrm{rbd}(\mathcal{S})$ its relative interior and boundary respectively. The smallest linear subspace of $\mathcal{H}$ that contains $\mathcal{S}$ is denoted by $\mathrm{span}(\mathcal{S})$. The smallest affine subspace that contains $\mathcal{S}$ is denoted by $\mathrm{aff}(\mathcal{S})$ which is also called the affine hull of $\mathcal{S}$, then $\mathrm{par}(\mathcal{S}) = \mathbb{R}_+(\mathcal{S} - \mathcal{S})$ denotes the subspace parallel to $\mathrm{aff}(\mathcal{S})$.

2.1 Convex analysis

We recall some important concepts and results that will be crucial to our exposition. The proofs are classical and can be found for instance in [16], which is comprehensive account on convex analysis in Hilbert spaces.

2.1.1 Convex sets and functions

Definition 2.1.1 (Convex set). A set $\mathcal{S}$ of $\mathcal{H}$ is *convex* if,

$$\forall x, x' \in \mathcal{S},\ \forall t \in]0,1[,\ tx + (1-t)x' \in \mathcal{S}.$$

Let $\mathcal{S} \subseteq \mathcal{H}$ be a nonempty set and function $F : \mathcal{S} \to \overline{\mathbb{R}}$, the *domain* of F is

$$\mathrm{dom}(F) \stackrel{\text{def}}{=} \big\{x \in \mathcal{S} : F(x) < +\infty\big\},$$

and its *epigraph* is

$$\mathrm{epi}(F) \stackrel{\text{def}}{=} \big\{(x,u) \in \mathcal{S} \times \mathbb{R} : F(x) \le u\big\}.$$

Function F is called *proper* if $-\infty \notin F(\mathcal{S})$ and $\mathrm{dom}(F) \neq \emptyset$.

Definition 2.1.2 (Convex function). A function $F : \mathcal{H} \to \overline{\mathbb{R}}$ is convex if $\mathrm{dom}(F)$ is convex and

$$\forall x, x' \in \mathrm{dom}(F), \forall t \in [0,1],\ F(tx + (1-t)x') \le tF(x) + (1-t)F(x').$$

It is moreover *strongly convex* with modulus δ, if $\delta > 0$ and

$$F(tx + (1-t)x') \le tF(x) + (1-t)F(x') - \frac{\delta}{2}t(1-t)\|x - x'\|^2.$$

Convexity can be preserved by various operations, such as pointwise maximum and supremum, sum by non-negative weights, composition with an affine operator, etc.

Definition 2.1.3 (Lower semi-continuity). Given a function $F : \mathcal{H} \to \overline{\mathbb{R}}$ and a point $x \in \mathcal{H}$. F is *lower semi-continuous (lsc)* at x if,

$$\liminf_{x' \to x} F(x') \ge F(x).$$

The class of *proper, convex and lsc functions* on $\mathcal{H}$ is denoted as $\Gamma_0(\mathcal{H})$.

Example 2.1.4 (Indicator function). Let $\mathcal{S} \subseteq \mathcal{H}$ be a convex non-empty closed set, the *indicator function* of $\mathcal{S}$, $\iota_{\mathcal{S}} \in \Gamma_0(\mathcal{H})$, is defined by

$$\iota_{\mathcal{S}}(x) = \begin{cases} 0, & \text{if } x \in \mathcal{S}, \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.1.1}$$

2.1.2 Infimal convolution and proximal mapping

Definition 2.1.5 (Infimal convolution). Let functions $F, G \in \Gamma_0(\mathcal{H})$, the *infimal convolution* of F and G is defined by

$$(F \mathbin{\triangledown} G)(x) \stackrel{\text{def}}{=} \inf_{y \in \mathcal{H}} \big(F(y) + G(x-y)\big), \tag{2.1.2}$$

which is convex.

The infimal convolution of indicator function $\iota_{\mathcal{S}}(\cdot)$ and the quadratic function $\|\cdot\|^2$ leads to the *distance function*, which finds the points in $\mathcal{S}$ closest to a given point x.

Example 2.1.6 (Distance function). Let $\mathcal{S} \subset \mathcal{H}$ be a convex set and a point $x \in \mathcal{H}$, define $\mathrm{dist}(x, \mathcal{S})$ the distance of x to $\mathcal{S}$ by the function

$$\mathrm{dist}(x, \mathcal{S}) \stackrel{\text{def}}{=} \inf_{y \in \mathcal{S}} \|x - y\|.$$

Then we have,

$$\begin{aligned} \mathrm{dist}^2(x, \mathcal{S}) &\stackrel{\text{def}}{=} \inf_{y \in \mathcal{S}} \|x - y\|^2 \\ &= \inf_{y \in \mathcal{S}} \big(\iota_{\mathcal{S}}(y) + \|x - y\|^2\big) = (\iota_{\mathcal{S}} \mathbin{\triangledown} \|\cdot\|^2)(x). \end{aligned}$$

One very important instance of infimal convolution is the *Moreau envelope (or Moreau-Yosida regularization)*.

Definition 2.1.7. Let function $F \in \Gamma_0(\mathcal{H})$, and $\gamma > 0$. The *Moreau envelope* of F parameterised by γ is the infimal convolution

$$ {}^{\gamma}F(x) \stackrel{\text{def}}{=} \big(F \,\dot{\vee}\, \big(\frac{1}{2\gamma}\|\cdot\|^2\big)\big)(x) = \inf_{y\in\mathcal{H}} \big(F(y) + \frac{1}{2\gamma}\|x-y\|^2\big). \tag{2.1.3} $$

It is obvious that the distance function is the Moreau envelope of the indicator function parameterised by $\gamma = 1/2$.

The Moreau envelope is continuous and strongly convex, given any $x \in \mathcal{H}$, the infimum of (2.1.3) is uniquely attained. The unique minimizer determined by (2.1.3) is called the *proximal point* of x, and the evaluation to get such point is called the *proximity operator*, see the definition below.

Definition 2.1.8 (Proximity operator). Let function $F \in \Gamma_0(\mathcal{H})$, parameter $\gamma > 0$ and a point $x \in \mathcal{H}$, the *proximity operator (or proximal mapping)* of F parameterised by γ is defined by

$$ \operatorname{prox}_{\gamma F}(x) \stackrel{\text{def}}{=} \operatorname{Argmin}_{y\in\mathcal{H}} F(y) + \frac{1}{2\gamma}\|y-x\|^2, $$

where $\operatorname{Argmin}(F)$ the set of minimizers of F, and it is a *singleton* for this case.

A very simple and widely used example of proximity operator is the *projection operator*.

Example 2.1.9 (Projection operator). Let $\mathcal{S}$ be a non-empty closed convex set, define its *projection operator* as, $\forall x \in \mathcal{H}$

$$ \mathcal{P}_{\mathcal{S}}(x) \stackrel{\text{def}}{=} \operatorname{Argmin}_{y\in\mathcal{S}} \|x-y\|, $$

then we have $\operatorname{prox}_{\iota_{\mathcal{S}}} = \mathcal{P}_{\mathcal{S}}$, and ${}^{\gamma}\iota_{\mathcal{S}}(x) = \operatorname{dist}^2(x,\mathcal{S}), \forall\gamma > 0$.

The Moreau envelope is not only as mentioned continuous, but moreover has Lipschitz continuous gradient.

Lemma 2.1.10. *Let $F \in \Gamma_0(\mathcal{H})$, and $\gamma > 0$, then ${}^{\gamma}F$ is differentiable with gradient reading*

$$ \nabla({}^{\gamma}F) \stackrel{\text{def}}{=} \frac{1}{\gamma}(\operatorname{Id} - \operatorname{prox}_{\gamma F}), $$

which is $(1/\gamma)$-Lipschitz continuous.

2.1.3 Differentiability

Definition 2.1.11. Let function $J \in \Gamma_0(\mathcal{H})$, the sub-differential of J is a set-valued operator, $\forall x \in \mathcal{H}$

$$ \partial J(x) \stackrel{\text{def}}{=} \big\{g \in \mathcal{H} : \langle x'-x,\, g\rangle + J(x) \le J(x'),\ \forall x' \in \mathcal{H}\big\}. $$

J is *sub-differentiable* at x if $\partial J(x) \neq \emptyset$, and an element of $\partial J(x)$ is called a *sub-gradient*.

The sub-differential is a generalization of derivative to non-differentiable functions, and it is always a convex closed set.

Example 2.1.12 (Normal cone). Let $\mathcal{S} \subseteq$ be a non-empty closed convex set, the sub-differential of the indicator function of $\mathcal{S}$ is define by

$$ \partial\iota_{\mathcal{S}}(x) = \begin{cases} \mathcal{N}_{\mathcal{S}}(x) \stackrel{\text{def}}{=} \big\{g \in \mathcal{H} : \langle x'-x,\, g\rangle \le 0,\ \forall x' \in \mathcal{S}\big\}, & \text{if } x \in \mathcal{S}, \\ \emptyset, & \text{otherwise,} \end{cases} $$

where $\mathcal{N}_{\mathcal{S}}(\cdot)$ denotes the *normal cone* operator, see *e.g.* [152, Section 23].

Given a function, the Fermat's theorem adequately characterises its extrema. For convex functions, we have the following.

Theorem 2.1.13 (Fermat's theorem). *Let $J : \mathcal{H} \to \overline{\mathbb{R}}$ be proper convex, then $x^\star$ is a minimizer of J if, and only if*

$$0 \in \partial J(x^\star). \tag{2.1.4}$$

Condition (2.1.4) is also called the first-order (necessary and sufficient) optimality condition.

2.1.4 Conjugacy and duality

Definition 2.1.14 (Conjugate). Let $F : \mathcal{H} \to \overline{\mathbb{R}}$, the *Fenchel conjugate* of F is defined by

$$F^*(v) \stackrel{\text{def}}{=} \sup_{x \in \mathcal{H}} \left(\langle x, v \rangle - F(x) \right).$$

Fenchel conjugate, or simply conjugate, is also called *Legendre transform* or *Legendre-Fenchel transform*. For indicator function $\iota_{\mathcal{S}}(x)$, the conjugate turns out to be the support function of $\mathcal{S}$.

Example 2.1.15 (Support function). Let $\mathcal{S} \subseteq \mathcal{H}$ be a non-empty convex set, the *support function* of $\mathcal{S}$ is defined by

$$\sigma_{\mathcal{S}}(v) \stackrel{\text{def}}{=} \sup_{x \in \mathcal{S}} \langle x, v \rangle = \iota^*_{\mathcal{S}}(y).$$

The *biconjugate* of F is defined by $F^{**} \stackrel{\text{def}}{=} (F^*)^*$, and by Definition 2.1.14 there holds $F^{**} \leq F$, moreover we have the following Fenchel-Moreau theorem [152, Corollary 12.2.1].

Theorem 2.1.16 (Fenchel-Moreau). *Let $F : \mathcal{H} \to \overline{\mathbb{R}}$ be a proper function, then $F = F^{**}$ if and only if F is convex and lower semi-continuous.*

The proximity operators of a $\Gamma_0(\mathcal{H})$ function and its conjugate are related by the *Moreau's identity*.

Theorem 2.1.17 (Moreau's identity). *Let function $F \in \Gamma_0(\mathcal{H})$ and $\gamma > 0$, then*

$$\mathrm{Id} = \mathrm{prox}_{\gamma F}(\cdot) + \gamma \, \mathrm{prox}_{F^*/\gamma}\left(\frac{\cdot}{\gamma}\right). \tag{2.1.5}$$

2.1.4.1 Fenchel-Rockafellar duality

Definition 2.1.18 (Primal and dual problems). Let $\mathcal{H}, \mathcal{G}$ be two real Hilbert spaces, $F \in \Gamma_0(\mathcal{H}), J \in \Gamma_0(\mathcal{G})$ and $L : \mathcal{H} \to \mathcal{G}$ be a bounded linear operator. The *primal problem* is defined by

$$\min_{x \in \mathcal{H}} F(x) + J(Lx), \tag{2.1.6}$$

and the Fenchel-Rockafellar *dual problem* is

$$\min_{v \in \mathcal{G}} F^*(-L^*v) + J^*(v), \tag{2.1.7}$$

where L^* is the adjoint of L. Let $x^\star, v^\star$ be the minimizers of the primal and dual problems respectively, then the *duality gap* is defined by

$$\mathcal{D}(x^\star, v^\star) \stackrel{\text{def}}{=} \big(F(x^\star) + J(Lx^\star)\big) - \big(-F^*(-L^*v^\star) - J^*(v^\star)\big).$$

Theorem 2.1.19 (Fenchel-Rockafellar). *For the primal and dual problems defined in Definition 2.1.18, we have*

$$F(x) + J(Lx) \geq -F^*(-L^*v) - J^*(v), \quad \forall x \in \mathcal{H} \quad \text{and} \quad \forall v \in \mathcal{G}.$$

If $\mathcal{G}$ is finite dimensional and $\mathrm{ri}(\mathrm{dom}(F)) \cap L\,\mathrm{ri}(\mathrm{dom}(J)) \neq \emptyset$, then

$$\inf_{x \in \mathcal{H}} \big(F(x) + J(Lx)\big) = -\min_{v \in \mathcal{G}} \big(F^*(-L^*v) + J^*(v)\big).$$

Proof. A result of combining Proposition 15.24(viii) and Theorem 15.23 in [16]. □

2.2 Variational analysis

We here collect some important results from variational analysis which will be used in Chapter 4 and Chapter 6.

Let $J : \mathcal{H} \to \overline{\mathbb{R}}$ be a lower semi-continuous function. Given $x \in \mathrm{dom}(J)$, the Fréchet subdifferential $\partial^F J(x)$ of J at x, is the set of vectors $v \in \mathcal{H}$ that satisfies

$$\liminf_{z \to x,\, z \neq x} \frac{1}{\|x - z\|}\big(J(z) - J(x) - \langle v,\, z - x\rangle\big) \geq 0.$$

If $x \notin \mathrm{dom}(J)$, then $\partial^F J(x) = \emptyset$. The limiting-subdifferential (or simply subdifferential) of J at x, written as $\partial J(x)$, is defined as,

$$\partial J(x) \stackrel{\text{def}}{=} \big\{v \in \mathcal{H} : \exists x_k \to x, J(x_k) \to J(x), v_k \in \partial^F J(x_k) \to v\big\}. \tag{2.2.1}$$

Both $\partial^F J(x)$ and $\partial J(x)$ are closed, with $\partial^F J(x)$ convex and $\partial^F J(x) \subset \partial J(x)$ [153, Proposition 8.5]. Since J is lsc, it is (sub-differentially) regular at x if and only if $\partial^F J(x) = \partial J(x)$ [153, Corollary 8.11]. These generalized notions of differentiation allow us to define the critical point. A necessary (but not sufficient) condition for x to be a minimizer of J is $0 \in \partial J(x)$. The set of critical points of J is defined as $\mathrm{crit}(J) \stackrel{\text{def}}{=} \{x \in \mathcal{H} : 0 \in \partial J(x)\}$.

Definition 2.2.1 (Prox-regularity [143]). An lsc function J is r-prox-regular at $\bar{x} \in \mathrm{dom}(J)$ for $\bar{v} \in \partial J(\bar{x})$ if $\exists r > 0$ such that

$$J(x') > J(x) + \langle v,\, x' - x\rangle - \frac{1}{2r}\|x - x'\|^2,$$

whenever x and x' are near $\bar{x}$, with $J(x)$ near $J(\bar{x})$ and $v \in \partial J(x)$ is near $\bar{v}$. J is prox-regular at $\bar{x}$ if it is prox-regular for every $\bar{v} \in \partial J(\bar{x})$.

2.3 Monotone and non-expansive operators

Given 2 non-empty sets $\mathcal{X}, \mathcal{U} \subseteq \mathcal{H}$, an operator (or mapping) $A : \mathcal{X} \rightrightarrows \mathcal{U}$ is called set-valued operator if A maps every point in $\mathcal{X}$ to a subset of $\mathcal{U}$, *i.e.*

$$A : \mathcal{X} \rightrightarrows \mathcal{U},\ x \in \mathcal{X} \mapsto A(x) \subseteq \mathcal{U}.$$

The *graph* of A is defined by $\mathrm{gra}(A) \stackrel{\text{def}}{=} \big\{(x, u) \in \mathcal{X} \times \mathcal{U} : u \in A(x)\big\}$. The *domain* and *range* of A are $\mathrm{dom}(A) \stackrel{\text{def}}{=} \big\{x \in \mathcal{X} : A(x) \neq \emptyset\big\}$ and $\mathrm{ran}(A) \stackrel{\text{def}}{=} A(\mathcal{X})$ respectively. The *inverse* of A defined through its graph $\mathrm{gra}(A^{-1}) \stackrel{\text{def}}{=} \big\{(u, x) \in \mathcal{U} \times \mathcal{X} : u \in A(x)\big\}$. The set of zeros of A are the points such that $\mathrm{zer}(A) \stackrel{\text{def}}{=} A^{-1}(0) = \big\{x \in \mathcal{X} : 0 \in A(x)\big\}$.

2.3.1 Monotone operators

Definition 2.3.1 (Monotone operator). Let $\mathcal{X}, \mathcal{U} \subseteq \mathcal{H}$ be two non-empty convex sets, a set-valued operator $A : \mathcal{X} \rightrightarrows \mathcal{U}$ is *monotone* if

$$\langle x - x',\, u - u'\rangle \geq 0,\ \forall (x, u), (x', u') \in \mathrm{gra}(A).$$

It is moreover *maximal monotone* if $\mathrm{gra}(A)$ is not strictly contained in the graph of any other monotone operators.

If there exists a constant $\delta > 0$ such that

$$\langle x - x',\, u - u'\rangle \geq \delta\|x - x'\|^2.$$

Then A is called *strongly monotone*.

A very important source of maximal monotone operators is the sub-differential of $\Gamma_0(\mathcal{H})$ functions [152], and we have the following lemma.

Lemma 2.3.2. *Let function $J \in \Gamma_0(\mathcal{H})$, then ∂J is maximal monotone.*

Definition 2.3.3 (Cocoercive operator). An operator $B : \mathcal{H} \to \mathcal{H}$ is called β-cocoercive if there exists $\beta > 0$ such that

$$\beta\|B(x) - B(x')\|^2 \leq \langle B(x) - B(x'),\, x - x'\rangle, \quad \forall x, x' \in \mathcal{H}.$$

The above equation implies that B is $(1/\beta)$-Lipschitz continuous.

Owing to the [15, Baillon-Haddad theorem], the Lipschitz continuous gradient of a convex function is cocoercive.

Theorem 2.3.4 (Baillon-Haddad [15]). *Let function $F \in \Gamma_0(\mathcal{H})$ be a convex differentiable function whose gradient ∇F is $(1/\beta)$-Lipschitz continuous, then ∇F is β-cocoercive.*

The next lemma shows the connection between strong monotonicity and cocoercivity.

Lemma 2.3.5. *Let $C : \mathcal{H} \rightrightarrows \mathcal{H}$ be a strongly monotone operator for some $\beta > 0$, then its inverse C^{-1} is β-cocoercive.*

2.3.2 Non-expansive operators

Definition 2.3.6 (Non-expansive operator). An operator $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ is called *non-expansive* if it is 1-Lipschitz continuous, *i.e.*

$$\|\mathcal{T}(x) - \mathcal{T}(x')\| \leq \|x - x'\|, \quad \forall x, x' \in \mathcal{H}.$$

For any $\alpha \in]0, 1[$, $\mathcal{T}$ is α-averaged if there exists a non-expansive operator $\mathcal{T}'$ such that

$$\mathcal{T} = \alpha\mathcal{T}' + (1 - \alpha)\mathrm{Id}.$$

$\mathcal{A}(\alpha)$ denotes the class of α-averaged operators on $\mathcal{H}$, in particular $\mathcal{A}(\frac{1}{2})$ is the class of firmly non-expansive operators. The following lemmas collect some properties of the α-averaged non-expansive operators and the firmly non-expansive operators.

Lemma 2.3.7. *Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator and $\alpha \in]0, 1[$. The following statements are equivalent:*

(i) *$\mathcal{T}$ is α-averaged non-expansive.*

(ii) *$(1 - 1/\alpha)\mathrm{Id} + (1/\alpha)\mathcal{T}$ is non-expansive.*

(iii) $\|\mathcal{T}(x) - \mathcal{T}(x')\|^2 \leq \|x - x'\|^2 - \frac{1-\alpha}{\alpha}\|(\mathrm{Id} - \mathcal{T})(x) - (\mathrm{Id} - \mathcal{T})(x')\|^2$, $\forall x, x' \in \mathcal{H}$.

Lemma 2.3.8. *Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator. The following statements are equivalent:*

(i) *$\mathcal{T}$ is firmly non-expansive.*

(ii) *$2\mathcal{T} - \mathrm{Id}$ is non-expansive.*

(iii) $\|\mathcal{T}(x) - \mathcal{T}(x')\|^2 \leq \langle \mathcal{T}(x) - \mathcal{T}(x'),\, x - x'\rangle, \forall x, x' \in \mathcal{H}$.

Lemma 2.3.9. *Let operator $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive for some $\beta > 0$. Then*

(i) *$\beta B \in \mathcal{A}(\frac{1}{2})$, i.e. is firmly non-expansive.*

(ii) *$\mathrm{Id} - \gamma B \in \mathcal{A}(\frac{\gamma}{2\beta})$ for $\gamma \in]0, 2\beta[$.*

The following lemmas collect some properties of the class of α-averaged non-expansive operators, that it is closed under relaxations, convex combinations and compositions.

Lemma 2.3.10. *Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be an α-averaged operator, then*

(i) $\mathrm{Id} + \lambda(\mathcal{T} - \mathrm{Id}) \in \mathcal{A}(\lambda\alpha)$, $\lambda \in]0, \frac{1}{\alpha}[$.

(ii) $\frac{1}{2\alpha}(\mathrm{Id} - \mathcal{T}) \in \mathcal{A}(\frac{1}{2})$.

Lemma 2.3.11. *Let $m \in \mathbb{N}_+$, $(\mathcal{T}_i)_{i\in\{1,\dots,m\}}$ be a finite family of non-expansive operators from $\mathcal{H}$ to $\mathcal{H}$, $(\omega_i)_i \in]0,1]^m$ and $\sum_i \omega_i = 1$, and let $(\alpha_i)_i \in]0,1]^m$ such that, for each $i \in \{1,\dots,m\}, \mathcal{T}_i \in \mathcal{A}(\alpha_i)$. Then,*

(i) $\sum_i \omega_i \mathcal{T}_i \in \mathcal{A}(\alpha)$ *with* $\alpha = \max_i \alpha_i$.

(ii) $\mathcal{T}_1 \cdots \mathcal{T}_m \in \mathcal{A}(\alpha)$ *with* $\alpha = \frac{m}{m-1+1/\max_{i\in\{1,\dots,m\}} \alpha_i}$.

Remark 2.3.12. A sharper bound of α can be obtained for $m = 2$ [64, Proposition 2.4],

$$\alpha = \frac{\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2}{1 - \alpha_1\alpha_2} \in]0,1[.$$

See also [138, Theorem 3].

2.3.2.1 Fixed-point iteration

Definition 2.3.13 (Fixed point). Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator, $x \in \mathcal{H}$ is called the *fixed point* of $\mathcal{T}$ if

$$x = \mathcal{T}(x).$$

We denote the set of fixed points of $\mathcal{T}$ as $\mathrm{fix}(\mathcal{T})$.

The set of fixed point of a non-expansive operator may be empty, such as translation by a non-zero vector. The existence of the fixed point is guaranteed by the theorem below.

Theorem 2.3.14. *Let $\mathcal{X}$ be a non-empty bounded closed convex subset of $\mathcal{H}$ and $\mathcal{T} : \mathcal{X} \to \mathcal{X}$ be a non-expansive operator, then* $\mathrm{fix}(\mathcal{T}) \neq \emptyset$.

Lemma 2.3.15. *Let $\mathcal{X}$ be a non-empty closed convex subset of $\mathcal{H}$ and $\mathcal{T} : \mathcal{X} \to \mathcal{H}$ be a non-expansive operator, then* $\mathrm{fix}(\mathcal{T})$ *is closed and convex.*

The fixed points of non-expansive operators, in most cases, cannot be computed in closed form, and one has to apply certain recursive procedures to find them, from which one of the most-known is the Krasnosel'skiĭ-Mann iteration [101, 123] is an important example.

Definition 2.3.16 (Krasnosel'skiĭ-Mann iteration). Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator such that $\mathrm{fix}(\mathcal{T}) \neq \emptyset$. Let $\lambda_k \in [0,1]$ and choose x_0 arbitrarily from $\mathcal{H}$, then the Krasnosel'skiĭ-Mann iteration of $\mathcal{T}$ is

$$x_{k+1} = x_k + \lambda_k(\mathcal{T}(x_k) - x_k). \tag{2.3.1}$$

Conditions for the convergence of the above iteration can be found in the next subsection.

2.3.2.2 Fejér monotonicity

Definition 2.3.17 (Fejér monotonicity). Let $\mathcal{S} \subseteq \mathcal{H}$ be a non-empty set and $\{x_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. Then

(i) $\{x_k\}_{k\in\mathbb{N}}$ is *Fejér monotone* with respect to $\mathcal{S}$ if

$$\|x_{k+1} - x\| \leq \|x_k - x\|, \quad \forall x \in \mathcal{S}, \forall k \in \mathbb{N}.$$

(ii) $\{x_k\}_{k\in\mathbb{N}}$ is *quasi-Fejér monotone* with respect to $\mathcal{S}$, if there exists a summable sequence $\{\epsilon_k\}_{k\in\mathbb{N}} \in \ell^1_+$ such that

$$\forall k \in \mathbb{N}, \quad \|x_{k+1} - x\| \leq \|x_k - x\| + \epsilon_k, \quad \forall x \in \mathcal{S}.$$

Quasi-Fejér monotonicity is a weaker condition than Fejér monotonicity, and it allows to analyze the convergence of the inexact version of Krasnosel'skiĭ-Mann iteration which is the main topic of Chapter 3.

Example 2.3.18. Let $\mathcal{X} \subseteq \mathcal{H}$ be a non-empty convex set, and $\mathcal{T} : \mathcal{X} \to \mathcal{X}$ be a non-expansive operator such that $\operatorname{fix}(\mathcal{T}) \neq \emptyset$. The sequence $\{x_k\}_{k\in\mathbb{N}}$ generated by $x_{k+1} = \mathcal{T}(x_k)$ is Fejér monotone with respect to $\operatorname{fix}(\mathcal{T})$.

Lemma 2.3.19. *Let $\mathcal{S} \subseteq \mathcal{H}$ be a non-empty set and $\{x_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. Assume the $\{x_k\}_{k\in\mathbb{N}}$ is quasi-Fejér monotone with respect to $\mathcal{S}$, then the following holds*

(i) *$\{x_k\}_{k\in\mathbb{N}}$ is bounded.*

(ii) *$\|x_k - x\|$ is bounded for any $x \in \mathcal{S}$.*

(iii) *$\{\operatorname{dist}(x_k, \mathcal{S})\}_{k\in\mathbb{N}}$ is decreasing and convergent.*

If every weak sequential cluster point of $\{x_k\}_{k\in\mathbb{N}}$ belongs to $\mathcal{S}$, then $\{x_k\}_{k\in\mathbb{N}}$ converges weakly to a point in $\mathcal{S}$.

The next theorem shows the convergence of Krasnosel'skiĭ-Mann iteration.

Theorem 2.3.20. *Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator such that $\operatorname{fix}(\mathcal{T}) \neq \emptyset$. Consider the Krasnosel'skiĭ-Mann iteration of $\mathcal{T}$, and choose $\lambda_k \in [0, 1]$ such that $\sum_{k\in\mathbb{N}} \lambda_k(1 - \lambda_k) = +\infty$, then the following holds*

(i) *$\{x_k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{fix}(\mathcal{T})$.*

(ii) *$\{x_k - \mathcal{T}(x_k)\}_{k\in\mathbb{N}}$ converges strongly to 0.*

(iii) *$\{x_k\}_{k\in\mathbb{N}}$ converges weakly to a point in $\operatorname{fix}(\mathcal{T})$.*

When $\mathcal{T}$ is α-averaged, then the condition for $\{\lambda_k\}_{k\in\mathbb{N}}$ becomes

$$\lambda_k \in [0, 1/\alpha] \text{ such that } \sum_{k\in\mathbb{N}} \lambda_k(1/\alpha - \lambda_k) = +\infty. \tag{2.3.2}$$

The convergence of the inexact Krasnosel'skiĭ-Mann iteration will be developed Chapter 3.

2.4 Operator splitting methods

The Krasnosel'skiĭ-Mann iteration (2.3.1) is of paramount importance to the design and analysis of first-order operator splitting methods. In this section, we provide a brief overview of several classical operator splitting methods which can be cast as special cases of (2.3.1). These methods include the Proximal Point Algorithm, Forward–Backward splitting, Douglas–Rachford splitting and several Primal–Dual splitting methods.

2.4.1 Resolvents of monotone operators

Definition 2.4.1 (Resolvent [38]). Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator and $\gamma > 0$, the *resolvent* of A is defined by

$$\mathcal{J}_{\gamma A} \overset{\text{def}}{=} (\operatorname{Id} + \gamma A)^{-1}. \tag{2.4.1}$$

The *reflection* of $\mathcal{J}_{\gamma A}$ is defined by

$$\mathcal{R}_{\gamma A} \overset{\text{def}}{=} 2\mathcal{J}_{\gamma A} - \operatorname{Id}. \tag{2.4.2}$$

Given a function $J \in \Gamma_0(\mathcal{H})$ and its sub-differential ∂J, combining together (2.4.1), Definition 2.1.8, and Lemma 2.3.2, we have

$$\operatorname{prox}_{\gamma J} = \mathcal{J}_{\gamma \partial J},$$

where $\gamma > 0$. See also [130]. Then in terms of the set of fixed points, we have

$$\operatorname{fix}(\operatorname{prox}_{\gamma J}) = \operatorname{fix}(\mathcal{J}_{\gamma \partial J}) = \operatorname{zer}(\partial J).$$

Lemma 2.4.2. *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator, and $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be firmly non-expansive. Then*

(i) *$\mathcal{J}_A$ is firmly non-expansive.*

(ii) *There exists a maximal monotone operator $A' : \mathcal{H} \rightrightarrows \mathcal{H}$ such that $\mathcal{T} = \mathcal{J}_{A'}$.*

The next theorem shows the relation between the resolvents of a maximal monotone operator A and its inverse A^{-1}, which thus generalises Theorem 2.1.17.

Theorem 2.4.3 (Yosida approximation). *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator and $\gamma > 0$, the* Yosida approximation *of A with γ is*

$$ {}^{\gamma}A \stackrel{\text{def}}{=} \frac{1}{\gamma}(\mathrm{Id} - \mathcal{J}_{\gamma A}) = (\gamma \mathrm{Id} + A^{-1})^{-1} = \mathcal{J}_{A^{-1}/\gamma}(\cdot/\gamma). \tag{2.4.3} $$

Moreover, we have from Moreau's identity that

$$ \mathrm{Id} = \mathcal{J}_{\gamma A}(\cdot) + \gamma \mathcal{J}_{A^{-1}/\gamma}\Big(\frac{\cdot}{\gamma}\Big). $$

2.4.2 Proximal point algorithm

The Fermat's theorem 2.1.13 indicates that, minimizing a convex function $J \in \Gamma_0(\mathcal{H})$ is equivalent to finding the zeros of its sub-differential operator ∂J. Therefore, a minimizer of J is a fixed point of $\mathcal{J}_{\partial J}$.

In general, given a maximal monotone operator A such that $\mathrm{zer}(A) \neq \emptyset$, the problem

$$ \text{find } \; x \in \mathcal{H} \; \text{ such that } \; 0 \in A(x). \tag{2.4.4} $$

i.e. finding an element of $\mathrm{zer}(A)$, is called a monotone inclusion problem. The best-known method for solving (2.4.4) is the Proximal Point Algorithm (PPA) [124, 151].

Algorithm 2.4.4 (Proximal Point Algorithm). Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator such that $\mathrm{zer}(A) \neq \emptyset$, $\{\gamma_k\}_{k\in\mathbb{N}}$ in $]0, +\infty]$ and $\{\lambda_k\}_{k\in\mathbb{N}}$ in $[0, 2]$. Choose $x_0 \in \mathcal{H}$ arbitrarily and for $k \in \mathbb{N}$, compute x_{k+1} based on the following rule,

$$ x_{k+1} = x_k + \lambda_k\big(\mathcal{J}_{\gamma_k A}(x_k) - x_k\big). \tag{2.4.5} $$

2.4.3 Operator splitting methods

In practice, we often encounter with problems which have more structures than (2.4.4), for instance finding the zeros of the sum of monotone operators.

Problem 2.4.5. Let $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive for some $\beta > 0$, $m > 1$ be a positive integer, and for every $i \in \{1, ..., m\}$, let $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone. Consider the problem

$$ \text{find } \; x \in \mathcal{H} \; \text{ such that } \; 0 \in B(x) + \textstyle\sum_{i=1}^{m} A_i(x). \tag{2.4.6} $$

There are also situations where A_i is composed with linear operators, or even parallel sum of maximal monotone operators are involved (*e.g.* Problem (2.4.13)). In principal, PPA can still be applied to solving such problems, however, even if the resolvent of B and each A_i can be computed , the resolvent of $B + \sum_i A_i$ in most cases is not accessible.

To circumvent this difficulty, a wise strategy is to design iterative schemes such that the resolvents of A_i's are computed separately, and use the cocoercivity of B. This is the reason that these schemes are dubbed *operator splitting methods*.

2.4.3.1 Forward–Backward splitting

For Problem 2.4.5, let $m = 1$.

Problem 2.4.6. Consider finding the zeros of the sum of two maximal monotone operators

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A + B)(x). \tag{2.4.7}$$

where

(i) $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone.

(ii) $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive for some $\beta > 0$.

(iii) $\operatorname{zer}(A + B) \neq \emptyset$.

Consider the following operator composed with A and B.

Lemma 2.4.7. *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone and $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive for some $\beta > 0$. Choose $\gamma \in]0, 2\beta[$, and define*

$$\mathcal{F}_{\text{FB}} = J_{\gamma A}(\text{Id} - \gamma B).$$

Then

(i) $\mathcal{F}_{\text{FB}}$ *is* $\frac{2\beta}{4\beta-\gamma}$*-averaged non-expansive.*

(ii) $\operatorname{zer}(A + B) = \operatorname{fix}(\mathcal{F}_{\text{FB}})$.

Proof. Lemma 2.3.9, 2.4.2 and Remark 2.3.12. □

The Forward–Backward splitting (FB) [119, 140] for solving (2.4.7) is described below.

Algorithm 2.4.8 (Forward–Backward splitting). For Problem (2.4.6), let $\gamma \in]0, 2\beta[$ and $\{\lambda_k\}_{k\in\mathbb{N}}$ in $[0, \frac{4\beta-\gamma}{2\beta}]$ such that $\sum_k \lambda_k(\frac{4\beta-\gamma}{2\beta} - \lambda_k) = +\infty$. Choose $x_0 \in \mathcal{H}$ arbitrarily and for $k \in \mathbb{N}$ apply the following iteration,

$$x_{k+1} = x_k + \lambda_k\big(\mathcal{F}_{\text{FB}}(x_k) - x_k\big). \tag{2.4.8}$$

The step-size γ can be varying along iterations, and this results in the non-stationary version of FB which is studied in [57, 63]. Problem (2.4.6) with $m \geq 2$ is considered in [148], in which the authors develop a *Generalized Forward–Backward splitting (GFB)*, see Section 3.4 for more details about GFB.

2.4.3.2 Douglas–Rachford splitting

For Problem 2.4.5, now let $B = 0$ and $m = 2$.

Problem 2.4.9. Consider finding the zeros of the sum of two maximal monotone operators

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A_1 + A_2)(x). \tag{2.4.9}$$

where

(i) $A_1, A_2 : \mathcal{H} \rightrightarrows \mathcal{H}$ are maximal monotone.

(ii) $\operatorname{zer}(A_1 + A_2) \neq \emptyset$.

The Douglas–Rachford splitting (DR) is an efficient algorithm to solve (2.4.9). Originally, the DR was proposed in [76] to solve a system of linear equations arising from the discretization of a partial differential equation. The extension of DR to maximal monotone operators is due to Lions and Mercier [119]. The relaxed form of DR is considered in [78, 55, 57]. Before describing the DR, we first consider the following non-expansive operator.

Lemma 2.4.10. *Let $A_1, A_2 : \mathcal{H} \rightrightarrows \mathcal{H}$ be two maximal monotone operators, $\gamma > 0$, and define*

$$\mathcal{F}_{\mathrm{DR}} \overset{\text{def}}{=} \tfrac{1}{2}(\mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2} + \mathrm{Id}).$$

Then

(i) $\mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2}$ *is non-expansive.*

(ii) $\mathcal{F}_{\mathrm{DR}} = \mathcal{J}_{\gamma A_1}(2\mathcal{J}_{\gamma A_2} - \mathrm{Id}) - \mathcal{J}_{\gamma A_2} + \mathrm{Id} \in \mathcal{A}(\frac{1}{2})$.

(iii) $\mathrm{zer}(A_1 + A_2) = \mathcal{J}_{\gamma A_2}(\mathrm{fix}(\mathcal{F}_{\mathrm{DR}}))$.

Algorithm 2.4.11 (Douglas–Rachford splitting). For Problem (2.4.9), let $\gamma > 0$ and $\{\lambda_k\}_{k\in\mathbb{N}}$ in $[0, 2]$ such that $\sum_k \lambda_k(2 - \lambda_k) = +\infty$. Choose $z_0 \in \mathcal{H}$ arbitrarily and let $x_0 = \mathcal{J}_{\gamma A_2}(z_0)$, for $k \in \mathbb{N}$ apply the following iteration,

$$\begin{cases} u_{k+1} = \mathcal{J}_{\gamma A_1}(2x_k - z_k), \\ z_{k+1} = (1 - \lambda_k)z_k + \lambda_k(z_k + u_{k+1} - x_k), \\ x_{k+1} = \mathcal{J}_{\gamma A_2}(z_{k+1}). \end{cases} \tag{2.4.10}$$

The DR iteration (2.4.10) can be written as the fixed-point iteration of z_k and the operator $\mathcal{F}_{\mathrm{DR}}$, which reads

$$z_{k+1} = z_k + \lambda_k\big(\mathcal{F}_{\mathrm{DR}}(z_k) - z_k\big).$$

When $\lambda_k \equiv 2$, the corresponding scheme is called Peaceman–Rachford splitting [141].

By using a product space trick ([158], see also Section 3.4), the DR can be extended to the case $m \geq 2$. In the context of convex optimization, DR is closely related to the alternating direction method of multipliers (ADMM), see Chapter 7 for more details about ADMM.

2.4.3.3 A Primal–Dual splitting algorithm

We finish this section by introducing a Primal–Dual splitting algorithm.

Definition 2.4.12 (Parallel sum). Let $C, D : \mathcal{H} \rightrightarrows \mathcal{H}$ be two set-valued operators, the *parallel sum* of C and D is defined by

$$C\square D \overset{\text{def}}{=} (C^{-1} + D^{-1})^{-1}. \tag{2.4.11}$$

Problem 2.4.13. Let $\mathcal{G}, \mathcal{H}$ be two real Hilbert spaces, and

(i) $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone, $B : \mathcal{H} \to \mathcal{H}$ is β_B-cocoercive for some $\beta_B > 0$.

(ii) $C, D : \mathcal{G} \rightrightarrows \mathcal{G}$ are maximal monotone, moreover D is β_D-strongly monotone for some $\beta_D > 0$.

(iii) $L : \mathcal{H} \to \mathcal{G}$ is a bounded linear operator.

Consider the *primal* monotone inclusion problem,

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A + B)(x) + L^*\big((C\square D)(Lx)\big), \tag{2.4.12}$$

the corresponding *dual* problem reads,

$$\text{find } v \in \mathcal{G} \text{ such that } (\exists x \in \mathcal{H}) \begin{cases} 0 \in (A + B)(x) + L^*v, \\ 0 \in (C^{-1} + D^{-1})(v) - Lx. \end{cases} \tag{2.4.13}$$

denote by $\mathcal{X}$ and $\mathcal{V}$ the solution sets of (2.4.12) and (2.4.13) respectively.

(iv) The set of minimizers of (2.4.12) and (2.4.13), *i.e.* $\mathcal{X}$ and $\mathcal{V}$, are both non-empty.

The study of Primal–Dual splitting dates back to the late 1950s, since then, various Primal–Dual splitting methods have been proposed in the literature. Below we introduce a Primal–Dual splitting algorithm proposed in [62], which covers [51, 66] as special cases.

Algorithm 2.4.14 (Primal–Dual splitting). For the Problem (2.4.13). Choose $\gamma_A, \gamma_C > 0$ such that

$$2\min\{\beta_B, \beta_D\}\min\{\tfrac{1}{\gamma_A}, \tfrac{1}{\gamma_C}\}\big(1 - \sqrt{\gamma_A\gamma_C\|L\|^2}\big) > 1.$$

Choose $x_0 \in \mathcal{H}$ and $v_0 \in \mathcal{G}$ arbitrarily, for $k \in \mathbb{N}$ apply the following iteration,

$$\begin{cases} x_{k+1} = \jmath_{\gamma_A A}\big(x_k - \gamma_A B(x_k) - \gamma_A L^* v_k\big), \\ \bar{x}_{k+1} = 2x_{k+1} - x_k, \\ v_{k+1} = \jmath_{\gamma_C C^{-1}}\big(v_k - \gamma_C D^{-1}(v_k) + \gamma_C L\bar{x}_{k+1}\big). \end{cases} \tag{2.4.14}$$

Fixed-point formulation For the Primal–Dual iteration (2.4.14), from the definition of the *resolvent*, we have that (2.4.14) is equivalent to

$$\begin{aligned} \tfrac{1}{\gamma_A}(x_k - x_{k+1}) - B(x_k) - L^* v_k &= A(x_{k+1}), \\ \bar{x}_{k+1} &= 2x_{k+1} - x_k, \\ \tfrac{1}{\gamma_C}(v_k - v_{k+1}) - D^{-1}(v_k) + L\bar{x}_{k+1} &= C^{-1}(v_{k+1}), \end{aligned}$$

which can be further written as

$$-\begin{bmatrix} B & 0 \\ 0 & D^{-1} \end{bmatrix}\begin{pmatrix} x_k \\ v_k \end{pmatrix} \in \begin{bmatrix} A & L^* \\ -L & C^{-1} \end{bmatrix}\begin{pmatrix} x_{k+1} \\ v_{k+1} \end{pmatrix} + \begin{bmatrix} \mathrm{Id}_{\mathcal{H}}/\gamma_A & -L^* \\ -L & \mathrm{Id}_{\mathcal{G}}/\gamma_C \end{bmatrix}\begin{pmatrix} x_{k+1} - x_k \\ v_{k+1} - v_k \end{pmatrix}, \tag{2.4.15}$$

where $\mathrm{Id}_{\mathcal{H}}, \mathrm{Id}_{\mathcal{G}}$ are the identity operators on $\mathcal{H}$ and $\mathcal{G}$ respectively. Define the product space $\boldsymbol{\mathcal{K}} = \mathcal{H}\times\mathcal{G}$, and let $\mathbf{Id}$ be the identity operator on $\boldsymbol{\mathcal{K}}$. Define the following variable and operators

$$\boldsymbol{z}_k \overset{\text{def}}{=} \begin{pmatrix} x_k \\ v_k \end{pmatrix}, \ \boldsymbol{A} \overset{\text{def}}{=} \begin{bmatrix} A & L^* \\ -L & C^{-1} \end{bmatrix}, \ \boldsymbol{B} \overset{\text{def}}{=} \begin{bmatrix} B & 0 \\ 0 & D^{-1} \end{bmatrix}, \ \boldsymbol{\mathcal{V}} \overset{\text{def}}{=} \begin{bmatrix} \mathrm{Id}_{\mathcal{H}}/\gamma_A & -L^* \\ -L & \mathrm{Id}_{\mathcal{G}}/\gamma_C \end{bmatrix}. \tag{2.4.16}$$

We have $\boldsymbol{A}$ is maximal monotone [41], $\boldsymbol{B}$ is $\min\{\beta_B, \beta_D\}$-cocoercive, and $\boldsymbol{\mathcal{V}}$ is self-adjoint and ν-positive definite for $\nu = (1 - \sqrt{\gamma_A\gamma_C\|L\|^2})\min\{\frac{1}{\gamma_A}, \frac{1}{\gamma_C}\}$ [173, 62].

Define $\boldsymbol{\mathcal{K}}_{\boldsymbol{\mathcal{V}}}$ the Hilbert space induced by $\boldsymbol{\mathcal{V}}$. Substitute the notions of (2.4.16) into (2.4.15) we get,

$$-\boldsymbol{B}(\boldsymbol{z}_k) \in \boldsymbol{A}(\boldsymbol{z}_{k+1}) + \boldsymbol{\mathcal{V}}(\boldsymbol{z}_{k+1} - \boldsymbol{z}_k),$$

which can be further written as

$$\boldsymbol{z}_{k+1} = (\boldsymbol{\mathcal{V}} + \boldsymbol{A})^{-1}(\boldsymbol{\mathcal{V}} - \boldsymbol{B})(\boldsymbol{z}_k) = (\mathbf{Id} + \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{A})^{-1}(\mathbf{Id} - \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{B})(\boldsymbol{z}_k). \tag{2.4.17}$$

It can be seen that, (2.4.17) is the Forward–Backward splitting in $\boldsymbol{\mathcal{K}}_{\boldsymbol{\mathcal{V}}}$ [62, 58].

2.5 Matrix analysis

In this section, we turn to linear algebra and present results on convergent matrices and their spectral properties. These results will be used in our local linear convergence analysis of first-order proximal splitting methods.

Let $M \in \mathbb{R}^{n\times n}$ be a square matrix, M is said to be *symmetric* if $M = M^T$, *skew symmetric* if $M = -M^T$, *normal* if $MM^T = M^TM$, *positive definite* if there exists $\delta > 0$ such that

$$\forall x \in \mathbb{R}^n, \ \langle x, Mx\rangle \geq \delta\|x\|^2,$$

and *positive semidefinite* if $\delta = 0$. The *kernel* and *range* of M are defined respectively by

$$\begin{aligned} \ker(M) &\overset{\text{def}}{=} \{x \in \mathbb{R}^n : Mx = 0\}, \\ \mathrm{ran}(M) &\overset{\text{def}}{=} \text{vector space spanned by the columns of } M. \end{aligned}$$

Let $\Theta_M \overset{\text{def}}{=} \{\eta_i\}_i, i = 1, ..., n$ be the set of *eigenvalues* of M, then the spectral radius of M is given by

$$\rho(M) \overset{\text{def}}{=} \max_{i \in \{1,...,n\}} |\eta_i|.$$

The operator (spectral) norm of M is defined as

$$\|M\| \overset{\text{def}}{=} \sup_{\|x\|=1} \|Mx\|.$$

Let $\Sigma_M \overset{\text{def}}{=} (\sigma_i)_i, i = 1, ..., n$ be the singular values of M sorted in descending order, then we have

$$\|M\| = \sigma_1 = \sqrt{\|M^T M\|}.$$

In general, $\rho(M) \leq \|M\|$, moreover we have the spectral radius formula [125]

$$\rho(M) \leq \|M^k\|^{1/k} \quad \text{and} \quad \rho(M) = \lim_{k \to +\infty} \|M^k\|^{1/k}. \tag{2.5.1}$$

If M is normal, then $\rho(M) = \|M\|$.

Convergent matrices The definition of convergent matrices is adopted from [125], denote M^k the k^{th} power of M.

Definition 2.5.1 (Convergent matrices). A matrix $M \in \mathbb{R}^{n \times n}$ is convergent to $M^\infty \in \mathbb{R}^{n \times n}$ if, and only if,

$$\lim_{k \to +\infty} \|M^k - M^\infty\| = 0.$$

M is said to be *linearly convergent* if there exists $\rho \in [0, 1[$, a constant $\mathcal{C} > 0$ and $K \in \mathbb{N}$ such that $\forall k \geq K$, there holds

$$\|M^k - M^\infty\| \leq \mathcal{C}\rho^k.$$

If the *infimum* of all convergence rates ρ above is also a convergence rate, then this *minimum* is called the *optimal convergence rate*.

Definition 2.5.2. For $M \in \mathbb{R}^{n \times n}$, $\eta \in \Theta_M$ is called *semisimple* if and only if $\ker(M - \eta\text{Id}) = \ker((M - \eta\text{Id})^2)$.

The definition of semisimplicity is equivalent to say that $\text{rank}(M - \eta\text{Id}) = \text{rank}((M - \eta\text{Id})^2)$. The following characterization of a convergent matrix is taken from [125].

Theorem 2.5.3 (Limits of powers). *For $M \in \mathbb{R}^{n \times n}$, the power of M is convergent to M^∞ if and only if*

$$\begin{aligned} &\rho(M) < 1, \\ &\textit{or else} \\ &\rho(M) = 1, \ \textit{and } 1 \textit{ is the only eigenvalue on the complex} \\ &\qquad \textit{unit circle, which is also semisimple.} \end{aligned} \tag{2.5.2}$$

When M^∞ exists,

$$M^\infty = \textit{the projector onto } \ker(\text{Id} - M) \textit{ along } \text{ran}(\text{Id} - M). \tag{2.5.3}$$

Lemma 2.5.4. *Suppose $M \in \mathbb{R}^{n \times n}$ is a convergent matrix, if $\rho(M) < 1$, then $M^\infty = 0$.*

Lemma 2.5.5. *Suppose $M \in \mathbb{R}^{n \times n}$ is convergent to some $M^\infty \in \mathbb{R}^{n \times n}$, then*

(i) *$M^\infty = \mathcal{P}_{\text{fix}(M)}$ if and only if $\text{fix}(M) = \text{fix}(M^T)$.*

(ii) *If M is non-expansive or normal, then $M^\infty = \mathcal{P}_{\text{fix}(M)}$.*

Proof. Corollary 2.7 of [18]. □

Whenever M is convergent, it converges linearly to M^∞, and we have the following lemma.

Lemma 2.5.6 (Convergence rate). *Suppose $M \in \mathbb{R}^{n\times n}$ is convergent to some $M^\infty \in \mathbb{R}^{n\times n}$,*

(i) *for any $k \in \mathbb{N}$,*

$$M^k - M^\infty = (M - M^\infty)^k \quad \text{and} \quad \|M^k - M^\infty\| \leq \|M - M^\infty\|^k.$$

The equality holds only when M is normal.

(ii) *We have $\rho(M - M^\infty) < 1$, and M is linearly convergent for any $\rho \in]\rho(M - M^\infty), 1[$.*

(iii) *$\rho(M - M^\infty)$ is the optimal convergence rate if one of the following holds*

(a) *M is normal.*

(b) *All the eigenvalues $\eta \in \Theta_M$ such that $|\eta| = \rho(M - M^\infty)$ are semisimple.*

Proof. See Theorems 2.12, 2.13, 2.15 and 2.16 of [18]. □

Convergence and non-expansiveness In general, there is no direction correspondence between the convergence of a matrix and its non-expansiveness, *i.e.* a matrix M can be convergent but expansive, or non-expansive but not convergent. For example, let $a \in]0, 1[$ and consider a 2×2 matrix

$$M = \begin{bmatrix} 1+a & -a \\ 1 & 0 \end{bmatrix}. \tag{2.5.4}$$

It can be shown that the two eigenvalues of M are 1 and a, hence M is convergent. However M is not non-expansive, since $\|M\| \geq 1 + a > 1$. Another examples is the matrix $-\mathrm{Id}$, which is non-expansive, yet not convergent.

The situation changes however, if the matrix is α-averaged. For the sake of simplicity, let M be a non-expansive matrix, and consider its relaxed version with fixed relaxation parameter λ such that $\lambda(1 - \lambda) > 0$. Then the iteration is convergent thanks to Theorem 2.3.20. Moreover, owing to the definition of averaged non-expansive operators, we have that matrix $\mathrm{Id} + \lambda(M - \mathrm{Id})$ is λ-averaged. This implies that the class of α-averaged matrices is also convergent.

2.6 Riemannian geometry

In this section, we collect some basic results on Riemannian geometry, and we refer to [103, 52] for comprehensive accounts on differential and Riemannian manifolds.

2.6.1 A glimpse of Riemannian geometry

We say that a set $\mathcal{M} \subset \mathbb{R}^n$ is a C^2-smooth manifold of codimension m around a point $\bar{x} \in \mathbb{R}^n$ if $\bar{x} \in \mathcal{M}$, and there is an open set $V \subset \mathbb{R}^n$ such that

$$\mathcal{M} \cap V = \big\{x \in V : \Phi_i(x) = 0,\ i = 1, ..., m\big\}$$

where Φ_i are C^2-smooth functions with $\Phi_i(\bar{x})$ linearly independent.

Let $\mathcal{M}$ be a C^2-smooth embedded submanifold of $\mathbb{R}^n$ around a point x. The natural embedding of a submanifold $\mathcal{M}$ into $\mathbb{R}^n$ permits to define a Riemannian structure and to introduce geodesics on $\mathcal{M}$. Denote $\mathcal{N}_\mathcal{M}(x)$ and $\mathcal{T}_\mathcal{M}(x)$ the normal space and tangent space of $\mathcal{M}$ at $x \in \mathcal{M}$ respectively.

Exponential map Geodesics generalize the concept of straight lines in $\mathbb{R}^n$, preserving the zero acceleration characteristic, to manifolds. Roughly speaking, a geodesic is locally the shortest path between two points on $\mathcal{M}$. We denote by $\mathfrak{g}(t; x, h)$ the value at $t \in \mathbb{R}$ of the geodesic starting at $\mathfrak{g}(0; x, h) = x \in \mathcal{M}$ with velocity $\dot{\mathfrak{g}}(t; x, h) = \frac{d\mathfrak{g}}{dt}(t; x, h) = h \in \mathcal{T}_\mathcal{M}(x)$ (which is uniquely defined). For

every $h \in \mathcal{T}_{\mathcal{M}}(x)$, there exists an interval I around 0 and a unique geodesic $\mathfrak{g}(t; x, h) : I \to \mathcal{M}$ such that $\mathfrak{g}(0; x, h) = x$ and $\dot{\mathfrak{g}}(0; x, h) = h$. The mapping

$$\mathrm{Exp}_x : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{M}, \ h \mapsto \mathrm{Exp}_x(h) = \mathfrak{g}(1; x, h),$$

is called *Exponential map*. Given $x, x' \in \mathcal{M}$, the direction $h \in \mathcal{T}_{\mathcal{M}}(x)$ we are interested in is such that

$$\mathrm{Exp}_x(h) = x' = \mathfrak{g}(1; x, h).$$

Parallel translation Given two points $x, x' \in \mathcal{M}$, let $\mathcal{T}_{\mathcal{M}}(x), \mathcal{T}_{\mathcal{M}}(x')$ be their corresponding tangent spaces. Define

$$\tau : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x'),$$

the parallel translation along the unique geodesic joining x to x', which is isomorphism and isometry w.r.t. the Riemannian metric.

Lemma 2.6.1. *Let $\mathcal{M}$ be a C^2-smooth manifold around x. Then for any $x' \in \mathcal{M} \cap \mathcal{N}$, where $\mathcal{N}$ is a neighbourhood of x, the projection operator $\mathcal{P}_{\mathcal{M}}(x')$ is uniquely valued and C^1 around x, and thus*

$$x' - x = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x' - x) + o(\|x' - x\|).$$

If moreover $\mathcal{M} = x + \mathcal{T}_{\mathcal{M}}(x)$ is an affine subspace, then $x' - x = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x' - x)$.

Proof. As $\mathcal{M}$ is a C^2–manifold around x, then $\mathcal{P}_{\mathcal{M}}(x')$ is uniquely valued [145] and moreover C^1 near x [110, Lemma 4]. Thus, continuous differentiability shows

$$x' - x = \mathcal{P}_{\mathcal{M}}(x') - \mathcal{P}_{\mathcal{M}}(x) = \mathrm{D}\mathcal{P}_{\mathcal{M}}(x)(x - x') + o(\|x - x'\|).$$

where $\mathrm{D}\mathcal{P}_{\mathcal{M}}(x)$ is the derivative of $\mathcal{P}_{\mathcal{M}}$ at x. By virtue of [110, Lemma 4], this derivative is given by

$$\mathrm{D}\mathcal{P}_{\mathcal{M}}(x) = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)},$$

The case where $\mathcal{M}$ is an affine subspace is immediate. We conclude the proof. □

Lemma 2.6.2. *Let $\mathcal{M}$ be a C^2-smooth manifold, a point $x \in \mathcal{M}$, and $\{x_k\}_{k\in\mathbb{N}}$ a sequence converging to x in $\mathcal{M}$. Define $\tau_k : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x_k)$ be the parallel translation along the unique geodesic joining x to x_k. Then, for any bounded vector $u \in \mathbb{R}^n$, we have*

$$(\tau_k^{-1}\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x_k)} - \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)})u = o(\|u\|).$$

Proof. From [1, Chapter 5], we deduce that for k sufficiently large,

$$\tau_k^{-1} = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)} + o(\|x_k - x\|).$$

In addition, locally near x along $\mathcal{M}$, the mapping $x \mapsto \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x)$ is C^1 [110, Lemma 4], hence,

$$\begin{aligned}\lim_{k\to\infty} \frac{\|(\tau_k^{-1}\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x_k)} - \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)})(u)\|}{\|u\|} &\leq \lim_{k\to\infty} \frac{\|\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x_k)} - \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)})\|\|u\|}{\|u\|} + o(\|x_k - x\|) \\ &\leq \lim_{k\to\infty} \|\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x_k)} - \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}\| + o(\|x_k - x\|) = 0.\end{aligned}$$

□

2.6.2 Riemannian gradient and Hessian

Let $\mathcal{M}$ be a C^2-smooth manifold, and suppose that function $R \in \Gamma_0(\mathbb{R}^n)$ is C^2-smooth along the manifold $\mathcal{M}$. In the following, we present the expressions of the Riemannian gradient and Hessian for R along $\mathcal{M}$.

For a vector $v \in \mathcal{N}_{\mathcal{M}}(x)$, the Weingarten map of $\mathcal{M}$ at x is the operator $\mathfrak{W}_x(\cdot, v) : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x)$ defined by

$$\mathfrak{W}_x(\cdot, v) = -\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}\mathrm{d}V[h], \tag{2.6.1}$$

where V is any local extension of v to a normal vector field on $\mathcal{M}$. The definition is independent of the choice of the extension V, and $\mathfrak{W}_x(\cdot, v)$ is a symmetric linear operator which is closely tied to the second fundamental form of $\mathcal{M}$, see [52, Proposition II.2.1].

Let R be a real-valued function which is C^2 along the $\mathcal{M}$ around x. The covariant gradient of R at $x' \in \mathcal{M}$ is the vector $\nabla_{\mathcal{M}} R(x') \in \mathcal{T}_{\mathcal{M}}(x')$ defined by

$$\langle \nabla_{\mathcal{M}} R(x'),\, h\rangle = \frac{d}{dt} R\big(\mathcal{P}_{\mathcal{M}}(x' + th)\big)\big|_{t=0},\ \forall h \in \mathcal{T}_{\mathcal{M}}(x'),$$

where $\mathcal{P}_{\mathcal{M}}$ is the projection operator onto $\mathcal{M}$. The covariant Hessian of R at x' is the symmetric linear mapping $\nabla^2_{\mathcal{M}} R(x')$ from $\mathcal{T}_{\mathcal{M}}(x')$ to itself which is defined as

$$\langle \nabla^2_{\mathcal{M}} R(x')h,\, h\rangle = \frac{d^2}{dt^2} R\big(\mathcal{P}_{\mathcal{M}}(x' + th)\big)\big|_{t=0},\ \forall h \in \mathcal{T}_{\mathcal{M}}(x'). \tag{2.6.2}$$

This definition agrees with the usual definition using geodesics or connections [126]. Now assume that $\mathcal{M}$ is a Riemannian embedded submanifold of $\mathbb{R}^n$, and that a function R has a C^2-smooth restriction on $\mathcal{M}$. This can be characterized by the existence of a C^2-smooth extension (representative) of R, *i.e.* a C^2-smooth function $\widetilde{R}$ on $\mathbb{R}^n$ such that $\widetilde{R}$ agrees with R on $\mathcal{M}$. Thus, the Riemannian gradient $\nabla_{\mathcal{M}} R(x')$ is also given by

$$\nabla_{\mathcal{M}} R(x') = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla \widetilde{R}(x'), \tag{2.6.3}$$

and $\forall h \in \mathcal{T}_{\mathcal{M}}(x')$, the Riemannian Hessian reads

$$\begin{aligned} \nabla^2_{\mathcal{M}} R(x')h &= \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \mathrm{d}(\nabla_{\mathcal{M}} R)(x')[h] = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \mathrm{d}\big(x' \mapsto \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla_{\mathcal{M}} \widetilde{R}\big)[h] \\ &= \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla^2 \widetilde{R}(x')(h) + \mathfrak{W}_{x'}\big(h, \mathcal{P}_{\mathcal{N}_{\mathcal{M}}(x')} \nabla \widetilde{R}(x')\big), \end{aligned} \tag{2.6.4}$$

where the last equality comes from [2, Theorem 1]. When $\mathcal{M}$ is an affine/linear subspace of $\mathbb{R}^n$, then obviously $\mathcal{M} = x + \mathcal{T}_{\mathcal{M}}(x)$ and $\mathfrak{W}_{x'}(h, \mathcal{P}_{\mathcal{N}_{\mathcal{M}}(x')} \nabla \widetilde{R}(x')) = 0$, hence (2.6.4) reduces to

$$\nabla^2_{\mathcal{M}} R(x') = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla^2 \widetilde{R}(x') \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')}.$$

Lemma 2.6.3 (Riemannian Taylor expansion). *Let R be C^2-smooth at x relative to manifold $\mathcal{M}$, and $x' \in \mathcal{M}$ be a point close to x. Denote $\tau : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x')$ the parallel translation along the unique geodesic joining x to x'. The Riemannian Taylor expansion of $R \in C^2(\mathcal{M})$ around x reads,*

$$\tau^{-1} \nabla_{\mathcal{M}} R(x') = \nabla_{\mathcal{M}} R(x) + \nabla^2_{\mathcal{M}} R(x) \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x' - x) + o(\|x' - x\|).$$

Proof. Since $x, x' \in \mathcal{M}$ are close, we have $x' = \mathrm{Exp}_x(h)$ for some $h \in \mathcal{T}_{\mathcal{M}}(x)$ small enough, and thus, the Taylor expansion [156, Remark 4.2] of $\nabla_{\mathcal{M}} R$ around x reads

$$\tau^{-1} \nabla_{\mathcal{M}} R(x') = \nabla_{\mathcal{M}} R(x) + \nabla^2_{\mathcal{M}} R(x)h + o(\|h\|). \tag{2.6.5}$$

Moreover, form the proof of [126, Theorem 4.9], one can show that

$$\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x') = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(\mathrm{Exp}_x(h)) = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x) + h + o(\|h\|^2).$$

Substituting back into (2.6.5) we get the claimed result. □

Part I

Global Convergence Guarantees

Chapter 3

Convergence Rates of Inexact Krasnosel'skiĭ-Mann Iteration

Main contributions of this chapter

- Global sub-linear convergence rates of inexact Krasnosel'skiĭ-Mann iteration (Theorem 3.2.1 and 3.2.4).
- Local linear convergence of the inexact Krasnosel'skiĭ-Mann iteration under metric sub-regularity (Theorem 3.3.3).
- Convergence of the non-stationary and inexact Krasnosel'skiĭ-Mann iteration (Theorem 3.5.2).

The content of this chapter appeared in [116].

Contents

We have seen from Section 2.4, that the PPA, DR/ADMM, FB and the Primal–Dual splitting methods can be cast as special instances of the Krasnosel'skiĭ-Mann iteration. In fact, the class of operator splitting methods that can be formulated into Krasnosel'skiĭ-Mann iteration is quite large, it includes for instance the methods proposed in [53, 166, 163, 41, 51, 148, 60, 66, 173] to name a few.

In this chapter, we present a convergence rate analysis for the inexact Krasnosel'skiĭ-Mann iteration built from non-expansive operators. Our results include two main parts: global pointwise and ergodic sub-linear convergence rates, and local linear convergence under a metric sub-regularity assumption. The obtained convergence rates can be applied to analyze the convergence rate of various operator splitting methods, including FB/GFB, DR/ADMM and Primal–Dual splitting to name a few. For these methods, we also develop easily verifiable termination criteria for finding an approximate solution, which can be seen as a generalization of the termination criterion for the classical gradient descent method. We finally develop a parallel analysis for the non-stationary and inexact Krasnosel'skiĭ-Mann iteration. The usefulness of our results is illustrated by applying them to a large class of structured monotone inclusion and convex optimization problems. Experiments on some large scale inverse problems in signal and image processing problems are shown.

3.1 Introduction

3.1.1 Problem statement

Let us now recall the Krasnosel'skiĭ-Mann iteration (2.3.1) defined in Section 2.3.2.1, and consider the *inexact* version of it, *i.e.* the computation of $\mathcal{F}(z_k)$ is carried out approximately.

Definition 3.1.1 (Inexact Krasnosel'skiĭ-Mann iteration). Let $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator whose set of fixed points $\mathrm{fix}(\mathcal{F})$ is non-empty. Let $\lambda_k \in]0, 1]$, and denote $\mathcal{F}_k = \lambda_k \mathcal{F} + (1-\lambda_k)\mathrm{Id}$. Then the inexact Krasnosel'skiĭ-Mann iteration of $\mathcal{F}$ is defined by

$$z_{k+1} = z_k + \lambda_k\big(\mathcal{F}(z_k) + \varepsilon_k - z_k\big) = \mathcal{F}_k z_k + \lambda_k \varepsilon_k, \tag{3.1.1}$$

where ε_k is the error of approximating $\mathcal{F}(z_k)$. The *residual* of the iteration is defined as

$$e_k = (\mathrm{Id} - \mathcal{F})(z_k) = \frac{z_k - z_{k+1}}{\lambda_k} + \varepsilon_k. \tag{3.1.2}$$

3.1.2 Preliminary results

Before presenting the convergence rates of inexact Krasnosel'skiĭ-Mann iteration, we collect some useful properties of it. Define the following two important notions

$$\mathcal{F}' \overset{\text{def}}{=} \mathrm{Id} - \mathcal{F} \quad \text{and} \quad \tau_k \overset{\text{def}}{=} \lambda_k(1 - \lambda_k). \tag{3.1.3}$$

Apparently, we have $e_k = \mathcal{F}'(z_k)$.

Proposition 3.1.2. *The following statements hold,*

(i) $\mathcal{F}_k \in \mathcal{A}(\lambda_k)$ *for* $\lambda_k \in]0, 1[$, *and if* $T \in \mathcal{A}(\alpha)$, *then* $\mathcal{F}_k \in \mathcal{A}(\lambda_k\alpha)$.

(ii) *For any* $z^\star \in \mathrm{fix}(\mathcal{F})$,

$$z^\star \in \mathrm{fix}(\mathcal{F}) \iff z^\star \in \mathrm{fix}(\mathcal{F}_k) \iff z^\star \in \mathrm{zer}(\mathcal{F}').$$

(iii) *If* $\sum_{k\in\mathbb{N}} \tau_k = +\infty$ *and* $\{\lambda_k\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$, *then,*

(a) $\{e_k\}_{k\in\mathbb{N}}$ *converges strongly to* 0.

(b) $\{z_k\}_{k\in\mathbb{N}}$ *is quasi-Fejér monotone with respect to* $\mathrm{fix}(\mathcal{F})$, *and converges weakly to a point* $z^\star \in \mathrm{fix}(\mathcal{F})$.

Proof. (i) A result of combining Definition 2.3.6 and Lemma 2.3.10. (ii) Straightforward. (iii) See [57, Lemma 5.1]. □

Let's now turn to the properties of e_k. Denote $\underline{\tau} = \inf_{k\in\mathbb{N}} \tau_k$, $\bar{\tau} = \sup_{k\in\mathbb{N}} \tau_k$. Since $\{z_k\}_{k\in\mathbb{N}}$ is bounded, so is $\{e_k\}_{k\in\mathbb{N}}$, we define the following two constants

$$\nu_1 = 2\sup_{k\in\mathbb{N}} \|\mathcal{F}_k(z_k) - z^\star\| + \sup_{k\in\mathbb{N}} \lambda_k\|\varepsilon_k\|, \quad \nu_2 = 2\sup_{k\in\mathbb{N}} \|e_k - e_{k+1}\|.$$

Lemma 3.1.3. *For the error term* e_k, *the following inequality holds*

$$\frac{1}{2\lambda_k}\|e_k - e_{k+1}\|^2 \leq \langle e_k - \varepsilon_k, e_k - e_{k+1}\rangle.$$

Proof. By Lemma 2.3.10, $\frac{1}{2}\mathcal{F}' \in \mathcal{A}(\frac{1}{2})$. It then follows from Lemma 2.3.8(iii) that $\forall p, q \in \mathcal{H}$,

$$\|\tfrac{1}{2}\mathcal{F}'(p) - \tfrac{1}{2}\mathcal{F}'(q)\|^2 \leq \langle p - q, \tfrac{1}{2}\mathcal{F}'(p) - \tfrac{1}{2}\mathcal{F}'(q)\rangle.$$

Letting $p = z_k, q = z_{k+1}$ and using the definition of e_k yield the desired result. □

Corollary 3.1.4. *If* $\mathcal{F}$ *is* α*-averaged, the inequality of Lemma 3.1.3 becomes*

$$\frac{1}{2\alpha\lambda_k}\|e_k - e_{k+1}\|^2 \leq \langle e_k - \varepsilon_k, e_k - e_{k+1}\rangle.$$

Lemma 3.1.5. *For* $z^\star \in \mathrm{fix}(\mathcal{F})$, $\lambda_k \in]0, 1]$, *we have*

$$\|z_{k+1} - z^\star\|^2 \leq \|z_k - z^\star\|^2 - \tau_k\|e_k\|^2 + \nu_1\lambda_k\|\varepsilon_k\|.$$

Proof. By virtue of [16, Corollary 2.14] we get

$$\begin{aligned}
\|z_{k+1} - z^\star\|^2 &= \|\mathcal{F}_k z_k - z^\star + \lambda_k\varepsilon_k\|^2 \\
&\leq \|(1 - \lambda_k)(z_k - z^\star) + \lambda_k(\mathcal{F}(z_k) - \mathcal{F}(z^\star))\|^2 + \nu_1\lambda_k\|\varepsilon_k\| \\
&= (1 - \lambda_k)\|z_k - z^\star\|^2 + \lambda_k\|\mathcal{F}(z_k) - \mathcal{F}(z^\star)\|^2 - \tau_k\|z_k - \mathcal{F}(z_k)\|^2 + \nu_1\lambda_k\|\varepsilon_k\| \\
&\leq \|z_k - z^\star\|^2 - \tau_k\|e_k\|^2 + \nu_1\lambda_k\|\varepsilon_k\|.
\end{aligned}$$

The non-expansiveness of $\mathcal{F}$ is used for the last inequality. □

This lemma indicates that $\{z_k\}_{k\in\mathbb{N}}$ is quasi-Fejér monotone with respect to $\mathrm{fix}(\mathcal{F})$ as stated in Proposition 3.1.2. Hence, $\{z_k\}_{k\in\mathbb{N}}$ is bounded, and the constant ν_2 makes sense.

Corollary 3.1.6. *If $\mathcal{F}$ is α-averaged, the inequality of Lemma 3.1.5 holds with*

$$\lambda_k \in]0, 1/\alpha] \quad \textit{and} \quad \tau_k = \lambda_k\big(\tfrac{1}{\alpha} - \lambda_k\big).$$

Lemma 3.1.7. *For $\lambda_k \in [0,1]$, $\{e_k\}_{k\in\mathbb{N}}$ obeys $\|e_{k+1}\|^2 - \nu_2\|\varepsilon_k\| \leq \|e_k\|^2$.*

Proof. We have

$$\begin{aligned}\|e_{k+1}\|^2 &= \|e_{k+1} - e_k + e_k\|^2 = \|e_k\|^2 - 2\langle e_k,\, e_k - e_{k+1}\rangle + \|e_k - e_{k+1}\|^2\\ &\leq \|e_k\|^2 - 2(1-\lambda_k)\langle e_k - \varepsilon_k,\, e_k - e_{k+1}\rangle - 2\langle \varepsilon_k,\, e_k - e_{k+1}\rangle\\ &\leq \|e_k\|^2 - \tfrac{1-\lambda_k}{\lambda_k}\|e_k - e_{k+1}\|^2 + \nu_2\|\varepsilon_k\| \leq \|e_k\|^2 + \nu_2\|\varepsilon_k\|\end{aligned}$$

where Lemma 3.1.3 is used twice in the second and third lines. □

3.2 Global convergence rates

In this section, we present the global convergence rate of the inexact Krasnosel'skiĭ-Mann iteration (3.1.1), the result consist of two aspects, piecewise and ergodic. Define d_0 as the distance of initial point z_0 to the set of fixed points, *i.e.* $d_0 = \mathrm{dist}(z_0, \mathrm{fix}(\mathcal{F}))$.

3.2.1 Pointwise convergence rates

Theorem 3.2.1 (Pointwise convergence rate). *For the inexact Krasnosel'skiĭ-Mann iteration (3.1.1), if there holds*

$$0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < 1 \quad \textit{and} \quad \{(k+1)\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell^1_+. \tag{3.2.1}$$

Denote $\mathcal{C}_1 = \nu_1 \sum_{k\in\mathbb{N}} \lambda_k\|\varepsilon_k\| + \nu_2\bar{\tau}\sum_{k\in\mathbb{N}}(k+1)\|\varepsilon_k\| < +\infty$, then we have

$$\|e_k\| \leq \sqrt{\frac{d_0^2 + \mathcal{C}_1}{\underline{\tau}(k+1)}}. \tag{3.2.2}$$

Proof. Condition (3.2.1) implies $\underline{\tau} > 0$, $\{\tau_k\}_{k\in\mathbb{N}} \notin \ell^1_+$ and $\{\lambda_k\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$. Therefore, sequence $\{z_k\}_{k\in\mathbb{N}}$ is quasi-Fejér monotone with respect to $\mathrm{fix}(\mathcal{F})$ ((iii) of Proposition 3.1.2). Thus, $\{\|e_k\|\}_{k\in\mathbb{N}}$ and $\{\|z_k - z^\star\|\}_{k\in\mathbb{N}}$ are bounded for any $z^\star \in \mathrm{fix}(\mathcal{F})$. Hence ν_1, ν_2 and $\mathcal{C}_1$ are bounded constants.

Choose $z^\star \in \mathrm{fix}(\mathcal{F})$ such that $d_0 = \|z_0 - z^\star\|$. From Lemma 3.1.5, we have $\forall k \in \mathbb{N}$,

$$\tau_k\|e_k\|^2 \leq \|z_k - z^\star\|^2 - \|z_{k+1} - z^\star\|^2 + \nu_1\lambda_k\|\varepsilon_k\|.$$

Summing up from $j = 0$ to k,

$$\textstyle\sum_{j=0}^k \tau_j\|e_j\|^2 \leq \|z_0 - z^\star\|^2 - \|z_{k+1} - z^\star\|^2 + \nu_1 \sum_{j=0}^k \lambda_j\|\varepsilon_j\|. \tag{3.2.3}$$

Owing to Lemma 3.1.7, we have $\forall j \leq k$,

$$\|e_k\|^2 - \nu_2 \textstyle\sum_{\ell=j}^{k-1} \|\varepsilon_\ell\| \leq \|e_j\|^2.$$

Substituting this back into (3.2.3) yields,

$$\begin{aligned}\big(\textstyle\sum_{j=0}^k\tau_j\big)\|e_k\|^2 &\leq \textstyle\sum_{j=0}^k\tau_j\|e_j\|^2 + \nu_2 \sum_{j=0}^k \tau_j \sum_{\ell=j}^{k-1} \|\varepsilon_\ell\|\\ &\leq d_0^2 + \nu_1 \textstyle\sum_{j=0}^k \lambda_j\|\varepsilon_j\| + \nu_2 \sum_{j=0}^k \tau_j \sum_{\ell=j}^{k-1} \|\varepsilon_\ell\|.\end{aligned} \tag{3.2.4}$$

Finally, since $(k+1)\underline{\tau} \leq \sum_{j=0}^k \tau_j$, we get,

$$\begin{aligned}(k+1)\underline{\tau}\|e_k\|^2 &\leq d_0^2 + \nu_1 \textstyle\sum_{j=0}^k \lambda_j\|\varepsilon_j\| + \nu_2\bar{\tau} \sum_{\ell=0}^{k-1} (\ell+1)\|\varepsilon_\ell\|\\ &\leq d_0^2 + \nu_1 \textstyle\sum_{k\in\mathbb{N}} \lambda_k\|\varepsilon_k\| + \nu_2\bar{\tau} \sum_{k\in\mathbb{N}} (k+1)\|\varepsilon_k\|,\end{aligned}$$

which leads to the desired result (3.2.2). □

Remark 3.2.2.

(i) After our work appeared, the $O(1/\sqrt{k})$ rate was improved to $o(1/\sqrt{k})$ by the authors of [72].

(ii) Since finding $z^\star \in \mathrm{fix}(\mathcal{F})$ is equivalent to finding a zero of $\mathcal{F}'$ ((ii) of Proposition 3.1.2), Theorem 3.2.1 tells us that $O(1/\epsilon)$ iterations are needed for (3.1.1) to reach an ϵ-accurate in terms of the error criterion $\|\mathcal{F}'(z_k)\|^2 \leq \epsilon$.

(iii) For the case of first-order methods for solving smooth optimization problems, *i.e.* the gradient descent where $\mathcal{F}'$ is just the gradient, the obtained pointwise convergence rate is the best-known complexity bound [133].

(iv) If the iteration (3.1.1) is exact, then the sequence $\{\|e_k\|\}_{k\in\mathbb{N}}$ is non-increasing (Lemma 3.1.7), hence we get $\|e_k\| \leq d_0/\sqrt{\sum_{j=0}^k \tau_j}$, which recovers the result of [65, Proposition 11]. Note that we provide a sharper monotonicity property compared to them.

(v) Obviously, our results hold true if we endow $\mathcal{H}$ with the following inner product and norm, $\forall x, y \in \mathcal{H}$

$$\langle x,\, y\rangle_{\mathcal{V}} = \langle x,\, \mathcal{V}y\rangle, \quad \|x\|_{\mathcal{V}} = \langle x,\, \mathcal{V}x\rangle,$$

where $\mathcal{V}$ is a bounded symmetric positive definite operator on $\mathcal{H}$.

When $\mathcal{F}$ is α-averaged, then Theorem 3.2.1 holds with an adjustment on the relaxation parameter λ_k.

Corollary 3.2.3. *If $\mathcal{F}$ is α-averaged, then condition (3.2.1) changes to*

$$0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < \frac{1}{\alpha},$$

and Theorem 3.2.1 still holds with $\tau_k = \lambda_k(\frac{1}{\alpha} - \lambda_k)$.

3.2.2 Ergodic convergence rates

We now turn to the ergodic convergence rate of (3.1.1). For this, let us define

$$\Lambda_k = \sum_{j=0}^k \lambda_j \quad \text{and} \quad \bar{e}_k = \frac{1}{\Lambda_k}\sum_{j=0}^k \lambda_j e_j.$$

Theorem 3.2.4 (Ergodic convergence rate). *Suppose that $\mathcal{C}_2 = \sum_{k\in\mathbb{N}} \lambda_k\|\varepsilon_k\| < +\infty$, then,*

$$\|\bar{e}_k\| \leq \frac{2(d_0 + \mathcal{C}_2)}{\Lambda_k}.$$

In particular, if $\inf_{k\in\mathbb{N}} \lambda_k > 0$, then $\|\bar{e}_k\| = O(1/k)$.

Proof. Again, let $z^\star \in \mathrm{fix}(\mathcal{F})$ such that $d_0 = \|z_0 - z^\star\|$. Since $\mathcal{F}_k$ is non-expansive, we have

$$\begin{aligned}\|z_{k+1} - z^\star\| &= \|\mathcal{F}_k(z_k) - \mathcal{F}_k(z^\star) + \lambda_k\varepsilon_k\| \leq \|z_k - z^\star\| + \lambda_k\|\varepsilon_k\| \\ &\leq \|z_{k-1} - z^\star\| + \sum_{j=k-1}^k \lambda_j\|\varepsilon_j\| \leq \|z_0 - z^\star\| + \sum_{j=0}^k \lambda_j\|\varepsilon_j\|.\end{aligned}$$

This together with the definition of $\bar{e}_k$ yields

$$\begin{aligned}\|\bar{e}_k\| &= \|\tfrac{1}{\Lambda_k}\textstyle\sum_{j=0}^k \lambda_j e_j\| = \tfrac{1}{\Lambda_k}\|\textstyle\sum_{j=0}^k (z_j - z_{j+1}) + \sum_{j=0}^k \lambda_j\varepsilon_j\| \\ &\leq \frac{1}{\Lambda_k}\big(\|z_0 - z^\star\| + \|z_{k+1} - z^\star\| + \textstyle\sum_{j=0}^k \lambda_j\|\varepsilon_j\|\big) \leq \frac{2(d_0 + \mathcal{C}_2)}{\Lambda_k},\end{aligned}$$

and we conclude the proof. □

Remark 3.2.5.

(i) As in the pointwise case, Theorem 3.2.4 holds when $\mathcal{F}$ is α-averaged, where λ_k can be allowed to vary in $[0, 1/\alpha]$.

(ii) When $\mathcal{F}$ is firmly non-expansive, *i.e.* the resolvent of a maximal monotone operator ((iv) of Lemma 2.3.8), an $O(1/k)$ ergodic convergence rate is also established in [128] with summable enlargement errors. For the methods which can also be cast in the HPE framework, our result coincides with the one in [128]. Note that in that work, they handled the non-stationary case (*i.e.* the parameter of the resolvent varies), see also our extension to the non-stationary case in Section 3.5. For the case without errors, we recover also the result in [14].

From Theorem 3.2.1 and 3.2.4, it is immediate to get the convergence rate bounds on the sequence $\{\|z_k - z_{k+1}\|\}_{k\in\mathbb{N}}$ in the exact case. To lighten the notation, let $v_k = z_k - z_{k+1}$ and $\bar{v}_k = \frac{1}{k+1}\sum_{j=0}^{k} v_j$.

Corollary 3.2.6. *Assume that $\varepsilon_k = 0$ for all $k \in \mathbb{N}$.*

(i) *If $0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < 1$, then $\|v_k\| \leq \sqrt{\frac{d_0^2}{\tau(k+1)}}$.*

(ii) *If $\underline{\lambda} = \inf_{k\in\mathbb{N}} \lambda_k > 0$, then $\|\bar{v}_k\| \leq \frac{2d_0}{k+1}$.*

Proof.

(i) By definition $v_k = \lambda_k e_k$, then from (3.2.2) we have $\|v_k\| = \|\lambda_k e_k\| \leq \sqrt{\frac{\lambda_k^2 d_0^2}{\tau(k+1)}} \leq \sqrt{\frac{d_0^2}{\tau(k+1)}}$.

(ii) A direct result of Theorem 3.2.4 by replacing Λ_k with $k+1$. □

3.3 Local linear convergence under metric sub-regularity

In the literature, for various splitting algorithms applied to a wide range of optimization problems, the following typical convergence profile has been observed in practice: globally the algorithm converges *sub-linearly*, and after a sufficiently large number of iterations, the algorithm turns to another regime where a *linear* convergence takes over. This has been for instance observed (and sometimes proved) for DR or FB when solving sparsity-enforcing minimization problems, see *e.g.* [73, 114].

In this section, we study the rationale underlying this local linear convergence behaviour. Our analysis relies on a metric sub-regularity assumption on the operator $\mathcal{F}'$ in (3.1.3).

Definition 3.3.1 (Metric sub-regularity [75]). A set-valued mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is called metrically sub-regular at $\tilde{z}$ for $\tilde{u} \in A(\tilde{z})$ if there exists $\kappa \geq 0$ along with neighbourhood $\mathcal{Z}$ of $\tilde{z}$ such that

$$\operatorname{dist}(z, A^{-1}(\tilde{u})) \leq \kappa \operatorname{dist}(\tilde{u}, A(z)), \ \forall z \in \mathcal{Z}. \tag{3.3.1}$$

The infimum of κ such that (3.3.1) holds is called the modulus of metric sub-regularity, and denoted by $\operatorname{subreg}(A; \tilde{z}|\tilde{u})$. The absence of metric regularity is signalled by $\operatorname{subreg}(A; \tilde{z}|\tilde{u}) = +\infty$.

Metric sub-regularity implies that, for any $z \in \mathcal{Z}$, $\operatorname{dist}(\tilde{u}, A(z))$ is bounded from below. The metric (sub-)regularity of multifunctions plays a crucial role in modern variational analysis and optimization. These properties are a key to study the stability of solutions of generalized equations, see the dedicated monograph [75].

Let us now specialize this notion to $\mathcal{F}'$ and $\tilde{u} = 0$. Since $\mathcal{F}'$ is single-valued and $\operatorname{zer}(\mathcal{F}') = \operatorname{fix}(\mathcal{F})$, from (3.3.1), metric sub-regularity of $\mathcal{F}'$ at some $z^\star \in \operatorname{fix}(\mathcal{F})$ for 0 is equivalent to

$$\operatorname{dist}(z, \operatorname{fix}(\mathcal{F})) \leq \kappa \|\mathcal{F}'(z)\|, \ \ \forall z \in \mathcal{Z}. \tag{3.3.2}$$

There are several concrete examples of operators $\mathcal{F}$ where $\mathcal{F}'$ fulfills (3.3.2).

Example 3.3.2 (Projection operator). Let $\mathcal{S} \subset \mathcal{H}$ be a non-empty closed convex set and $\mathcal{F} = \mathcal{P}_{\mathcal{S}}$ be the projection operator. Then $\operatorname{fix}(\mathcal{F}) = \mathcal{S}$, and

$$\operatorname{dist}(z, \mathcal{S}) = \|z - \mathcal{P}_{\mathcal{S}}(z)\| = \|\mathcal{F}'(z)\|.$$

Thus $\mathcal{F}'$ is metrically sub-regular at any $z^\star \in \mathcal{S}$ for 0 with $\mathcal{Z} = \mathcal{H}$ and modulus 1.

Using the relation between metric sub-regularity of $\mathcal{F}'$ and bounded linear regularity of $\mathcal{F}$ as defined in [22], other examples can be deduced for instance from [22, Examples 2.3 and 2.5]. Two other instructive examples, one on DR with two subspaces and the second on gradient descent will be discussed at the end of the section.

Another interesting situation is when $\mathcal{F}$ is firmly non-expansive, that is $\mathcal{F} = \mathcal{J}_A$ for some maximal monotone operator A ((iv) of Lemma 2.3.8), for which case (3.1.1) is the relaxed inexact PPA. Let $z^\star \in \operatorname{fix}(\mathcal{F}) = \operatorname{zer}(A)$ and suppose that $0 \in \operatorname{zer}(A)$. If A is metrically sub-regular at $z^\star$ for 0 with modulus γ, then

$$\begin{aligned}\operatorname{dist}(z, \operatorname{fix}(\mathcal{F})) = \operatorname{dist}(z, \operatorname{zer}(A)) &\leq \gamma \operatorname{dist}(0, A(z)) \\ &= \gamma \operatorname{dist}(0, \mathcal{F}^{-1}(z) - z) = \gamma \inf_{v \in \mathcal{F}^{-1}(z)} \|v - z\|, \quad \forall z \in \mathcal{Z}.\end{aligned}$$

Thus for all w such that $z = \mathcal{F}(w) \in \mathcal{Z}$, applying the previous inequality and using the triangle inequality, we get

$$\operatorname{dist}(\mathcal{F}(w), \operatorname{fix}(\mathcal{F})) \leq \gamma \|\mathcal{F}'(w)\| \quad \text{and} \quad \operatorname{dist}(w, \operatorname{fix}(\mathcal{F})) \leq (1+\gamma)\|\mathcal{F}'(w)\|.$$

Clearly, this is closely related, though not equivalent, to metric sub-regularity of $\mathcal{F}'$.

3.3.1 Main result

Metric sub-regularity implies that (3.3.2) gives an estimate for how far a point z is away from being the fixed-point of $\mathcal{F}$ in terms of the residual $\|z - \mathcal{F}(z)\|$. This is the rationale behind using such a regularity assumption on the operator $\mathcal{F}'$ to quantify the convergence rate on $\operatorname{dist}(z_k, \operatorname{fix}(\mathcal{F}))$. Thus, starting from $z_0 \in \mathcal{H}$, and by virtue of Theorem 3.2.1, one can recover a $o(1/\sqrt{k})$ rate on $\operatorname{dist}(z_k, \operatorname{fix}(\mathcal{F}))$. In fact, we can do even better as shown in the following theorem. We use the shorthand notation $d_k = \operatorname{dist}(z_k, \operatorname{fix}(\mathcal{F}))$.

Theorem 3.3.3 (Local convergence rate). *Let $z^\star \in \operatorname{fix}(\mathcal{F})$, suppose that $\mathcal{F}'$ is metrically sub-regular at $z^\star$ with neighbourhood $\mathcal{Z}$ of $z^\star$, let $\kappa > \operatorname{subreg}(\mathcal{F}'; z^\star|0)$, $\lambda_k \in [0,1]$. Choose $r \geq 0$ and define the ball around $z^\star$ as $\mathbb{B}_r(z^\star) \overset{\text{def}}{=} \{z \in \mathcal{H} : \|z - z^\star\| \leq r\} \subset \mathcal{Z}$, assume $\mathcal{C}_2 = \sum_{k\in\mathbb{N}} \lambda_k \|\varepsilon_k\|$ is small enough such that*

$$\mathbb{B}_{(r+\mathcal{C}_2)}(z^\star) \subseteq \mathcal{Z}.$$

Then for any starting point $z_0 \in \mathbb{B}_r(z^\star)$, we have for all $k \in \mathbb{N}$,

$$d_{k+1}^2 \leq \zeta_k\, d_k^2 + h_k, \quad \text{where} \quad \zeta_k = \begin{cases} 1 - \frac{\tau_k}{\kappa^2}, & \text{if } \tau_k/\kappa^2 \in]0,1] \\ \frac{\kappa^2}{\kappa^2+\tau_k}, & \text{otherwise} \end{cases} \in [0,1[, \tag{3.3.3}$$

and $h_k = \nu_1 \lambda_k \|\varepsilon_k\|$. Moreover,

(i) *d_k converges to 0 if $\{\tau_k\}_{k\in\mathbb{N}} \notin \ell_+^1$.*

(ii) *Let $\chi_k = \prod_{j=0}^k \zeta_j$, if $\chi = \limsup_{k\to+\infty} \sqrt[k]{\chi_k} < 1$, then $\{d_k^2\}_{k\in\mathbb{N}} \in \ell_+^1$. When $\varepsilon_k = 0$, then $\lim_{k\to+\infty} \sqrt[k]{d_k} < 1$, which is linear convergence.*

(iii) *If $0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < 1$, then there exists $\zeta \in]0,1[$ such that*

$$d_{k+1}^2 \leq \zeta^k \big(d_0^2 + \textstyle\sum_{j=0}^k \zeta^{-j+1} h_j\big).$$

Proof. Choose $r' > r \geq 0$, and make r smaller if necessary so that $\mathbb{B}_r(z^\star) \subset \mathbb{B}_{r'}(z^\star) \subseteq \mathcal{Z}$ and there holds

$$r + \mathcal{C}_2 \leq r'.$$

Pick arbitrarily $z_0 \in \mathbb{B}_r(z^\star)$ ($z_0 = z^\star$ will make no difference since we are considering the *inexact* case). Owing to Lemma 3.1.5, for any $\tilde{z} \in \operatorname{fix}(\mathcal{F})$, we have

$$\|z_{k+1} - \tilde{z}\| \leq \|z_k - \tilde{z}\| + \lambda_k\|\varepsilon_k\| \leq \cdots \leq \|z_0 - \tilde{z}\| + \textstyle\sum_{j=0}^k \lambda_j \|\varepsilon_j\| \leq r + \mathcal{C}_2 \leq r',$$

which implies that starting from any point $z_0 \in \mathbb{B}_r(z^\star)$, $z_k \in \mathbb{B}_{r'}(z^\star)$ holds for all $k \in \mathbb{N}$, hence $z_k \in \mathcal{Z}$. For $\forall k \in \mathbb{N}$, let $\tilde{z} \in \operatorname{fix}(\mathcal{F})$ be the fixed point such that $d_k = \|z_k - \tilde{z}\|$, and denote $h_k = \nu_1 \lambda_k \|\varepsilon_k\|$, then owing to the metric sub-regularity of $\mathcal{F}'$ and Lemma 3.1.5,

$$d_{k+1}^2 \leq \|z_{k+1} - \tilde{z}\|^2 \leq \|z_k - \tilde{z}\|^2 - \tau_k \|\mathcal{F}'(z_k) - \mathcal{F}'(\tilde{z})\|^2 + h_k \leq d_k^2 - \tfrac{\tau_k}{\kappa^2} d_k^2 + h_k \tag{3.3.4}$$

$$\leq d_k^2 - \tfrac{\tau_k}{\kappa^2}(d_{k+1}^2 - h_k) + h_k = d_k^2 - \tfrac{\tau_k}{\kappa^2} d_{k+1}^2 + \tfrac{1}{\zeta_k} h_k. \tag{3.3.5}$$

If $\tau_k/\kappa^2 \in]0,1[$, then from (3.3.4) we have $d_{k+1}^2 \leq (1 - \frac{\tau_k}{\kappa^2}) d_k^2 + h_k$, or if $1 \leq \tau_k/\kappa^2$, (3.3.5) yields $d_{k+1}^2 \leq \frac{\kappa^2}{\kappa^2+\tau_k} d_k^2 + h_k$, $\lambda_k \in]0,1]$ ensures $\kappa^2/(\kappa^2+\tau_k) \in]0,1]$. Therefore, we have

$$\zeta_k = \begin{cases} 1 - \frac{\tau_k}{\kappa^2}, & \text{if } \tau_k/\kappa^2 \in]0,1[\\ \frac{\kappa^2}{\kappa^2+\tau_k}, & \text{if } 1 \leq \tau_k/\kappa^2 \end{cases} \in]0,1].$$

Furthermore,

$$d_{k+1}^2 \leq \zeta_k\, d_k^2 + h_k \leq \cdots \leq \chi_k\, d_0^2 + \textstyle\sum_{j=0}^k \phi_{k-j} h_j \leq \chi_k\, d_0^2 + \sum_{j=0}^k h_j, \tag{3.3.6}$$

where $\chi_k = \prod_{j=0}^k \zeta_j$ and $\phi_{k-j} = \prod_{\ell=j+1}^k \zeta_\ell$.

(i) From (3.3.6) we have $d_{k+1}^2 \leq d_k^2 + h_k$, then the $d_k^2 \to d \geq 0$ ([147, Lemma 2.2.2]). If $\{\tau_k\}_{k\in\mathbb{N}} \notin \ell_+^1$, then $\|e_k\| \to 0$ (Theorem 3.2.1), and by metric sub-regularity we have $d_k \leq \kappa \|e_k\|$, therefore $d = 0$.

(ii) If $\chi = \limsup_{k\to+\infty} \sqrt[k]{\chi_k} < 1$, then $\lim_{k\to+\infty} \chi_k = 0$ and $\{\chi_k\}_{k\in\mathbb{N}}, \{\phi_k\}_{k\in\mathbb{N}} \in \ell_+^1$. Since also $\{h_k\}_{k\in\mathbb{N}} \in \ell_+^1$, hence the convolution $\{\sum_{j=0}^k \phi_{k-j} h_j\}_{k\in\mathbb{N}} \in \ell_+^1$, as a result, $\{\chi_k\, d_0^2 + \sum_{j=0}^k \phi_{k-j} h_j\}_{k\in\mathbb{N}} \in \ell_+^1$ and so is $\{d_k^2\}_{k\in\mathbb{N}}$.
If $\varepsilon_k = 0$, then from (3.3.6) we have $\lim_{k\to+\infty} \sqrt[k]{d_k} \leq \limsup_{k\to+\infty} \sqrt[k]{\chi_k} < 1$.

(iii) If $0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < 1$, then there exists $\zeta = \sup_{k\in\mathbb{N}} \kappa^2/(\kappa^2+\tau_k) < 1$ which concludes the result. □

Remark 3.3.4.

(i) If the iteration is exact, *i.e.* $\varepsilon_k \equiv 0$, then we have $d_{k+1} \leq (\sqrt{\zeta})^k d_0$.

(ii) When the fixed point is a singleton, Theorem 3.3.3 holds replacing d_k by $\|z_k - z^\star\|$.

(iii) For simplicity, suppose the iteration is exact, and let $z^\star \in \operatorname{fix}(\mathcal{F})$ such that $d_k = \|z_k - z^\star\|$. Then we have

$$\|e_k\|^2 = \|z_k - z^\star + \mathcal{F}z^\star - \mathcal{F}z_k\|^2 \leq 4 d_k^2,$$

which means locally, $\|e_k\|$ also converges linearly to 0 given that $\sum_{k\in\mathbb{N}} \tau_k = +\infty$.

(iv) As far as the claim in (iii) of Theorem 3.3.3 is concerned, if $\exists \xi \in]0,1[$ such that $h_k = O(\xi^k)$, then
  (a) If $\xi < \zeta$, then $d_{k+1}^2 = O(\zeta^k)$.
  (b) If $\xi = \zeta$, then $d_{k+1}^2 = O(k\zeta^k) = o\big((\zeta+\delta)^k\big)$, $\delta > 0$.
  (c) If $\xi > \zeta$, then $d_{k+1}^2 = o(\xi^k)$.

Theorem 3.3.3 extends readily to the α-averaged case.

Corollary 3.3.5. *If $\mathcal{F}$ is α-averaged, then Theorem 3.3.3 holds by substituting $\lambda_k\alpha$ for λ_k and $\kappa\alpha$ for κ, and moreover $\lambda_k \in]0, 1/\alpha]$.*

Remark 3.3.6.

(i) When $\lambda_k \equiv \lambda$ and $\varepsilon_k \equiv 0$, our second rate estimate in (3.3.3) encompasses that of [22, Lemma 3.8].

(ii) Equivalent characterizations of metric sub-regularity can be given, for instance in terms of derivative criteria. In particular, as $\mathcal{F}'$ is single-valued, metric sub-regularity of $\mathcal{F}'$ holds if $\mathcal{F}'$ is differentiable on a neighbourhood of $z^\star$ with non-singular derivatives at z around $z^\star$, and the operator

norms of their inverses are uniformly bounded [74, Theorem 1.2]. Computing the metric regularity modulus κ is however far from obvious in general even for the differentiable case.

3.3.2 Two examples

To make the above result easier to grasp, we now discuss two illustrating examples.

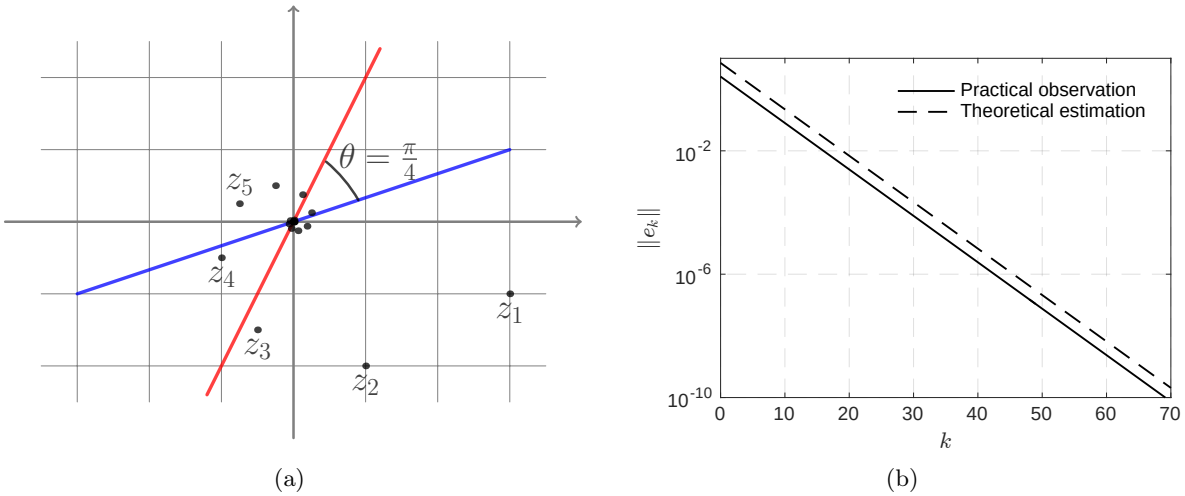
Douglas–Rachford with two subspaces in $\mathbb{R}^2$ Let T_1 and T_2 two subspaces in $\mathbb{R}^2$ forming an angle of $\theta \in]0, \pi/2[$. Consider the problem of finding $T_1 \cap T_2 = \{0\}$ using Douglas–Rachford splitting. Following [22, Example 2.3 and 2.5], the fixed-point operator of DR for this case reads

$$\mathcal{F}_{\text{DR}} = \begin{bmatrix} \cos^2(\theta) & -\sin(\theta)\cos(\theta) \\ \sin(\theta)\cos(\theta) & \cos^2(\theta) \end{bmatrix}.$$

It is then easy to check that $\mathcal{F}'_{\text{DR}} = \text{Id} - \mathcal{F}_{\text{DR}}$ is such that $\|\mathcal{F}'_{\text{DR}}(z)\| = \sin(\theta)\|z\|$, hence metrically sub-regular with modulus $1/\sin(\theta)$. It then follows from (3.3.3) that the rate estimate is

$$\zeta_k = 1 - (2-\lambda_k)\lambda_k \sin^2(\theta) = (1-\lambda_k)^2 + \lambda_k(2-\lambda_k)\cos^2(\theta) \in]0,1[, \tag{3.3.7}$$

for $\lambda_k \in]0,2[$. This is exactly the optimal rate estimate provided in [118, 18] (see also Chapter 7). Observe that as remarked in [118], the best rate $\zeta_k \equiv \cos^2(\theta)$ is obtained for $\lambda_k \equiv 1$ (*i.e.* no relaxation).



(a) (b)

Figure 3.1: Example of finding the intersection of two subspace in $\mathbb{R}^2$ using DR method. (a) two lines intersect with the angle $\pi/4$. (b) convergence profile of $\{\|z_k - z_{k+1}\|\}_{k\in\mathbb{N}}$.

Figure 3.1 shows an example of the above problem, where the angle between the two lines is $\theta = \pi/4$. The non-relaxed DR iteration is applied, and the dots in (a) of Figure 3.1 is the sequence $\{z_k\}_{k\in\mathbb{N}}$. The convergence profile of $\{\|z_k - z_{k+1}\|\}_{k\in\mathbb{N}}$ is shown in figure (b), where the solid line stands for the practical observation and dashed line for the theoretically predicted profile using (3.3.3). From (3.3.7) and (i) of Remark 3.3.4, we have $\zeta = \cos^2(\theta)$, hence our rate estimate of $\{\|z_k - z_{k+1}\|\}_{k\in\mathbb{N}}$ is $\sqrt{\zeta} = \sqrt{2}/2$.

Gradient descent The second example is the gradient descent. Consider minimizing a smooth convex function $F : \mathbb{R}^n \to \mathbb{R}$ whose gradient ∇F is Lipschitz continuous, and moreover F is locally C^2-smooth and strongly convex on a closed convex neighbourhood $\mathcal{S} \subset \mathbb{R}^n$ a minimizer. F admits a unique minimizer, and without loss of generality, we assume it is 0. Let $\underline{\eta}, \bar{\eta}$ be the local strong convexity modulus of F and the Lipschitz constant of ∇F respectively. Then $\underline{\eta} > 0$ is equivalent metric sub-regularity of ∇F at 0 for 0 with modulus $1/\underline{\eta}$ [6, Theorem 3.5].

For simplicity, consider the non-relaxed gradient descent [147] for minimizing F with constant step-size, *i.e.* $x_{k+1} = (\mathrm{Id} - \gamma\nabla F)(x_k)$ where $\gamma \in]0, 2/\bar{\eta}[$. Define $\mathcal{F}_{\mathrm{GD}} \overset{\mathrm{def}}{=} \mathrm{Id} - \gamma\nabla F$ which is $(\gamma\bar{\eta}/2)$-averaged (Lemma 2.3.9), then it is easy to see that the above iteration is a special case of (3.1.1) without error. Define

$$\mathcal{F}'_{\mathrm{GD}} \overset{\mathrm{def}}{=} \mathrm{Id} - \mathcal{F}_{\mathrm{GD}} = \gamma\nabla F.$$

By virtue of [75, Theorem 4B.1], $\mathcal{F}'_{\mathrm{GD}}$ is metrically regular, hence sub-regular, and the metric regularity modulus $\kappa = 1/(\gamma\underline{\eta})$. In fact, we could have anticipated this directly from the local strong monotonicity of ∇F. Specializing the rate of Theorem 3.3.3, we get

$$\zeta = 1 - \frac{t(2-t)}{\mathrm{cnd}^2} \in [0, 1[,$$

where we set $t = \gamma\bar{\eta} \in]0, 2[$, and $\mathrm{cnd} = \bar{\eta}/\underline{\eta}$ can be seen as the condition number of the Hessian of F along $\mathcal{S}$. It is obvious that the smallest ζ is attained for $t = 1$, *i.e.* $\gamma = 1/\bar{\eta}$.

An experiment is demonstrated in Figure 3.2, where the step-size is chosen as $\gamma = 1/(2\bar{\eta})$. The convergence profiles of $\|e_k\| = \|x_k - x_{k+1}\|$ is plotted. For the legends of the plots, "T" means theoretical predictions while "P" for practical observation. As predicted by our result, the convergence profile exhibits two regimes, a global sub-linear one, and a local linear one.

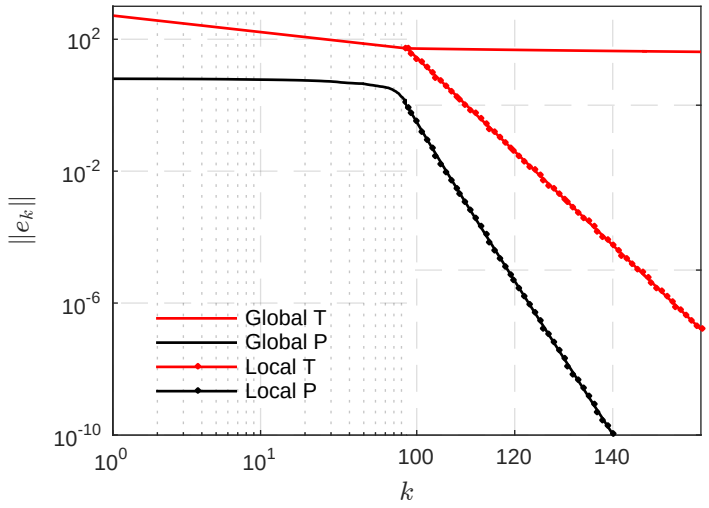


Figure 3.2: Global and local convergence profiles of gradient descent minimizing F with $(\underline{\eta}, \bar{\eta}) = (0.8, 1)$. The step-size is chosen as $\gamma = 1/(2\bar{\eta})$.

3.4 Applications to operator splitting methods

In this section, we apply the obtained results to conduct quantitative convergence analysis for a class of operator splitting methods in the literature, and mainly focus on the global convergence rates. For the methods introduced in Section 2.4 (PPA, FB, DR/ADMM, Primal–Dual splitting), out results can be adapted to them readily. In to following, we focus on mainly 3 operator splitting methods, the GFB, DR/ADMM and a Primal–Dual splitting method.

3.4.1 Generalized Forward–Backward splitting

The generalized Forward–Backward splitting method [148] considers solving the following monotone inclusion problem, let $m > 1$ be an integer,

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in \big(B + \textstyle\sum_{i=1}^{m} A_i\big)(x), \tag{3.4.1}$$

where $B : \mathcal{H} \to \mathcal{H}$ still is β-cocoercive for some $\beta > 0$, and

(i) $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone for each $i \in \{1, ..., m\}$.

(ii) $\operatorname{zer}(B + \sum_i A_i) \neq \emptyset$.

By lifting problem (3.4.1) into a product space, GFB achieves full splitting by applying the resolvent of each A_i's implicitly and B explicitly. The inexact form of GFB method is described in Algorithm 3.

Algorithm 3: Inexact generalized Forward–Backward splitting

Initial: Let $(\omega_i)_i \in]0,1[^m$ such that $\sum_{i=1}^m \omega_i = 1$, $\gamma \in]0, 2\beta[$.

repeat

Let $\lambda_k \in]0, \frac{4\beta-\gamma}{2\beta}[$, $\varepsilon_{1,k}, \varepsilon_{2,i,k} \in \mathcal{H}$:

For $i \in \{1, \cdots, m\}$

$$\left\lfloor \begin{array}{l} z_{i,k+1} = z_{i,k} + \lambda_k \big(\jmath_{\frac{\gamma}{\omega_i} A_i}(2x_k - z_{i,k} - \gamma B(x_k) + \varepsilon_{1,k}) + \varepsilon_{2,i,k} - x_k \big), \\ x_{k+1} = \sum_{i=1}^m \omega_i z_{i,k+1}. \end{array} \right. \tag{3.4.2}$$

$k = k + 1$;

until *convergence*;

GFB recovers FB when $m = 1$. If $B = 0$, then GFB recovers DR in the product space.

Product space of GFB Let $(\omega_i)_i \in]0, +\infty]^m$, let $\boldsymbol{\mathcal{H}} = \mathcal{H}^m$ be the product space endowed with the scalar product and norm defined by

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \boldsymbol{\mathcal{H}},\ \langle \boldsymbol{x}, \boldsymbol{x}' \rangle = \sum_{i=1}^m \omega_i \langle x_i, x_i' \rangle,\ \ \|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^m \omega_i \|x_i\|^2}.$$

Define $\mathbf{Id}$ the identity operator on $\boldsymbol{\mathcal{H}}$.

Let $\boldsymbol{\mathcal{S}} = \{\boldsymbol{x} = (x_i)_i \in \boldsymbol{\mathcal{H}} | x_1 = \cdots = x_m\}$ and its orthogonal complement $\boldsymbol{\mathcal{S}}^\perp = \big\{\boldsymbol{x} = (x_i)_i \in \boldsymbol{\mathcal{H}} | \sum_{i=1}^m \omega_i x_i = 0\big\} \subset \boldsymbol{\mathcal{H}}$. We also define the canonical isometry $\boldsymbol{C} : \mathcal{H} \to \boldsymbol{\mathcal{S}}$, $x \mapsto (x, \cdots, x)$. We have $\forall \boldsymbol{z} \in \boldsymbol{\mathcal{H}}$,

$$\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{z}) \overset{\text{def}}{=} \boldsymbol{C}\big(\textstyle\sum_{i=1}^m \omega_i z_i\big).$$

Clearly $\mathcal{P}_{\boldsymbol{\mathcal{S}}}$ is self-adjoint, and its reflection operator is $\mathcal{R}_{\boldsymbol{\mathcal{S}}} \overset{\text{def}}{=} 2\mathcal{P}_{\boldsymbol{\mathcal{S}}} - \mathbf{Id}$.

Let $\boldsymbol{\gamma} = (\gamma_i)_i \in]0, +\infty[^m$. For maximal monotone operators $A_i, i \in \{1, ..., m\}$ on $\mathcal{H}$, define the operator $\boldsymbol{\gamma A} : \boldsymbol{\mathcal{H}} \rightrightarrows \boldsymbol{\mathcal{H}}, \boldsymbol{x} = (x_i)_i \mapsto \times_{i=1}^m \gamma_i A_i(x_i)$, *i.e.* its graph is

$$\operatorname{gra}(\boldsymbol{\gamma A}) = \bigtimes_{i=1}^m \operatorname{gra}(\gamma_i A_i) = \big\{(\boldsymbol{x}, \boldsymbol{u}) \in \boldsymbol{\mathcal{H}}^2 | \boldsymbol{x} = (x_i)_i,\ \boldsymbol{u} = (u_i)_i,\ u_i \in \gamma_i A_i(x_i)\big\}.$$

For a single-valued maximal monotone operator B, denote $\boldsymbol{B} : \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}$, $\boldsymbol{x} = (x_i)_i \mapsto (B(x_i))_i$. It is immediate to check that both $\boldsymbol{\gamma A}$ and $\boldsymbol{B}$ are maximal monotone. We also define the operators $\boldsymbol{B}_{\boldsymbol{\mathcal{S}}} = \boldsymbol{B}\mathcal{P}_{\boldsymbol{\mathcal{S}}}$, $\jmath_{\boldsymbol{\gamma A}} = (\jmath_{\gamma_i A_i})_i$ and $\mathcal{R}_{\boldsymbol{\gamma A}} = 2\jmath_{\boldsymbol{\gamma A}} - \mathbf{Id}$.

3.4.1.1 Fixed point formulation of GFB

Follow the product space, define the following 2 operators that act on $\boldsymbol{\mathcal{H}}$,

$$\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}} = \frac{1}{2}(\mathcal{R}_{\boldsymbol{\gamma A}} \mathcal{R}_{\boldsymbol{\mathcal{S}}} + \mathbf{Id}),\ \ \boldsymbol{\mathcal{F}}_{2,\gamma} = \mathbf{Id} - \gamma \boldsymbol{B}_{\boldsymbol{\mathcal{S}}}. \tag{3.4.3}$$

Lemma 3.4.1. *For the two operators defined in (3.4.3), the following statements hold:*

(i) *The composed operator* $\boldsymbol{\mathcal{F}}_{\text{GFB}} = \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}} \boldsymbol{\mathcal{F}}_{2,\gamma}$ *is* $\frac{2\beta}{4\beta-\gamma}$*-averaged.*

(ii) $\forall \boldsymbol{\gamma} \in]0, +\infty[^m$, *there exists a maximal monotone operator* $\boldsymbol{A}'_{\boldsymbol{\gamma}} : \boldsymbol{\mathcal{H}} \rightrightarrows \boldsymbol{\mathcal{H}}$ *such that* $\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}} = \jmath_{\boldsymbol{A}'_{\boldsymbol{\gamma}}}$.

Proof. (i) Lemma 2.4.10 and 2.4.7. (ii) Lemma 2.3.8, or see [148, Proposition 4.15]. □

Define $u_{i,k+1} \overset{\text{def}}{=} \mathcal{J}_{\frac{\gamma}{\omega_i}A_i}(2x_k - z_{i,k} - \gamma B(x_k))$, $i \in \{1, ..., m\}$, then the iterates of Algorithm 3 can be reformulated to the following:

$$\left\lfloor \begin{array}{l} \text{For } i = 1, \cdots, m \\ \quad \left\lfloor \begin{array}{l} u_{i,k+1} = \mathcal{J}_{\frac{\gamma}{\omega_i}A_i}(2x_k - z_{i,k} - \gamma B(x_k)), \\ z_{i,k+1} = z_{i,k} + \lambda_k(u_{i,k+1} + \varepsilon_{i,k} - x_k), \end{array} \right. \\ x_{k+1} = \sum_{i=1}^m \omega_i z_{i,k+1}, \end{array} \right.$$

where $\varepsilon_{i,k} = (\mathcal{J}_{\frac{\gamma}{\omega_i}A_i}(2x_k - z_{i,k} - \gamma B(x_k) + \varepsilon_{1,k}) + \varepsilon_{2,i,k}) - u_{i,k+1}$. Let $\boldsymbol{\varepsilon}_{1,k} = \boldsymbol{C}(\varepsilon_{1,k})$, $\boldsymbol{\varepsilon}_{2,k} = (\varepsilon_{2,i,k})_i$, then we can obtain the fixed point iteration of GFB.

Lemma 3.4.2 (Fixed point formulation of GFB). *Let $\boldsymbol{\gamma} = \left(\frac{\gamma}{\omega_i}\right)_i$, then Algorithm 3 is equivalent to the following inexact relaxed fixed point iteration,*

$$\begin{cases} \boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \lambda_k\big(\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{2,k} - \boldsymbol{z}_k\big), \\ x_{k+1} = \sum_{i=1}^m \omega_i z_{i,k+1}. \end{cases} \tag{3.4.4}$$

Proof. Start from Algorithm 3, we have

$$\begin{aligned} \boldsymbol{z}_{k+1} &= \boldsymbol{z}_k + \lambda_k\big(\mathcal{J}_{\boldsymbol{\gamma}\boldsymbol{A}}\big((2\mathcal{P}_{\boldsymbol{\mathcal{S}}} - 2\gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}} - \mathbf{Id} + \gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}})(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}\big) \\ &\quad - \tfrac{1}{2}\big((2\mathcal{P}_{\boldsymbol{\mathcal{S}}} - 2\gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}} - \mathbf{Id} + \gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}})(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}\big) + \tfrac{1}{2}(\boldsymbol{z}_k - \gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{2,k} - \boldsymbol{z}_k\big) \\ &= \boldsymbol{z}_k + \lambda_k\big(\mathcal{J}_{\boldsymbol{\gamma}\boldsymbol{A}}\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) - \tfrac{1}{2}\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) \\ &\quad + \tfrac{1}{2}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{2,k} - \boldsymbol{z}_k\big) \\ &= \boldsymbol{z}_k + \lambda_k\big(\tfrac{1}{2}(2\mathcal{J}_{\boldsymbol{\gamma}\boldsymbol{A}} - \mathbf{Id})\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \tfrac{1}{2}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{2,k} - \boldsymbol{z}_k\big) \\ &= \boldsymbol{z}_k + \lambda_k\big(\tfrac{1}{2}\mathcal{R}_{\boldsymbol{\gamma}\boldsymbol{A}}\mathcal{R}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \tfrac{1}{2}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{2,k} - \boldsymbol{z}_k\big) \\ &= \boldsymbol{z}_k + \lambda_k\big(\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{2,k} - \boldsymbol{z}_k\big). \end{aligned}$$

Equivalence $\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) = \mathcal{R}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k})$ is applied above since $\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k} \in \boldsymbol{\mathcal{S}}$. □

To lighten the notation, let $\boldsymbol{\mathcal{F}}_{\text{GFB}} = \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}\boldsymbol{\mathcal{F}}_{2,\gamma}$, then (3.4.4) can be written as

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \lambda_k(\boldsymbol{\mathcal{F}}_{\text{GFB}}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_k - \boldsymbol{z}_k), \tag{3.4.5}$$

where $\boldsymbol{\varepsilon}_k = (\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}(\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}_k) + \boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{2,k}) - \boldsymbol{\mathcal{F}}_{\text{GFB}}(\boldsymbol{z}_k)$. Obviously, (3.4.4) is exactly in the form of (3.1.1). Therefore, the GFB iterates converge weakly [148, Theorem 4.1], and obey the convergence rates in Theorem 3.2.1 and 3.2.4. Moreover, we can establish certain convergence rates for the structured monotone inclusion (3.4.1).

3.4.1.2 Convergence rates of GFB

Define the following variables

$$g_{k+1} = \tfrac{1}{\gamma}x_k - B(x_k) - \tfrac{1}{\gamma}\textstyle\sum_i \omega_i u_{i,k+1},\ \bar{g}_{k+1} = \tfrac{1}{\gamma}\bar{x}_k - B(\bar{x}_k) - \tfrac{1}{\gamma}\sum_{i=1}^m \omega_i \bar{u}_{i,k+1},$$
$$\bar{x}_k = \tfrac{1}{k+1}\textstyle\sum_{j=0}^k x_j,\ \bar{u}_{i,k+1} = \tfrac{1}{k+1}\sum_{j=0}^k u_{i,j+1},\ i \in \{1, ..., m\}.$$

Proposition 3.4.3. *We have $g_{k+1} \in \sum_i A_i u_{i,k+1}$. Moreover,*

(i) *If $0 < \inf_{k\in\mathbb{N}}\lambda_k \le \sup_{k\in\mathbb{N}}\lambda_k < \frac{4\beta-\gamma}{2\beta}$, $\{(k+1)\|\varepsilon_{1,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$ and $\forall i \in \{1, ..., m\}$, $\{(k+1)\|\varepsilon_{2,i,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, then*

$$d\big(0, \textstyle\sum_i A_i u_{i,k+1} + B(\sum_i \omega_i u_{i,k+1})\big) \le \frac{1}{\gamma}\sqrt{\frac{d_0^2 + \mathcal{C}_1}{\underline{\tau}(k+1)}}.$$

(ii) *If* $\underline{\lambda} = \inf_{k\in\mathbb{N}} \lambda_k > 0$, $\{\lambda_k\|\varepsilon_{1,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$ *and* $\{\lambda_k\|\varepsilon_{2,i,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, $\forall i \in \{1,...,m\}$, *then*

$$\|\bar{g}_{k+1} + B(\textstyle\sum_i \omega_i \bar{u}_{i,k+1})\| \leq \frac{2(d_0 + \mathcal{C}_2)}{\gamma\underline{\lambda}(k+1)}.$$

$\mathcal{C}_1, \mathcal{C}_2$ *are the constants in Theorem 3.2.1 and 3.2.4 respectively.*

Remark 3.4.4. The rate in (i) above in fact is $o(1/\sqrt{k})$ ((i) of Remark 3.2.2).

Proof.

(i) From the definition of $u_{i,k+1}$ and the resolvent equation (2.4.1), we have

$$\begin{aligned} &2x_k - z_{i,k} - \gamma B(x_k) - u_{i,k+1} \in \tfrac{\gamma}{\omega_i} A_i(u_{i,k+1}) \\ \iff &\tfrac{\omega_i}{\gamma}\big(2x_k - z_{i,k} - \gamma B(x_k) - u_{i,k+1} + \gamma B(\textstyle\sum_i \omega_i u_{i,k+1})\big) \in A_i(u_{i,k+1}) + \omega_i B(\textstyle\sum_i \omega_i u_{i,k+1}). \end{aligned}$$

Then sum up over i to get

$$\tfrac{1}{\gamma}x_k - B(x_k) - \tfrac{1}{\gamma}\textstyle\sum_i \omega_i u_{i,k+1} + B(\sum_i \omega_i u_{i,k+1}) \in \sum_i A_i(u_{i,k+1}) + B(\sum_i \omega_i u_{i,k+1}).$$

For the summability of the errors, we have

$$\begin{aligned} \textstyle\sum_{k\in\mathbb{N}}(k+1)\|\boldsymbol{\varepsilon}_k\| &\leq \textstyle\sum_{k\in\mathbb{N}}(k+1)\|\boldsymbol{\varepsilon}_{1,k}\| + \sum_{k\in\mathbb{N}}(k+1)\|\boldsymbol{\varepsilon}_{2,k}\| \\ &\leq \textstyle\sum_{k\in\mathbb{N}}(k+1)\|\varepsilon_{1,k}\| + \sum_{i=1}^m w_i \sum_{k\in\mathbb{N}}(k+1)\|\varepsilon_{2,i,k}\| < +\infty, \end{aligned}$$

hence Theorem 3.2.1 applies. Combining this with the fact that $\mathrm{Id} - \gamma B$ is non-expansive (Lemma 2.3.9), we obtain

$$\begin{aligned} &d\big(0, \textstyle\sum_i A_i(u_{i,k+1}) + B(\sum_i \omega_i u_{i,k+1})\big) \\ &\leq \|g_{k+1} + \gamma B(\textstyle\sum_i \omega_i u_{i,k+1})\| = \tfrac{1}{\gamma}\|(\mathrm{Id} - \gamma B)x_k - (\mathrm{Id} - \gamma B)(\textstyle\sum_i \omega_i u_{i,k+1})\| \\ &\leq \tfrac{1}{\gamma}\|x_k - \textstyle\sum_i \omega_i u_{i,k+1}\| \leq \tfrac{1}{\gamma}\|\boldsymbol{x_k} - \boldsymbol{u}_{k+1}\| \leq \tfrac{1}{\gamma}\|\boldsymbol{e}_k\| \leq \tfrac{1}{\gamma}\sqrt{\tfrac{d_0^2 + \mathcal{C}_1}{\underline{\tau}(k+1)}}. \end{aligned}$$

(ii) Again, one can check that indeed $(\lambda_k\|\boldsymbol{\varepsilon}_k\|) \in \ell^1_+$. Thus, owing to Theorem 3.2.4 and the non-expansiveness of $\mathrm{Id} - \gamma B$, we have

$$\begin{aligned} \|\bar{g}_{k+1} + B(\textstyle\sum_i \omega_i \bar{u}_{i,k+1})\| &= \tfrac{1}{\gamma}\|(\mathrm{Id} - \gamma B)(\bar{x}_k) - (\mathrm{Id} - \gamma B)(\textstyle\sum_i \omega_i \bar{u}_{i,k+1})\| \\ &\leq \tfrac{1}{\gamma}\|\bar{x}_k - \textstyle\sum_i \omega_i \bar{u}_{i,k+1}\| \leq \tfrac{1}{\gamma}\|\bar{\boldsymbol{x}}_k - \bar{\boldsymbol{u}}_{k+1}\| \\ &\leq \tfrac{1}{\gamma(k+1)}\|\textstyle\sum_{j=0}^k (\boldsymbol{z}_{j+1} - \boldsymbol{z}_j - \lambda_j\boldsymbol{\varepsilon}_j)/\lambda_j\| \\ &\leq \tfrac{1}{\gamma\underline{\lambda}(k+1)}\big(\|\boldsymbol{z}_0 - \boldsymbol{z}_{k+1}\| + \textstyle\sum_{j=0}^k \lambda_j\|\boldsymbol{\varepsilon}_j\|\big) \leq \tfrac{2(d_0 + \mathcal{C}_2)}{\gamma\underline{\lambda}(k+1)}. \end{aligned}$$

Where $\boldsymbol{x_k} = \boldsymbol{C}(x_k)$, $\boldsymbol{u}_k = (u_{i,k})_i$ and $\bar{\boldsymbol{x}}_k, \bar{\boldsymbol{u}}_{k+1}$ are defined accordingly. □

Remark 3.4.5.

(i) Proposition 3.4.3 indicates that GFB provides an ϵ-accurate solution in at most $O(1/\epsilon)$ iterations for the criterion $d(0, \sum_i A_i(u_{i,k}) + B(\sum_i \omega_i u_{i,k}))^2 \leq \epsilon$, which also recovers the sub-differential stopping criterion of the FB method when $m = 1$. This can then be viewed as a generalization of the best-known complexity bounds of the gradient descent method [133].

(ii) In [40], the following problem is considered

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A + B + \mathcal{N}_{\mathcal{S}})(x), \tag{3.4.6}$$

where

(a) $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone, $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive with $\beta > 0$.

(b) $\mathcal{S}$ is a closed vector subspace of $\mathcal{H}$.

(c) $\operatorname{zer}(A+B+\mathcal{N}_{\mathcal{S}}) \neq \emptyset$.

A so-called Forward–Douglas–Rachford splitting (FDR) is proposed to solve the above problem whose iteration is, with proper choice of parameters

$$\begin{aligned} z_{k+1} &= (1-\lambda_k)z_k + \frac{\lambda_k}{2}(\mathrm{Id} + \mathcal{R}_{\gamma A}\mathcal{R}_{N_{\mathcal{S}}})(\mathrm{Id} - \gamma \mathcal{P}_{\mathcal{S}} B \mathcal{P}_{\mathcal{S}})(z_k), \\ x_{k+1} &= \mathcal{P}_{\mathcal{S}}(z_{k+1}). \end{aligned} \tag{3.4.7}$$

By form, FDR shares the same structure as GFB. Therefore, FDR also obeys the convergence rate established in Section 3.2.

(iii) Using the *transportation formula* of the cocoercive operators in [162, Lemma 2.8], one can write the *non-relaxed* and *exact* GFB into the HPE framework introduced there, and derive the convergence rate [128] as discussed in the related work. However, it should be pointed out that, to establish the convergence rates for GFB under the HPE framework, neither relaxation nor errors can handle, and the convergence rate is non-uniform [128][1].

3.4.2 Douglas–Rachford splitting and ADMM

For the monotone inclusion problem (3.4.1), let $B = 0$, $m = 2$, then (3.4.1) becomes the problem we presented in Section 2.4.3.2 for the Douglas–Rachford splitting. Different from the DR iteration in (2.4.10), let us consider the inexact version of it, which is described below.

Algorithm 4: Inexact Douglas–Rachford splitting

Initial: $\gamma > 0$, $z_0 \in \mathcal{H}$ and $x_0 = \mathcal{J}_{\gamma A_2}(z_0)$.

repeat

Let $\lambda_k \in]0, 2]$, let $\varepsilon_{1,k}, \varepsilon_{2,k+1} \in \mathcal{H}$:

$$\left\lfloor \begin{aligned} z_{k+1} &= (1-\lambda_k)z_k + \lambda_k\big(z_k + \mathcal{J}_{\gamma A_1}(2x_k - z_k) + \varepsilon_{1,k} - x_k\big), \\ x_{k+1} &= \mathcal{J}_{\gamma A_2}(z_{k+1}) + \varepsilon_{2,k+1}. \end{aligned} \right. \tag{3.4.8}$$

$k = k + 1$;

until *convergence*;

If we compose A_1 with some bounded linear operator $L : \mathcal{H} \to \mathcal{G}$, then the problem would become

$$\text{find } \; x \in \mathcal{H} \; \text{ such that } \; 0 \in (L^* A_1 L + A_2)(x). \tag{3.4.9}$$

One well-known method to solve the above problem is the ADMM which is DR method applied to the dual of (3.4.9) [78], see Section 7.6 for more details. Therefore, in the following we discuss only the DR method.

DR (3.4.8) takes exactly the form (3.1.1) with

$$\begin{aligned} \mathcal{F}_{\text{DR}} &= \tfrac{1}{2}(\mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2} + \mathrm{Id}), \\ \varepsilon_k &= \big(\tfrac{1}{2}(\mathcal{R}_{\gamma A_1}(\mathcal{R}_{\gamma A_2}(z_k) + 2\varepsilon_{2,k}) + z_k) - \mathcal{F}_{\text{DR}}(z_k)\big) + \varepsilon_{1,k}, \end{aligned} \tag{3.4.10}$$

see *e.g.* [57]. Moreover, we have $\mathcal{F}_{\text{DR}} \in \mathcal{A}(\frac{1}{2})$ and $(\varepsilon_k)_{k\in\mathbb{N}} \in \ell^1_+$ owing to the summability of $\{\varepsilon_{1,k}\}_{k\in\mathbb{N}}, \{\varepsilon_{2,k}\}_{k\in\mathbb{N}}$. There if $\{\lambda_k(2-\lambda_k)\}_{k\in\mathbb{N}} \notin \ell^1_+$, the DR iterates then obey the convergence rates in Section 3.2.

[1] Let $(x_k, v_k) \in \operatorname{gra}(A)$ be a sequence generated by an iterative method for solving the monotone inclusion problem $0 \in A(x)$, then the non-uniform convergence rate means that for every $k \in \mathbb{N}$, there exists a $j \leq k$ such that $\|v_j\| = O(1/\sqrt{k})$.

Next, we turn to the corresponding monotone inclusion problem (2.4.9), and develop a criterion similar to Proposition 3.4.3. Define

$$u_{k+1} = \mathcal{J}_{\gamma A_1}(2x_k - z_k),\ v_{k+1} = \mathcal{J}_{\gamma A_2}(z_{k+1}),$$
$$g_{k+1} = \tfrac{1}{\gamma}(2x_k - z_k - u_{k+1} + z_{k+1} - v_{k+1}).$$

Proposition 3.4.6. *We have $g_{k+1} \in A_1(u_{k+1}) + A_2(v_{k+1})$. Moreover, if $0 < \inf_{k\in\mathbb{N}} \lambda_k \leq \sup_{k\in\mathbb{N}} \lambda_k < 2$, $\{(k+1)\|\varepsilon_{1,k}\|\}_{k\in\mathbb{N}}, \{(k+1)\|\varepsilon_{2,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, then*

$$d\big(0, A_1(u_{k+1}) + A_2(v_{k+1})\big) \leq \frac{1+\lambda_k}{\gamma}\sqrt{\frac{d_0^2 + \mathcal{C}_1}{\tau(k+1)}} + o(1/(k+1)), \tag{3.4.11}$$

where d_0 and $\mathcal{C}_1$ are those defined in Theorem 3.2.1.

Remark 3.4.7. The rate can be improved to $o(1/\sqrt{k})$ as pointed in (i) of Remark 3.2.2.

Proof. From (3.4.8), we have

$$2x_k - z_k - u_{k+1} \in \gamma A_1(u_{k+1}) \quad \text{and} \quad z_{k+1} - v_{k+1} \in \gamma A_2(v_{k+1}),$$

whose sum leads to

$$\begin{aligned}
g_{k+1} &= \tfrac{1}{\gamma}(2x_k - z_k - u_{k+1} + z_{k+1} - v_{k+1}) \\
&= \tfrac{1}{\gamma}\big(2x_k - z_k - \big(\tfrac{1}{\lambda_k}(z_{k+1} - z_k) + x_k - \varepsilon_{1,k}\big) + z_{k+1} - v_{k+1}\big) \\
&= \tfrac{1}{\gamma}\big(x_k - z_k - \tfrac{1}{\lambda_k}(z_{k+1} - z_k) + \varepsilon_{1,k} + z_{k+1} - v_{k+1}\big) \\
&= \tfrac{1}{\gamma}\big((v_k - z_k) - (v_{k+1} - z_{k+1}) + \tfrac{1}{\lambda_k}(z_k - z_{k+1}) + \varepsilon_{1,2,k}\big) \in A_1(u_{k+1}) + A_2(v_{k+1}),
\end{aligned}$$

where $\varepsilon_{1,2,k} = \varepsilon_{1,k} + \varepsilon_{2,k}$. Note that $v_k - z_k = (\mathrm{Id} - \mathcal{J}_{\gamma A_2})(z_k)$ and $(\mathrm{Id} - \mathcal{J}_{\gamma A_2})$ is firmly non-expansive (Lemma 2.3.10), whence we get

$$\begin{aligned}
&d\big(0, A_1(u_{k+1}) + A_2(v_{k+1})\big) \leq \|g_{k+1}\| \\
&= \tfrac{1}{\gamma}\|(v_k - z_k) - \tfrac{1}{\lambda_k}(v_{k+1} - z_{k+1}) + (z_k - z_{k+1}) + \varepsilon_k + \varepsilon_{1,2,k} - \varepsilon_k\| \\
&\leq \tfrac{1}{\gamma}(\|z_k - z_{k+1}\| + \|e_k\| + \|\varepsilon_{1,2,k} - \varepsilon_k\|) \leq \tfrac{1}{\gamma}(\|z_k - z_{k+1} + \lambda_k\varepsilon_k - \lambda_k\varepsilon_k\| + \|e_k\|) \\
&\leq \tfrac{1}{\gamma}(\lambda_k\|e_{k+1}\| + \lambda_k\|\varepsilon_k\| + \|e_k\| + \|\varepsilon_{1,2,k} - \varepsilon_k\|) \leq \tfrac{1+\lambda_k}{\gamma}\|e_k\| + \tfrac{1}{\gamma}\big(\lambda_k\|\varepsilon_k\| + \|\varepsilon_{1,2,k} - \varepsilon_k\|\big).
\end{aligned} \tag{3.4.12}$$

For the error $\|\varepsilon_k\|$, we have

$$\|\varepsilon_k\| \leq \tfrac{1}{2}\|\mathcal{R}_{\gamma A_1}(\mathcal{R}_{\gamma A_2}(z_k) + 2\varepsilon_{2,k}) - \mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2}(z_k)\| + \|\varepsilon_{1,k}\| \leq \|\varepsilon_{1,k}\| + \|\varepsilon_{2,k}\|, \tag{3.4.13}$$

and similarly

$$\|\varepsilon_k - \varepsilon_{1,2,k}\| \leq 2\|\varepsilon_{2,k}\|. \tag{3.4.14}$$

Denote $c_k = \frac{1}{\gamma}((2+\lambda_k)\|\varepsilon_{2,k}\| + \|\varepsilon_{1,k}\|)$, since $\{(k+1)\|\varepsilon_{1,k}\|\}_{k\in\mathbb{N}}, \{(k+1)\|\varepsilon_{2,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, then we have $\|\varepsilon_{1,k}\| = o(1/(k+1))$ and $\|\varepsilon_{2,k}\| = o(1/(k+1))$, which means $c_k = o(1/(k+1))$. Combining this with (3.4.12) leads to the first claim of the proposition. □

Remark 3.4.8. The obtained convergence rate can be readily adapted to the ADMM method, by exploiting the fact that ADMM is nothing but DR applied to the Fenchel dual of the problem. In particular, one can then show that the pointwise convergence rate of ADMM is indeed $O(1/\sqrt{k})$. A similar result in the exact case is presented in [93].

3.4.3 A Primal–Dual splitting method

In [173], a more general monotone problem which involves parallel sums is considered. Let m be a strictly positive integer and index $i \in \{1, ..., m\}$, and $(\omega_i)_i \in]0,1[^m$ such that $\sum_i \omega_i = 1$. Consider the following *primal* monotone inclusion problem,

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A+B)(x) + \textstyle\sum_i \omega_i L_i^*\big((C_i \square D_i)(L_i x - r_i)\big), \tag{3.4.15}$$

where

(i) $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone, $B : \mathcal{H} \to \mathcal{H}$ is $\beta_{_B}$-cocoercive for some $\beta_{_B} > 0$.

(ii) $L_i : \mathcal{H} \to \mathcal{G}_i$ is a non-zero bounded linear operator.

(iii) $C_i, D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ are maximal monotone, D_i is $\beta_{_{D_i}}$-strongly monotone for some $\beta_{_{D_i}} > 0$.

The *dual* problem of (3.4.15) reads,

$$\text{find } (v_1 \in \mathcal{G}_1, ..., v_m \in \mathcal{G}_m) \text{ such that } (\exists x \in \mathcal{H}) \begin{cases} 0 \in (A+B)(x) + \sum_i \omega_i L_i^* v, \\ 0 \in (C_i^{-1} + D_i^{-1})(v_i) - L_i x + r_i, \end{cases} \tag{3.4.16}$$

denote by $\mathcal{P}$ and $\mathcal{D}$ the solution sets of (3.4.15) and (3.4.16) respectively.

(iv) The set of minimizers of (3.4.15) and (3.4.16), *i.e.* $\mathcal{P}$ and $\mathcal{D}$, are both non-empty.

In [173], the author proposed a Primal–Dual splitting method described in Algorithm 5 to solve the above problem, whose inexact iteration is given below. Given $\gamma_{_A}, (\gamma_{_{C_i}})_i > 0$, define

$$\beta = \min\{\beta_{_B}, \beta_{_{D_1}}, ..., \beta_{_{D_m}}\} \min\left\{\frac{1}{\gamma_{_A}}, \frac{1}{\gamma_{_{C_1}}}, ..., \frac{1}{\gamma_{_{C_m}}}\right\}\left(1 - \sqrt{\gamma_{_A} \textstyle\sum_i \gamma_{_{C_i}} \omega_i \|L_i\|^2}\right).$$

Algorithm 5: An inexact Primal–Dual splitting

Initial: Choose $\gamma_{_A}, (\gamma_{_{C_i}})_i > 0$ such that $2\beta > 1$, and $x_0 \in \mathcal{H}, v_{i,0} \in \mathcal{G}_i$;

repeat

Let $\lambda_k \in]0, \frac{4\beta-1}{2\beta}]$, $(\varepsilon_{1,k}, \varepsilon_{2,k}) \in \mathcal{H}$ and $(\varepsilon_{3,i,k}, \varepsilon_{4,i,k}) \in \mathcal{G}_i$ for $i \in \{1, ..., m\}$:

$$\begin{aligned} &p_{k+1} = \partial_{\gamma_{_A} A}\big(x_k - \gamma_{_A}(\textstyle\sum_i \omega_i L_i^* v_{i,k} + B(x_k) + \varepsilon_{1,k})\big) + \varepsilon_{2,k}, \\ &\bar{x}_{k+1} = 2p_{k+1} - x_k, \\ &x_{k+1} = x_k + \lambda_k(p_{k+1} - x_k), \\ &\text{For } i = 1, \cdots, m \\ &\quad q_{i,k+1} = \partial_{\gamma_{_{C_i}} C_i^{-1}}\big(v_{i,k} + \gamma_{_{C_i}}(L_i \bar{x}_{k+1} - D_i^{-1}(v_{i,k}) - \varepsilon_{3,i,k} - r_i)\big) + \varepsilon_{4,i,k}, \\ &\quad v_{i,k+1} = v_{i,k} + \lambda_k(q_{i,k+1} - v_{i,k}). \end{aligned} \tag{3.4.17}$$

$k = k + 1$;

until *convergence*;

A very similar problem to (3.4.15) is considered in [62], where the authors consider the same problem but without the weight parameter ω_i. The main difference of the two algorithms is the product space which will be introduced shortly, and they are the same when $m = 1$, see Chapter 8 for the method proposed in [62]. Algorithm 5 recovers the method in [66] when $m = 1, r = 0, D = 0$, and the method proposed in [51] if moreover $B = 0$ and $\lambda_k \equiv 1$.

Fixed point formulation In this part, we recall briefly the fixed-point formulation corresponding to (5) whose detailed derivation can be found in [173, Section 3]. Define the product space $\boldsymbol{\mathcal{G}} = \mathcal{G}_1 \times \cdots \times \mathcal{G}_m$ endowed with the scalar product $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle_{\boldsymbol{\mathcal{G}}} = \sum_{i=1}^m \omega_i \langle v_{1,i}, v_{2,i} \rangle_{\mathcal{G}_i}$ and associated norm $\|\cdot\|_{\boldsymbol{\mathcal{G}}}$. Let $\boldsymbol{\mathcal{K}} = \mathcal{H} \oplus \boldsymbol{\mathcal{G}}$ be the Hilbert direct sum with the scalar product $\langle (x_1, \boldsymbol{v}_1), (x_2, \boldsymbol{v}_2) \rangle_{\boldsymbol{\mathcal{K}}} = \langle x_1, x_2 \rangle + \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle_{\boldsymbol{\mathcal{G}}}$ and norm $\|\cdot\|_{\boldsymbol{\mathcal{K}}}$.

Define the following operators on $\boldsymbol{\mathcal{K}}$,

$$\begin{aligned} &\boldsymbol{C} : \boldsymbol{\mathcal{K}} \rightrightarrows \boldsymbol{\mathcal{K}},\ (x, \boldsymbol{v}) \mapsto (A(x)) \times (r_1 + C_1^{-1}(v_1)) \times \cdots \times (r_m + C_m^{-1}(v_m)), \\ &\boldsymbol{D} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}},\ (x, \boldsymbol{v}) \mapsto (\textstyle\sum_i \omega_i L_i^* v_i, -L_1 x, \cdots, -L_m x), \\ &\boldsymbol{E} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}},\ (x, \boldsymbol{v}) \mapsto (B(x), D_1^{-1}(v_1), \cdots, D_m^{-1}(v_m)), \\ &\boldsymbol{\mathcal{V}} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}},\ (x, \boldsymbol{v}) \mapsto (\tfrac{1}{\gamma_{_A}} x - \textstyle\sum_i \omega_i L_i^* v_i, \tfrac{1}{\gamma_{_{C_i}}} v_1 - L_1 x, \cdots, \tfrac{1}{\gamma_{_{C_m}}} v_m - L_m x). \end{aligned}$$

It can be shown that $\boldsymbol{C}$ and $\boldsymbol{D}$ are maximal monotone, $\boldsymbol{E}$ is β-cocoercive, and $\boldsymbol{\mathcal{V}}$ is self-adjoint and η-strongly positive. Then the fixed point equation of (5) is [173]

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \lambda_k\big(\jmath_{\boldsymbol{A}}(\boldsymbol{z}_k - \boldsymbol{B}(\boldsymbol{z}_k) - \boldsymbol{\varepsilon}_{2,k}) + \boldsymbol{\varepsilon}_{1,k} - \boldsymbol{z}_k\big), \tag{3.4.18}$$

where $\boldsymbol{A} = \boldsymbol{\mathcal{V}}^{-1}(\boldsymbol{C}+\boldsymbol{D})$, $\boldsymbol{B} = \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{E}$, $\boldsymbol{z}_k = (x_k, v_{1,k}, \cdots, v_{m,k})$, and the errors are

$$\boldsymbol{\varepsilon}_{1,k} = (\varepsilon_{2,k}, \varepsilon_{4,1,k}, \cdots, \varepsilon_{4,m,k}),\ \ \boldsymbol{\varepsilon}_{3,k} = (\varepsilon_{1,k}, \varepsilon_{3,1,k}, \cdots, \varepsilon_{3,n,k}),$$
$$\boldsymbol{\varepsilon}_{4,k} = (\tfrac{1}{\gamma_A}\varepsilon_{2,k}, \tfrac{1}{\gamma_{C_1}}\varepsilon_{4,1,k}, \cdots, \tfrac{1}{\gamma_{C_m}}\varepsilon_{4,m,k}),\ \ \boldsymbol{\varepsilon}_{2,k} = \boldsymbol{\mathcal{V}}^{-1}\big((\boldsymbol{D}+\boldsymbol{E})(\boldsymbol{\varepsilon}_{1,k}) + \boldsymbol{\varepsilon}_{3,k} - \boldsymbol{\varepsilon}_{4,k}\big).$$

Iteration (3.4.18) has the structure of FB, and by Remark 2.3.12, the fixed-point operator

$$\boldsymbol{\mathcal{F}}_{\text{PD}} = \jmath_{\boldsymbol{A}}(\mathbf{Id} - \boldsymbol{B}) \in \mathcal{A}(\tfrac{2\beta}{4\beta-1}).$$

Thus, (3.4.18) is a special instance of (3.1.1). In addition, since $\varepsilon_{1,k}$ and $\varepsilon_{2,k}$ (resp. $\varepsilon_{3,i,k}$ and $\varepsilon_{4,i,k}$) are summable in $\mathcal{H}$ (resp. in $\mathcal{G}_i$), so are $\boldsymbol{\varepsilon}_{1,k}$ and $\boldsymbol{\varepsilon}_{2,k}$ in $\boldsymbol{\mathcal{K}}$. Therefore, the iterates (3.4.17) obey the convergence rates in Section 3.2.

Observing that $\operatorname{fix}(\boldsymbol{\mathcal{F}}_{\text{PD}}) = \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{zer}(\boldsymbol{C}+\boldsymbol{D}+\boldsymbol{E})$, we also have the following bounds. We will denote $\delta = \max\{\frac{1}{\gamma_A}, \frac{1}{\gamma_{C_1}}, \ldots, \frac{1}{\gamma_{C_m}}\}$, $\underline{\tau} = \inf_{k\in\mathbb{N}} \lambda_k(\frac{4\beta-1}{2\beta} - \lambda_k)$, and

$$\boldsymbol{g}_{k+1} = \tfrac{1}{\lambda_k}(\boldsymbol{z}_{k+1} - (1-\lambda_k)\boldsymbol{z}_k) - \boldsymbol{\varepsilon}_k,$$

where $\boldsymbol{\varepsilon}_k = (\jmath_{\boldsymbol{A}}(\boldsymbol{z}_k - \boldsymbol{B}(\boldsymbol{z}_k) - \boldsymbol{\varepsilon}_{2,k}) + \boldsymbol{\varepsilon}_{1,k}) - \jmath_{\boldsymbol{A}}(\boldsymbol{z}_k - \boldsymbol{B}(\boldsymbol{z}_k))$.

Proposition 3.4.9. *Suppose* $0 < \inf_{k\in\mathbb{N}} \lambda_k \le \sup_{k\in\mathbb{N}} \lambda_k < \frac{2\beta}{4\beta-1}$, $\{(k+1)\|\varepsilon_{j,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, $j = 1, 2$, *and* $\{(k+1)\|\varepsilon_{j,i,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, $j = 3, 4$. *Then*

$$d\big(0, (\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{g}_{k+1})\big) \le \frac{2\delta}{\eta}\sqrt{\frac{d_0^2 + \mathcal{C}_1}{\underline{\tau}(k+1)}},$$

where $\mathcal{C}_1 < +\infty$ *and* $d_0 = \inf_{\boldsymbol{z}\in\operatorname{fix}(\boldsymbol{\mathcal{F}}_{\text{PD}})} \|\boldsymbol{z}_0 - \boldsymbol{z}\|_{\boldsymbol{\mathcal{K}}}$.

The proof is similar to that of Proposition 3.4.3 and will not be included here, moreover it can be improved to $o(1/\sqrt{k})$. Again, Proposition 3.4.9 can serve as a stopping criterion for the primal-dual monotone inclusion (3.4.15)-(3.4.16). An ergodic bound can also be derived. However, for the sake of brevity, we choose to omit the presentation of the result.

3.5 The non-stationary case

The fixed-point iteration discussed in Section 3.2 is stationary, namely, the fixed-point operator $\mathcal{F}$ of (3.1.1) is fixed along the iterations. In this section, we study the non-stationary version of it, and show that, the non-stationary case can be seen as a perturbation of the stationary one. Moreover the iterates are convergent if the extra perturbation error is absolutely summable.

3.5.1 General convergence analysis

The definition of non-stationary Krasnosel'skiĭ-Mann iteration is given below.

Definition 3.5.1 (Non-stationary Krasnosel'skiĭ-Mann iteration). Let $\mathcal{F}_\Gamma : \mathcal{H} \to \mathcal{H}$ be a non-expansive operator depending on a parameter Γ. Let $\lambda_k \in]0, 1]$. Then the non-stationary Krasnosel'skiĭ-Mann iteration is defined by

$$z_{k+1} = z_k + \lambda_k\big(\mathcal{F}_{\Gamma_k}(z_k) + \varepsilon_k - z_k\big) = \mathcal{F}_{\Gamma_k,\lambda_k}(z_k) + \lambda_k\varepsilon_k, \tag{3.5.1}$$

with $\mathcal{F}_{\Gamma_k,\lambda_k} = \lambda_k\mathcal{F}_{\Gamma_k} + (1-\lambda_k)\mathrm{Id}$. If we define $\varepsilon_{\Gamma_k} = (\mathcal{F}_{\Gamma_k} - \mathcal{F}_\Gamma)(z_k)$, $\pi_k = \varepsilon_{\Gamma_k} + \varepsilon_k$, then (3.5.1) can be rewritten as

$$z_{k+1} = \big(\lambda_k\mathcal{F}_\Gamma + (1-\lambda_k)\mathrm{Id}\big)(z_k) + \lambda_k\pi_k = \mathcal{F}_{\Gamma,\lambda_k} z_k + \lambda_k\pi_k, \tag{3.5.2}$$

where $\mathcal{F}_{\Gamma,\lambda_k} = \lambda_k \mathcal{F}_\Gamma + (1-\lambda_k)\mathrm{Id}$. The corresponding residual e_k of (3.5.1) becomes

$$e_k = \frac{z_k - z_{k+1}}{\lambda_k} + \pi_k.$$

Comparing (3.5.2) to Definition 3.1.1, a new error sequence π_k is introduced. To obtain convergence of the non-stationary iteration, we adapt arguments from [104] (Banach spaces endowed with an appropriate compatible topology) and [105] (real Hilbert spaces). For convenience, we recall that $\underline{\tau} = \inf_{k\in\mathbb{N}} \lambda_k(1-\lambda_k)$.

Theorem 3.5.2 (Convergence of (3.5.1)). *Assume the following holds:*

(A.1) $\mathrm{fix}(\mathcal{F}_\Gamma) \neq \emptyset$.

(A.2) $\forall k \in \mathbb{N}$, $\mathcal{F}_{\Gamma_k,\lambda_k}$ *is* $(1+\beta_k)$*-Lipschitz with* $\beta_k \geq 0$, *and* $\{\beta_k\}_{k\in\mathbb{N}} \in \ell^1_+$.

(A.3) $\lambda_k \in]0,1[$ *such that* $\underline{\tau} > 0$.

(A.4) $\{\lambda_k \|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$.

(A.5) $\forall \rho \in [0,+\infty[$, *the sequence* $\{\lambda_k \Delta_{k,\rho}\}_{k\in\mathbb{N}} \in \ell^1_+$, *where*

$$\Delta_{k,\rho} = \sup_{\|z\|\leq\rho} \|\mathcal{F}_{\Gamma_k}(z) - \mathcal{F}_\Gamma(z)\|. \tag{3.5.3}$$

Then $\{e_k\}_{k\in\mathbb{N}}$ *converges strongly to* 0, *and* $\{z_k\}_{k\in\mathbb{N}}$ *converges weakly to a point* $z^\star \in \mathrm{fix}(\mathcal{F}_\Gamma)$.

Proof. For the sequence $\{z_k\}_{k\in\mathbb{N}}$ generated by (3.5.1), and $z^\star \in \mathrm{fix}(\mathcal{F}_\Gamma)$, we have

$$\begin{aligned}
\|z_{k+1} - z^\star\| &= \|\mathcal{F}_{\Gamma_k,\lambda_k}(z_k) + \lambda_k\varepsilon_k - \mathcal{F}_{\Gamma,\lambda_k}(z^\star)\| \\
&\leq \|\mathcal{F}_{\Gamma_k,\lambda_k}(z_k) - \mathcal{F}_{\Gamma_k,\lambda_k}(z^\star)\| + \|\mathcal{F}_{\Gamma_k,\lambda_k}(z^\star) - \mathcal{F}_{\Gamma,\lambda_k}(z^\star)\| + \lambda_k\|\varepsilon_k\| \\
&\leq (1+\beta_k)\|z_k - z^\star\| + \lambda_k \Delta_{k,\|z^\star\|} + \lambda_k\|\varepsilon_k\|.
\end{aligned}$$

As $\{\beta_k\}_{k\in\mathbb{N}}$, $\{\lambda_k\|\varepsilon_k\|\}_{k\in\mathbb{N}}$ and $\{\lambda_k\Delta_{k,\|z^\star\|}\}_{k\in\mathbb{N}}$ are summable by assumptions (A.2), (A.4) and (A.5), it follows from [147, Lemma 2.2.2] that the sequence $\{\|z_k - z^\star\|\}_{k\in\mathbb{N}}$ converges, hence bounded. Therefore, z_k is bounded in norm by some $\rho \in [0,+\infty[$. This implies that

$$\|z_{k+1} - \mathcal{F}_{\Gamma,\lambda_k}(z_k)\| = \|\mathcal{F}_{\Gamma_k,\lambda_k}(z_k) + \lambda_k\varepsilon_k - \mathcal{F}_{\Gamma,\lambda_k}(z_k)\| \leq \lambda_k(\Delta_{k,\rho} + \|\varepsilon_k\|).$$

In other words, the (inexact) non-stationary iteration (3.5.1) can be seen as a perturbed version of the (inexact) stationary one with an extra-error term which is summable owing to (A.5). The rest of the proof follows by applying [104, Proposition 2.1 and Remark 2.2] (see also [39, Remark 14]) using (A.1) and (A.3). □

If $\mathcal{F}_{\Gamma_k}$ were non-expansive, then obviously (A.2) is in force, and in turn, Theorem 3.5.2 holds. In the specific scenario where $\forall k \in \mathbb{N}$, $\mathcal{F}_{\Gamma_k}$ is α_k-averaged, $\alpha_k \in]0,1]$, we can further refine the choice of λ_k. Here we take a different route from the one in [57]. By definition, $\forall k \in \mathbb{N}$, there exists a non-expansive operator $\mathcal{F}_{\Gamma_k} : \mathcal{H} \to \mathcal{H}$ such that $\mathcal{F}_{\Gamma_k} = \alpha_k \mathcal{F}_{\Gamma_k} + (1-\alpha_k)\mathrm{Id}$. Let $\mathcal{F}_\Gamma$ be a non-expansive operator, and $\lambda'_k = \alpha_k\lambda_k$. We have the following corollary.

Corollary 3.5.3. *Assume that*

(A′.1) $\mathrm{fix}(\mathcal{F}_\Gamma) \neq \emptyset$.

(A′.2) $\forall k \in \mathbb{N}$, $\mathcal{F}_{\Gamma_k}$ *is* α_k*-averaged,* $\alpha_k \in]0,1]$.

(A′.3) $\lambda_k \in]0, \frac{1}{\alpha_k}]$ *such that* $\inf_{k\in\mathbb{N}} \lambda'_k(1-\lambda'_k) > 0$.

(A′.4) $\{\lambda' e_k\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$.

(A′.5) $\forall \rho \in [0,+\infty[$, $\{\lambda'_k\Delta_{k,\rho}\}_{k\in\mathbb{N}} \in \ell^1_+$, *where* $\Delta_{k,\rho} = \sup_{\|z\|\leq\rho} \|\mathcal{F}_{\Gamma_k}(z) - \mathcal{F}_\Gamma(z)\|$.

Then $\{e_k\}_{k\in\mathbb{N}}$ *converges strongly to* 0, *and* $\{z_k\}_{k\in\mathbb{N}}$ *converges weakly to a point* $z^\star \in \mathrm{fix}(\mathcal{F}_\Gamma)$.

Assumptions (A.4)-(A.5) of Theorem 3.5.2 imply that

$$\sum_{k\in\mathbb{N}}\lambda_k\|\pi_k\| \leq \sum_{k\in\mathbb{N}}\lambda_k(\|\varepsilon_{\Gamma_k}\| + \|\varepsilon_k\|) < +\infty.$$

Therefore, if we can further impose a stronger summability assumption on $\{\pi_k\}_{k\in\mathbb{N}}$ as in Theorem 3.2.1, then we can obtain the iteration-complexity bounds for the non-stationary iteration (3.5.1) as well. This is stated in the following result. Recall $\Lambda_k = \sum_{j=0}^{k}\lambda_j$, $\bar{e}_k = \frac{1}{\Lambda_k}\sum_{j=0}^{k}\lambda_j e_j$, and $d_0 = \inf_{z\in\mathrm{fix}(\mathcal{F}_\Gamma)}\|z_0 - z\|$.

Proposition 3.5.4. *Assume that (A.1) holds, and that $\forall k\in\mathbb{N}, \mathcal{F}_{\Gamma_k}$ is non-expansive.*

(i) *Suppose that (A.3) is verified, $\{(k+1)\|\varepsilon_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$ and $\{(k+1)\Delta_{k,\rho}\}_{k\in\mathbb{N}}\ell^1_+$, $\forall\rho\in[0,+\infty[$, where $\Delta_{k,\rho}$ is given by (3.5.3). Then,*

$$\|e_k\| \leq \sqrt{\frac{d_0^2 + \mathcal{C}_1}{\underline{\tau}(k+1)}},$$

where $\mathcal{C}_1$ is a bounded constant (see Theorem 3.2.1).

(ii) *Suppose that $\lambda_k \in]0,1]$ such that (A.4)-(A.5) are verified. Then*

$$\|\bar{e}_k\| \leq \frac{2(d_0 + \mathcal{C}_2)}{\Lambda_k}$$

where $\mathcal{C}_2$ is a bounded constant (see Theorem 3.2.4). In particular, if $\inf_{k\in\mathbb{N}}\lambda_k > 0$, we get $O(1/k)$ ergodic convergence rate.

Proof.

(i) By assumption, we have $\sum_{k\in\mathbb{N}}(k+1)\|\pi_k\| < +\infty$. All assumptions of Theorem 3.2.1 are then fulfilled and the result follows.

(ii) Similarly all required assumptions to apply Theorem 3.2.4 are in force. □

Remark 3.5.5.

(i) The convergence rate of $\|e_k\|$ can be improved to $o(1/\sqrt{k})$ as pointed out in (i) of Remark 3.2.2.

(ii) If metric sub-regularity assumption is imposed on $\mathrm{Id} - \mathcal{F}_\Gamma$, a result similar to Theorem 3.3.3 can be stated. But now, we have $c_k = \nu_1\lambda_k(\|\varepsilon_k\| + \Delta_{k,\rho})$. So the actual local convergence behaviour depends also on the additional perturbation error brought by non-stationarity as captured by $\Delta_{k,\rho}$, even in the exact case. Thus, similarly to Remark 3.3.4, if $\|\varepsilon_k\|$ and $\Delta_{k,\rho}$ converge linearly, then so does $d_k = d(z_k, \mathrm{fix}(\mathcal{F}_\Gamma))$ locally. If the non-stationary error $\Delta_{k,\rho}$ decays sub-linearly, then it dominates.

3.5.2 Application to the non-stationary GFB

As discussed in Section 3.4.1, the fixed-point operator $\boldsymbol{\mathcal{F}}_{\mathrm{GFB}}$ of GFB depends on a parameter γ, for the sake of appearance, we denote $\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}} \overset{\text{def}}{=} \boldsymbol{\mathcal{F}}_{\mathrm{GFB}}$ in this section. Now let γ vary along the iteration, and denote the following operators

$$\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}_k} = \tfrac{1}{2}(\mathcal{R}_{\boldsymbol{\gamma}_k\boldsymbol{A}}\mathcal{P}_{\boldsymbol{\mathcal{S}}} + \mathbf{Id}),\quad \boldsymbol{\mathcal{F}}_{2,\boldsymbol{\gamma}_k} = \mathbf{Id} - \boldsymbol{\gamma}_k\boldsymbol{B}_{\boldsymbol{\mathcal{S}}} \quad\text{and}\quad \boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k} = \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}_k}\boldsymbol{\mathcal{F}}_{2,\boldsymbol{\gamma}_k}.$$

Recall that for $\{\gamma_k\}_{k\in\mathbb{N}} \in]0, 2\beta[$, $\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k}$ is α_k-averaged with $\alpha_k = \frac{2\beta}{4\beta-\gamma_k}$, and $\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}}$ is α-averaged with $\alpha = \frac{2\beta}{4\beta-\gamma}$, for $\gamma \in]0, 2\beta[$. Moreover, there exists non-expansive operators $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ and $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}_k}$ such that $\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}} = \alpha\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}} + (1-\alpha)\mathbf{Id}$ and $\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k} = \alpha_k\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}_k} + (1-\alpha_k)\mathbf{Id}$.

Let $\boldsymbol{\varepsilon}_{\boldsymbol{\gamma}_k} = (\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k} - \boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}})(\boldsymbol{z}_k)$, $\boldsymbol{\pi}_k = \boldsymbol{\varepsilon}_{\boldsymbol{\gamma}_k} + \boldsymbol{\varepsilon}_k$. The the non-stationary version of (3.4.4) is defined by

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \lambda_k(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}}(\boldsymbol{z}_k) + \boldsymbol{\pi}_k - \boldsymbol{z}_k). \tag{3.5.4}$$

Theorem 3.5.6. *For the non-stationary iteration (3.5.4), if the following assumptions hold*

(A″.1) $\mathrm{zer}(B + \sum_i A_i) \neq \emptyset$.

(A″.2) $\lambda_k \in]0, \frac{1}{\alpha_k}[$, *such that* $\inf_{k\in\mathbb{N}} \lambda_k(\frac{1}{\alpha_k} - \lambda_k) > 0$.

(A″.3) $\{\lambda_k \|b_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$ *and* $\{\lambda_k \|a_{i,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, $\forall i \in 1n$.

(A″.4) $\gamma_k, \gamma \in]0, 2\beta[$ *such that* $0 < \underline{\gamma} \leq \gamma_k \leq \bar{\gamma} < 2\beta$, *and* $\{\lambda_k|\gamma_k - \gamma|\}_{k\in\mathbb{N}} \in \ell^1_+$.

then, the sequence $\{\boldsymbol{z}_k\}_{k\in\mathbb{N}}$ *generated by* (3.5.4) *converges weakly to a point in* $\operatorname{fix}(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}})$.

If we further assume that $\{(k+1)|\gamma_k - \gamma|\}_{k\in\mathbb{N}} \in \ell^1_+$, $\{(k+1)\|b_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$, *and* $\{(k+1)\|a_{i,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$, $\forall i \in 1n$, *then we obtain the pointwise iteration-complexity bound for the non-stationary version of GFB algorithm as stated in (i) of Proposition 3.5.4.*

Proof. It is sufficient to verify the conditions of Corollary 3.5.3 to conclude.

- Assumption (A′.1) is fulfilled thanks to (A″.1) since $\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\operatorname{fix}(\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}})) = \mathcal{P}_{\boldsymbol{\mathcal{S}}}(\operatorname{fix}(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}})) = \operatorname{zer}(B + \sum_i A_i) \neq \emptyset$.
- As $\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k} \in \mathcal{A}(\alpha_k)$, $\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k,\lambda_k} \in \mathcal{A}(\alpha_k\lambda_k)$, and thus assumption (A′.2) is in force.
- (A′.3) holds thanks to (A″.2).
- (A′.4) follows from (A″.3).
- It remains to check that (A″.4) implies (A′.5).

By definition of $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}$ and $\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}_k}$, we have

$$\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}} = \big(1 - \tfrac{1}{\alpha}\big)\mathrm{Id} + \tfrac{1}{\alpha}\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}} \quad \text{and} \quad \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}_k} = \big(1 - \tfrac{1}{\alpha_k}\big)\mathrm{Id} + \tfrac{1}{\alpha_k}\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k}.$$

It then follows that

$$\begin{aligned}
\|\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}_k}(\boldsymbol{z}) - \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}(\boldsymbol{z})\| &= \|\tfrac{1}{\alpha_k}(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k} - \mathrm{Id})(\boldsymbol{z}) - \tfrac{1}{\alpha}(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}} - \mathrm{Id})(\boldsymbol{z})\| \\
&\leq |\tfrac{1}{\alpha_k} - \tfrac{1}{\alpha}|\|(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}} - \mathrm{Id})(\boldsymbol{z})\| + \frac{1}{\alpha_k}\|(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k} - \mathrm{Id})(\boldsymbol{z}) - \big(\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}} - \mathrm{Id}\big)(\boldsymbol{z})\| \\
&\leq \frac{|\gamma_k - \gamma|}{2\beta}(2\rho + \|\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}}(0)\|) + \frac{1}{\alpha_k}\|\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}}(\boldsymbol{z})\|.
\end{aligned} \tag{3.5.5}$$

Now, non-expansiveness of $\boldsymbol{\mathcal{F}}_{1,\gamma_k}$ yields

$$\begin{aligned}
\|\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}_k}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}}(\boldsymbol{z})\| &\leq \|\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}_k}\boldsymbol{\mathcal{F}}_{2,\gamma_k}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}_k}\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z})\| + \|\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}_k}\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z})\| \\
&\leq \underbrace{\|\boldsymbol{\mathcal{F}}_{2,\gamma_k}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z})\|}_{\text{Term 1}} + \underbrace{\|\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}_k}\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z})\|}_{\text{Term 2}}.
\end{aligned} \tag{3.5.6}$$

We first bound the first term in (3.5.6),

$$\begin{aligned}
\|\boldsymbol{\mathcal{F}}_{2,\gamma_k}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z})\| &\leq |\gamma_k - \gamma|\|\boldsymbol{B}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{z})\| \\
\text{(Triangle inequality and } B \text{ is } \beta^{-1}\text{-Lipschitz)} &\leq (\beta^{-1}\rho + \|B(0)\|)|\gamma_k - \gamma|,
\end{aligned} \tag{3.5.7}$$

where $B(0)$ is bounded. Let's now turn to the second term of (3.5.6). Denote $\boldsymbol{z}_{\boldsymbol{\mathcal{S}}} = \mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{z})$ and $\boldsymbol{z}_{\boldsymbol{\mathcal{S}}^\perp} = \boldsymbol{z} - \boldsymbol{z}_{\boldsymbol{\mathcal{S}}}$, then

$$\boldsymbol{v} = \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}\boldsymbol{\mathcal{F}}_{2,\gamma'}(\boldsymbol{z}) \Leftrightarrow \boldsymbol{v} = \boldsymbol{z}_{\boldsymbol{\mathcal{S}}^\perp} + \boldsymbol{\jmath}_{\boldsymbol{\gamma A}}(\boldsymbol{z}_{\boldsymbol{\mathcal{S}}} - \boldsymbol{z}_{\boldsymbol{\mathcal{S}}^\perp} - \gamma'\boldsymbol{B}(\boldsymbol{z}_{\boldsymbol{\mathcal{S}}})),$$

Let $\boldsymbol{y} = \boldsymbol{z}_{\boldsymbol{\mathcal{S}}} - \boldsymbol{z}_{\boldsymbol{\mathcal{S}}^\perp} - \gamma\boldsymbol{B}(\boldsymbol{z}_{\boldsymbol{\mathcal{S}}})$, then we have

$$\boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}_k}\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}) - \boldsymbol{\mathcal{F}}_{1,\boldsymbol{\gamma}}\boldsymbol{\mathcal{F}}_{2,\gamma}(\boldsymbol{z}) = \boldsymbol{\jmath}_{\boldsymbol{\gamma}_k\boldsymbol{A}}(\boldsymbol{y}) - \boldsymbol{\jmath}_{\boldsymbol{\gamma A}}(\boldsymbol{y}).$$

Denote $\boldsymbol{u}_k = \boldsymbol{\jmath}_{\boldsymbol{\gamma}_k\boldsymbol{A}}\boldsymbol{y}$, $\boldsymbol{u} = \boldsymbol{\jmath}_{\boldsymbol{\gamma A}}\boldsymbol{y}$. By definition of the resolvent, this is equivalent to

$$\Big(\boldsymbol{u}, \frac{\boldsymbol{y} - \boldsymbol{u}}{\gamma}\Big) \in \operatorname{gra}(\boldsymbol{A}) \quad \text{and} \quad \Big(\boldsymbol{u}_k, \frac{\boldsymbol{y} - \boldsymbol{u}_k}{\gamma_k}\Big) \in \operatorname{gra}(\boldsymbol{A}).$$

Since $\boldsymbol{A}$ is monotone, and by assumptions on γ_k and γ, it follows that

$$\langle \tfrac{\boldsymbol{y} - \boldsymbol{u}}{\gamma_k} - \tfrac{\boldsymbol{y} - \boldsymbol{u}_k}{\gamma}, \boldsymbol{u} - \boldsymbol{u}_k\rangle \geq 0 \Longleftrightarrow \|\boldsymbol{u} - \boldsymbol{u}_k\|^2 \leq \tfrac{\gamma_k - \gamma}{\gamma}\langle \boldsymbol{y} - \boldsymbol{u}, \boldsymbol{u} - \boldsymbol{u}_k\rangle.$$

Therefore, using Cauchy-Schwartz inequality and the fact that $\mathbf{Id} - \boldsymbol{\jmath}_{\boldsymbol{\gamma A}} \in \mathcal{A}(\frac{1}{2})$ is firmly non-expansive,

$$\|\boldsymbol{u} - \boldsymbol{u}_k\| \leq \frac{|\gamma_k - \gamma|}{\underline{\gamma}}\|(\mathbf{Id} - \boldsymbol{\jmath}_{\boldsymbol{\gamma A}})\boldsymbol{y}\| \leq \frac{|\gamma_k - \gamma|}{\underline{\gamma}}(\|\boldsymbol{y}\| + \|\boldsymbol{\jmath}_{\boldsymbol{\gamma A}}(\mathbf{0})\|), \tag{3.5.8}$$

where $\|\mathcal{J}_{\boldsymbol{\gamma A}}(0)\|^2 = \sum_i \omega_i \|\mathcal{J}_{\frac{\gamma}{\omega_i}}(0)\|^2$. Using the triangle inequality, the Pythagorean theorem and non-expansiveness of $\beta \boldsymbol{B}_{\boldsymbol{\mathcal{S}}}$, we obtain

$$\begin{aligned} \|\boldsymbol{y}\| &\leq \|\boldsymbol{z}_{\boldsymbol{\mathcal{S}}} - \boldsymbol{z}_{\boldsymbol{\mathcal{S}}^\perp}\| + \gamma \|\boldsymbol{B}(\boldsymbol{z}_{\boldsymbol{\mathcal{S}}})\| \leq \rho + \gamma \|\boldsymbol{B}(\boldsymbol{z}_{\boldsymbol{\mathcal{S}}}) - \boldsymbol{B}_{\boldsymbol{\mathcal{S}}}(0)\| + \gamma \|\boldsymbol{B}_{\boldsymbol{\mathcal{S}}}(0)\| \\ &\leq \rho + \gamma\beta^{-1}\|\boldsymbol{z}\| + \gamma \|B(0)\| \leq \rho + \bar{\gamma}\beta^{-1}\rho + \bar{\gamma}\|B(0)\|. \end{aligned} \tag{3.5.9}$$

Putting together (3.5.5), (3.5.7), (3.5.8) and (3.5.9), we get $\forall \rho \in [0, +\infty[$

$$\begin{aligned} \textstyle\sum_{k\in\mathbb{N}} \lambda_k \alpha_k \Delta_{k,\rho} &= \textstyle\sum_{k\in\mathbb{N}} \lambda_k \alpha_k \sup_{\|\boldsymbol{z}\|\leq\rho} \|\boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}_k}(\boldsymbol{z}) - \boldsymbol{\mathcal{T}}_{\boldsymbol{\gamma}}(\boldsymbol{z})\| \\ &\leq \mathcal{C} \textstyle\sum_{k\in\mathbb{N}} \lambda_k |\gamma_k - \gamma| < +\infty, \end{aligned}$$

where

$$\mathcal{C} = \frac{2\rho + \|\boldsymbol{\mathcal{F}}_{\boldsymbol{\gamma}}(0)\|}{4\beta - \bar{\gamma}} + \frac{\rho}{\beta}\Big(1 + \frac{\beta}{\underline{\gamma}} + \frac{\bar{\gamma}}{\underline{\gamma}}\Big) + \Big(1 + \frac{\bar{\gamma}}{\underline{\gamma}}\Big)\|B(0)\| + \frac{1}{\underline{\gamma}}\|\mathcal{J}_{\boldsymbol{\gamma A}}(\boldsymbol{0})\| < +\infty.$$

Consequently, (A.5) is fulfilled.

The last statement of the theorem is a simple application of Theorem 3.2.1. □

Remark 3.5.7.

(i) For the non-stationary versions of the methods discussed in Section 3.4, for instance, the DR and FDR methods, whose fixed-point operators also depend on γ_k, Theorem 3.5.6 is also applicable. In general, the summability assumption (A″.4) of Theorem 3.5.6 is hard to remove, except for some special cases. For instance for the FBS method, as stated in [63, Theorem 3.4], where $\gamma_k \in]0, 2\beta[$, but $\lambda_k \in]0, 1[$ instead of $]0, \frac{4\beta-\gamma_k}{2\beta}[$ here.

(ii) Recall the term c_k in the bound (3.3.3) of Theorem 3.3.3. Clearly, for the non-stationary GFB iteration, even if the approximation error $\boldsymbol{\varepsilon}_k = 0$, we still have $c_k = \nu_1 \lambda_k \boldsymbol{\varepsilon}_{\gamma_k} \neq 0$. Therefore, under metric sub-regularity of $\mathbf{Id} - \boldsymbol{\mathcal{F}}_\gamma$, a bound similar to (3.3.3) can be obtained, whose performance will depend on how fast $|\gamma_k - \gamma|$ converges to 0.

(iii) In Chapter 7, the non-stationary Douglas–Rachford splitting is considered (Algorithm 13), to ensure the global convergence, summability condition $\{\lambda_k|\gamma_k - \gamma|\}_{k\in\mathbb{N}} \in \ell^1_+$ similar to the non-stationary GFB is needed.

3.6 Numerical experiments

In order to demonstrate the established convergence rates, in this section we consider 3 numerical examples:

(i) anisotropic total variation (TV) based image deconvolution with intensity constraint.

(ii) matrix completion with non-negativity constraints (NMC).

(iii) principal component pursuit problem (PCP) with application to video background and foreground decomposition.

All the problems are solved by both GFB and the Primal–Dual splitting (3.4.17) which is denoted as "PD". It should be noted that, since we do not have errors for the presented experiments, the pointwise convergence rates obtained are actually $o(1/\sqrt{k})$.

Linear inverse problems Suppose we have an object $x_{\text{ob}} \in \mathbb{R}^n$, instead of accessing it directly, we can obtain only the observation indirectly through some bounded linear operator $\mathcal{K}$, and possibly contaminated by noise,

$$f = \mathcal{K}x_{\text{ob}} + w, \tag{3.6.1}$$

where $\mathcal{K} : \mathbb{R}^n \to \mathbb{R}^m$ is the bounded linear operator, $w \in \mathbb{R}^m$ stands for noise, and $f \in \mathbb{R}^m$ is the observation. Typical examples of (3.6.1) include the compressed sensing problem where $\mathcal{K}$ is Gaussian

random matrix [46], and the image deblurring problem where $\mathcal{K}$ is the point spread function as we have seen in Section 3.6. Then common types of noise w consists of sparse noise, uniform noise and noise with bounded ℓ_2-norm.

Recovering x_{ob} from f in general is a difficult problem, since $\mathcal{K}$ usually is either under-determined or square yet badly conditioned, or the noise w needs to be modeled properly. The situation changes if certain prior information of x_{ob} is available (*e.g.* sparsity), and a popular approach to recover/approximate x_{ob} is through the following regularization model

$$\min_{x\in\mathbb{R}^n} F(x) + \sum\nolimits_{i=1}^{m} R_i(x), \tag{3.6.2}$$

where $m \in \mathbb{N}_+$ is an integer,

- for each $i = 1, ..., m$, R_i is the regularization term based on the prior information(s), such as (group) sparsity, piecewise constant, and low-rank, *etc.*.
- F is the data fidelity term depends on the noise model, for instance when w has bounded ℓ_2-norm, then F takes the form $F = \frac{1}{2}\|\mathcal{K}x - f\|^2$.

A wide range of problems, arising from fields like inverse problems, signal/image processing, and machine learning *etc.*, can be covered by model (3.6.2).

Anisotropic TV deconvolution For this example, let $x_{\text{ob}} \in \mathbb{R}^{n\times n}$ is an image, $\mathcal{K} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be the linear operator corresponding to a point spread function (*i.e.* blur kernel), and w be an additive white Gaussian noise. Then $f \in \mathbb{R}^{n\times n}$ is the noised blur observation of x_{ob}. The deconvolution procedure is to provably recover or approximate x_{ob} from f, here we consider the anisotropic TV [154] based deconvolution model which is

$$\min_{x\in\mathbb{R}^{n\times n}} \frac{1}{2}\|\mathcal{K}x - f\|^2 + \mu\|\mathrm{D}_{\text{DIF}}(x)\|_1 + \iota_\Omega(x), \tag{3.6.3}$$

where $\mu > 0$ is the regularization parameter determined based on the noise level, D_{DIF} denotes a finite difference approximation of the derivative [154], and $\iota_\Omega(\cdot)$ is the indicator function of box constraint, for instance $\Omega = [0, 255]^{n^2}$ if x is gray scale image. The problem can be solved by both GFB and PD methods, where the proximity operator of $\iota_\Omega(\cdot)$ is the projection onto set Ω, and for GFB, the proximity operator of $\|\mathrm{D}_{\text{DIF}}(\cdot)\|_1$ is computed by graph-cut [48, 97].

Figure 3.3 displays the observed pointwise and ergodic rates of $\|e_k\|$ and the theoretical bounds given by Theorem 3.2.1 and 3.2.4. Pointwise convergence rate is shown in subfigure (a) and (c), whose left half is log-log plot while the right half is semi-log plot. As predicted by Theorem 3.2.1, globally $\|e_k\|$ converges at the rate of $O(1/\sqrt{k})$. Then for a sufficiently large iteration number, a linear convergence regime takes over as clearly seen from the semi-log plot, which is in consistent with the result of Theorem 3.3.3. Let us mention that the local linear convergence curve (dashed line) is *fitted* to the observed one, since the regularity modulus necessary to compute the theoretical rate in Theorem 3.3.3 is not easy to estimate. For the ergodic convergence, subfigure (b) and (d) of Figure 3.3, $O(1/k)$ convergence rates are observed which coincides with Theorem 3.2.4.

Non-negative matrix completion For this example, suppose $x_{\text{ob}} \in \mathbb{R}^{n\times n}$ in (3.6.1) is a low rank matrix with non-negative entries, $\mathcal{K} : \mathbb{R}^{n\times n} \to \mathbb{R}^m$ is a measurement operator which selects m entries of its argument uniformly at random. The matrix completion problem consists in recovering x_{ob}, or finding an approximation of it, by solving a convex optimization problem, namely the minimization of the nuclear norm [45, 47, 150]. In penalised form, the problem reads

$$\min_{x\in\mathbb{R}^{n\times n}} \frac{1}{2}\|f - \mathcal{K}x\|^2 + \mu\|x\|_* + \iota_{P_+}(x), \tag{3.6.4}$$

where $\iota_{P_+}(\cdot)$ is the indicator function of the non-negative orthant accounting for the non-negativity constraint, and $\mu > 0$ is a regularization parameter typically chosen proportional to the noise level.

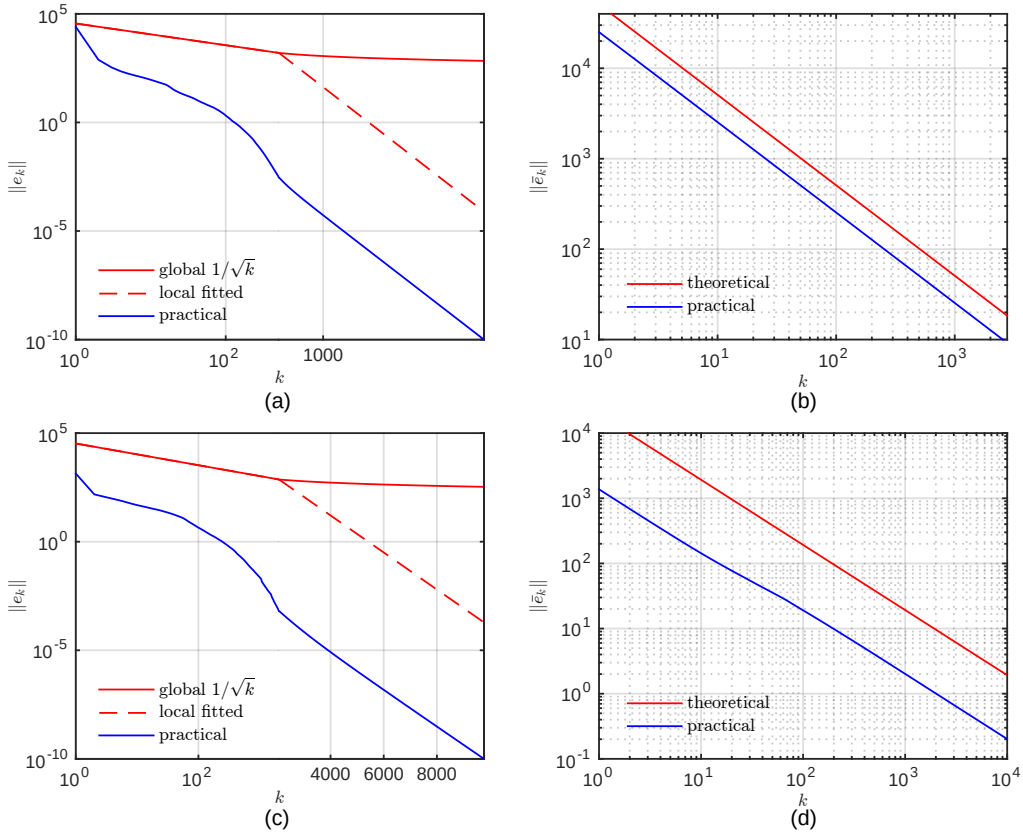


Figure 3.3: TV deconvolution of the cameraman image. (a) pointwise convergence of GFB, (b) ergodic convergence of GFB, (c) pointwise convergence of PD, (d) ergodic convergence of PD. Both methods achieve local linear convergence. Note that $(\|e_k\|)_{k\in\mathbb{N}}$ is non-increasing, which coincides with Lemma 3.1.7 when $\varepsilon_k = 0$.

The proximity operator of both $\|\cdot\|_*$, $\iota_{P_+}(\cdot)$ have explicit forms, as $\mathrm{prox}_{\|\cdot\|_*}(x)$ is soft–thresholding to the singular values of x and $\mathrm{prox}_{\iota_{P_+}}(x)$ is the projector on the non-negative orthant.

Figure 3.4 displays the observed pointwise and ergodic rates of $\|e_k\|$ and the theoretical bounds computed given by Theorem 3.2.1 and 3.2.4. Both global and local convergence behaviours are similar to those observed from Figure 3.3.

Principal component pursuit In this experiment, we consider the PCP problem [44]. Different from (3.6.1), the forward observation model of PCP problem reads,

$$f = x_{\mathrm{ob},L} + x_{\mathrm{ob},S} + w,$$

where $x_{\mathrm{ob},L}$ is low-rank, $x_{\mathrm{ob},S}$ is sparse, and f, w are the observation and noise respectively. The PCP proposed in [44] attempts to provably recover $(x_{\mathrm{ob},L}, x_{\mathrm{ob},S})$ to a good approximation, by solving a convex optimization. Here, we also add a non-negativity constraint to the low-rank component, which leads to the following convex problem

$$\min_{x_L, x_S \in \mathbb{R}^{n\times n}} \frac{1}{2}\|f - x_L - x_S\|^2 + \mu_1\|x_S\|_1 + \mu_2\|x_L\|_* + \iota_{P_+}(x_L). \tag{3.6.5}$$

Observe that for given an x_L, the minimizer of (3.6.5) is $x_S^\star = \mathrm{prox}_{\mu_1\|\cdot\|_1}(f - x_L)$. Thus, (3.6.5) is equivalent to

$$\min_{x_L \in \mathbb{R}^{n\times n}} {}^1\big(\mu_1\|\cdot\|_1\big)(f - x_L) + \mu_2\|x_L\|_* + \iota_{P_+}(x_L), \tag{3.6.6}$$

where ${}^1\big(\mu_1\|\cdot\|_1\big)(f - x_L)$ is the Moreau Envelope of $\mu_1\|\cdot\|_1$.

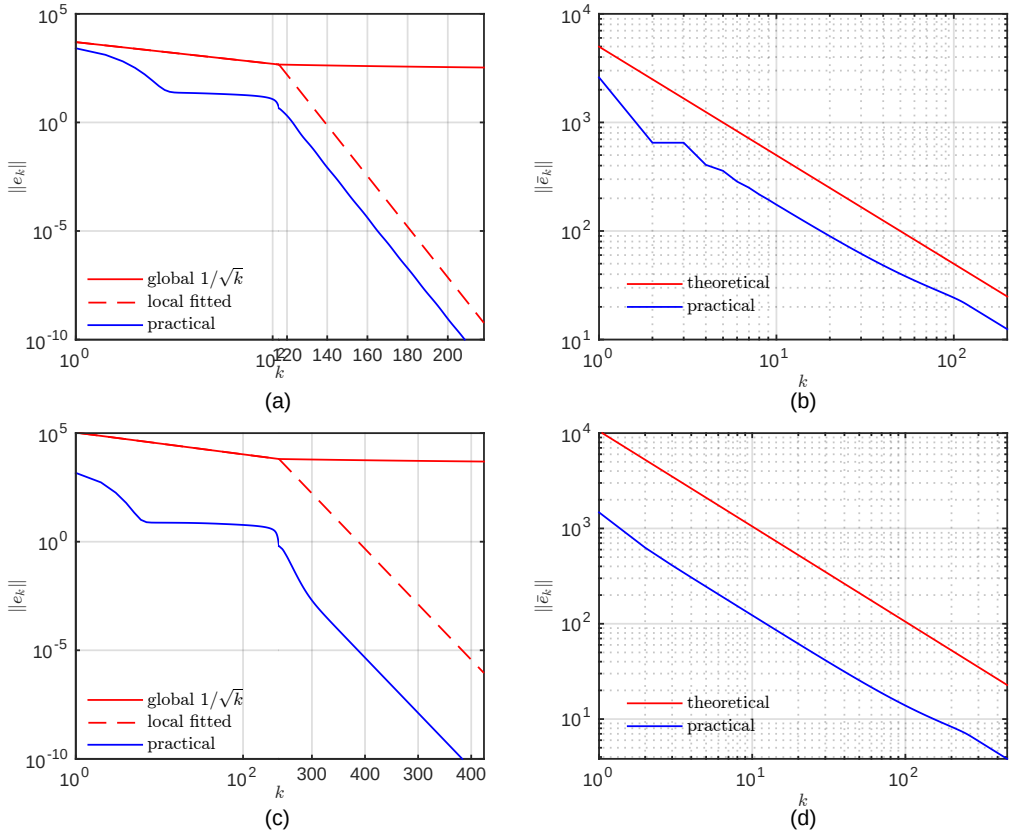


Figure 3.4: The size of the matrix x is 400×300, $\operatorname{rank}(x) = 20$ and the operator $\mathcal{K}$ is random projection mask. (a) pointwise convergence of GFB, (b) ergodic convergence of GFB, (c) pointwise convergence of PD, (d) ergodic convergence of PD.

A synthetic example is designed to demonstrate the comparison of the two methods, as shown in Figure 3.5. Pointwise convergence rate of $\|e_k\|$ is shown in subfigure (a) and (c). Then in subfigure (b) and (d), we display the convergence behaviour of the criteria provided in Proposition 3.4.9.

Non-stationary iteration Now we illustrate the non-stationary iteration of GFB applied to PCP problem. The above comparison indicates that in practice we can choose the relaxation parameter λ_k as $\lambda_1 = 1$ and $\lambda_k \approx \frac{1}{\alpha}$, $\forall k \geq 2$. Next, we compare this setting with the non-stationary GFB, for the stationary case, we let $\gamma = 1.5\beta, \lambda_k = \{1, \frac{1}{1.05\alpha}, \cdots\}$, for the non-stationary case, let $\gamma_0 = 1.5\beta$, then 3 scenarios of γ_k are considered,

$$\gamma_{1,k} = \gamma_0 + \frac{1.9\beta - \gamma_0}{1.1^k}, \ \gamma_{2,k} = \gamma_0 + \frac{1.9\beta - \gamma_0}{k^2}, \ \gamma_{3,k} = \gamma_0 + \frac{1.9\beta - \gamma_0}{k}.$$

The result is given in Figure 3.6. We can conclude from the observation that

- Globally, the non-stationary iterations are slower than the stationary one. Then locally, the local linear rates are also slower than the stationary one, moreover, for $\gamma_{2,k}$ and $\gamma_{3,k}$, the iterations eventually become sub-linear, implying that, the iteration is controlled by the perturbation error.
- The non-summability of $(|\gamma_{3,k} - \gamma_0|)_{k\in\mathbb{N}} \notin \ell^1_+$ shows that our assumption on the convergence of the non-stationary iteration is appropriate.

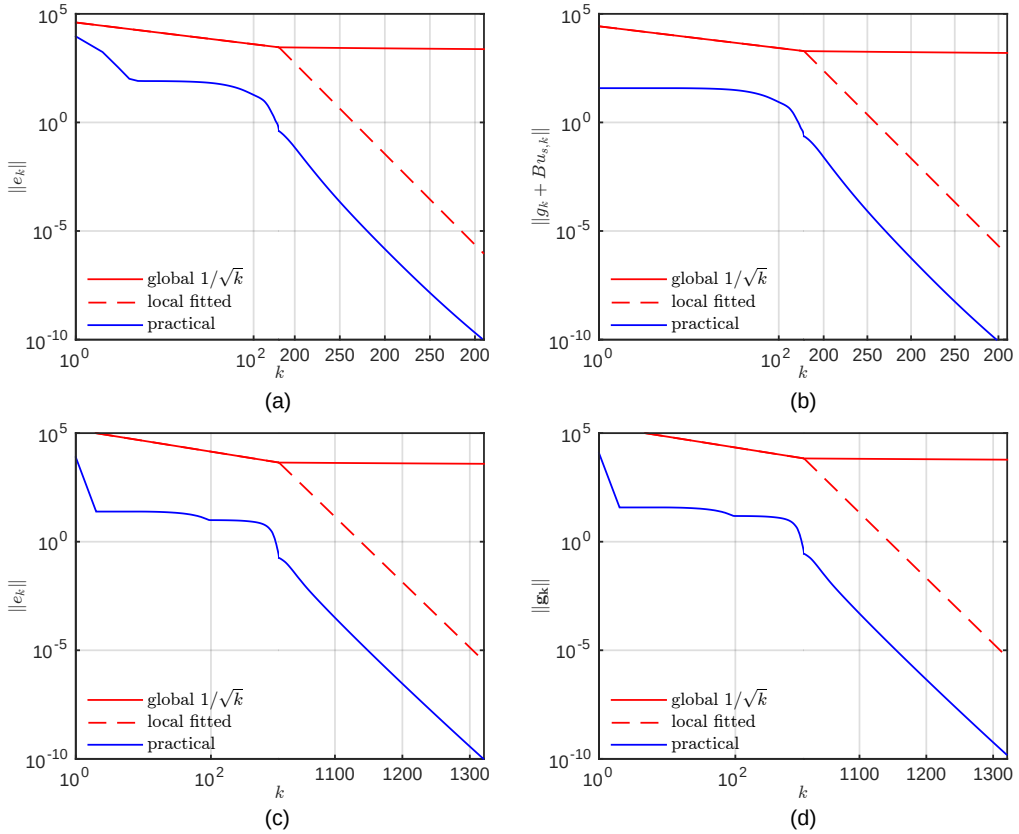


Figure 3.5: The size of the matrix f is 400×300, $\text{rank}(x_{\text{ob},L}) = 20$ and the sparsity of $x_{\text{ob},S}$ is 25%, namely, 25% of the elements of $x_{\text{ob},L}$ are non-zero. (a) pointwise convergence of $\|e_k\|$ of GFB, (b) pointwise convergence of $\|g_k + B(\sum_i \omega_i u_{i,k})\|$ of GFB, (c) pointwise convergence of $\|e_k\|$ of PD, (d) pointwise convergence $\|\boldsymbol{g}_k\|$ of PD.

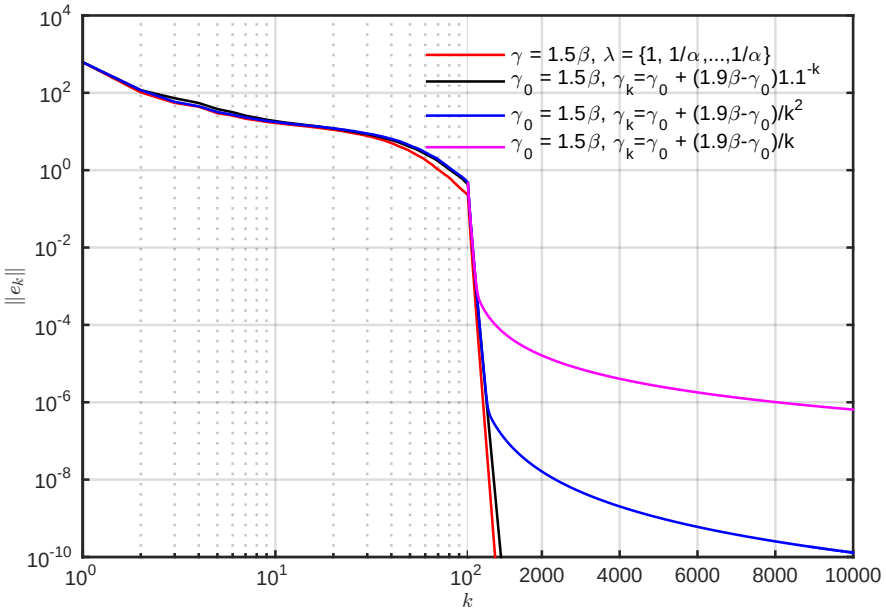


Figure 3.6: Comparison of stationary and non-stationary iterations of GFB.

Chapter 4

Multi-step Inertial Operator Splitting Algorithms

Main contributions of this chapter

- We propose a variable metric multi-step inertial operator splitting method (Algorithm 7) and prove its global convergence.
- We propose a multi-step inertial Forward–Backward for non-convex optimization (Algorithm 11) and prove the global convergence of the generated sequence.

The content on multi-step inertial FB for non-convex optimization will appear in [117].

Contents

In this chapter, we propose the following two multi-step inertial operator splitting methods:

(i) A variable metric multi-step inertial Forward–Backward splitting method, dubbed as "MUSTARD", for general monotone inclusion problems.

(ii) A multi-step inertial Forward–Backward splitting method, which is called "ncvx-MiFB", for non-convex optimization.

For the MUSTARD algorithm, global convergence properties of the generated sequence are provided in Section 4.2. Moreover, since it is proposed for solving general monotone inclusion problems, extensions to various existing operator splitting methods (*e.g.* non-relaxed DR, GFB and Primal–Dual splitting) are discussed in Section 4.3. One common property of these operator splitting methods is that their fixed-point iteration can be written as finding the zeros of a monotone inclusion problem, which is different from the original one, for instance the Primal–Dual splitting method (2.4.14) we have seen in Section 2.4.3.3.

For the ncvx-MiFB, under the assumption that the objective function obeys the Kurdyka-Łojasiewicz property (see Definition 4.4.2), we prove global convergence of its generated sequence to a critical point in Section 4.4.

4.1 Introduction

4.1.1 Problem statement

Let's recall monotone inclusion problem introduced in Section 2.4 for the Forward–Backward splitting method, *i.e.* finding the zeros of the sum of two maximal monotone operators,

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in A(x) + B(x), \tag{$\mathcal{P}_{\text{inc}}$}$$

where $\mathcal{H}$ is real Hilbert space, and

(**A.1**) $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a set-valued maximal monotone operator.

(**A.2**) $B : \mathcal{H} \to \mathcal{H}$ is maximal monotone and β-cocoercive for some $\beta > 0$.

(**A.3**) $\operatorname{zer}(A + B) \neq \emptyset$, *i.e.* the set of zeros of $A + B$ is non-empty.

Owing to Lemma 2.3.2 and 2.3.9, we can link problem ($\mathcal{P}_{\text{inc}}$) to finding the minimizers of the non-smooth optimization problem

$$\min_{x \in \mathcal{H}} \ \Phi(x) = F(x) + R(x), \tag{$\mathcal{P}_{\text{opt}}$}$$

where

(**A.4**) $R \in \Gamma_0(\mathcal{H})$ is a non-smooth function.

(**A.5**) F is convex and differentiable, and ∇F is $(1/\beta)$-Lipschitz continuous.

(**A.6**) $\operatorname{Argmin}(\Phi) \neq \emptyset$, *i.e.* the set of minimizers is non-empty.

Under assumptions (**A.4**)-(**A.6**), the first-order optimality condition of problem ($\mathcal{P}_{\text{opt}}$) (Theorem 2.1.13), reads

$$x^\star \in \operatorname{Argmin}(\Phi) \iff 0 \in \partial R(x^\star) + \nabla F(x^\star). \tag{4.1.1}$$

Clearly, if we let $A = \partial R$ and $B = \nabla F$, then ($\mathcal{P}_{\text{opt}}$) is simply a special case of ($\mathcal{P}_{\text{inc}}$).

As introduced in Section 2.4, both problems above can be handled by the Forward–Backward splitting method [119, 140]. We have seen from Chapter 3 that the convergence rate of the sequence generated by FB method is at best $o(1/\sqrt{k})$ (*i.e.* $\|x_k - x_{k-1}\| = o(1/\sqrt{k})$), which is not fast. See also [134, 36, 24] for the $O(1/k)$ convergence rate of objective function of the FB, *i.e.* $\Phi(x_k) - \Phi(x^\star) = O(1/k)$.

Due to the slow convergence property, acceleration techniques of the FB method have been a hot topic over the past years, and several authors have developed very efficient variants of the FB. Before presenting the proposed algorithms, we think it is necessary to go through a brief overview of the literature, about the development of several popular acceleration techniques: from the acceleration of the gradient descent, then PPA, and eventually the FB method.

4.1.2 Related work

For the sake of brevity, here we consider the non-relaxed version of the FB iteration

$$x_{k+1} = \operatorname{J}_{\gamma_k A}(\operatorname{Id} - \gamma_k B)(x_k), \ \gamma_k \in]0, 2\beta[.$$

It is straightforward to see that FB reduces to the PPA (2.4.5) when $B = 0$, and the gradient descent when $A = 0$.

4.1.2.1 Gradient descent and heavy ball method

Consider the optimization problem ($\mathcal{P}_{\text{opt}}$) and let $R = 0$, then FB becomes the following gradient descent method,

$$x_{k+1} = x_k - \gamma_k \nabla F(x_k), \ \gamma_k \in]0, 2\beta[. \tag{4.1.2}$$

Gradient descent is probably the best-known optimization method, and can be dated back to the 19th century or even earlier [142].

As for the FB method, owing to the result of Chapters 3, without any stronger assumptions (*e.g.* strong convexity), the global convergence rate of $\|x_k - x_{k-1}\|$ is at best $o(1/k)$. In terms of the objective function, an $O(1/k)$ rate is proved in [134] for $\gamma \in]0, \beta]$, then in [36] same result is obtained for $\gamma \in]0, 2\beta[$. In the mid-sixties, Polyak introduced several methods that can speed up gradient descent

[146] and proposed the so-called "heavy ball method", which takes the following iterative form

$$y_k = x_k + a_k(x_k - x_{k-1}),\ a_k \in [0, 1],$$
$$x_{k+1} = y_k - \gamma_k \nabla F(x_k),\ \gamma_k \in]0, 2\beta[.$$

Unlike gradient descent, the heavy ball method first "extrapolates" in the direction of $x_k - x_{k-1}$, getting a new point y_k, and then replaces x_k in (4.1.2) with y_k and keeps $\nabla F(x_k)$ unchanged. Such a modification can greatly boost the speed of gradient descent, especially when the objective is *strongly convex* (see Chapter 6 for more details). Later in his book [147], Polyak suggested the idea of using more than 2 steps, *i.e.*

$$y_k = \mathcal{E}(x_k, x_{k-1}, ..., x_{k-s}),$$

where $s \geq 1$ is an integer, and $\mathcal{E}(\cdot)$ is the function which perform "extrapolation" onto those points. However, neither convergence nor rate result is provided, and we will partly answer this open question (see Theorem 4.2.3).

Meanwhile, in his seminal work [132], Nesterov introduced an "optimal first-order method" for smooth optimization, which achieves $O(1/k^2)$ convergence speed on the objective. To make a long story short, the work by these 2 pioneers lays down the foundations for most of nowadays acceleration techniques.

4.1.2.2 Dynamical system and inertial PPA

An alternative perspective of gradient descent method is through the dynamical system approach, in the continuous setting. Let $x(t)$ be the trajectory of x over time t, gradient descent can be seen as an explicit (Euler forward) discretization of the following first-order dynamical system

$$0 = \dot{x}(t) + \nabla F(x(t)),$$

where $\dot{x}$ stands for the first-order derivative. The heavy ball method, as shown in [146, 147], can be seen as the second-order dynamical system with friction, that reads

$$0 = \ddot{x}(t) + \lambda \dot{x}(t) + \nabla F(x(t)),$$

with proper initial conditions.

Motivated by this dynamical system perspective, in [4], Alvarez generalized the above formulation to the differential inclusion of a non-smooth convex function,

$$0 \in \ddot{x}(t) + \lambda \dot{x}(t) + \partial R(x(t)),$$

and proposed an inertial PPA scheme, which reads

$$y_k = x_k + a_k(x_k - x_{k-1}),\ a_k \in [0, 1[,$$
$$x_{k+1} = \mathrm{prox}_{\gamma_k \partial R}(y_k),\ \gamma_k > 0.$$

Later on in [5], the authors replaced the sub-differential ∂R with a maximal monotone set-valued operator in the dynamical system, and derived an inertial PPA method which has the same iteration as above, except that replacing the proximity operator with the resolvent of monotone operator.

4.1.2.3 Inertial FB and FISTA

Inspired by the "heavy ball method" and inertial PPA, in [131] Moudafi and Oliny further extended the dynamical system to the case of the sum of two operators, *i.e.* $A + B$ and

$$0 \in \ddot{x}(t) + \lambda \dot{x}(t) + (A + B)(x(t)),$$

where A and B satisfy assumptions (**A.1**) and (**A.2**) respectively. They proposed the following inertial Forward–Backward splitting method

$$\begin{aligned} y_k &= x_k + a_k(x_k - x_{k-1}),\ a_k \in [0,1[,\\ x_{k+1} &= \jmath_{\gamma_k A}(y_k - \gamma_k B(x_k)),\ \gamma_k \in]0, 2\beta[. \end{aligned}$$

Another inertial FB scheme was proposed later in [120] where the authors further replaced the $\gamma_k B(x_k)$ with $\gamma_k B(y_k)$.

On the other hand, in the context of convex optimization for solving ($\mathcal{P}_{\text{opt}}$), based upon Nesterov's work [132], Beck and Teboulle proposed the accelerated FISTA method (fast iterative shrinkage-thresholding algorithm) whose iteration is, let $t_0 = 1$,

$$\begin{aligned} t_k &= \frac{1+\sqrt{1+4t_{k-1}^2}}{2},\ a_k = \frac{t_{k-1}-1}{t_k},\\ y_k &= x_k + a_k(x_k - x_{k-1}),\\ x_{k+1} &= \operatorname{prox}_{\gamma_k \partial R}(\mathrm{Id} - \gamma_k \nabla F)(y_k),\ \gamma_k \in]0, \beta]. \end{aligned} \tag{4.1.3}$$

Like the smooth case in [132], the above scheme also achieves $O(1/k^2)$ convergence rate for the objective function, *i.e.* $\Phi(x_k) - \Phi(x^\star) = O(1/k^2)$ where $x^\star \in \operatorname{Argmin}(\Phi)$ is a global minimizer. It should be noted that, such a rate is also the lower complexity bound for first-order algorithms solving this class of problems [133], *i.e.* no first-order algorithm can do better uniformly over this class of problems.

While the sequences generated by the (inertial) FB are convergent, the convergence of FISTA iterates has been an open problem until recently. This question was settled in [50] where the authors proved the convergence of the FISTA by a modification of the inertial parameter, which is

$$t_k = \frac{k+d-1}{d},\ d > 2,\ \ a_k = \frac{t_{k-1}-1}{t_k} = \frac{k-1}{k+d}. \tag{4.1.4}$$

Such modification of the inertial parameter not only preserves the $O(1/k^2)$ convergence rate of the objective function, but also enabled the authors to prove the weak convergence of the generated sequence. Later on in [13], Attouch *et al.* proposed the same convergence result in the continuous dynamical system case. The convergence rate of the objective is actually even $o(1/k^2)$ as proved in [12].

The above different inertial Forward–Backward splitting methods and FISTA can be further summarized into the following algorithm, which is proposed in [115].

Algorithm 6: A generalized inertial Forward–Backward splitting

Initial: $\bar{a}, \bar{b} \in]0,1]$. $x_0 \in \mathcal{H}$, $x_{-1} = x_0$.

repeat

Let $a_k \in [0, \bar{a}]$, $b_k \in [0, \bar{b}]$, $\gamma_k \in]0, 2\beta[$:

$$\left\lfloor \begin{aligned} y_{a,k} &= x_k + a_k(x_k - x_{k-1}),\\ y_{b,k} &= x_k + b_k(x_k - x_{k-1}),\\ x_{k+1} &= \jmath_{\gamma_k A}\big(y_{a,k} - \gamma_k B(y_{b,k})\big). \end{aligned} \right.$$

$k = k + 1$;

until *convergence*;

4.1.2.4 Variable metric FB

Another branch of FB variants is the *variable metric Forward–Backward* developed by Combettes and Vũ in [62].

Define $\mathcal{S}(\mathcal{H}) = \{\mathcal{V} : \mathcal{H} \to \mathcal{H} | \mathcal{V}^* = \mathcal{V}, \|\mathcal{V}\| < +\infty\}$ the set of self-adjoint bounded linear operators act on $\mathcal{H}$, the Loewner partial ordering on $\mathcal{S}(\mathcal{H})$ is defined by

$$\forall \mathcal{V}_1, \mathcal{V}_2 \in \mathcal{S}(\mathcal{H}), \quad \mathcal{V}_1 \succcurlyeq \mathcal{V}_2 \iff \forall x \in \mathcal{H}, \langle \mathcal{V}_1 x, x\rangle \geq \langle \mathcal{V}_2 x, x\rangle.$$

Given any positive constant $\nu \in]0, +\infty[$, denote $\mathcal{M}_\nu$ by

$$\mathcal{M}_\nu \stackrel{\text{def}}{=} \big\{\mathcal{V} \in \mathcal{S}(\mathcal{H}) \quad \text{s.t.} \ \ \mathcal{V} \succcurlyeq \nu \mathrm{Id}\big\}, \tag{4.1.5}$$

i.e. the set of *symmetric positive definite operators* whose spectrum is bounded from below by ν. For any $\mathcal{V} \in \mathcal{M}_\nu$, define the following induced scalar product and norm,

$$\forall x, x' \in \mathcal{H}, \ \ \langle x, x'\rangle_{\mathcal{V}} = \langle x, \mathcal{V}x'\rangle_{\mathcal{V}}, \ \ \|x\|_{\mathcal{V}} = \sqrt{\langle x, \mathcal{V}x\rangle_{\mathcal{V}}}.$$

By endowing the Hilbert space $\mathcal{H}$ with the above scalar product and norm, we obtain the real Hilbert space which is denoted by $\mathcal{H}_{\mathcal{V}}$.

Let $\mathcal{V} \in \mathcal{M}_\nu$, it is easy to see that for ($\mathcal{P}_{\text{inc}}$), let $x^\star \in \operatorname{zer}(A + B)$, then there also holds

$$0 \in \mathcal{V}^{-1}(A + B)(x^\star).$$

Based on this observation, in [62] Combettes and Vũ proposed a variable metric Forward–Backward splitting method,

$$x_{k+1} = \mathcal{J}_{\gamma_k \mathcal{V}_k^{-1} A}(\mathrm{Id} - \gamma_k \mathcal{V}_k^{-1} B)(x_k), \ \ \gamma_k \in]0, 2\beta\nu[, \tag{4.1.6}$$

where $\mathcal{J}_{\gamma_k \mathcal{V}_k^{-1} A} \stackrel{\text{def}}{=} (\mathrm{Id} + \gamma_k \mathcal{V}_k^{-1} A)^{-1}$. The convergence of the above iteration is guaranteed if

$$\mu = \sup_{k \to +\infty} \|\mathcal{V}_k\| < +\infty \quad \text{and} \quad (1 + \eta_k)\mathcal{V}_{k+1}^{-1} \succcurlyeq \mathcal{V}_k^{-1},$$

and $\{\eta_k\}_{k \in \mathbb{N}} \in \ell^1_+$.

The summability condition on $\{\eta_k\}_{k \in \mathbb{N}}$ implies that $\mathcal{V}_k$ converges pointwisely to some $\mathcal{V} \in \mathcal{M}_\nu$ [61, Lemma 2.3]. The operators $\mathcal{V}_k$ can be seen as *preconditioning*, and its ideal case is when $\mathcal{V}_k \to \mathcal{V} = \nabla B$ where ∇B is the Jacobian of B if B is differentiable and ∇B is self-adjoint positive definite. However, one has to keep in mind that the variable metric FB has two drawbacks: (i) one has to design and store appropriate operators $\mathcal{V}_k$. (ii) the resolvent $\mathcal{J}_{\gamma_k \mathcal{V}_k^{-1} A}$ in general is very difficult to compute even if $\mathcal{J}_{\gamma_k A}$ is available (in fact might be as difficult as solving the original problem). One possible approach is to consider quasi-Newton-type metrics, see for instance [25] and the references therein.

Another important extension of variable metric FB is that it can be applied to analyse the convergence properties of several Primal–Dual splitting methods. This viewpoint has been initiated by He and Yuan and subsequently adapted by other authors, see *e.g.* [92, 173, 66, 62].

4.2 Variable metric multi-step inertial operator splitting

4.2.1 The MUSTARD algorithm

Putting together the ideas of multi-step inertial methods and variable metric ones, we here propose a new operator splitting algorithm: *variable metic MUlti-Step inerTial operAtoR splitting methoD*, which is dubbed as "MUSTARD" for solving ($\mathcal{P}_{\text{inc}}$) and ($\mathcal{P}_{\text{opt}}$). The details of the algorithm for solving ($\mathcal{P}_{\text{inc}}$) are given below in Algorithm 7.

Algorithm 7: The MUSTARD algorithm

Initial: $s \in \mathbb{N}_+$ and $\mathcal{S} \overset{\text{def}}{=} \{0, ..., s-1\}$, let $\mathcal{M}_\nu$ as defined in (4.1.5). $x_0 \in \mathcal{H}$, $x_{-i} = x_0, i \in \mathcal{S}$. Choose $\bar{\epsilon}, \underline{\epsilon} > 0$ such that $\underline{\epsilon} \leq 2\beta\nu - \bar{\epsilon}$.

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}}, \{b_{i,k}\}_{i\in\mathcal{S}} \in]-1, 2]^s$, $\mathcal{V}_k \in \mathcal{M}_\nu$ and $\gamma_k \in [\underline{\epsilon}, 2\beta\nu - \bar{\epsilon}]$:

$$\left\lfloor \begin{aligned} y_{a,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}), \\ y_{b,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}), \\ x_{k+1} &= \mathcal{J}_{\gamma_k \mathcal{V}_k^{-1} A}\big(y_{a,k} - \gamma_k \mathcal{V}_k^{-1} B(y_{b,k})\big). \end{aligned} \right. \tag{4.2.1}$$

$k = k + 1$;

until *convergence*;

Remark 4.2.1. The main reasons of considering the monotone inclusion ($\mathcal{P}_{\text{inc}}$) instead of the optimization problem ($\mathcal{P}_{\text{opt}}$), and involving variable metric are as followings:

(i) Problem ($\mathcal{P}_{\text{opt}}$) is only a special case of ($\mathcal{P}_{\text{inc}}$). In particular, for various operators splitting methods (*e.g.* non-relaxed DR, GFB, and Primal–Dual splitting), their fixed-point iterations can be written as certain monotone inclusion problems, while the involved monotone operators are *not* sub-differential of convex functions.

(ii) As we have seen from the examples of Primal–Dual splitting methods (*e.g.* (2.4.14) in Section 2.4.3.3 and (3.4.17) in Section 3.4.3), their corresponding monotone inclusion formulations are under different metric (*i.e.* metric $\boldsymbol{\mathcal{V}}$ for the two examples).

As a result, considering the monotone inclusion problem and involving the variable metric allow us to extend the MUSTARD algorithm to a broad class of operator splitting methods.

Relation to previous work By form, the MUSTARD Algorithm 7 is the most general variable metric Forward–Backward splitting we are aware of, it is brand new to the literature for $s \geq 2$. For the case $s = 1$, it is more general than Algorithm 6 as variable metric is considered, and recovers the variable metric Forward–Backward proposed in [62] if $s = 0$.

If we choose $s = 1$ and $\mathcal{V}_k = \mathrm{Id}$, then based on the choice of the inertial parameters a_k and b_k, the relations between Algorithm 7 with the aforementioned work are as following,

- $a_k = 0$, $b_k = 0$: this is the original FB method [119, 140].
- $a_k \in [0, \bar{a}], b_k = 0$: this is the case studied in [131] for ($\mathcal{P}_{\text{inc}}$). In the context of optimization with $R = 0$, one recovers the heavy ball method [146].
- $a_k \in [0, \bar{a}]$, $b_k = a_k$: this corresponds to the work of [120] for solving ($\mathcal{P}_{\text{inc}}$). If moreover restrict $\gamma_k \in]0, \beta]$ and let $a_k \to 1$, then Algorithm 6 specializes to FISTA-type methods [24, 50, 13, 12] developed for optimization.
- $a_k \in [0, \bar{a}]$, $b_k \in]0, \bar{b}]$, $a_k \neq b_k$: the general inertial FB scheme Algorithm 6.

Below we also highlight the several important characteristics of the MUSTARD algorithm.

(i) Similarly to some existing inertial methods [131, 120, 115], the choice of the step-size γ_k allowed by MUSTARD is $]0, 2\beta[$ if $\mathcal{V}_k \equiv \mathrm{Id}$;

(ii) The algorithm allows multiple steps, which is characterized by s. In particular for $s = 2$, we show that very promising practical results can be obtained (see Section 4.5).

(iii) We allow to use *negative* inertial parameters. For $s = 1$, the inertial parameters should be positive and lie in $[0, 1[$ to ensure convergence. However, for $s \geq 2$, we will show that one can benefit from negative choices of the inertial parameters. In particular, for $s = 2$, the numerical experiments

of Section 4.5 implies that a good choice of the inertial parameters should be

$$a_{0,k}, b_{0,k} \in]0, 2] \quad \text{and} \quad a_{1,k}, b_{1,k} \in]-1, 0].$$

Such an inertial setting can be investigated through the dynamical system perspective, see below for a short introduction.

(iv) As the problem we are considering is the general monotone inclusion problem, we can generalize the MUSTARD algorithm to the methods whose iterations are related to some monotone inclusion problem, for example the non-relaxed DR, GFB and several Primal–Dual splitting methods.

MUSTARD as a discretised dynamical system Now consider the metric free case of MUSTARD algorithm for the optimization problem ($\mathcal{P}_{\text{opt}}$) with $\mathcal{V} = \text{Id}$. Consider the following second-order dynamical system

$$\ddot{x}(t) + c_0(t)\dot{x}(t) + (\partial R + \nabla F)(x(t)) = 0, \tag{4.2.2}$$

where $c_0(t) \geq 0$ is an asymptotically vanishing viscous damping function. Typically, $c_0(t)$ moderately decreases to 0, *i.e.* $\lim_{t\to+\infty} c_0(t) = 0$ and $\int_t c_0(t) = +\infty$.

Let $0 < \omega_2 < \omega_1$ be two weights such that $\omega_1 + \omega_2 = 1$, $h > 0$ be the time step-size, $t_k = kh$ and $x_k = x(t_k)$. Consider an implicit (Euler backward) discretization w.r.t. ∂R and an explicit (Euler forward) discretization w.r.t. ∇F, and a weighted sum of explicit and implicit discretization of $\ddot{x}(t)$, *i.e.*

$$0 \in \frac{\omega_1}{h^2}(x_{k+1} - 2x_k + x_{k-1}) + \frac{\omega_2}{h^2}(x_k - 2x_{k-1} + x_{k-2}) + \frac{c_0(kh)}{h}(x_k - x_{k-1}) + \partial R(x_{k+1}) + \nabla F(y_{b,k}),$$

then we obtain the following inclusion

$$x_k + a_{0,k}(x_k - x_{k-1}) + a_{1,k}(x_{k-1} - x_{k-2}) - \gamma\nabla F(y_{b,k}) \in x_{k+1} + \gamma\partial R(x_{k+1}), \tag{4.2.3}$$

where we have

$$a_{0,k} = 1 - \frac{\omega_2}{\omega_1} - \frac{hc_0(kh)}{\omega_1}, \quad a_{1,k} = \frac{\omega_2}{\omega_1} \quad \text{and} \quad \gamma = \frac{h^2}{\omega_1}.$$

If we moreover set

$$y_{b,k} = x_k + b_{0,k}(x_k - x_{k-1}) + b_{1,k}(x_{k-1} - x_{k-2}),$$

with $b_{0,k}, b_{1,k}$ being properly chosen, then we obtain the MUSTARD scheme for the case $s = 2$ and $\mathcal{V}_k \equiv \text{Id}$.

If we choose $c_0(kh) = \frac{d}{kh}, d > 3$, then (4.2.3) simplifies to the following inclusion

$$x_k + \big(1 - \frac{d}{\omega_1 k}\big)(x_k - x_{k-1}) - \frac{\omega_2}{\omega_1}(x_k - 2x_{k-1} + x_{k-2}) - \gamma\nabla F(y_{b,k}) \in x_{k+1} + \gamma\partial R(x_{k+1}).$$

If we further let

$$\omega_1 = 1, \ \omega_2 = 0 \quad \text{and} \quad y_{b,k} = x_k + b_k(x_k - x_{k-1}), \ b_k \in [0, 1],$$

then we recover a special case of Algorithm 6. If one moreover sets $b_k = a_k = (1 - d/k)$, then we obtain the FISTA scheme as studied in [160, 11, 50].

4.2.2 Global convergence of MUSTARD

In this section, we present the global convergence analysis for the MUSTARD algorithm. We summarize our results as follows:

(i) $s \geq 2$: In Theorem 4.2.3, we establish conditional convergence of the sequence $\{x_k\}_{k\in\mathbb{N}}$ for fixed metric $\mathcal{V}$, where the terminology "conditional convergence" refers to the fact that for the convergence of the sequence to occur, the sequences $\{a_{i,k}\}_{i\in\mathcal{S}}, \{b_{i,k}\}_{i\in\mathcal{S}}$ has to be chosen depending (conditionally) on the sequence $\{x_k\}_{k\in\mathbb{N}}$ in such a way that an appropriate condition holds, *e.g.* (4.2.5). Unfortunately, so far for the case $s \geq 2$ we only have a result for the case of fixed metric $\mathcal{V}_k \equiv \mathcal{V}$. However, it is sufficient to cover many algorithms of interest.

(ii) $s = 1$:

(a) in Theorem 4.2.5 we manage to prove conditional convergence of $\{x_k\}_{k\in\mathbb{N}}$ for a variable metric $\mathcal{V}_k$.

(b) We also devise choices of the inertial parameters and metrics that are independent of $\{x_k\}_{k\in\mathbb{N}}$ and still guarantee global convergence (see Theorem 4.2.9). We dub this unconditional convergence.

All the proofs of the above results are gathered in Section 4.6.

For the sake of generality, we consider the inexact version of the MUSTARD algorithm. The following definition is needed.

Definition 4.2.2 (ε-enlargement). Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a set-valued maximal monotone operator and $\varepsilon \geq 0$. Then the ε-enlargement of A is defined by,

$$A^{\varepsilon}(x) \stackrel{\text{def}}{=} \big\{v \in \mathcal{H},\ \langle u - v,\ y - x\rangle \geq -\varepsilon,\ \forall y \in \mathcal{H},\ u \in A(y)\big\}.$$

From the definition, it is easy to verify that for $0 \leq \varepsilon_1 \leq \varepsilon_2$ we have $A^{\varepsilon_1}(x) \subset A^{\varepsilon_2}(x)$ and $A^0(x) = A(x)$. Thus A^{ε} is an enlargement of A.

For the updating step of x_{k+1} in (4.2.1), consider the following inexact form

$$y_{a,k} - \gamma_k\big(\mathcal{V}_k^{-1}B(y_{b,k}) + \xi_k\big) - x_{k+1} \in \gamma_k\mathcal{V}_k^{-1}A^{\varepsilon_k}(x_{k+1}),$$

where $\xi_k \in \mathcal{H}$ is the error in the evaluation of the gradient operator B, and ε_k is the enlargement error. Then we obtain the inexact form of MUSTARD

$$\begin{aligned} y_{a,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}),\\ y_{b,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}),\\ y_{a,k} - \gamma_k(\mathcal{V}_k^{-1}B(y_{b,k}) + \xi_k) - x_{k+1} &\in \gamma_k\mathcal{V}_k^{-1}A^{\varepsilon_k}(x_{k+1}). \end{aligned} \tag{4.2.4}$$

4.2.2.1 Conditional convergence

We present first the conditional convergence of the inexact MUSTARD algorithm. For each $i \in \mathcal{S}$, define $\zeta_{i,k} \stackrel{\text{def}}{=} a_{i,k} - \frac{\gamma_k b_{i,k}}{2\beta\nu}$ and $\bar{a}_i \stackrel{\text{def}}{=} \sup_{k\in\mathbb{N}} |a_{i,k}|$.

Theorem 4.2.3 (Conditional convergence $s \geq 2$). *For the inexact MUSTARD iteration (4.2.4), let conditions (A.1)-(A.3) hold, fix the metric $\mathcal{V}_k \equiv \mathcal{V} \in \mathcal{M}_\nu$, and let $\xi_k \equiv 0$. Suppose that the following two conditions hold*

(i) *the error $\{\varepsilon_k\}_{k\in\mathbb{N}} \in \ell^1_+$;*

(ii) *the inertial parameters $\{a_{i,k}\}_{i\in\mathcal{S}}$ are such that $\sum_{i\in\mathcal{S}} \bar{a}_i < 1$.*

Then the generated sequence $\{x_k\}_{k\in\mathbb{N}}$ is bounded. If moreover the following summability condition holds

$$\textstyle\sum_{k\in\mathbb{N}} \max\Big\{\max_{i\in\mathcal{S}} \zeta_{i,k}^2, \max_{i\in\mathcal{S}} |b_{i,k}|, \max_{i\in\mathcal{S}} |a_{i,k}|\Big\} \sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2 < +\infty, \tag{4.2.5}$$

then there exists an $x^\star \in \operatorname{zer}(A + B)$ such that the sequence $\{x_k\}_{k\in\mathbb{N}}$ weakly converges to $x^\star$.

The proof of the theorem can be found in Section 4.6 from page 83.

Remark 4.2.4. If the inertial parameters $\{a_{i,k}\}_{i\in\mathcal{S}}, \{b_{i,k}\}_{i\in\mathcal{S}}$ are chosen in $[0, 1]$ such that

$$\zeta_{i,k}^2 = \big(a_{i,k} - \tfrac{\gamma_k b_{i,k}}{2\beta\nu}\big)^2 \leq a_{i,k}^2$$

then condition (4.2.5) simplifies to

$$\textstyle\sum_{k\in\mathbb{N}} \max\Big\{\max_{i\in\mathcal{S}} |b_{i,k}|, \max_{i\in\mathcal{S}} |a_{i,k}|\Big\} \sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2 < +\infty.$$

Condition (4.2.5) can be enforced by a simple online updating rule such as, for each $i \in \mathcal{S}$, given $a_i, b_i \in [0,1]$,

$$a_{i,k} = \min\left\{a_i, c_{a,i,k}\right\}, \quad b_{i,k} = \min\left\{b_i, c_{b,i,k}\right\}, \tag{4.2.6}$$

where $c_{a,i,k}, c_{b,i,k} > 0$, and $\max\{c_{a,i,k}, c_{b,i,k}\} \sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2$ is summable. For instance, one can choose

$$c_{a,i,k} = \frac{c_{a,i}}{k^{1+\delta} \sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2}, \ c_{a,i} > 0, \ \delta > 0,$$

and similarly for $c_{b,k}$.

When $s = 1$, then we have the following theorem with a variable metric $\mathcal{V}_k$. To lighten the notations, for $s = 1$, we denote $a_k = a_{0,k}$ and $b_k = b_{0,k}$.

Theorem 4.2.5 (Conditional convergence $s = 1$). *For the inexact MUSTARD iteration (4.2.4), let conditions (A.1)-(A.3) hold. Suppose that the following conditions are satisfied*

(i) *For the metric sequence $\{\mathcal{V}_k\}_{k\in\mathbb{N}} \in \mathcal{M}_\nu$ with $\nu > 0$, suppose that there exists a non-negative sequence $\{\eta_k\}_{k\in\mathbb{N}} \in \ell^1_+$ such that*

$$\mu = \sup_{k\in\mathbb{N}} \|\mathcal{V}_k\| < +\infty \quad \text{and} \quad (1+\eta_k)\mathcal{V}_k \succcurlyeq \mathcal{V}_{k+1}.$$

(ii) *the inertial parameter is $a_k \in [0,1]$ such that $\bar{c} \overset{\text{def}}{=} \sup_{k\in\mathbb{N}} a_k(1+\eta_{k-1}) < 1$, and*

$$\sup_{k\in\mathbb{N}} \frac{1}{1-\bar{c}} \sum_{m=1}^{k} \frac{(1-\bar{c}^{k-m+1})\eta_m}{1+\eta_m} < 1. \tag{4.2.7}$$

(iii) *the errors $\{\varepsilon_k\}_{k\in\mathbb{N}} \in \ell^1_+$ and $\{\|\xi_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$.*

Then the generated sequence $\{x_k\}_{k\in\mathbb{N}}$ is bounded. If moreover the following summability condition holds

$$\sum_{k\in\mathbb{N}} \max\{a_k, b_k\}\|x_k - x_{k-1}\|^2 < +\infty, \tag{4.2.8}$$

then there exists an $x^\star \in \operatorname{zer}(A+B)$ such that the sequence $\{x_k\}_{k\in\mathbb{N}}$ weakly converges to $x^\star$.

The proof of the theorem can be found in Section 4.6 from page 86.

Remark 4.2.6.

(i) If the sequence $\{\eta_k\}_{k\in\mathbb{N}}$ satisfies $\sum_{k\in\mathbb{N}} \eta_k < 1 - \bar{c}$, then condition (4.2.7) is in force. Given $\bar{c} \in [0,1[$, let $\delta, \kappa > 0$, and set η_k as

$$\eta_k = \frac{\kappa}{k^{1+\delta}}.$$

Then for fixed δ, (4.2.7) can be met with a proper choice of κ;

(ii) If $a_k \geq b_k$, then (4.2.8) recovers the conditions in [131, 120] for the conditional convergence of $\{x_k\}_{k\in\mathbb{N}}$.

An empirical choice of the inertial parameters We introduce two empirical ways to set up the inertial parameters. For the sake of simplicity, let $\mathcal{V}_k \equiv \mathrm{Id}$, hence $\nu = 1$. Consider the constant parameter setting,

$$\gamma \in]0, 2\beta[\quad \text{and} \quad b_i = a_i \in]-1, 2[, i \in \mathcal{S}.$$

Moreover, let $(a_i)_{i\in\mathcal{S}}$ be monotone non-increasing, *i.e.* $a_0 \geq a_1 \geq \cdots \geq a_{s-1}$.

Summarizing from multi-step inertial Forward–Backward and gradient descent, we obtain the following two empirical bounds for the summand $\sum_{i\in\mathcal{S}} a_i$:

$$\begin{aligned} &\text{“Upper bound 1”: } \textstyle\sum_i a_i \in \left]0, \min\left\{1, \frac{2\beta-\gamma}{\gamma}\right\}\right[, \\ &\text{“Upper bound 2”: } \textstyle\sum_i a_i \in \left]0, \min\left\{1, \frac{2\beta-\gamma}{2|\beta-\gamma|}\right\}\right[. \end{aligned} \tag{4.2.9}$$

In practice, to ensure the convergence of the generated sequence $\{x_k\}_{k\in\mathbb{N}}$, these two bounds should be applied together with the online updating rule of inertial parameters (4.2.6). Most of the time, with proper choice of each a_i, (4.2.6) may never be triggered.

Remark 4.2.7. Compare (4.2.9) with (ii) of Theorem 4.2.3, the main difference is that here we consider the summand with signs. This means that we can choose positive inertial parameters bigger than 1, and then compensate with negative ones. As a matter of fact, as we will see in the numerical experiment, negative inertial parameter would make the convergence even faster. For instance, for the case $s = 2$ with $\sum_i a_i$ being fixed, then the choice $a_1 < 0 < a_0$ may outperform the one with $a_1, a_0 \geq 0$.

The two upper bounds are shown graphically in Figure 4.1. It can be observed that for $\gamma \leq \beta$, the largest value that can be allowed is 1, which corresponds to the choice of FISTA method whose inertial parameter tends to 1 as $k \to +\infty$.

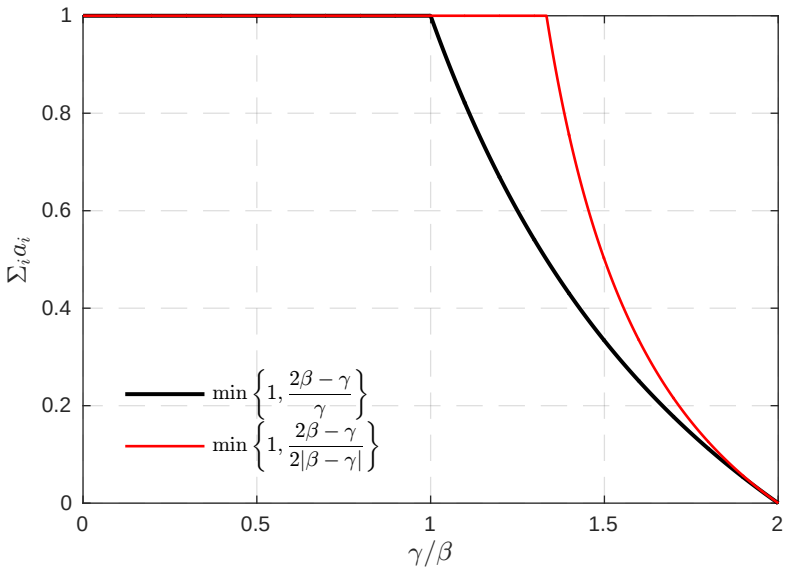


Figure 4.1: Two empirical upper bounds for the sum of inertial parameters $\sum_i a_i$: "Upper bound 1", $\sum_{i\in\mathcal{S}} a_i \in \left[0, \min\left\{1, \frac{2\beta-\gamma}{\gamma}\right\}\right]$; "Upper bound 2", $\sum_{i\in\mathcal{S}} a_i \in \left[0, \min\left\{1, \frac{2\beta-\gamma}{2|\beta-\gamma|}\right\}\right]$.

Remark 4.2.8.

(i) Between the two bounds in (4.2.9), "Upper bound 2" is much less stringent than "Upper bound 1".

(ii) For inertial Forward–Backward, $\sum_i a_i$ too close to 1 is not a good choice. Such an observation (*e.g.* from the numerical experiments in Section 4.5) coincides with the existing studies on FISTA (*e.g.* the local oscillation). A theoretical explanation for such behaviours of $\sum_i a_i$ too close to 1 is left to Chapter 6.

All the above remarks will be made clear in the numerical experiment section, typically for the multi-step inertial Forward–Backward splitting method.

Lastly, it should be emphasised that the two empirical bounds in (4.2.9) are designed for multi-step inertial FB, gradient descent, and the original PPA method. They may not work for the other inertial schemes. As a matter of fact, as we will see in the numerical experiments of Section 4.5, the choices of inertial parameters for inertial Douglas–Rachford and Chambolle-Pock Primal–Dual splitting method [51] are rather limited. Moreover, compared to inertial Forward–Backward and gradient descent, the gains of inertia in DR and Primal–Dual splitting are very small.

The reasons underlying such differences on the acceleration brought by inertia to different algorithms is quite complicated to justify in general. However, it can be explained partly through the local linear convergence analysis as we will describe in Chapters 6-8.

4.2.2.2 Unconditional convergence

Besides the conditional convergence, we can devise choices of $\{\{a_{i,k}\}_{i\in\mathcal{S}}\}_{k\in\mathbb{N}}$ and $\{\{b_{i,k}\}_{i\in\mathcal{S}}\}_{k\in\mathbb{N}}$ that are independent of $\{x_k\}_{k\in\mathbb{N}}$, and still guarantee the global convergence. We dub this *unconditional convergence*. The following result generalizes those in [5, 131, 120, 115].

For the unconditional convergence of Algorithm 7, we restrict ourselves to the case $s = 1$.

Theorem 4.2.9 (Unconditional convergence). *For the inexact MUSTARD iteration (4.2.4), let conditions (A.1)-(A.3) hold. Suppose that the following conditions are satisfied*

(i) *For the metric sequence $\{\mathcal{V}_k\}_{k\in\mathbb{N}} \in \mathcal{M}_\nu$ for $\nu > 0$, suppose that there exists a non-negative sequence $\{\eta_k\}_{k\in\mathbb{N}} \in \ell^1_+$ such that*

$$\mu = \sup_{k\in\mathbb{N}} \|\mathcal{V}_k\| < +\infty \quad \text{and} \quad (1+\eta_k)\mathcal{V}_k \succcurlyeq \mathcal{V}_{k+1}.$$

(ii) *choose the inertial parameters $a_k, b_k \in [0,1]$, such that (4.2.7) holds and moreover there exists $\tau > 0$ and*

$$\begin{cases} 1 + a_k - \frac{\gamma_k}{2\beta\nu}\big((1+b_k)^2 + \eta_{k-1}b_k(b_k+1)\big) \geq \tau : a_k \leq \frac{\gamma_k}{2\beta\nu}b_k, \\ 1 - (3+2\eta_k)a_k - \frac{\gamma_k}{2\beta\nu}\big((1+b_k)^2 + \eta_{k-1}b_k(b_k-1)\big) \geq \tau : \begin{cases} b_k \leq a_k, \ \text{or} \\ \frac{\gamma_k}{2\beta\nu}b_k \leq a_k < b_k, \end{cases} \end{cases} \tag{4.2.10}$$

(iii) *the errors are $\{\varepsilon_k\}_{k\in\mathbb{N}} \in \ell^1_+$ and $\{\|\xi_k\|\}_{k\in\mathbb{N}} \in \ell^1_+$.*

Then $\sum_{k\in\mathbb{N}} \|x_k - x_{k-1}\|^2 < +\infty$, and there exists $x^\star \in \operatorname{zer}(A+B)$ such that the sequence $\{x_k\}_{k\in\mathbb{N}}$ converges weakly to $x^\star$.

See Section 4.6 for the proof from page 90. When the metric $\mathcal{V}_k$ is fixed, *i.e.* $\mathcal{V}_k \equiv \mathcal{V} \in \mathcal{M}_\nu$, then $\eta_k \equiv 0$ and condition (4.2.10) simplifies to

$$\begin{cases} 1 + a_k - \frac{\gamma_k}{2\beta\nu}(1+b_k)^2 \geq \tau : a_k \leq \frac{\gamma_k}{2\beta\nu}b_k, \\ 1 - 3a_k - \frac{\gamma_k}{2\beta\nu}(1+b_k)^2 \geq \tau : b_k \leq a_k \text{ or } \frac{\gamma_k}{2\beta\nu}b_k \leq a_k < b_k. \end{cases} \tag{4.2.11}$$

Figure 4.2 shows graphically the conditions (4.2.11). We choose $\tau = 0.01$ and two different choices of γ are considered. It can be observed that with γ becoming bigger, the range of a, b in (4.2.10) becomes smaller. Moreover, compared to the empirical choice of inertial parameters, the allowed choices by (4.2.10) are quite conservative. For instance, for the case $b_k \equiv a_k \equiv a$ and $\gamma = \beta\nu$, the biggest value can be allowed is $a \equiv \sqrt{5} - 2$. In comparison, when $B = 0$, b_k vanishes and the upper bound of a_k is $1/3$ which coincides with the result of [4, 5].

4.3 Extensions of MUSTARD algorithm

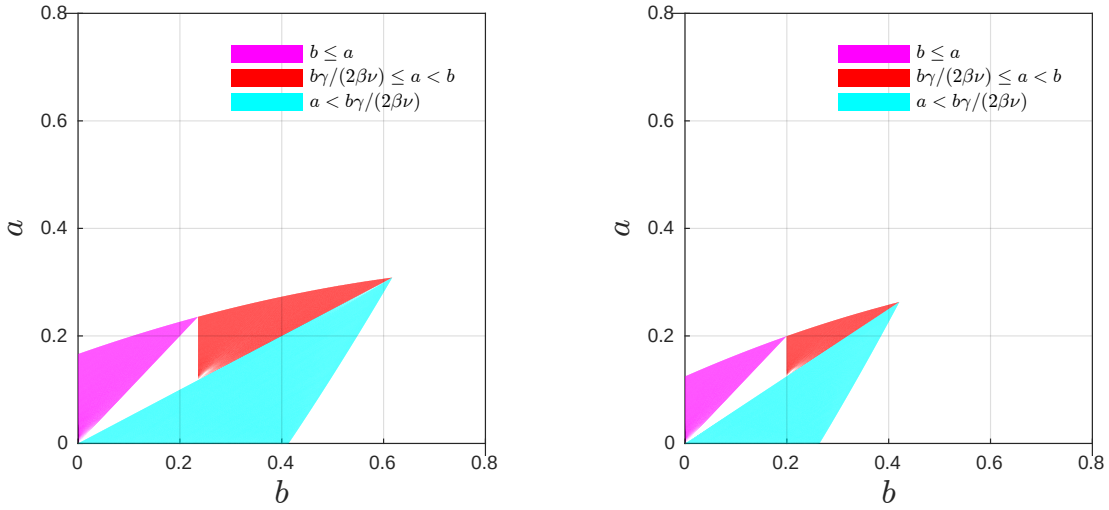
In this section, we demonstrate the broad applicabilities of our MUSTARD Algorithm 7, presenting the inertial versions of various operator splitting methods.

The extension to a multi-step inertial gradient descent, PPA and FB is rather straightforward, since we only need to make the metric $\mathcal{V}_k \equiv \mathrm{Id}$. Hence, in the following, we consider extending the MUSTARD algorithm to DR, GFB and a Primal–Dual splitting method. For the sake of simplicity, we present only the exact version of them, and the convergence of these extensions can be derived immediately from the theorems in Section 4.2.

4.3.1 Inertial Douglas–Rachford/ADMM

Recall the monotone inclusion problem addressed by DR splitting

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A_1 + A_2)(x), \tag{4.3.1}$$



(a) $\gamma = \beta\nu$ (b) $\gamma = 1.25\beta\nu$

Figure 4.2: Sets of allowable (a, b) ensuring the convergence for a given γ. (a) $\gamma = \beta\nu$. (b) $\gamma = 1.25\beta\nu$. We set the value of τ in (4.2.11) as 0.01. Each colour shaded region corresponds to a different condition appearing in (4.2.11), *i.e.* the cyan one corresponds to the first inequality of (4.2.11), while the magenta and red ones correspond to the two conditions of the second inequality of (4.2.11) respectively.

where $A_1, A_2 : \mathcal{H} \rightrightarrows \mathcal{H}$ are maximal monotone.

Consider the non-relaxed DR iteration, owing to (3.4.10), the fixed-point iteration with respect to z_k can be written as

$$\begin{aligned} z_{k+1} &= \tfrac{1}{2}(\mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2} + \mathrm{Id})(z_k), \\ x_{k+1} &= \mathcal{J}_{\gamma A_2}(z_{k+1}), \end{aligned}$$

where the fixed-point operator $\mathcal{F}_{\mathrm{DR}} = \frac{1}{2}(\mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2} + \mathrm{Id})$ is *firmly non-expansive*, by virtue of (ii) of Lemma 2.4.10. Then, owing to Lemma 2.4.2, we know that there exists a maximal monotone operator $A_3 : \mathcal{H} \rightrightarrows \mathcal{H}$, such that

$$\mathcal{J}_{A_3} = \tfrac{1}{2}(\mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2} + \mathrm{Id}),$$

and moreover we have

$$\mathrm{zer}(A_3) = \mathrm{fix}(\mathcal{F}_{\mathrm{DR}}) \quad \text{and} \quad \mathrm{zer}(A_1 + A_2) = \mathcal{J}_{A_3}\big(\mathrm{zer}(A_3)\big).$$

As a result, the non-relaxed DR iteration can be seen as the PPA method for solving the monotone inclusion of

$$\text{find } \; z \in \mathcal{H} \; \text{ such that } \; 0 \in A_3(z).$$

Specialize the MUSTARD algorithm to this specific problem, we obtain a multi-step inertial DR splitting method, which is given below in Algorithm 8.

Algorithm 8: A multi-step inertial Douglas–Rachford (MiDR)

Initial: $\gamma > 0$, $s \in \mathbb{N}_+$ and $\mathcal{S} = \{0, ..., s-1\}$. $z_0 \in \mathcal{H}$ and $x_0 = \mathcal{J}_{\gamma A_2}(z_0)$.

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}} \in \,]-1, 2]^s$:

$$\left\lfloor \begin{aligned} y_k &= z_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(z_{k-i} - z_{k-i-1}), \\ z_{k+1} &= \tfrac{1}{2}(\mathcal{R}_{\gamma A_1}\mathcal{R}_{\gamma A_2} + \mathrm{Id})(y_k). \end{aligned} \right. \tag{4.3.2}$$

$k = k + 1$;

until *convergence*;

Whence we obtain the solution $z^\star$, then the solution to the original problem is simply $x^\star = \jmath_{\gamma A_2}(z^\star)$.

Remark 4.3.1.

(i) Algorithm 8 can be extended to the ADMM method , since ADMM is nothing but applying DR to the dual problem of (4.3.1). See Section 7.6 for the relation between ADMM and DR.

(ii) When $s = 1$, the MUSTARD-DR algorithm recovers the inertial DR scheme proposed in [32] without relaxation.

(iii) A pre-conditioned version of Algorithm 8 can be considered by inserting the metric $\mathcal{V}$ from the very beginning. For $s = 1$, the preconditioned DR was considered in [37].

4.3.2 Inertial Generalized Forward–Backward

In this part, we present a multi-step inertial GFB algorithm. Recall from Section 3.4 the problem

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in \big(B + \textstyle\sum_{j=1}^m A_j\big)(x),$$

where $B : \mathcal{H} \to \mathcal{H}$ is β-cocoercive, and $A_j : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone for each j. The corresponding fixed-point operator if GFB method, given in (3.4.3), that

$$\mathcal{F}_{\text{GFB}} = \tfrac{1}{2}(\mathcal{R}_{\boldsymbol{\gamma A}}\mathcal{R}_{\boldsymbol{\mathcal{S}}} + \mathbf{Id})\big(\mathbf{Id} - \gamma \boldsymbol{B}_{\boldsymbol{\mathcal{S}}}\big). \tag{4.3.3}$$

Similarly to the argument used for deriving MiDR, by virtue of Lemma 3.4.1, there exists a maximal monotone operator $\boldsymbol{A}' : \boldsymbol{\mathcal{H}} \rightrightarrows \boldsymbol{\mathcal{H}}$ such that $\jmath_{\boldsymbol{A}'} = \frac{1}{2}(\mathcal{R}_{\boldsymbol{\gamma A}}\mathcal{R}_{\boldsymbol{\mathcal{S}}} + \mathbf{Id})$. Then given $\boldsymbol{z}^\star \in \operatorname{fix}(\mathcal{F}_{\text{GFB}})$, we have

$$\boldsymbol{z}^\star = \mathcal{F}_{\text{GFB}}(\boldsymbol{z}^\star) = (\mathbf{Id} + \boldsymbol{A}')^{-1}(\mathbf{Id} - \gamma \boldsymbol{B}_{\boldsymbol{\mathcal{S}}})(\boldsymbol{z}^\star),$$

which means that GFB is equivalent to solve the following monotone inclusion problem on the product space $\boldsymbol{\mathcal{H}}$,

$$\text{find } \boldsymbol{z} \in \boldsymbol{\mathcal{H}} \text{ such that } 0 \in (\boldsymbol{A}' + \gamma \boldsymbol{B}_{\boldsymbol{\mathcal{S}}})(\boldsymbol{z}), \tag{4.3.4}$$

where $\gamma \boldsymbol{B}_{\boldsymbol{\mathcal{S}}}$ is (β/γ)-cocoercive, and $\beta/\gamma \in]1/2, +\infty[$ since the GFB chooses $\gamma \in]0, 2\beta[$.

To this end, apply the MUSTARD algorithm to the monotone inclusion (4.3.4) with constant step-size 1, we get the following iteration,

$$\begin{aligned} \boldsymbol{y}_{a,k} &= \boldsymbol{z}_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(\boldsymbol{z}_{k-i} - \boldsymbol{z}_{k-i-1}), \\ \boldsymbol{y}_{b,k} &= \boldsymbol{z}_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(\boldsymbol{z}_{k-i} - \boldsymbol{z}_{k-i-1}), \\ \boldsymbol{z}_{k+1} &= (\mathbf{Id} + \boldsymbol{A}')^{-1}\big(\boldsymbol{y}_{a,k} - \gamma \boldsymbol{B}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{y}_{b,k})\big). \end{aligned} \tag{4.3.5}$$

Translate the above inertial scheme to the original monotone inclusion problem on $B + \sum_j A_j$, we obtain a multi-step inertial GFB method described in Algorithm 9.

Remark 4.3.2.

(i) When the cocoercive operator vanishes, *i.e.* $B = 0$, then Algorithm 9 becomes the multi-step inertial DR splitting method in the product space.

(ii) As for the inertial DR method, since the operator $\boldsymbol{A}'$ is implicit, it is difficult to put in the metric for Algorithm 9. In [149], a preconditioned GFB method is proposed for which the structure of fixed-point operator (4.3.3) is kept, hence we can adapt the inertial scheme to the that method.

Inertial Forward–Douglas–Rachford Recall the Forward–Douglas–Rachford splitting (3.4.7) of [40] introduced in Section 3.4,

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A + B + N_{\mathcal{S}})(x),$$

Algorithm 9: A multi-step inertial generalized Forward–Backward (MiGFB)

Input: $(\omega_j)_j \in]0,1[^m$ such that $\sum_{j=1}^m \omega_j = 1$, $\gamma \in]0, 2\beta[$. $s \in \mathbb{N}_+$ and $\mathcal{S} = \{0, ..., s-1\}$.

Initial: $k = 0$, $\boldsymbol{z}_0 = (z_{j,0})_j \in \boldsymbol{\mathcal{H}}$, $\boldsymbol{z}_{-i} = \boldsymbol{z}_0, i \in \mathcal{S}$;

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}}, (b_{i,k})_{i\in\mathcal{S}} \in]-1, 2]^s$:

$$y_{a,j,k} = z_{j,k} + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(z_{j,k-i} - z_{j,k-i-1}),$$
$$x_{a,k} = \textstyle\sum_{j=1}^m \omega_j y_{a,j,k}, \quad x_{b,k} = x_k + \sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}),$$

for $j = 1$ **to** m **do**

$$z_{j,k+1} = y_{a,j,k} + \jmath_{\frac{\gamma}{\omega_j}A_j}\big(2x_{a,k} - y_{a,j,k} - \gamma_k B(x_{b,k})\big) - x_{a,k},$$

$$x_{k+1} = \textstyle\sum_{j=1}^m \omega_j z_{j,k+1};$$

$k = k + 1$;

until *convergence*;

where A is maximal monotone, B is β-cocoercive, $\mathcal{S}$ is a closed vector subspace of $\mathcal{H}$, and $\operatorname{zer}(A + B + N_{\mathcal{S}}) \neq \emptyset$. Specializing the above result to this problem, we obtain the following multi-step inertial Forward–Douglas–Rachford splitting method,

$$y_{a,k} = z_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(z_{k-i} - z_{k-i-1}),$$
$$y_{b,k} = z_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(z_{k-i} - z_{k-i-1}),$$
$$z_{k+1} = \frac{1}{2}(\mathrm{Id} + \mathcal{R}_{\gamma A}\mathcal{R}_{N_{\mathcal{S}}})\big(y_{a,k} - \gamma\mathcal{P}_{\mathcal{S}}B\mathcal{P}_{\mathcal{S}}(y_{b,k})\big).$$

Whence we get $z^\star$, and recover a solution of the problem $x^\star = \mathcal{P}_{\mathcal{S}}(z^\star)$.

4.3.3 Inertial Primal–Dual splitting

The last extension we present is the multi-step inertial Primal–Dual splitting methods. This extension can be applied to the Primal–Dual splitting methods proposed in [51, 173, 66, 62], but is not limited to them.

Recall the problem 2.4.13, the primal formulation is

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A + B)(x) + L^*\big((C\square D)(Lx)\big),$$

where $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone, $B : \mathcal{H} \to \mathcal{H}$ is β_B-cocoercive. $C, D : \mathcal{G} \rightrightarrows \mathcal{G}$ are maximal monotone, and D is β_D-strongly monotone. $L : \mathcal{H} \to \mathcal{G}$ is a bounded linear operator.

Recall that the Forward–Backward structure (2.4.17) of the Primal–Dual splitting method in Section 2.4.3.3. The extension of the MUSTARD algorithm to (8.1.3) is rather straightforward, and we get the following multi-step inertial iteration

$$\boldsymbol{y}_{a,k} = \boldsymbol{z}_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(\boldsymbol{z}_{k-i} - \boldsymbol{z}_{k-i-1}),$$
$$\boldsymbol{y}_{b,k} = \boldsymbol{z}_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(\boldsymbol{z}_{k-i} - \boldsymbol{z}_{k-i-1}),$$
$$\boldsymbol{z}_{k+1} = (\mathbf{Id} + \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{A})^{-1}\big(\boldsymbol{y}_{a,k} - \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{B}(\boldsymbol{y}_{b,k})\big).$$

The detailed iteration corresponding to the above scheme is given in Algorithm 10.

Algorithm 10 can also be adapted to the case of multiple parallel sums as the problem (3.4.15) discussed in Section 3.4, since it is rather straightforward, we shall skip the detailed description here.

Remark 4.3.3.

(i) When $s = 1$ and $b_{0,k} \equiv a_{0,k}$, the above inertial scheme recovers the inertial Primal–Dual splitting method proposed in [120].

Algorithm 10: A multi-step inertial Primal–Dual (MiPD)

Initial: $s \in \mathbb{N}_+$ and $\mathcal{S} = \{0, ..., s-1\}$, $x_{-i} = x_0, i \in \mathcal{S}$.

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}}, (b_{i,k})_{i\in\mathcal{S}} \in]-1, 2]^s$,

$$\left\lfloor \begin{aligned} y_{x,a,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}),\\ y_{v,a,k} &= v_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(v_{k-i} - v_{k-i-1}),\\ y_{x,b,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}),\\ y_{v,b,k} &= v_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(v_{k-i} - v_{k-i-1}),\\ x_{k+1} &= \mathcal{J}_{\gamma_A A}(y_{x,a,k} - \gamma_A B(y_{x,b,k}) - \gamma_A L^* y_{v,a,k}),\\ \bar{x}_{k+1} &= 2x_{k+1} - y_{x,a,k},\\ v_{k+1} &= \mathcal{J}_{\gamma_C C^{-1}}(y_{v,a,k} - \gamma_C D^{-1}(y_{v,b,k}) + \gamma_C L\bar{x}_{k+1}), \end{aligned} \right. \tag{4.3.6}$$

$k = k + 1$;

until *convergence*;

(ii) If $B = 0$ and $D = 0$, then Algorithm 10 becomes an inertial PPA under different metric.

4.4 Multi-step inertial FB for non-convex optimization

In this section, we consider generalizing the MUSTARD algorithm to the non-convex optimization problems, *i.e.* the sum of two non-necessarily convex functions, one of which is proper lower semi-continuous while the other is differentiable with a Lipschitz continuous gradient. For the sake of brevity, we abuse the notations in ($\mathcal{P}_{\text{opt}}$)

$$\min_{x\in\mathcal{H}} \big\{\Phi(x) \overset{\text{def}}{=} R(x) + F(x)\big\}, \tag{$\mathcal{P}_{\text{ncvx}}$}$$

where now

(**A.7**) $R : \mathcal{H} \to \overline{\mathbb{R}}$ is proper lower semi-continuous (lsc), and bounded from below.

(**A.8**) $F : \mathcal{H} \to \mathbb{R}$ is finite-valued, differentiable, and ∇F is $(1/\beta)$-Lipschitz continuous.

As we can see, no convexity is imposed neither on F nor on R.

4.4.1 The ncvx-MiFB algorithm

Different from the convex case, when R is non-convex, then its proximity operator $\text{prox}_R(x)$ in general is a set-valued mapping or even empty if R is not well defined (*e.g.* not bounded from below). Nevertheless, the FB splitting for solving ($\mathcal{P}_{\text{ncvx}}$) reads

$$x_{k+1} \in \text{prox}_{\gamma_k R}\big(x_k - \gamma_k \nabla F(x_k)\big), \tag{4.4.1}$$

where $\gamma_k > 0$ is the step-size. $\text{prox}_{\gamma R}(x)$ is non-empty under (**A.7**) and is set-valued in general. Lower-boundedness of R can be relaxed by requiring *e.g.* coercivity of the objective in Definition 2.1.8.

We propose the following multi-step inertial Forward–Backward splitting algorithm (Algorithm 11), which we call as "ncvx-MiFB".

Remark 4.4.1. Though Algorithm 11 is very close to the MUSTARD algorithm, it should be pointed out that it is *not* a simple extension of MUSTARD to the non-convex case with $\mathcal{V}_k \equiv \text{Id}$. There are several reasons justifying this claim:

Algorithm 11: Multi-step Inertial Forward–Backward (ncvx-MiFB)

Initial: $s \in \mathbb{N}_+$ and $\mathcal{S} \overset{\text{def}}{=} \{0, ..., s-1\}$. $x_0 \in \mathcal{H}$, $x_{-i} = x_0, i \in \mathcal{S}$. Choose $\bar{\epsilon}, \underline{\epsilon} > 0$ such that $0 < \underline{\gamma} \leq \gamma_k \leq \bar{\gamma} < \beta$.

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}}, \{b_{i,k}\}_{i\in\mathcal{S}} \in]-1, 2]^s$ and $\gamma_k \in [\underline{\gamma}, \bar{\gamma}]$:

$$\left\lfloor \begin{array}{l} y_{a,k} = x_k + \sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}), \\ y_{b,k} = x_k + \sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}), \\ x_{k+1} \in \operatorname{prox}_{\gamma_k R}\big(y_{a,k} - \gamma_k \nabla F(y_{b,k})\big). \end{array} \right. \tag{4.4.2}$$

$k = k + 1$;

until *convergence*;

(i) Many convex functions satisfy the Kurdyka-Łojasiewicz property, though not all of them. For example, in [30, Section 4.3] the authors designed a counter-example where the function is convex but not KL.

(ii) The convergence results of Section 4.2 for the MUSTARD algorithm can not be applied to the non-convex case, since the sub-differentials of non-convex functions are no longer maximal monotone. Hence a new convergence result is needed.

As a consequence, it necessary and important to study the convergence property of multi-step inertial Forward–Backward splitting to the case of non-convex optimization.

Related work The convergence property of (4.4.1) was first established in [10] under the assumption that the objective Φ satisfies the Kurdyka-Łojasiewicz (KL) property (Definition 4.4.2). Following their footprints, [33, 136] established convergence of the special inertial schemes of [131] in the non-convex setting. In [81], the authors also proposed an variable metric Forward–Backward splitting for the non-convex optimization. Our proposed method can also handle the variable metric, however for the same of simplicity, we choose to not include it here.

4.4.2 Global convergence of ncvx-MiFB

Let $R : \mathcal{H} \to \overline{\mathbb{R}}$ be a proper lsc function. For η_1, η_2 such that $-\infty < \eta_1 < \eta_2 < +\infty$, define the set

$$[\eta_1 < R < \eta_2] \overset{\text{def}}{=} \{x \in \mathcal{H} : \eta_1 < R(x) < \eta_2\}.$$

The definition below is taken from [9].

Definition 4.4.2 (Kurdyka-Łojasiewicz property). $R : \mathcal{H} \to \overline{\mathbb{R}}$ is said to have the Kurdyka-Łojasiewicz property at $\bar{x} \in \operatorname{dom}(R)$ if there exists $\eta \in]0, +\infty]$, a neighbourhood U of $\bar{x}$ and a continuous concave function $\varphi : [0, \eta[\to \mathbb{R}_+$ such that

(i) $\varphi(0) = 0$, φ is C^1 on $]0, \eta[$, and for all $t \in]0, \eta[$, $\varphi'(t) > 0$.

(ii) for all $x \in U \cap [R(\bar{x}) < R < R(\bar{x}) + \eta]$, the Kurdyka-Łojasiewicz inequality holds

$$\varphi'\big(R(x) - R(\bar{x})\big)\operatorname{dist}\big(0, \partial R(x)\big) \geq 1. \tag{4.4.3}$$

Proper lsc functions which satisfy the Kurdyka-Łojasiewicz property at each point of $\operatorname{dom}(\partial R)$ are called KL functions.

Roughly speaking, KL functions become sharp up to reparameterization via φ, called a desingularizing function for R. Among real-extended-valued lower semi-continuous functions, typical KL functions are semi-algebraic functions or more generally functions definable in an o-minimal structure, see

[27, 29, 28, 30]. References on functions definable in an o-minimal structure can be found in [68, 171]. For instance, the ℓ_0 pseudo-norm and the rank function are indeed KL, see Section 5.2 for more details about these two functions.

Global convergence of ncvx-MiFB In this part we present the global convergence of the sequence generated by Algorithm 11. Let $\mu, \nu > 0$ be two constants. For $i \in \mathcal{S}$ and $k \in \mathbb{N}$, define the following quantities,

$$\zeta_k \overset{\text{def}}{=} \frac{1-\gamma_k/\beta-\mu-\nu\gamma_k}{2\gamma_k},\ \underline{\zeta} \overset{\text{def}}{=} \inf_{k\in\mathbb{N}} \zeta_k \quad \text{and} \quad \lambda_{i,k} \overset{\text{def}}{=} \frac{sa_{i,k}^2}{2\gamma_k\mu} + \frac{sb_{i,k}^2}{2\nu\beta^2},\ \bar{\lambda}_i \overset{\text{def}}{=} \sup_{k\in\mathbb{N}} \lambda_{i,k}. \tag{4.4.4}$$

Theorem 4.4.3 (Convergence of ncvx-MiFB (Algorithm 11)). *For problem ($\mathcal{P}_{\text{ncvx}}$), suppose that (A.7)-(A.8) hold. If moreover Φ is a proper lsc KL function, and R is bounded from below. For Algorithm 11, choose $\mu, \nu, \gamma_k, a_{i,k}, b_{i,k}$ such that*

$$\tau \overset{\text{def}}{=} \underline{\zeta} - \textstyle\sum_{i\in\mathcal{S}} \bar{\lambda}_i > 0. \tag{4.4.5}$$

Then each bounded sequence $\{x_k\}_{k\in\mathbb{N}}$ satisfies

(i) *$\{x_k\}_{k\in\mathbb{N}}$ has finite length, i.e. $\sum_{k\in\mathbb{N}} \|x_k - x_{k-1}\| < +\infty$.*

(ii) *There exists a critical point $x^\star \in \operatorname{crit}(\Phi)$ such that $\lim_{k\to+\infty} x_k = x^\star$.*

(iii) *If Φ has the KL property at a global minimizer $x^\star$, then starting sufficiently close from $x^\star$, any sequence $\{x_k\}_{k\in\mathbb{N}}$ converges to a global minimum of Φ and satisfies (i).*

See from page 92 for the detailed proof.

Remark 4.4.4.

(i) Boundedness of the sequence is automatically ensured under standard assumptions such as coercivity of Φ.

(ii) As long as (4.4.5) is satisfied, Theorem 4.4.3 holds true for arbitrary choice of $s \in \mathbb{N}_+$, which is on the contrary to the convergence result of the MUSTARD algorithm, as we only have conditional convergence for $s \geq 2$ (Theorem 4.2.3).

(iii) When $a_{i,k} \equiv 0$ and $b_{i,k} \equiv 0$, *i.e.* the case of FB splitting, condition (4.4.5) holds naturally as long as $\bar{\gamma} < \beta$ which recovers the case of [10].

From (4.4.4) and (4.4.5), we conclude the following results for the inertial parameters, see also Figure 4.3 for a graphical illustration:

(i) $s = 1$: let $b_{0,k} \equiv b$ and $a_{0,k} \equiv a$, then (4.4.5) implies that,

$$\frac{a^2}{2\underline{\gamma}\mu} + \frac{b^2\beta^2}{2\nu} < \underline{\zeta} = \frac{1-\bar{\gamma}/\beta-\mu-\nu\bar{\gamma}}{2\bar{\gamma}},$$

which is an ellipsoid. See Figure 4.3 (a).

(ii) When $s \geq 2$, for each $i \in \mathcal{S}$, let $b_{i,k} = a_{i,k} \equiv a_i$ (*i.e.* constant symmetric inertial parameters), then (4.4.5) indicates

$$\left(\frac{1}{2\underline{\gamma}\mu} + \frac{\beta^2}{2\nu}\right) \textstyle\sum_{i\in\mathcal{S}} a_i^2 < \underline{\zeta},$$

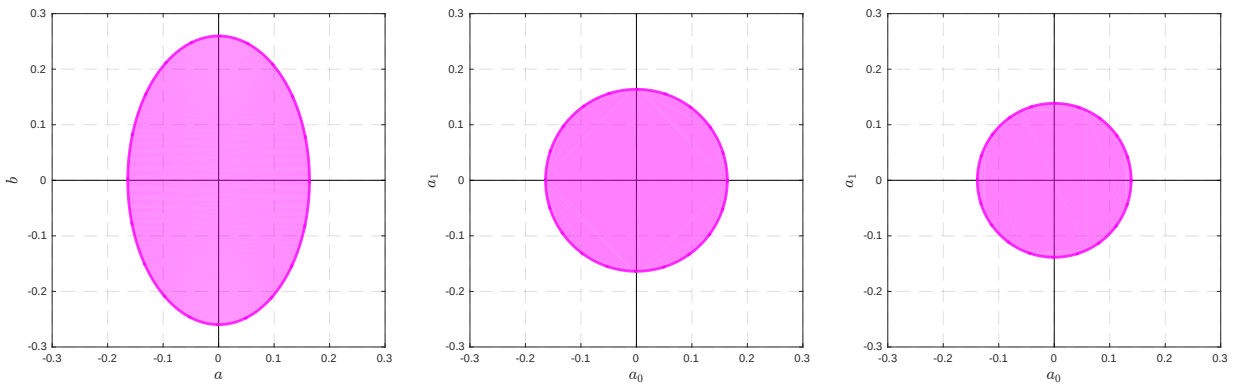
which is a ball. Similar result if $b_{i,k} \equiv 0$. See Figure 4.3 (b) and (c) for these two situations.

An empirical approach for inertial parameters Similarly to the MUSTARD algorithm, we also have an empirical bound for the inertial parameters. Consider the parameters setting:

$$\gamma_k \equiv \gamma \in]0, \beta[\quad \text{and} \quad b_{i,k} = a_{i,k} \equiv a_i \in]-1, 2[, i \in \mathcal{S}.$$

We have the following empirical bound for the summand $\sum_{i\in\mathcal{S}} a_i$:

$$\textstyle\sum_i a_i \in \left]0, \min\left\{1, \frac{\beta-\gamma}{|2\gamma-\beta|}\right\}\right[. \tag{4.4.6}$$



(a) $s = 1$ (b) $s = 2 : b_{i,k} \equiv 0$ (c) $s = 2 : b_{i,k} \equiv a_{i,k}$

Figure 4.3: Permitted inertial parameters by (4.4.5), where we set $\beta = 1, \mu = 0.3, \nu = 0.3$, $\gamma_k \equiv 0.4\beta$ and $\tau = 10^{-3}$.

To ensure the convergence $\{x_k\}_{k\in\mathbb{N}}$, an online updating rule should be applied together with the empirical bound. More precisely, choose a_i according to (4.4.6). Then for each $k \in \mathbb{N}$, let $b_{i,k} = a_{i,k}$ and choose $a_{i,k}$ such that

$$\sum_i a_{i,k} = \min\Big\{\sum_i a_i, c_k\Big\}, \tag{4.4.7}$$

where $c_k > 0$ is such that $\{c_k \sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|\}_{k\in\mathbb{N}}$ is summable. For instance, one can choose $c_k = \frac{c}{k^{1+q}\sum_{i\in I}\|x_{k-i}-x_{k-i-1}\|}, c > 0, q > 0$.

Remark 4.4.5. The allowed choices of the summand $\sum_i a_i$ by (4.4.6) is larger than those of Theorem 4.4.3. For instance, (4.4.6) allows $\sum_i a_i = 1$ for $\gamma \in]0, \frac{2}{3}\beta]$. While for Theorem 4.4.3, $\sum_i a_i = 1$ can be reached only when $\gamma \to 0$.

4.5 Numerical experiments

In this section, we consider several concrete examples of linear inverse problems to demonstrate the performance of the proposed algorithms. First for MUSTARD, 3 extensions of the algorithm are considered: multi-step inertial Forward–Backward (MiFB, *i.e.* MUSTARD with $\mathcal{V}_k \equiv \mathrm{Id}$), multi-step inertial Douglas–Rachford (MiDR) and a multi-step inertial Primal–Dual splitting. The ncvx-MiFB algorithm is demonstrated at the end of the section.

4.5.1 Inertial Forward–Backward splitting

Recall the linear inverse problem (3.6.2) in Section 3.6. For the forward observation model (3.6.1), let $\mathcal{K}$ be generated from the standard random Gaussian ensemble, and w have bounded ℓ_2-norm. Moreover let $m = 1$, then (3.6.2) becomes

$$\min_{x\in\mathbb{R}^n} R(x) + \frac{1}{2}\|\mathcal{K}x - f\|^2. \tag{4.5.1}$$

In this experiment, 4 different cases of R are considered: sparsity promoting ℓ_1-norm, group sparsity promoting $\ell_{1,2}$-norm, anti-sparsity ℓ_∞-norm, and low-rank promoting nuclear norm. The detailed settings of each example are

ℓ_1-norm $(m, n) = (48, 128)$, $\|x_{\mathrm{ob}}\|_0 = 8$, *i.e.* x_{ob} has 8 non-zero elements.
$\ell_{1,2}$-norm $(m, n) = (48, 128)$, x_{ob} has 3 non-zero blocks of size 4.
ℓ_∞-norm $(m, n) = (63, 64)$, $|\mathcal{I}_{x_{\mathrm{ob}}}| = 8$ where $\mathcal{I}_{x_{\mathrm{ob}}} = \{i : |x_i| = \|x_{\mathrm{ob}}\|_\infty\}$.
Nuclear norm $(m, n) = (640, 1024)$, $x_{\mathrm{ob}} \in \mathbb{R}^{32\times 32}$ and $\mathrm{rank}(x_{\mathrm{ob}}) = 4$.

The proximity operators of all the above functions can be computed efficiently, and more properties of these functions, together with total variation, are presented in Chapter 5.

We apply the multi-step inertial Forward–Backward splitting (MiFB) to solve problem (4.5.1). To highlight the performance of MiFB, we present the following comparisons

- Comparison among FB, FISTA and MiFB for $s = 1, 2$.
- The effects of negative inertial parameters for $s = 2$.

For the FB and MiFB, two fixed choices of step-size are considered: $\gamma = \beta$ and $\gamma = 1.5\beta$. Then for the FISTA, the sequence convergent one [50] is applied applied, *i.e.* the inertial parameter is computed through

$$t_k = \frac{k+d-1}{d},\ d > 2 \quad \text{and} \quad a_k = \frac{t_{k-1}-1}{t_k} = \frac{k-1}{k+d},$$

and we set $d = 3$ for the numerical experiments.

Comparison of FB, FISTA and MiFB For the MiFB, we consider the symmetric inertial parameters setting, *i.e.* $b_i = a_i, i = 0, ..., s-1$, and moreover

(i) $s = 1$:
- For $\gamma = \beta$, choose $b_0 = a_0 = \sqrt{5} - 2 - 10^{-3}$ such that Theorem 4.2.9 applies, and the sequence is guaranteed to be convergent.
- For $\gamma = 1.5\beta$, choose a_0 according to the empirical bounds (4.2.9) and the online updating rule (4.2.6).

(ii) $s = 2$: only $\gamma = 1.5\beta$ is considered, and choose a_0, a_1 according to the empirical bounds (4.2.9) and the online updating rule (4.2.6).

The comparisons of the methods are presented below in Figure 4.4, the MiFB with $s = 1, 2$ are marked as "1-MiFB" and "2-MiFB" respectively. "Opt-1" and "Opt-2" repensent the "upper bound 1" and "upper bound 2" of (4.2.9) respectively. We can conclude the following observations:

(i) Under the same choice of γ, MiFB is faster than FB.

(ii) For the case $\gamma = \beta$, FISTA is the slowest if very high precision of the solution is needed (higher than 10^{-8} for instance), and the original FISTA [24] has almost the same performance as $d = 3$ of FISTA [50]. This is mainly due to the fact the inertial parameter of FISTA tends to 1 as $k \to +\infty$. Similarly for the MiFB, if the value of $\sum_i a_i$ is too close to 1, MiFB will also be very slow and oscillate locally. Such behaviour of too big inertial value will be analyzed at length in Chapter 6.

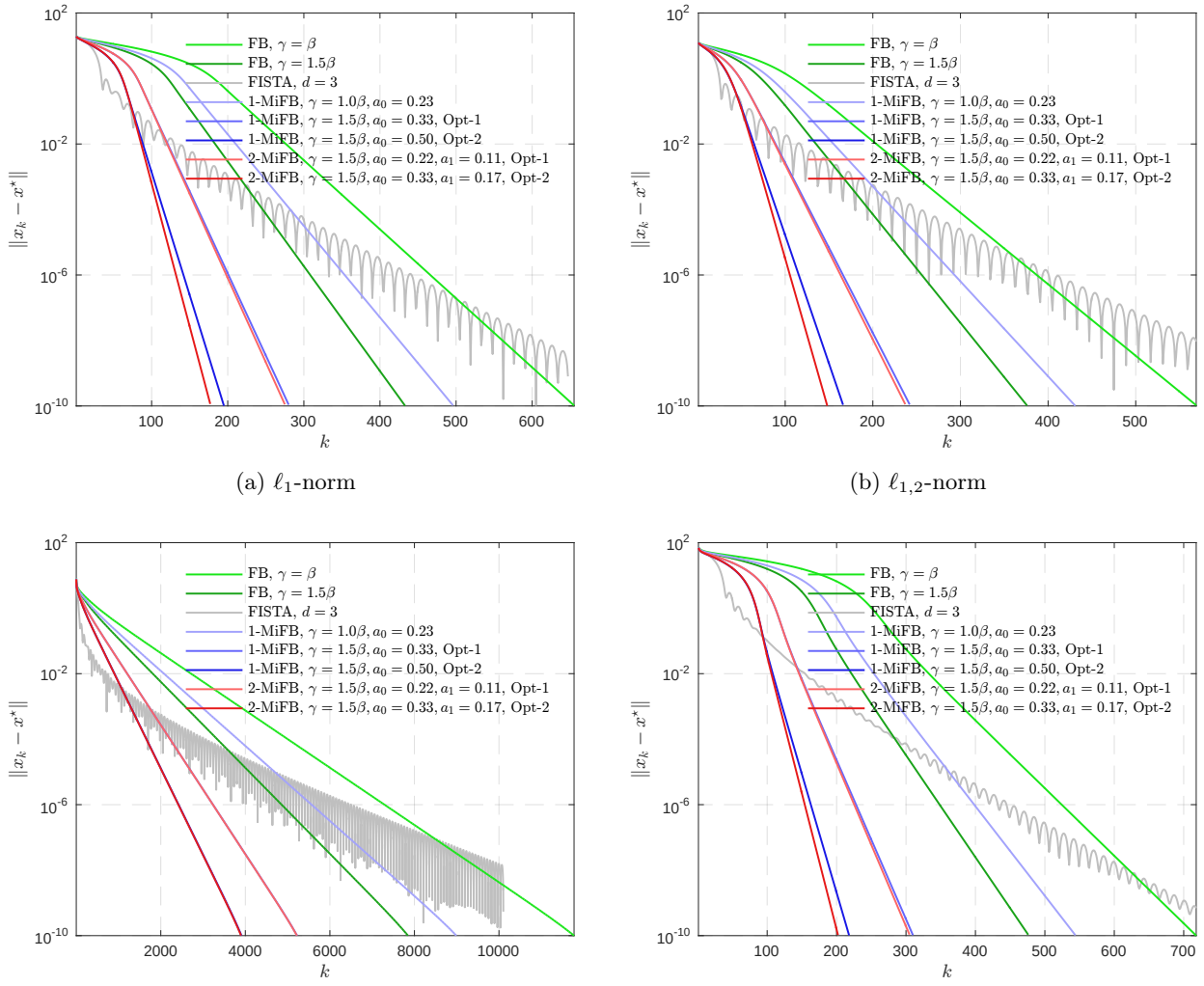
(iii) For the MiFB, under the same value of $\sum_i a_i$, "1-MiFB" and "2-MiFB" have very similar performance, with "2-MiFB" being the slightly faster one.

Negative inertial parameter In this part we compare the effects of negative inertial parameters for $s = 2$. For this experiment, we will present only for ℓ_1-norm and nuclear norm, as the other two examples yield very similar observations.

Given the summand $\sum_i a_i$, the effects of negative a_1 depends on the oscillation of the sequence $\{\|x_k - x^\star\|\}_{k\in\mathbb{N}}$ (or $\{\|x_k - x_{k-1}\|\}_{k\in\mathbb{N}}$). If $\{\|x_k - x^\star\|\}_{k\in\mathbb{N}}$ oscillates, then negative a_1 will speed up the convergence, and slow down the convergence otherwise, this is supported by the following two experiments.

For the first experiment, we choose $\gamma = \beta$ and $\sum_i a_i = 0.9$. The result is shown in Figure 4.5. As we can see, oscillation appears for both examples, and negative value a_1 indeed speeds up the convergence. The reasons are that

(i) First, the oscillation is caused by too big inertial value $\sum_i a_i$, which results in a high momentum of $\sum_i a_i(x_{k-i} - x_{k-i-1})$.



(a) ℓ_1-norm

(b) $\ell_{1,2}$-norm

(c) ℓ_∞-norm

(d) Nuclear norm

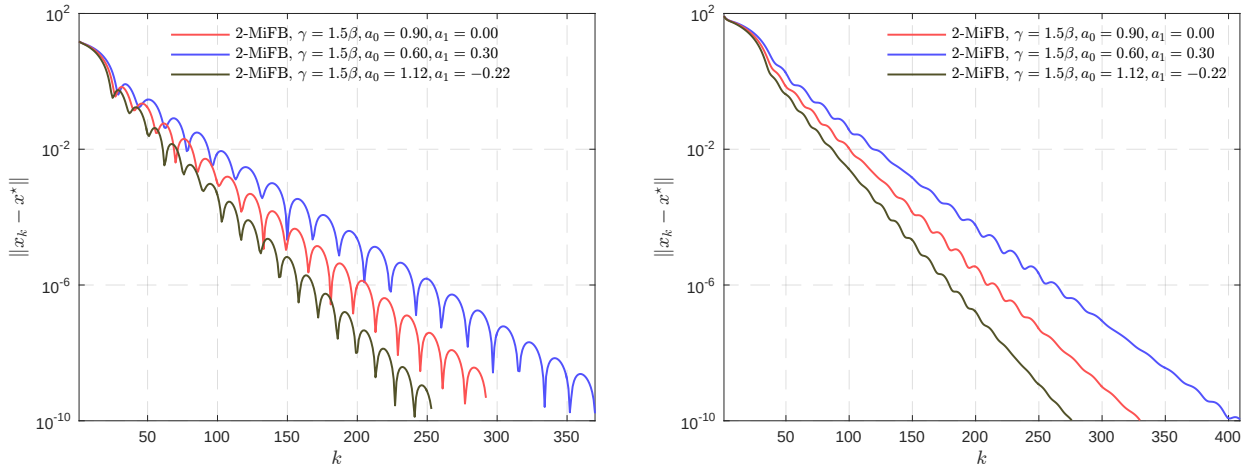
Figure 4.4: Comparison of FB, FISTA and MiFB. "1-MiFB" and "2-MiFB" stand for MiFB with the inertial steps $s = 1$ and $s = 2$, and "Opt-1" and "Opt-2" represents the "upper bound 1" and "upper bound 2" of (4.2.9) respectively. For the ℓ_∞-norm, the almost invisible dark blue line "1-MiFB, $\gamma = 1.5\beta, a_0 = 0.5$, Opt-1" overlaps with the dark red line "2-MiFB, $\gamma = 1.5\beta, a_0 = 0.33, a_1 = 0.17$, Opt-2".

(ii) In general, $x_{k-1} - x_{k-2}$ has higher momentum than $x_k - x_{k-1}$. Therefore, positive a_1 increases the momentum of $\sum_i a_i(x_{k-i} - x_{k-i-1})$ which is already large enough, while negative a_1 reduces the momentum, hence leads to a faster convergence speed.

Adapt the above reasoning, we can the explain the observation obtained from the second experiment, for which we set $\gamma = 1.5\beta$ and $\sum_i a_i = 0.5$. See Figure 4.6 for the plots.

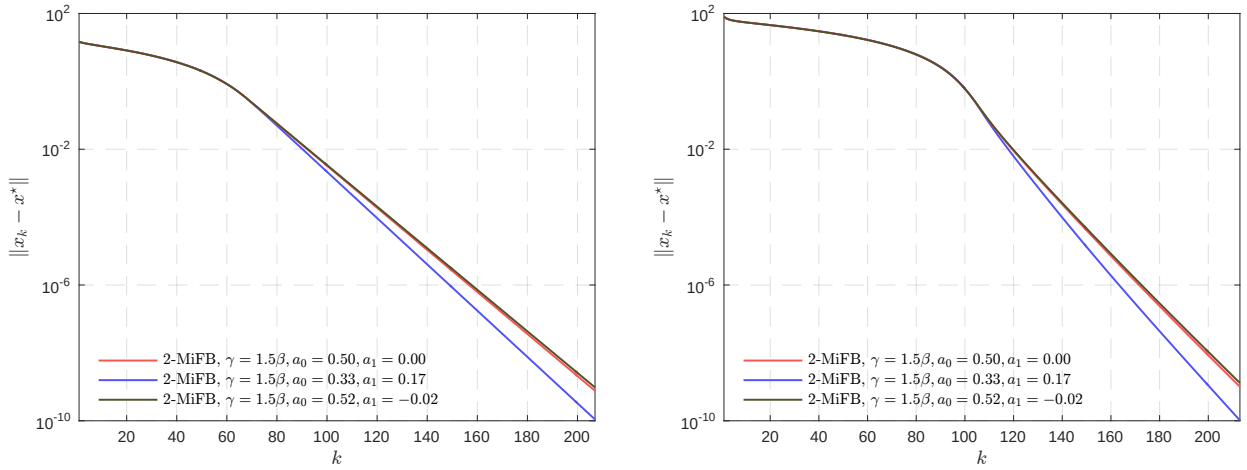
(i) For both the examples, no oscillation behaviours observed, which means that the momentum of $\sum_i a_i(x_{k-i} - x_{k-i-1})$ is not large enough.

(ii) As $x_{k-1} - x_{k-2}$ has higher momentum, then negative a_1 reduces the momentum of $\sum_i a_i(x_{k-i} - x_{k-i-1})$ which is not large enough, hence slow down the convergence.



(a) ℓ_1-norm (b) Nuclear norm

Figure 4.5: Effect of negative inertial parameter, $\gamma = \beta$ and $\sum_i a_i = 0.9$



(a) ℓ_1-norm (b) Nuclear norm

Figure 4.6: Effect of negative inertial parameter, $\gamma = 1.5\beta$ and $\sum_i a_i = 0.5$

4.5.2 Inertial Douglas–Rachford splitting

Now suppose that there is no noise in the observation model (3.6.1), *i.e.* $w = 0$. Then instead of (3.6.2), the following equality constrained problem should be considered

$$\min_{x\in\mathbb{R}^n} R(x) \quad \text{subject to} \quad \mathcal{K}x = \mathcal{K}x_{\text{ob}}.$$

Furthermore, the above constrained problem can be ban be formulated as

$$\min_{x\in\mathbb{R}^n} R(x) + J(x), \tag{4.5.2}$$

where $J = \iota_{\mathcal{S}}(\cdot)$ is the indicator function of the constraint $\mathcal{S} \overset{\text{def}}{=} \{x \in \mathbb{R}^n : \mathcal{K}x_{\text{ob}} = \mathcal{K}x\} = x_{\text{ob}} + \ker(\mathcal{K})$. As both functions R and J are non-smooth, a proper choice to solve (4.5.2) is the Douglas–Rachford splitting. The proximity operator of J is the projection operator onto $\mathcal{S}$, which reads $\text{prox}_{\gamma J}(x) = x + \mathcal{K}^+(f - \mathcal{K}x)$ where $\mathcal{K}^+ = \mathcal{K}^T(\mathcal{K}\mathcal{K}^T)^{-1}$ is the Moore-Penrose pseudo-inverse of $\mathcal{K}$. For R, 4 examples are considered: $\ell_1, \ell_{1,2}, \ell_\infty$-norms and nuclear norm, and the setting of each example is the same as the MiFB experiments.

The multi-step inertial Douglas–Rachford splitting (MiDR) is applied to solve (4.5.2), and the ob-

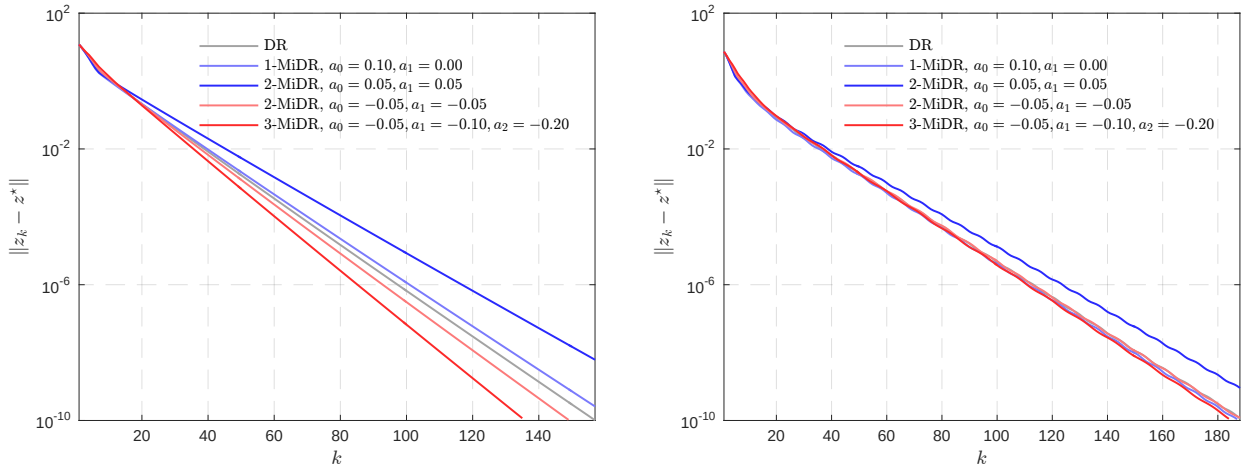
tained result is shown in Figure 4.7, the parameter γ of the DR is set as 1. Unlike the case of MiFB, the gains of inertial to MiDR is rather limited, moreover it depends on the properties of the involved functions.

For the considered examples, J is polyhedral. For R, ℓ_1 and ℓ_∞-norms are polyhedral, while $\ell_{1,2}$-norm and nuclear norm are not. We have the following observation from Figure 4.7:

(i) When both functions J, R are polyhedral, negative $\sum_i a_i$ gives faster convergence, and the 3-step inertial DR "3-MiDR" is the fastest, though the advantage is not big. See Figure 4.7 (a) and (c).

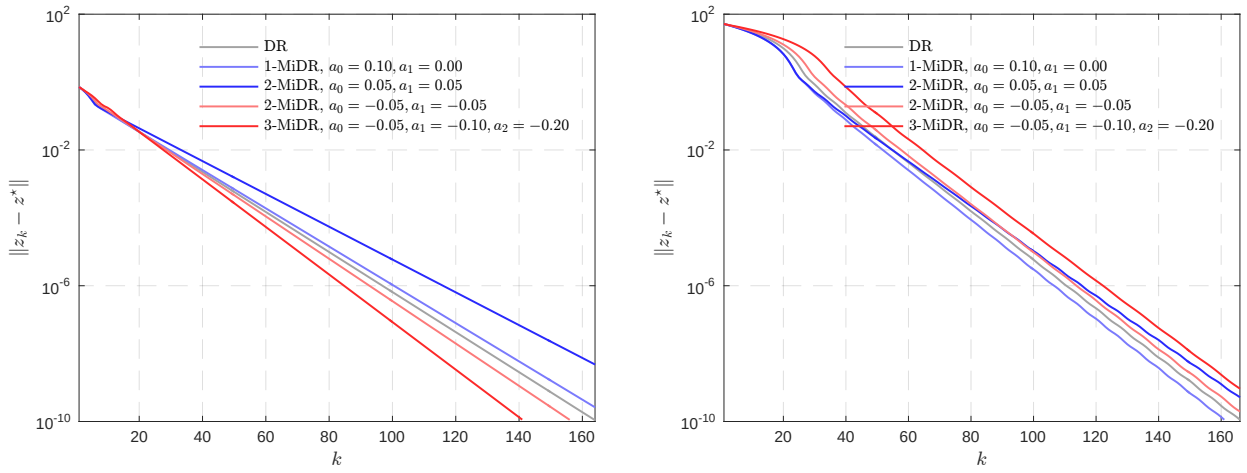
(ii) When R is not polyhedral (*e.g.* $\ell_{1,2}$-norm and nuclear norm), the performance of different inertial settings are very close, with "1-MiDR" being the slightly faster one. See Figure 4.7 (b) and (d).

Conclude from the above comparison, in practice the inertial parameters for MiDR should be chosen based on the properties of the involved functions, and the (absolute) value of the inertial parameters should not be too big. A theoretical explanation about this behaviour is presented in Chapter 7.



(a) ℓ_1-norm

(b) $\ell_{1,2}$-norm



(c) ℓ_∞-norm

(d) Nuclear norm

Figure 4.7: Comparison of DR and MiDR for problem $\min_{x\in\mathbb{R}^n} R(x)$ s.t. $\mathcal{K}x = \mathcal{K}x_{\text{ob}}$ with R being $\ell_1, \ell_{1,2}, \ell_\infty$-norms and nuclear norm. The parameter γ in DR algorithm is set to be 1 for all examples.

4.5.3 Inertial Primal–Dual splitting

In this section, we consider the same problems as in MiDR experiments, and solve it with multi-step inertial Primal–Dual splitting (MiPD). By reformulating the constraint as $J(x) = \iota_{\{0\}}(\mathcal{K}x - f), f =$

$\mathcal{K}x_{\text{ob}}$, we obtain the following saddle-point problem,

$$\min_{x\in\mathbb{R}^n}\max_{v\in\mathbb{R}^m} R(x) + \langle \mathcal{K}x - b,\, v\rangle - \iota^*_{\{0\}}(v), \tag{4.5.3}$$

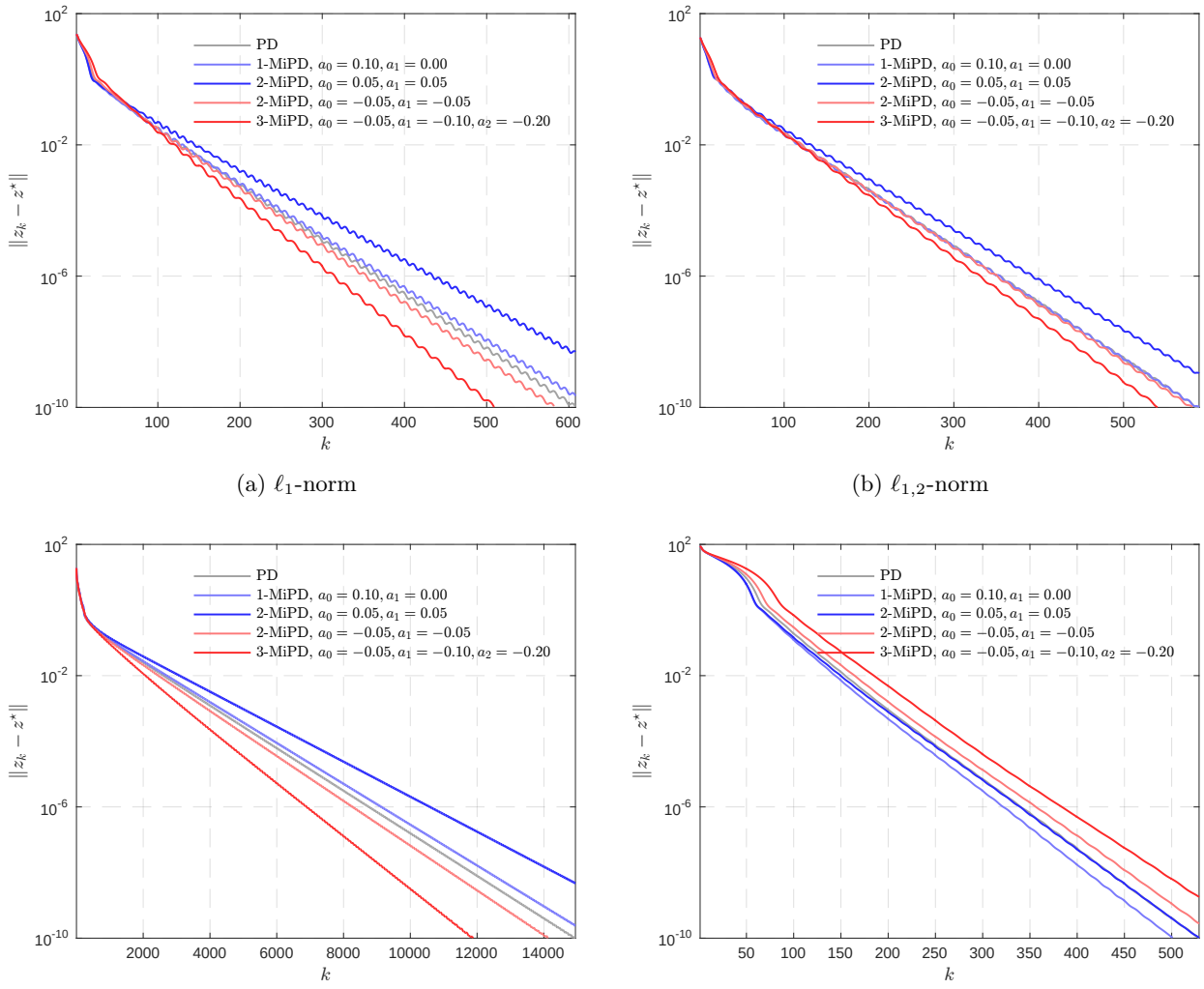
where $\iota^*_{\{0\}}(\cdot)$ is the support function of $\{0\}$. The Chambolle-Pock Primal–Dual splitting method [51] is applied, and for this case, the 1-step inertial Primal–Dual splitting coincides with the one proposed in [120]. Moreover, for the step-sizes γ_A, γ_C in Algorithm 10, we choose $\gamma_A\gamma_C\|\mathcal{K}\|^2 = 0.99$ and $\gamma_A = \gamma_C$.

The experiment results are shown in Figure 4.8, from which we obtain the following observations which are almost the same as MiDR experiment:

(i) When both functions J, R are polyhedral, then negative $\sum_i a_i$ gives faster convergence, and the "3-MiPD" is the fastest one. See Figure 4.8 (a) and (c).

(ii) When R is not polyhedral (*e.g.* $\ell_{1,2}$-norm and nuclear norm), the performance of different inertial parameters are very close, with "1-MiPD" being slightly faster. See Figure 4.8 (b) and (d).

A theoretical explanation about this behaviour is presented in Chapter 8.



(a) ℓ_1-norm

(b) $\ell_{1,2}$-norm

(c) ℓ_∞-norm

(d) Nuclear norm

Figure 4.8: Comparison of Primal–Dual (PD) and MiPD for solving (4.5.3) with R being $\ell_1, \ell_{1,2}, \ell_\infty$-norms and nuclear norm.

4.5.4 The non-convex case

For this demonstration of ncvx-MiFB, we consider the linear inverse problem (4.5.1) with R being ℓ_0 pseudo-norm (*i.e.* counting number of non-zero of a given vector $x \in \mathbb{R}^n$) and the rank function. The detailed settings of the two examples are

ℓ_0 **pseudo-norm** $(m, n) = (48, 128)$, $\|x_{\text{ob}}\|_0 = 8$, *i.e.* x_{ob} has 8 non-zero elements.
Rank function $(m, n) = (640, 1024)$, $x_{\text{ob}} \in \mathbb{R}^{32\times 32}$ and $\text{rank}(x_{\text{ob}}) = 4$.

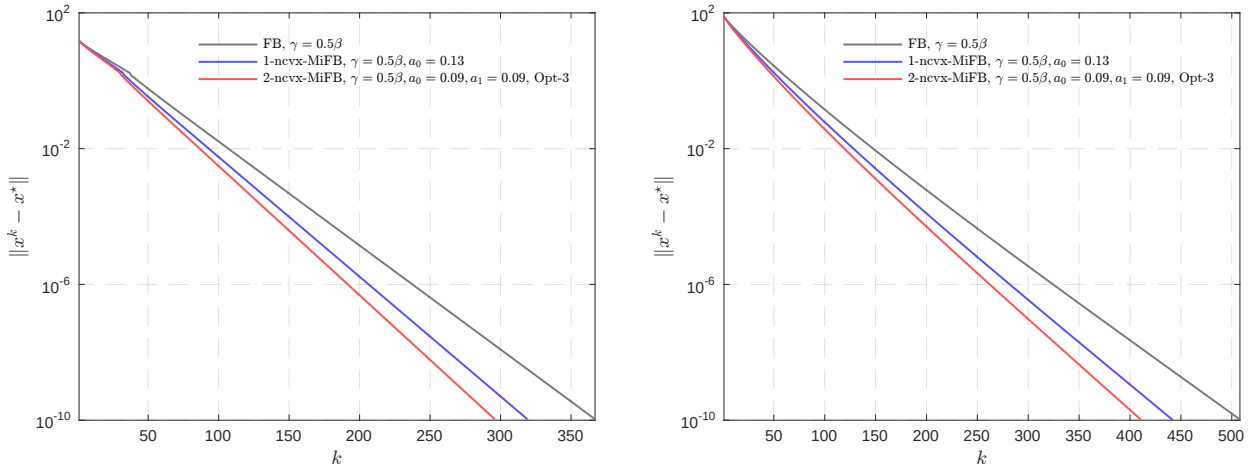
The proximity operator of ℓ_0 pseudo-norm is given by hard-thresholding, while the proximity operator of the rank function is applying hard-thresholding to the singular values.

For the numerical experiments, 3 different schemes are compared: FB, 1-step ncvx-MiFB (denoted as "1-ncvx-MiFB") and 2-step ncvx-MiFB (denoted as "2-ncvx-MiFB"). Only one choice of γ is considered $\gamma_k \equiv 0.5\beta$, and the inertial parameters of ncvx-MiFB are chosen as $b_k = a_k \equiv a_i, i = 0, 1$ such that Theorem 4.4.3 applies. The reason of choosing only one γ is that, $\{x_k\}_{k\in\mathbb{N}}$ in general converges to different critical points under different values of γ, hence makes it difficult to compare the effects of inertia between different values of γ.

The numerical result is shown in Figure 4.9, from which we obtain very similar observations as the ones from the convex case (*i.e.* Figure 4.4):

(i) The ncvx-MiFB scheme is faster than FB both globally and locally.

(ii) Comparing the two ncvx-MiFB inertial schemes, "2-ncvx-MiFB" outperforms "1-ncvx-MiFB", showing the advantages of a 2-step inertial scheme over the 1-step one.



(a) ℓ_0 pseudo-norm (b) Rank function

Figure 4.9: Comparison of Forward–Backward splitting and ncvx-MiFB with R being ℓ_0 pseudo-norm and the rank function.

The empirical bound (4.4.6) and inertial steps s We now present a short comparison of the empirical bound (4.4.6) of inertial parameters and different choices of s under bigger choice of $\gamma = 0.8\beta$. MiFB with 3 inertial steps, *i.e.* $s = 3$, is added which is noted as "3-ncvx-MiFB", see the *magenta* line in Figure 4.10.

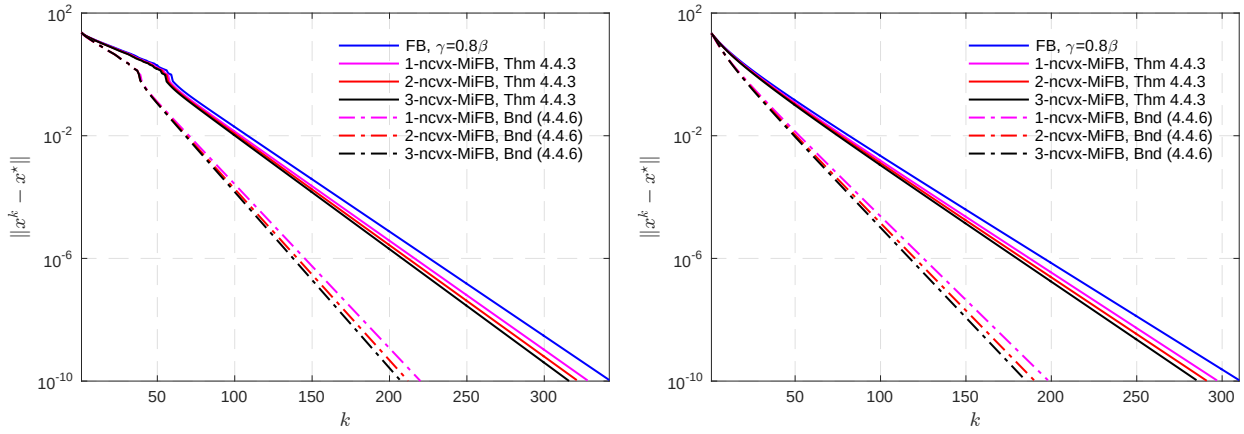
Similar to the above experiments, we choose $b_{i,k} = a_{i,k} \equiv a_i, i \in \mathcal{S}$, and "Thm 4.4.3" means that a_i's are chosen according to Theorem 4.4.3, while "Bnd (4.4.6)" means that a_i's are chosen based on the empirical bound (4.4.6). We can infer from Figure 4.10 that:

(i) Compared to the results in Figure 4.9, a bigger choice of γ leads to faster convergence. Yet still, under the same choice of γ, ncvx-MiFB is faster than FB both locally and globally;

(ii) For either "Thm 4.4.3" or "Bnd (4.4.6)", the performance of the three ncvx-MiFB schemes are close, this is mainly due to the fact that values of the sum $\sum_{i\in\mathcal{S}} a_i$ for each scheme are close.

(iii) Then between "Thm 4.4.3" and "Bnd (4.4.6)", "Bnd (4.4.6)" shows faster convergence result, since the allowed value of $\sum_{i\in\mathcal{S}} a_i$ of (4.4.6) is bigger than that of Theorem 4.4.3.

It should be noted that, when $\gamma \in]0, \frac{2}{3}\beta]$, the largest value of $\sum_{i\in\mathcal{S}} a_i$ allowed by (4.4.6) is 1. If we choose $\sum_{i\in I} a_i$ equal or very close to 1, then it can be observed in practice that ncvx-MiFB locally oscillates, which is a well-known property of the FISTA scheme [24, 50]. We refer to Chapter 6 for discussions of the properties of such oscillation behaviour.



(a) ℓ_0 pseudo-norm (b) Rank function

Figure 4.10: Comparison of ncvx-MiFB under different inertial settings. We fix $\gamma_k \equiv 0.8\beta$ for all tests. For the three inertial schemes, the inertial parameters were chosen such that (4.4.5) holds.

4.6 Proofs of main theorems

4.6.1 Proofs of Section 4.2

We need the following two lemmas before presenting the proof of the main theorems.

Lemma 4.6.1. *Let the operators $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone, $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive, and $\mathcal{V} \in \mathcal{S}_\nu$. Then for $\gamma \in]0, 2\beta\nu[$,*

(i) $\mathcal{V}^{-1}A : \mathcal{H}_\mathcal{V} \rightrightarrows \mathcal{H}_\mathcal{V}$ *is maximal monotone.*

(ii) $(\mathrm{Id} + \gamma\mathcal{V}^{-1}A)^{-1} : \mathcal{H}_\mathcal{V} \to \mathcal{H}_\mathcal{V}$ *is* firmly non-expansive.

(iii) $(\mathrm{Id} - \gamma\mathcal{V}^{-1}B) : \mathcal{H}_\mathcal{V} \to \mathcal{H}_\mathcal{V}$ *is* $\frac{\gamma}{2\beta\nu}$-averaged non-expansive.

(iv) *The operator* $(\mathrm{Id} + \gamma\mathcal{V}^{-1}A)^{-1}(\mathrm{Id} - \gamma\mathcal{V}^{-1}B)$ *is* $\frac{2\beta\nu}{2\beta\nu-\gamma}$-averaged non-expansive.

Proof. (i)-(ii) See [62, Lemma 3.7]. (iv) See [138, Theorem 3]. For (iii), since $B : \mathcal{H} \to \mathcal{H}$ is β-cocoercive, given any $x, x' \in \mathcal{H}$, we have

$$\begin{aligned}\langle x - x',\, \mathcal{V}^{-1}Bx - \mathcal{V}^{-1}Bx'\rangle_\mathcal{V} &\geq \beta\|Bx - Bx'\|^2 \\ &= \beta\mathcal{V}\langle \mathcal{V}^{-1}Bx - \mathcal{V}^{-1}Bx',\, \mathcal{V}^{-1}Bx - \mathcal{V}^{-1}Bx'\rangle_\mathcal{V} \\ &\geq \beta\nu\|\mathcal{V}^{-1}Bx - \mathcal{V}^{-1}Bx'\|_\mathcal{V}^2,\end{aligned}$$

which means $\mathcal{V}^{-1}B : \mathcal{H}_\mathcal{V} \to \mathcal{H}_\mathcal{V}$ is $(\beta\nu)$-cocoercive. The rest of the proof follows [16, Proposition 4.33]. □

Lemma 4.6.2. *Let $\{d_k\}_{k\in\mathbb{N}}, \{\delta_k\}_{k\in\mathbb{N}}$ be two non-negative sequences, and $\omega \in \mathbb{R}^s$ such that*

$$d_{k+1} \leq \sum_{i\in\mathcal{S}} \omega_i d_{k-i} + \delta_k, \tag{4.6.1}$$

for all $k \geq s$. If $\sum_i \omega_i \in [0,1[$ and $\sum_{k\in\mathbb{N}} \delta_k < +\infty$, then

$$\sum_{k\in\mathbb{N}} d_k < +\infty.$$

Remark 4.6.3. A special case of Lemma 4.6.2 appears in [33, Lemma 3]. It should be noted that in our case, non-negativity is *not* imposed to the weight ω_i's, but only the sum of them. In fact, we can even afford all ω_i's to be negative, as long as $\sum_{i\in\mathcal{S}} \omega_i d_{k-i} + \delta_k$ is positive for all $k \in \mathbb{N}$.

Proof. From (4.6.1), suppose that $d_{-1} = d_{-2} = d_{-s+1} = 0$, then sum up for both sides from $k = 0$,

$$\begin{aligned}\sum_{k\in\mathbb{N}} d_{k+1} \leq \sum_{k\in\mathbb{N}} \sum_{i\in\mathcal{S}} \omega_i d_{k-i} + \sum_{k\in\mathbb{N}} \delta_k &\Longrightarrow \sum_{k\in\mathbb{N}} d_k \leq d_0 + \sum_{i\in\mathcal{S}} \omega_i \sum_{k\in\mathbb{N}} d_k + \sum_{k\in\mathbb{N}} \delta_k \\ &\Longrightarrow \big(1 - \sum_{i\in\mathcal{S}} \omega_i\big) \sum_{k\in\mathbb{N}} d_k \leq d_0 + \sum_{k\in\mathbb{N}} \delta_k.\end{aligned}$$

Since we assume $\sum_{i\in\mathcal{S}} \omega_i < 1$ and δ_k is summable, then we have

$$\sum_{k\in\mathbb{N}} d_k \leq \big(1 - \sum_{i\in\mathcal{S}} \omega_i\big)^{-1} \big(d_0 + \sum_{k\in\mathbb{N}} \delta_k\big) < +\infty,$$

which concludes the proof. □

The following lemma is inspired by a result by Attouch *et al.* [11, Lemma A.9].

Lemma 4.6.4. *Let $\{\alpha_k\}_{k\in\mathbb{N}}, \{\beta_k\}_{k\in\mathbb{N}}$ be two summable sequences in $[0, +\infty[$, and $C < +\infty$ be a positive constant. Consider the non-negative sequence $\{\theta_k\}_{k\in\mathbb{N}}$, which is defined by, for all $k \geq s$,*

$$\theta_{k+1}^2 \leq C^2 + \sum_{j=1}^k \alpha_j \theta_j^2 + \sum_{j=1}^k \beta_j \theta_j.$$

If there holds $\sum_{k\in\mathbb{N}} \alpha_k < 1$, then we have

$$\theta_{k+1} \leq \frac{1}{1 - \sum_{k\in\mathbb{N}} \alpha_k} \big(C + \sum_{k\in\mathbb{N}} \beta_k\big).$$

holds for all $k \in \mathbb{N}$.

Proof. Given $k \in \mathbb{N}$, denote $\Theta_k = \max_{j=1,\dots,k} \theta_j$. Then for $1 \leq j \leq k$, we have

$$\begin{aligned}\theta_k^2 &\leq C^2 + \Theta_k^2 \sum_{j=1}^k \alpha_j + \Theta_k \sum_{j=1}^k \beta_j \\ &\leq C^2 + \Theta_k^2 \sum_{j\in\mathbb{N}} \alpha_j + \Theta_k \sum_{j\in\mathbb{N}} \beta_j.\end{aligned}$$

Taking the maximum over $1 \leq j \leq k$, we get

$$\Theta_k^2 \leq C^2 + \Theta_k^2 \sum_{j\in\mathbb{N}} \alpha_j + \Theta_k \sum_{j\in\mathbb{N}} \beta_j,$$

which leads to

$$\big(1 - \sum_{j\in\mathbb{N}} \alpha_j\big)\Theta_k^2 - \Theta_k \sum_{j\in\mathbb{N}} \beta_j - C^2 \leq 0.$$

Since $\sum_{j\in\mathbb{N}} \alpha_j < 1$, then the quadratic equation $(1 - \sum_{j\in\mathbb{N}} \alpha_j)\Theta_k^2 - \Theta_k \sum_{j\in\mathbb{N}} \beta_j - C^2 = 0$ admits two real roots, and Θ_k is then bounded from above by the larger one. □

Since the metric $\mathcal{V} \in \mathcal{M}_\nu$ is fixed with $\nu > 0$, without the loss of generality we present the proof of Theorem 4.2.3 on $\mathcal{H}$.

Proof of Theorem 4.2.3. Let $x^\star \in \operatorname{zer}(A + B)$, *i.e.* a solution, then

$$\begin{aligned}-B(x^\star) &\in A(x^\star), \\ (y_{a,k} - x_{k+1}) - \gamma_k B(y_{b,k}) &\in \gamma_k A^{\varepsilon_k}(x_{k+1}).\end{aligned} \tag{4.6.2}$$

Define the following quantities

$$\varphi_k = \frac{1}{2}\|x_k - x^\star\|^2, \; E_{x,k} = \frac{1}{2}\|x_k - x_{k-1}\|^2 \quad \text{and} \quad E_{b,k+1} = \frac{1}{2}\|y_{b,k} - x_{k+1}\|^2. \tag{4.6.3}$$

Given the expression of φ_k we have

$$\begin{aligned}\varphi_k - \varphi_{k+1} &= \tfrac{1}{2}\langle x_k - x^\star,\, x_k - x^\star\rangle - \tfrac{1}{2}\langle x_{k+1} - x^\star,\, x_{k+1} - x^\star\rangle \\ &= \tfrac{1}{2}\langle x_k - x_{k+1} - x^\star + 2x_{k+1} - x^\star,\, x_k - x_{k+1}\rangle \\ &= \tfrac{1}{2}\|x_k - x_{k+1}\|^2 + \langle x_k - y_{a,k} + y_{a,k} - x_{k+1},\, x_{k+1} - x^\star\rangle \\ &= E_{x,k+1} + \langle y_{a,k} - x_{k+1},\, x_{k+1} - x^\star\rangle - \textstyle\sum_{i\in\mathcal{S}} a_{i,k}\langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x^\star\rangle.\end{aligned} \tag{4.6.4}$$

Meanwhile, by virtue of the monotonicity of A and (4.6.2), given any $u_{k+1} \in A^{\varepsilon_k}(x_{k+1})$ and $u^\star \in A(x^\star)$, we have $\langle \gamma_k u_{k+1} - \gamma_k u^\star,\, x_{k+1} - x^\star\rangle \geq -\gamma_k\varepsilon_k$, hence

$$\langle (y_{a,k} - x_{k+1}) - \gamma_k B(y_{b,k}) + \gamma_k B(x^\star),\, x_{k+1} - x^\star\rangle \geq -\gamma_k\varepsilon_k.$$

Combine this with (4.6.4) we get

$$\begin{aligned}\varphi_k - \varphi_{k+1} \geq E_{x,k+1} &- \textstyle\sum_{i\in\mathcal{S}} a_{i,k}\langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x^\star\rangle \\ &+ \gamma_k\langle B(y_{b,k}) - B(x^\star),\, x_{k+1} - x^\star\rangle - \gamma_k\varepsilon_k.\end{aligned} \tag{4.6.5}$$

For each inner product $\langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x^\star\rangle$, we have

$$\begin{aligned}&\langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x^\star\rangle \\ &= \langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x_k + x_k - x_{k-i}\rangle + \langle x_{k-i} - x_{k-i-1},\, x_{k-i} - x^\star\rangle \\ &= \langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x_k\rangle + \langle x_{k-i} - x_{k-i-1},\, x_k - x_{k-i}\rangle + \langle x_{k-i} - x_{k-i-1},\, x_{k-i} - x^\star\rangle \\ &= \langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x_k\rangle + \tfrac{1}{2}\big(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2\big) + (\varphi_{k-i} - \varphi_{k-i-1}),\end{aligned} \tag{4.6.6}$$

where the Pythagoras relation below is applied to the inner products $\langle x_{k-i} - x_{k-i-1},\, x_k - x_{k-i}\rangle$ and $\langle x_{k-i} - x_{k-i-1},\, x_{k-i} - x^\star\rangle$ respectively,

$$2\langle c_1 - c_2,\, c_1 - c_3\rangle = \|c_1 - c_2\|^2 + \|c_1 - c_3\|^2 - \|c_2 - c_3\|^2. \tag{4.6.7}$$

Putting (4.6.6) back into (4.6.5) yields,

$$\begin{aligned}\varphi_{k+1} - \varphi_k &\leq -E_{x,k+1} - \gamma_k\langle B(y_{b,k}) - B(x^\star),\, x_{k+1} - x^\star\rangle + \gamma_k\varepsilon_k \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}\langle x_{k-i} - x_{k-i-1},\, x_{k+1} - x^\star\rangle \\ &= -E_{x,k+1} - \gamma_k\langle B(y_{b,k}) - B(x^\star),\, x_{k+1} - x^\star\rangle + \gamma_k\varepsilon_k \\ &\quad + \langle \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}),\, x_{k+1} - x_k\rangle \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2) + \sum_{i\in\mathcal{S}} a_{i,k}(\varphi_{k-i} - \varphi_{k-i-1}),\end{aligned}$$

and further leads to

$$\begin{aligned}&\varphi_{k+1} - \varphi_k - \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(\varphi_{k-i} - \varphi_{k-i-1}) \\ &\leq -E_{x,k+1} - \gamma_k\langle B(y_{b,k}) - B(x^\star),\, x_{k+1} - x^\star\rangle + \gamma_k\varepsilon_k \\ &\quad + \langle \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}),\, x_{k+1} - x_k\rangle + \sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2),\end{aligned} \tag{4.6.8}$$

Since B is β-cocoercive, then

$$\begin{aligned}&\langle B(y_{b,k}) - B(x^\star),\, x_{k+1} - x^\star\rangle \\ &= \langle B(y_{b,k}) - B(x^\star),\, x_{k+1} - y_{b,k} + y_{b,k} - x^\star\rangle \\ &\geq \beta\|B(y_{b,k}) - B(x^\star)\|^2 + \langle B(y_{b,k}) - B(x^\star),\, x_{k+1} - y_{b,k}\rangle \\ &\geq \beta\|B(y_{b,k}) - B(x^\star)\|^2 - \beta\|B(y_{b,k}) - B(x^\star)\|^2 - \tfrac{1}{4\beta}\|x_{k+1} - y_{b,k}\|^2 = -\tfrac{1}{2\beta}E_{b,k+1}.\end{aligned} \tag{4.6.9}$$

Combine (4.6.9) and (4.6.8), we get

$$\begin{aligned}&\varphi_{k+1} - \varphi_k - \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(\varphi_{k-i} - \varphi_{k-i-1}) \\ &\leq -E_{x,k+1} + \tfrac{\gamma_k}{2\beta}E_{b,k+1} + \langle \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}),\, x_{k+1} - x_k\rangle \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2) + \gamma_k\varepsilon_k.\end{aligned} \tag{4.6.10}$$

For $E_{b,k+1}$, we have

$$\begin{aligned} E_{b,k+1} = \tfrac{1}{2}\|x_{k+1} - y_{a,k}\|^2 &= \tfrac{1}{2}\|x_{k+1} - x_k - \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 \\ &= \tfrac{1}{2}\|x_{k+1} - x_k\|^2 + \tfrac{1}{2}\|\textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}\langle x_k - x_{k+1},\, x_{k-i} - x_{k-i-1}\rangle. \end{aligned} \tag{4.6.11}$$

Substitute into (4.6.10) we get,

$$\begin{aligned} &\varphi_{k+1} - \varphi_k - \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(\varphi_{k-i} - \varphi_{k-i-1}) \\ &\le -E_{x,k+1} + \langle \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}),\, x_{k+1} - x_k\rangle + \tfrac{\gamma_k}{2\beta}E_{b,k+1} \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}\big(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2\big) + \gamma_k\varepsilon_k \\ &= -E_{x,k+1} + \tfrac{\gamma_k}{2\beta}E_{x,k+1} + \langle \textstyle\sum_{i\in\mathcal{S}} (a_{i,k} - \frac{\gamma_k b_{i,k}}{2\beta})(x_{k-i} - x_{k-i-1}),\, x_{k+1} - x_k\rangle \\ &\quad + \tfrac{\gamma_k}{4\beta}\|\textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 + \textstyle\sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}\big(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2\big) + \gamma_k\varepsilon_k \\ &= -\lambda_k E_{x,k+1} + \langle \textstyle\sum_{i\in\mathcal{S}} \zeta_{i,k}(x_{k-i} - x_{k-i-1}),\, x_{k+1} - x_k\rangle + \tfrac{\gamma_k}{4\beta}\|\textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}\big(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2\big) + \gamma_k\varepsilon_k. \end{aligned} \tag{4.6.12}$$

Define the following parameters

$$\lambda_k = 1 - \tfrac{\gamma_k}{2\beta} \in \big[\tfrac{\bar{\epsilon}}{2\beta}, 1 - \tfrac{\epsilon}{2\beta}\big] \quad \text{and} \quad \zeta_{i,k} = a_{i,k} - \tfrac{\gamma_k b_{i,k}}{2\beta},\ \forall i \in \mathcal{S},$$

and vector

$$v_k = x_{k+1} - x_k - \frac{1}{\lambda_k}\sum_{i\in\mathcal{S}} \frac{\zeta_{i,k}}{\lambda_k}(x_{k-i} - x_{k-i-1}).$$

Substituting (4.6.9) back into (4.6.8), and combine with (4.6.12),

$$\begin{aligned} &\varphi_{k+1} - \varphi_k - \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(\varphi_{k-i} - \varphi_{k-i-1}) \\ &\le -\lambda_k E_{x,k+1} + \langle \textstyle\sum_{i\in\mathcal{S}} \zeta_{i,k}(x_{k-i} - x_{k-i-1}),\, x_{k+1} - x_k\rangle + \tfrac{\gamma_k}{4\beta}\|\textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2) + \gamma_k\varepsilon_k \\ &= -\tfrac{\lambda_k}{2}\|v_k\|^2 + \tfrac{1}{2\lambda_k}\|\textstyle\sum_{i\in\mathcal{S}} \zeta_{i,k}(x_{k-i} - x_{k-i-1})\|^2 + \tfrac{\gamma_k}{4\beta}\|\textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{a_{i,k}}{2}(\|x_k - x_{k-i-1}\|^2 - \|x_k - x_{k-i}\|^2) + \gamma_k\varepsilon_k \\ &\le -\tfrac{\lambda_k}{2}\|v_k\|^2 + \tfrac{1}{2\lambda_k}\|\textstyle\sum_{i\in\mathcal{S}} \zeta_{i,k}(x_{k-i} - x_{k-i-1})\|^2 + \tfrac{\gamma_k}{4\beta}\|\textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{|a_{i,k}|}{2}(\|x_k - x_{k-i-1}\|^2 + \|x_k - x_{k-i}\|^2) + \bar{\gamma}\varepsilon_k. \end{aligned} \tag{4.6.13}$$

Let's collect the quadratic terms in (4.6.13) in δ_k, and apply the Jensen's inequality

$$\begin{aligned} \delta_k &= \tfrac{1}{2\lambda_k}\|\textstyle\sum_{i\in\mathcal{S}} \zeta_{i,k}(x_{k-i} - x_{k-i-1})\|^2 + \tfrac{\gamma_k}{4\beta}\|\textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1})\|^2 \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{|a_{i,k}|}{2}\big(\|x_k - x_{k-i-1}\|^2 + \|x_k - x_{k-i}\|^2\big) \\ &\le \tfrac{s}{2\lambda_k}\textstyle\sum_{i\in\mathcal{S}} \zeta_{i,k}^2\|x_{k-i} - x_{k-i-1}\|^2 + \tfrac{s\gamma_k}{4\beta}\textstyle\sum_{i\in\mathcal{S}} b_{i,k}^2\|x_{k-i} - x_{k-i-1}\|^2 \\ &\quad + \textstyle\sum_{i\in\mathcal{S}} \frac{|a_{i,k}|}{2}\big(\|x_k - x_{k-i-1}\|^2 + \|x_k - x_{k-i}\|^2\big) \\ &\le \tfrac{s}{2\lambda_k}\textstyle\sum_{i\in\mathcal{S}} \zeta_{i,k}^2\|x_{k-i} - x_{k-i-1}\|^2 + \tfrac{s\gamma_k}{4\beta}\textstyle\sum_{i\in\mathcal{S}} |b_{i,k}|\|x_{k-i} - x_{k-i-1}\|^2 \\ &\quad + s\textstyle\sum_{i\in\mathcal{S}} |a_{i,k}| \sum_{j\in\mathcal{S}} \|x_{k-j} - x_{k-j-1}\|^2 \\ &\le \tfrac{s}{2\lambda_k}\max_{i\in\mathcal{S}} \zeta_{i,k}^2 \textstyle\sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2 + \tfrac{s\gamma_k}{4\beta}\max_{i\in\mathcal{S}} |b_{i,k}| \textstyle\sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2 \\ &\quad + s^2 \max_{i\in\mathcal{S}} |a_{i,k}| \textstyle\sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2 \\ &\le \max\big\{\tfrac{s}{2\underline{\lambda}}, \tfrac{s\bar{\gamma}}{4\beta}, s^2\big\} \max\big\{\max_{i\in\mathcal{S}} \zeta_{i,k}^2, \max_{i\in\mathcal{S}} |b_{i,k}|, \max_{i\in\mathcal{S}} |a_{i,k}|\big\} \textstyle\sum_{i\in\mathcal{S}} \|x_{k-i} - x_{k-i-1}\|^2. \end{aligned}$$

Therefore, as

$$\max\left\{\tfrac{s}{2\underline{\lambda}}, \tfrac{s\bar{\gamma}}{4\beta}, s^2\right\} < +\infty$$

is a constant, the sequence $\{\delta_k\}_{k\in\mathbb{N}}$ is summable if condition (4.2.5) holds. Now define $\theta_k = \varphi_k - \varphi_{k-1}$, then we get the following key estimate from (4.6.13)

$$\theta_{k+1} \leq -\tfrac{\lambda_k}{2}\|v_k\|^2 + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}\theta_{k-i} + \delta_k + \bar{\gamma}\varepsilon_k. \tag{4.6.14}$$

(i) If $\sum_i \bar{a}_i < 1$, define $[\theta]_+ = \max\{\theta, 0\}$, then (4.6.14) gives

$$\begin{aligned}[\theta_{k+1}]_+ &\leq \textstyle\sum_{i\in\mathcal{S}} a_{i,k}[\theta_{k-i}]_+ + \delta_k + \bar{\gamma}\varepsilon_k \\ &\leq \textstyle\sum_{i\in\mathcal{S}} \bar{a}_i[\theta_{k-i}]_+ + \delta_k + \bar{\gamma}\varepsilon_k.\end{aligned}$$

Owing to Lemma 4.6.2, $\{[\theta_k]_+\}_{k\in\mathbb{N}}$ is summable since $\sum_{i\in\mathcal{S}} \bar{a}_i < 1$. In turn,

$$\begin{aligned}\varphi_{k+1} - \textstyle\sum_{j=1}^{k+1}[\theta_j]_+ &\leq \varphi_{k+1} - \theta_{k+1} - \textstyle\sum_{j=1}^{k}[\theta_j]_+ \\ &= \varphi_k - \textstyle\sum_{j=1}^{k}[\theta_j]_+.\end{aligned}$$

It then follows that the sequence $\{\varphi_k - \sum_{j=1}^{k}[\theta_j]_+\}_{k\in\mathbb{N}}$ is decreasing and bounded from below, hence convergent, and we deduce that φ_k is also convergent.

(ii) If $a_{i,k} \equiv 0$ for all $i = 0, ..., s-1$, then (4.6.14) entails

$$\varphi_{k+1} \leq \varphi_k + \delta_k + \bar{\gamma}\varepsilon_k,$$

which means that the sequence $\{x_k\}_{k\in\mathbb{N}}$ is quasi-Fejér monotone relative to $\operatorname{zer}(A+B)$ [56, Definition 1.1(3)], thus φ_k is convergent [56, Proposition 3.6].

In summary, for $\sum_{i\in\mathcal{S}} \bar{a}_i < 1$, $\lim_{k\to\infty}\|x_k - x^\star\|$ exists for any $x^\star \in \operatorname{zer}(A+B)$.

By assumption (4.2.5), $a_{i,k}(x_{k-i} - x_{k-i-1}) \to 0$ and $b_{i,k}(x_{k-i} - x_{k-i-1}) \to 0$, and thus for each $i \in \mathcal{S}$,

$$\tfrac{\zeta_{i,k}}{\lambda_k}(x_{k-i} - x_{k-i-1}) \to 0, \tag{4.6.15}$$

since $\lambda_k \geq \frac{\bar{\epsilon}}{2\beta} > 0$. Moreover, from (4.6.14), we obtain

$$\textstyle\sum_{k\in\mathbb{N}}\|v_k\|^2 \leq \frac{4\beta}{\bar{\epsilon}}\big(\varphi_0 + \sum_{k\in\mathbb{N}}(\sum_{i\in\mathcal{S}}\bar{a}_i[\theta_{k-i}]_+ + e_k)\big) < +\infty.$$

Consequently, $v_k \to 0$. Combining this with (4.6.15), we get that $x_{k+1} - x_k \to 0$. In turn, $y_{a,k} - x_{k+1} \to 0$ and $y_{b,k} - x_{k+1} \to 0$.

Let $\bar{x}$ be a weak cluster point of $(x_k)_{k\in\mathbb{N}}$, and let us fix a subsequence, say $x_{k_j} \rightharpoonup \bar{x}$. We get from (4.6.2) that

$$u_{k_j} \overset{\text{def}}{=} \tfrac{1}{\gamma_{k_j}}(y_{a,k_j} - x_{k_j+1}) - B(y_{b,k_j}) - \xi_{k_j} \in A^{\varepsilon_{k_j}}(x_{k_j+1})$$

Since B is cocoercive and $y_{b,k_j} \rightharpoonup \bar{x}$, we have $B(y_{b,k_j}) \to B(\bar{x})$. In turn, $u_{k_j} \to -B(\bar{x})$ since $\gamma_k \geq \underline{\epsilon} > 0$. Since $(x_{k_j+1}, u_{k_j}) \in \operatorname{gra}(A^{\varepsilon_{k_j}})$, and the graph of the enlargement of A is sequentially weakly-strongly closed in $\mathbb{R}_+ \times \mathcal{H} \times \mathcal{H}$ [164, Proposition 3.4(b)], we get that $-B(\bar{x}) \in A(\bar{x})$, *i.e.* $\bar{x}$ is a solution of ($\mathcal{P}_{\text{inc}}$). That is, every weak sequential cluster point of $\{x_k\}_{k\in\mathbb{N}}$ lies in $\operatorname{zer}(A+B)$. Applying Opial's Theorem [139] concludes the proof. □

Proof of Theorem 4.2.5. Define

$$A_{v_k} = \mathcal{V}_k^{-1}A \quad \text{and} \quad B_{v_k} = \mathcal{V}_k^{-1}B.$$

Let $x^\star \in \operatorname{zer}(A+B)$, then

$$\begin{aligned}-B_{v_k}(x^\star) &\in A_{v_k}(x^\star), \\ (y_{a,k} - x_{k+1}) - \gamma_k B_{v_k}(y_{b,k}) - \gamma_k\xi_k &\in \gamma_k A^{\varepsilon_k}_{v_k}(x_{k+1}).\end{aligned} \tag{4.6.16}$$

Define the following quantities

$$\varphi_k = \frac{1}{2}\|x_k - x^\star\|^2_{\mathcal{V}_k},\ E_{x,k} = \frac{1}{2}\|x_k - x_{k-1}\|^2_{\mathcal{V}_k} \quad \text{and} \quad E_{b,k+1} = \frac{1}{2}\|y_{b,k} - x_{k+1}\|^2_{\mathcal{V}_k}. \tag{4.6.17}$$

Given the expression of φ_k we have

$$\begin{aligned}
\varphi_k - \frac{1}{1+\eta_k}\varphi_{k+1} &= \frac{1}{2}\langle x_k - x^\star,\, x_k - x^\star\rangle_{\mathcal{V}_k} - \frac{1}{1+\eta_k}\frac{1}{2}\langle x_{k+1} - x^\star,\, x_{k+1} - x^\star\rangle_{\mathcal{V}_{k+1}} \\
&\geq \frac{1}{2}\langle x_k - x^\star,\, x_k - x^\star\rangle_{\mathcal{V}_k} - \frac{1}{2}\langle x_{k+1} - x^\star,\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} \\
&= \frac{1}{2}\langle x_k - x_{k+1} - x^\star + 2x_{k+1} - x^\star,\, x_k - x_{k+1}\rangle_{\mathcal{V}_k} \\
&= \frac{1}{2}\|x_k - x_{k+1}\|^2_{\mathcal{V}_k} + \langle x_k - y_{a,k} + y_{a,k} - x_{k+1},\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} \\
&\geq \frac{1}{1+\eta_k}\frac{1}{2}\|x_k - x_{k+1}\|^2_{\mathcal{V}_{k+1}} + \langle x_k - y_{a,k} + y_{a,k} - x_{k+1},\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} \\
&= \frac{1}{1+\eta_k}E_{x,k+1} + \langle y_{a,k} - x_{k+1},\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} - a_k\langle x_k - x_{k-1},\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k}.
\end{aligned} \tag{4.6.18}$$

Meanwhile, by virtue of the monotonicity of $A_{\mathcal{V}_k}$ on $\mathcal{H}_{\mathcal{V}_k}$ and (4.6.16), given any $u_{k+1} \in A^{\varepsilon_k}_{\mathcal{V}_k}(x_{k+1})$ and $u^\star \in A_{\mathcal{V}_k}(x^\star)$, we have $\langle \gamma_k u_{k+1} - \gamma_k u^\star,\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} \geq -\gamma_k\varepsilon_k$, hence back to (4.6.16) we get

$$\langle (y_{a,k} - x_{k+1}) - \gamma_k B_{\mathcal{V}_k}(y_{b,k}) + \gamma_k B_{\mathcal{V}_k}(x^\star) - \gamma_k\xi_k,\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} \geq -\gamma_k\varepsilon_k.$$

Combine this with (4.6.18) we get

$$\begin{aligned}
\varphi_k - \frac{1}{1+\eta_k}\varphi_{k+1} \geq \frac{1}{1+\eta_k}E_{x,k+1} - a_k\langle x_k - x_{k-1},\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} \\
+ \gamma_k\langle B_{\mathcal{V}_k}(y_{b,k}) - B_{\mathcal{V}_k}(x^\star) + \xi_k,\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} - \gamma_k\varepsilon_k.
\end{aligned} \tag{4.6.19}$$

For the inner product $\langle x_k - x_{k-1},\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k}$, apply the Pythagoras relation (4.6.7),

$$\begin{aligned}
\langle x_k - x_{k-1},\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} &= \langle x_k - x_{k-1},\, x_{k+1} - x_k\rangle_{\mathcal{V}_k} + \langle x_k - x_{k-1},\, x_k - x^\star\rangle_{\mathcal{V}_k} \\
&= \langle x_k - x_{k-1},\, x_{k+1} - x_k\rangle_{\mathcal{V}_k} + E_{x,k} + \frac{1}{2}\|x_k - x^\star\|^2_{\mathcal{V}_k} - \frac{1}{2}\|x_{k-1} - x^\star\|^2_{\mathcal{V}_k} \\
&\geq \langle x_k - x_{k-1},\, x_{k+1} - x_k\rangle_{\mathcal{V}_k} + E_{x,k} + \big(\varphi_k - (1+\eta_{k-1})\varphi_{k-1}\big),
\end{aligned} \tag{4.6.20}$$

Putting (4.6.20) back into (4.6.19) yields,

$$\begin{aligned}
\frac{1}{1+\eta_k}\varphi_{k+1} - \varphi_k \leq -\frac{1}{1+\eta_k}E_{x,k+1} - \gamma_k\langle B_{\mathcal{V}_k}(y_{b,k}) - B_{\mathcal{V}_k}(x^\star) + \xi_k,\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} + \gamma_k\varepsilon_k \\
+ a_k\langle x_k - x_{k-1},\, x_{k+1} - x_k\rangle_{\mathcal{V}_k} + a_k E_{x,k} + a_k\big(\varphi_k - (1+\eta_{k-1})\varphi_{k-1}\big),
\end{aligned}$$

and further leads to

$$\begin{aligned}
&\varphi_{k+1} - \varphi_k - a_k\big(\varphi_k - (1+\eta_{k-1})\varphi_{k-1}\big) \\
&= \varphi_{k+1} - \varphi_k - a_k(1+\eta_{k-1})\big(\frac{1}{1+\eta_{k-1}}\varphi_k - \varphi_{k-1}\big) \\
&\leq -E_{x,k+1} - \gamma_k\langle B_{\mathcal{V}_k}(y_{b,k}) - B_{\mathcal{V}_k}(x^\star) + \xi_k,\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} + \gamma_k\varepsilon_k \\
&\quad + a_k\langle x_k - x_{k-1},\, x_{k+1} - x_k\rangle_{\mathcal{V}_k} + a_k E_{x,k},
\end{aligned} \tag{4.6.21}$$

Since $B_{\mathcal{V}_k}$ is $(\beta\nu)$-cocoercive, then similar to (4.6.9) we have

$$\langle B_{\mathcal{V}_k}(y_{b,k}) - B_{\mathcal{V}_k}(x^\star),\, x_{k+1} - x^\star\rangle_{\mathcal{V}_k} \geq -\frac{1}{2\beta\nu}E_{b,k+1}. \tag{4.6.22}$$

Define the following parameters

$$c_k = a_k(1+\eta_{k-1}),\ \ \lambda_k = 1 - \frac{\gamma_k}{2\beta\nu} \in \Big[\frac{\bar{\epsilon}}{2\beta\nu}, 1 - \frac{\underline{\epsilon}}{2\beta\nu}\Big] \quad \text{and} \quad \zeta_k = a_k - \frac{\gamma_k b_k}{2\beta\nu},$$

and $v_k = x_{k+1} - x_k - \frac{\zeta_k}{\lambda_k}(x_k - x_{k-1})$. Substituting (4.6.22) back into (4.6.21), and since

$$E_{b,k+1} = \frac{1}{2}\|x_k - x_{k+1}\|_{\mathcal{V}_k} + b_k^2 E_{x,k} + b_k\langle x_k - x_{k+1},\, x_k - x_{k-1}\rangle_{\mathcal{V}_k},$$

we get from (4.6.21) that

$$\begin{aligned}
&\tfrac{1}{1+\eta_k}\varphi_{k+1}-\varphi_k-c_k\big(\tfrac{1}{1+\eta_{k-1}}\varphi_k-\varphi_{k-1}\big)\\
&\le -\tfrac{1}{2}\|x_k-x_{k+1}\|_{\mathcal{V}_k}+a_k\langle x_k-x_{k-1},\,x_{k+1}-x_k\rangle_{\mathcal{V}_k}+a_kE_{x,k}\\
&\quad+\tfrac{\gamma_k}{2\beta\nu}E_{b,k+1}+\gamma_k\varepsilon_k-\gamma_k\langle\xi_k,\,x_{k+1}-x^\star\rangle_{\mathcal{V}_k}\\
&=\big(\tfrac{\gamma_k}{2\beta\nu}-1\big)\tfrac{1}{2}\|x_k-x_{k+1}\|^2_{\mathcal{V}_k}+\big(a_k-\tfrac{\gamma_kb_k}{2\beta\nu}\big)\langle x_k-x_{k-1},\,x_{k+1}-x_k\rangle_{\mathcal{V}_k}+\big(a_k+\tfrac{\gamma_kb_k^2}{2\beta\nu}\big)E_{x,k}\\
&\quad+\gamma_k\varepsilon_k-\gamma_k\langle\xi_k,\,x_{k+1}-x^\star\rangle_{\mathcal{V}_k}\\
&=-\tfrac{\lambda_k}{2}\|x_k-x_{k+1}\|^2_{\mathcal{V}_k}+\zeta_k\langle x_k-x_{k-1},\,x_{k+1}-x_k\rangle_{\mathcal{V}_k}+\big(a_k+\tfrac{\gamma_kb_k^2}{2\beta\nu}\big)E_{x,k}\\
&\quad+\gamma_k\varepsilon_k-\gamma_k\langle\xi_k,\,x_{k+1}-x^\star\rangle_{\mathcal{V}_k}\\
&=\big(-\tfrac{\lambda_k}{2}\|x_{k+1}-x_k-\tfrac{\zeta_k}{\lambda_k}(x_k-x_{k-1})\|^2_{\mathcal{V}_k}+\tfrac{\zeta_k^2}{\lambda_k}E_{x,k}\big)+\big(a_k+\tfrac{\gamma_kb_k^2}{2\beta\nu}\big)E_{x,k}\\
&\quad+\gamma_k\varepsilon_k-\gamma_k\langle\xi_k,\,x_{k+1}-x^\star\rangle_{\mathcal{V}_k}\\
&=-\tfrac{\lambda_k}{2}\|v_k\|^2_{\mathcal{V}_k}+\big(a_k+\tfrac{\zeta_k^2}{\lambda_k}+\tfrac{\gamma_kb_k^2}{2\beta\nu}\big)E_{x,k}+\gamma_k\varepsilon_k-\gamma_k\langle\xi_k,\,x_{k+1}-x^\star\rangle_{\mathcal{V}_k}\\
&\le-\tfrac{\lambda_k}{2}\|v_k\|^2_{\mathcal{V}_k}+\big(\tfrac{2a_k}{\lambda_k}+\tfrac{\gamma_kb_k}{2\beta\nu}\big)E_{x,k}+\gamma_k\varepsilon_k-\gamma_k\langle\xi_k,\,x_{k+1}-x^\star\rangle_{\mathcal{V}_k}\\
&\le-\tfrac{\lambda_k}{2}\|v_k\|^2_{\mathcal{V}_k}+\big(\tfrac{4\beta\nu}{\bar\epsilon}a_k+(1-\tfrac{\bar\epsilon}{2\beta\nu})b_k\big)E_{x,k}+\bar\gamma\big(\varepsilon_k+\|\xi_k\|_{\mathcal{V}_k}\|x_{k+1}-x^\star\|_{\mathcal{V}_k}\big),
\end{aligned}\tag{4.6.23}$$

Now define $\theta'_k=\frac{1}{1+\eta_{k-1}}\varphi_k-\varphi_{k-1}$ and $\delta_k=\big(\frac{4\beta\nu}{\bar\epsilon}a_k+(1-\frac{\bar\epsilon}{2\beta\nu})b_k\big)E_{x,k}$. Then we get from (4.6.23)

$$\theta'_{k+1}\le-\tfrac{\lambda_k}{2}\|v_k\|^2_{\mathcal{V}_k}+c_k\theta'_k+\delta_k+\bar\gamma(\varepsilon_k+\|\xi_k\|_{\mathcal{V}_k}\|x_{k+1}-x^\star\|_{\mathcal{V}_k}).\tag{4.6.24}$$

Owing to the assumption on $\{\mathcal{V}_k\}_{k\in\mathbb{N}}$, we have $\mu\mathrm{Id}\succcurlyeq\mathcal{V}_k\succcurlyeq\nu\mathrm{Id}$. Then (4.6.24) can be further written as

$$\theta'_{k+1}\le-\tfrac{\lambda_k}{2}\|v_k\|^2_{\mathcal{V}_k}+c_k\theta'_k+\delta_k+\bar\gamma\big(\varepsilon_k+\tfrac{\mu}{\sqrt{\nu}}\|\xi_k\|\sqrt{2\varphi_{k+1}}\big).\tag{4.6.25}$$

For simplicity, we denote $C=\sqrt{2}\mu/\sqrt{\nu}$. Now define $\theta_k=\varphi_k-\varphi_{k-1}=\frac{\eta_{k-1}}{1+\eta_{k-1}}\varphi_k+\theta'_k$, we have $\theta_1=\varphi_1-\varphi_0=0$ since $x_1=x_0$. Then from (4.6.25) we get

$$\begin{aligned}
\theta_{k+1}&=\theta'_{k+1}+\tfrac{\eta_k}{1+\eta_k}\varphi_{k+1}\\
&\le-\tfrac{\lambda_k}{2}\|v_k\|^2_{\mathcal{V}_k}+c_k\theta_k+\tfrac{\eta_k}{1+\eta_k}\varphi_{k+1}+\delta_k+\bar\gamma(\varepsilon_k+C\|\xi_k\|\sqrt{\varphi_{k+1}})\\
&\le\textstyle\prod_{j=1}^k c_j\theta_1+\sum_{j=1}^k\big(\prod_{l=j}^k c_{l-j}\big)\big(\frac{\eta_j}{1+\eta_j}\varphi_{j+1}+\delta_j+\bar\gamma(\varepsilon_j+C\|\xi_j\|\sqrt{\varphi_{j+1}})\big)\\
&\le\textstyle\sum_{j=1}^k\bar c^{k-j}\big(\frac{\eta_j}{1+\eta_j}\varphi_{j+1}+\delta_j+\bar\gamma(\varepsilon_j+C\|\xi_j\|\sqrt{\varphi_{j+1}})\big).
\end{aligned}\tag{4.6.26}$$

(i) For $\bar c\in]0,1[$. Sum up (4.6.26) from $m=1$ to k we get

$$\begin{aligned}
\textstyle\sum_{m=1}^k\theta_{m+1}&=\varphi_{k+1}-\varphi_1\\
&\le\textstyle\sum_{m=1}^k\sum_{j=1}^m\bar c^{m-j}\big(\frac{\eta_j}{1+\eta_j}\varphi_{j+1}+\delta_j+\bar\gamma(\varepsilon_j+C\|\xi_j\|\sqrt{\varphi_{j+1}})\big)\\
&=\textstyle\sum_{m=1}^k\big(\sum_{j=1}^{k-m}\bar c^j\big)\big(\frac{\eta_m}{1+\eta_m}\varphi_{m+1}+\delta_m+\bar\gamma(\varepsilon_m+C\|\xi_m\|\sqrt{\varphi_{m+1}})\big)\\
&=\textstyle\sum_{m=1}^k\frac{1-\bar c^{k-m+1}}{1-\bar c}\big(\frac{\eta_m}{1+\eta_m}\varphi_{m+1}+\delta_m+\bar\gamma(\varepsilon_m+C\|\xi_m\|\sqrt{\varphi_{m+1}})\big),
\end{aligned}$$

which leads to

$$\begin{aligned}
\varphi_{k+1}&\le\big(\varphi_1+\tfrac{1}{1-\bar c}\textstyle\sum_{m\in\mathbb{N}}(\delta_m+\bar\gamma\varepsilon_m)\big)\\
&\quad+\tfrac{1}{1-\bar c}\textstyle\sum_{m=1}^k\frac{(1-\bar c^{k-m+1})\eta_m}{1+\eta_m}\varphi_{m+1}+\frac{\bar\gamma C}{1-\bar c}\sum_{m=1}^k\|\xi_m\|\sqrt{\varphi_{m+1}}.
\end{aligned}$$

Since $\{\delta_k\}_{k\in\mathbb{N}}$ is summable owing to (4.2.8), $\{\varepsilon_k\}_{k\in\mathbb{N}}$ and $\{\xi_k\}_{k\in\mathbb{N}}$ are summable, and $S=\sup_{k\in\mathbb{N}}\frac{1}{1-\bar c}\sum_{m=1}^k\frac{(1-\bar c^{k-m+1})\eta_m}{1+\eta_m}<1$ by (4.2.7). Then, in the view of Lemma 4.6.4 we have that

$\{\varphi_k\}_{k\in\mathbb{N}}$ is bounded from above by F^2 where F is

$$F = \frac{1}{1-S}\Big(\sqrt{\varphi_1 + \frac{1}{1-\bar{c}}\sum_{m\in\mathbb{N}}(\delta_m + \bar{\gamma}\varepsilon_m)} + \frac{\bar{\gamma}C}{1-\bar{c}}\sum_{k\in\mathbb{N}}\|\xi_k\|\Big).$$

Then define

$$e_k \overset{\text{def}}{=} \eta_k F^2 + \delta_k + \bar{\gamma}(\varepsilon_k + CF\|\xi_k\|),$$

then $\{e_k\}_{k\in\mathbb{N}}$ is summable. Back to (4.6.26), and let $[\theta]_+ = \max\{\theta, 0\}$, then

$$[\theta_{k+1}]_+ \leq \bar{c}[\theta_k]_+ + e_k,$$

Therefore, using that $\bar{c} < 1$, $[\theta_k]_+$ is summable (Lemma 4.6.2). In turn,

$$\begin{aligned}\varphi_{k+1} - \sum_{j=1}^{k+1}[\theta_j]_+ &\leq \varphi_{k+1} - \theta_{k+1} - \sum_{j=1}^{k}[\theta_j]_+\\ &= \varphi_k - \sum_{j=1}^{k}[\theta_j]_+.\end{aligned}$$

We then infer that the sequence $(\varphi_k - \sum_{j=1}^{k}[\theta_j]_+)_{k\in\mathbb{N}}$ is decreasing and bounded from below, hence convergent, whence we deduce that φ_k is also convergent.

(ii) If $a_k \equiv 0$ (hence $c_k \equiv 0$). Summability of $\{\eta_k\}_{k\in\mathbb{N}}$ is equivalent to

$$D = \prod_{k\in\mathbb{N}}(1+\eta_k) < +\infty.$$

Thus (4.6.26) yields

$$\begin{aligned}\varphi_{k+1} &\leq (1+\eta_k)\varphi_k + (1+\eta_k)\big(\delta_k + \bar{\gamma}(\varepsilon_k + C\|\xi_k\|\sqrt{\varphi_{k+1}})\big)\\ &\leq \prod_{m=1}^{k}(1+\eta_m)\varphi_1 + \sum_{m=1}^{k}\Big(\prod_{j=m}^{k}(1+\eta_j)\Big)\big(\delta_m + \bar{\gamma}(\varepsilon_m + C\|\xi_m\|\sqrt{\varphi_{m+1}})\big)\\ &\leq D\varphi_1 + D\sum_{m=1}^{k}\big(\delta_m + \bar{\gamma}\varepsilon_m\big) + \bar{\gamma}DC\sum_{m=1}^{k}\|\xi_m\|\sqrt{\varphi_{m+1}}.\end{aligned}$$

Again, owing to the summability of $\{\delta_k\}_{k\in\mathbb{N}}$, $\{\varepsilon_k\}_{k\in\mathbb{N}}$ and $\{\|\xi_k\|\}_{k\in\mathbb{N}}$, Lemma 4.6.4 asserts that φ_k is bounded. Now define F as

$$F = \sqrt{D(\varphi_1 + \textstyle\sum_{m\in\mathbb{N}}(\delta_m + \bar{\gamma}\varepsilon_m))} + \bar{\gamma}DC\sum_{k\in\mathbb{N}}\|\xi_k\|,$$

then $\sqrt{\varphi_k} \leq F$ holds for all $k \in \mathbb{N}$. Let $\bar{\eta} = \sup_{k\in\mathbb{N}}\eta_k$, and define

$$f_k \overset{\text{def}}{=} (1+\bar{\eta})\big(\delta_k + \bar{\gamma}(\varepsilon_k + CF\|\xi_k\|)\big),$$

then $\{f_k\}_{k\in\mathbb{N}}$ is summable. Therefore we have

$$\varphi_{k+1} \leq (1+\eta_k)\varphi_k + f_k,$$

which implies that $\{x_k\}_{k\in\mathbb{N}}$ is variable metric quasi-Fejér monotone relative to $\operatorname{zer}(A+B)$ [61, Definition 3.1], thus φ_k is convergent [61, Proposition 3.2].

In summary, we have $\lim_{k\to\infty}\|x_k - x^\star\|_{\mathcal{V}_k}$ exists for any $x^\star \in \operatorname{zer}(A+B)$ and $\{x_k\}_{k\in\mathbb{N}}$ is bounded.

By assumption (4.2.8), $a_k(x_k - x_{k-1}) \to 0$ and $b_k(x_k - x_{k-1}) \to 0$, and thus

$$\frac{\zeta_k}{\lambda_k}(x_k - x_{k-1}) \to 0, \tag{4.6.27}$$

since $\lambda_k \geq \frac{\bar{\epsilon}}{2\beta\nu} > 0$. Moreover, from (4.6.26), we obtain

$$\nu\sum_{k\in\mathbb{N}}\|v_k\|^2 \leq \sum_{k\in\mathbb{N}}\|v_k\|_{\mathcal{V}_k}^2 \leq \frac{4\beta\nu}{\bar{\epsilon}}\big(\varphi_0 + \sum_{k\in\mathbb{N}}(\bar{c}[\theta_k]_+ + e_k)\big) < +\infty.$$

Consequently, $v_k \to 0$. Combining this with (4.6.27), we get that $x_{k+1} - x_k \to 0$. In turn, $y_{a,k} - x_{k+1} \to 0$ and $y_{b,k} - x_{k+1} \to 0$.

Let $\bar{x}$ be a weak cluster point of $(x_k)_{k\in\mathbb{N}}$, and let us fix a subsequence, say $x_{k_j} \rightharpoonup \bar{x}$. We get from (4.6.16) that

$$u_{k_j} \overset{\text{def}}{=} \frac{1}{\gamma_{k_j}}\mathcal{V}_{k_j}(y_{a,k_j} - x_{k_j+1}) - B(y_{b,k_j}) - \mathcal{V}_{k_j}\xi_{k_j} \in A^{\varepsilon_{k_j}}(x_{k_j+1})$$

Since B is β-cocoercive and $y_{b,k_j} \rightharpoonup \bar{x}$, we have $B(y_{b,k_j}) \to B(\bar{x})$. In turn, since $\gamma_k \geq \underline{\epsilon} > 0$, we have

$$\begin{aligned}\|u_{k_j} + B(\bar{x})\| &= \|\tfrac{1}{\gamma_{k_j}}\mathcal{V}_{k_j}(y_{a,k_j} - x_{k_j+1}) + B(\bar{x}) - B(y_{b,k_j}) - \mathcal{V}_{k_j}\xi_{k_j}\| \\ &\leq \|\tfrac{1}{\gamma_{k_j}}\mathcal{V}_{k_j}(y_{a,k_j} - x_{k_j+1})\| + \|B(\bar{x}) - B(y_{b,k_j})\| + \|\mathcal{V}_{k_j}\xi_{k_j}\| \\ &\leq \tfrac{1}{\underline{\epsilon}}\mu\|y_{a,k_j} - x_{k_j+1}\| + \tfrac{1}{\beta}\|\bar{x} - y_{b,k_j}\| + \mu\|\xi_{k_j}\| \to 0.\end{aligned}$$

Hence, $u_{k_j} \to -B(\bar{x})$. As $(x_{k_j+1}, u_{k_j}) \in \operatorname{gra}(A^{\varepsilon_{k_j}})$, and the graph of the enlargement of A is sequentially weakly-strongly closed in $\mathbb{R}_+ \times \mathcal{H} \times \mathcal{H}$ [164, Proposition 3.4(b)], we get that $-B(\bar{x}) \in A(\bar{x})$, *i.e.* $\bar{x}$ is a solution of ($\mathcal{P}_{\text{inc}}$). That is, every weak sequential cluster point of $\{x_k\}_{k\in\mathbb{N}}$ lies in $\operatorname{zer}(A+B)$. Applying [61, Theorem 3.3] we conclude to the weak convergence of $\{x_k\}_{k\in\mathbb{N}}$ to a solution $x^\star \in \operatorname{zer}(A+B)$. □

Proof of Theorem 4.2.9. From the proof of Theorem 4.2.5, $\{x_k\}_{k\in\mathbb{N}}$ is bounded, with the imposed assumptions of the theorem. Let $C \in]0, +\infty[$ be the constant such that

$$\sup_{k\in\mathbb{N}} \max\left\{\tfrac{1}{2}\|x_k - x^\star\|^2_{\mathcal{V}_k}, \|x_k - x^\star\|_{\mathcal{V}_k}\right\} \leq C.$$

Then from (4.6.23), we apply Young's inequality to get

$$\begin{aligned}&\tfrac{1}{1+\eta_k}\varphi_{k+1} - \varphi_k - c_k\big(\tfrac{1}{1+\eta_{k-1}}\varphi_k - \varphi_{k-1}\big) \\ &\leq \big(\tfrac{\gamma_k}{2\beta\nu} - 1\big)\tfrac{1}{2}\|x_{k+1} - x_k\|_{\mathcal{V}_k} + \big(a_k - \tfrac{\gamma_k b_k}{2\beta\nu}\big)\langle x_k - x_{k-1},\, x_{k+1} - x_k\rangle_{\mathcal{V}_k} \\ &\quad + \big(a_k + \tfrac{\gamma_k b_k^2}{2\beta\nu}\big)\tfrac{1}{2}\|x_k - x_{k-1}\|_{\mathcal{V}_k} + \bar{\gamma}\big(\varepsilon_k + C\|\xi_k\|_{\mathcal{V}_k}\big) \\ &\leq \big(\tfrac{\gamma_k}{2\beta\nu} - 1\big)\tfrac{1}{2}\|x_{k+1} - x_k\|_{\mathcal{V}_k} + |a_k - \tfrac{\gamma_k b_k}{2\beta\nu}|\big(\tfrac{1}{2}\|x_{k+1} - x_k\|_{\mathcal{V}_k} + \tfrac{1}{2}\|x_k - x_{k-1}\|_{\mathcal{V}_k}\big) \\ &\quad + \big(\tfrac{\gamma_k b_k^2}{2\beta\nu} + a_k\big)\tfrac{1}{2}\|x_k - x_{k-1}\|_{\mathcal{V}_k} + \bar{\gamma}\big(\varepsilon_k + C\|\xi_k\|_{\mathcal{V}_k}\big) \\ &= S_k\tfrac{1}{2}\|x_{k+1} - x_k\|_{\mathcal{V}_k} + T_k\tfrac{1}{2}\|x_k - x_{k-1}\|_{\mathcal{V}_k} + \bar{\gamma}\big(\varepsilon_k + C\|\xi_k\|_{\mathcal{V}_k}\big),\end{aligned} \tag{4.6.28}$$

where the coefficients are defined as

$$S_k = \tfrac{\gamma_k}{2\beta\nu} - 1 + |a_k - \tfrac{\gamma_k b_k}{2\beta\nu}|, \quad T_k = \tfrac{\gamma_k}{2\beta\nu}b_k^2 + a_k + |a_k - \tfrac{\gamma_k b_k}{2\beta\nu}|.$$

Apply the boundedness of $\{x_k\}_{k\in\mathbb{N}}$ again, from (4.6.28) we get

$$\begin{aligned}&\varphi_{k+1} - \varphi_k - c_k\big(\varphi_k - \varphi_{k-1}\big) \\ &\leq \tfrac{1}{1+\eta_k}\varphi_{k+1} - \varphi_k - c_k\big(\tfrac{1}{1+\eta_{k-1}}\varphi_k - \varphi_{k-1}\big) + \tfrac{\eta_k}{1+\eta_k}\varphi_{k+1} \\ &\leq S_k\tfrac{1}{2}\|x_{k+1} - x_k\|_{\mathcal{V}_k} + T_k\tfrac{1}{2}\|x_k - x_{k-1}\|_{\mathcal{V}_k} + \bar{\gamma}\big(\varepsilon_k + C\|\xi_k\|_{\mathcal{V}_k}\big) + \tfrac{\eta_k}{1+\eta_k}\varphi_{k+1} \\ &\leq S_k\tfrac{1}{2}\|x_{k+1} - x_k\|_{\mathcal{V}_k} + T_k\tfrac{1}{2}\|x_k - x_{k-1}\|_{\mathcal{V}_k} + \bar{\gamma}\big(\varepsilon_k + C\|\xi_k\|_{\mathcal{V}_k}\big) + C\eta_k.\end{aligned} \tag{4.6.29}$$

Suppose a_k, b_k, γ_k and η_k are chosen such that

$$c_k = a_k(1+\eta_{k-1}),\ S_k \text{ and } T_k \text{ are non-decreasing.}$$

Define

$$\phi_k \overset{\text{def}}{=} \varphi_k - c_k\varphi_{k-1} + T_k\tfrac{1}{2}\|x_k - x_{k-1}\|_{\mathcal{V}_k} \quad \text{and} \quad \delta_k = \bar{\gamma}(\varepsilon_k + C\|\xi_k\|_{\mathcal{V}}) + C\eta_k.$$

Then from (4.6.29),

$$\begin{aligned}
&\phi_{k+1}-\phi_k\\
&=(\varphi_{k+1}-c_{k+1}\varphi_k)-(\varphi_k-c_k\varphi_{k-1})+T_{k+1}\tfrac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_{k+1}}-T_k\tfrac{1}{2}\|x_k-x_{k-1}\|_{\mathcal{V}_k}+\delta_k\\
&\leq(\varphi_{k+1}-c_k\varphi_k)-(\varphi_k-c_k\varphi_{k-1})+T_{k+1}\tfrac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_{k+1}}-T_k\tfrac{1}{2}\|x_k-x_{k-1}\|_{\mathcal{V}_k}+\delta_k\\
&=\varphi_{k+1}-\varphi_k-c_k(\varphi_k-\varphi_{k-1})+T_{k+1}\tfrac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_{k+1}}-T_k\tfrac{1}{2}\|x_k-x_{k-1}\|_{\mathcal{V}_k}+\delta_k\\
&\leq S_k\tfrac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_k}+T_k\tfrac{1}{2}\|x_k-x_{k-1}\|_{\mathcal{V}_k}+T_{k+1}\tfrac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_{k+1}}-T_k\tfrac{1}{2}\|x_k-x_{k-1}\|_{\mathcal{V}_k}+\delta_k\\
&\leq(S_{k+1}+(1+\eta_k)T_{k+1})\tfrac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_k}+\delta_k\\
&=\big(\tfrac{\gamma_{k+1}}{2\beta\nu}-1+(1+\eta_k)\big(\tfrac{\gamma_{k+1}b_{k+1}^2}{2\beta\nu}+a_{k+1}\big)+(2+\eta_k)|a_{k+1}-\tfrac{\gamma_{k+1}b_{k+1}}{2\beta\nu}|\big)\tfrac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_k}+\delta_k.
\end{aligned}\tag{4.6.30}$$

Sequence $\{\delta_k\}_{k\in\mathbb{N}}$ is summable owing to our assumptions. We have the following discussions based on the sign of $|a_{k+1}-\frac{\gamma_{k+1}b_{k+1}}{2\beta\nu}|$, define $F_{x,k}=\frac{1}{2}\|x_{k+1}-x_k\|_{\mathcal{V}_k}$:

(i) $a_k\in[0,\bar{a}]$, $b_k\in[0,\bar{b}]$, $b_k\leq a_k$. We have $\frac{\gamma_k}{2\beta\nu}b_k<a_k$, then from (4.6.30), and under the second condition in (4.2.10),

$$\begin{aligned}
&\phi_{k+1}-\phi_k\\
&\leq\big(\tfrac{\gamma_{k+1}}{2\beta\nu}-1+(1+\eta_k)\big(\tfrac{\gamma_{k+1}b_{k+1}^2}{2\beta\nu}+a_{k+1}\big)+(2+\eta_k)(a_{k+1}-\tfrac{\gamma_{k+1}b_{k+1}}{2\beta\nu})\big)F_{x,k}+\delta_k\\
&=\big(\big((3+2\eta_k)a_{k+1}-1\big)+\tfrac{\gamma_{k+1}}{2\beta\nu}\big(1+(1+\eta_k)b_{k+1}^2-(2+\eta_k)b_{k+1}\big)\big)F_{x,k}+\delta_k\\
&=\big(\big((3+2\eta_k)a_{k+1}-1\big)+\tfrac{\gamma_{k+1}}{2\beta\nu}\big((1-b_{k+1})^2+\eta_kb_{k+1}(b_{k+1}-1)\big)\big)F_{x,k}+\delta_k\\
&\leq-\tau F_{x,k}+\delta_k.
\end{aligned}\tag{4.6.31}$$

(ii) $a_k\in[0,\bar{a}]$, $b_k\in]0,\bar{b}]$, $a_k<b_k$. We need to discuss again the relationship between a_{k+1} and $\frac{\gamma_{k+1}}{2\beta\nu}b_{k+1}$, which splits into two sub-cases.

(a) If $\frac{\gamma_{k+1}}{2\beta\nu}b_{k+1}\leq a_{k+1}$, $k\in\mathbb{N}$, then from the second condition in (4.2.10) and (4.6.31),

$$\begin{aligned}
\phi_{k+1}-\phi_k&=\big(\big((3+2\eta_k)a_{k+1}-1\big)+\tfrac{\gamma_{k+1}}{2\beta\nu}\big((1-b_{k+1})^2+\eta_kb_{k+1}(b_{k+1}-1)\big)\big)F_{x,k}+\delta_k\\
&\leq-\tau F_{x,k}+\delta_k.
\end{aligned}\tag{4.6.32}$$

(b) If $a_{k+1}<\frac{\gamma_{k+1}}{2\beta\nu}b_{k+1}$, $k\in\mathbb{N}$, then from the first condition of (4.2.10), we have

$$\begin{aligned}
\phi_{k+1}-\phi_k&\leq\big(-\big(1+a_{k+1}\big)+\tfrac{\gamma_{k+1}}{2\beta\nu}\big(1+(1+\eta_k)b_{k+1}^2+(2+\eta_k)b_{k+1}\big)\big)F_{x,k}+\delta_k\\
&=\big(-\big(1+a_{k+1}\big)+\tfrac{\gamma_{k+1}}{2\beta\nu}\big((1+b_{k+1})^2+\eta_kb_{k+1}(b_{k+1}+1)\big)\big)F_{x,k}+\delta_k\\
&\leq-\tau F_{x,k}+\delta_k.
\end{aligned}\tag{4.6.33}$$

Under the assumption (i), we have from (4.6.31) (respectively (4.6.32) or (4.6.33)), that

$$\phi_{k+1}-\phi_1\leq-\tau\sum_{j=0}^{k}E_{x,j+1}+\sum_{j=1}^{k}\delta_j,$$

which leads to

$$\begin{aligned}
\sum_{j=0}^{k}E_{x,j}&\leq\tfrac{1}{\tau}(\phi_1-\phi_{k+1})+\tfrac{1}{\tau}\sum_{j=1}^{k}\delta_j\\
&\leq\tfrac{1}{\tau}(\phi_1+\bar{c}\varphi_k)+\tfrac{1}{\tau}\sum_{j=1}^{k}\delta_j<+\infty,
\end{aligned}$$

If the errors vanish, then from (4.6.31) (respectively (4.6.32) or (4.6.33)) we have

$$\varphi_k-\bar{c}\varphi_{k-1}\leq\phi_k\leq\phi_1\Longrightarrow\varphi_k\leq\bar{c}^k\varphi_0+\phi_1\sum_{j=0}^{k-1}\bar{c}^j\leq\bar{c}^k\varphi_0+\tfrac{\phi_1}{1-\bar{c}}.$$

and moreover

$$\sum_{j=0}^{k}E_{x,j}\leq\tfrac{1}{\tau}(\phi_1-\phi_{k+1})\leq\tfrac{1}{\tau}(\phi_1+\bar{c}\varphi_k)\leq\tfrac{1}{\tau}\big(\bar{c}^{k+1}\varphi_0+\tfrac{\phi_1}{1-\bar{c}}\big)<+\infty,$$

which means that the summability condition in (4.2.8) is satisfied. The rest of the proof follows the same arguments as in those in the last part of the proof of Theorem 4.2.5. □

4.6.2 Proofs of Section 4.4

Define $\Delta_k \overset{\text{def}}{=} \|x_k - x_{k-1}\|$.

Lemma 4.6.5. *For the update of x_{k+1} in (4.4.2), given any $k \in \mathbb{N}$, define*

$$g_{k+1} \overset{\text{def}}{=} \frac{1}{\gamma_k}(y_{a,k} - x_{k+1}) - \nabla F(y_{b,k}) + \nabla F(x_{k+1}).$$

We have $g_{k+1} \in \partial\Phi(x_{k+1})$, and moreover,

$$\|g_{k+1}\| \leq \big(\tfrac{1}{\underline{\gamma}} + \tfrac{1}{\beta}\big)\Delta_{k+1} + \textstyle\sum_{i\in\mathcal{S}}\big(\frac{|a_{i,k}|}{\underline{\gamma}} + |b_{i,k}|\big)\Delta_{k-i}. \tag{4.6.34}$$

Proof. From the definition of proximity operator and the update of x_{k+1} (4.4.2), we have $y_{a,k} - \gamma_k\nabla F(y_{b,k}) - x_{k+1} \in \gamma_k\partial R(x_{k+1})$, add $\gamma_k\nabla F(x_{k+1})$ to both sides, then

$$g_{k+1} = \frac{y_{a,k} - \gamma_k\nabla F(y_{b,k}) - x_{k+1} + \gamma_k\nabla F(x_{k+1})}{\gamma_k} \in \partial\Phi(x_{k+1}).$$

Then, apply the triangle inequality and the Lipschitz continuity of ∇F, we get

$$\begin{aligned}
\|g_{k+1}\| &= \|\tfrac{1}{\gamma_k}(y_{a,k} - x_{k+1}) - \nabla F(y_{b,k}) + \nabla F(x_{k+1})\| \\
&\leq \frac{1}{\gamma_k}\|y_{a,k} - x_{k+1}\| + \frac{1}{\beta}\|y_{b,k} - x_{k+1}\| \\
&\leq \frac{1}{\gamma_k}\big(\Delta_{k+1} + \textstyle\sum_{i\in\mathcal{S}}|a_{i,k}|\Delta_{k-i}\big) + \frac{1}{\beta}\big(\Delta_{k+1} + \textstyle\sum_{i\in\mathcal{S}}|b_{i,k}|\Delta_{k-i}\big) \\
&\leq \big(\tfrac{1}{\underline{\gamma}} + \tfrac{1}{\beta}\big)\Delta_{k+1} + \textstyle\sum_{i\in\mathcal{S}}\big(\frac{|a_{i,k}|}{\underline{\gamma}} + |b_{i,k}|\big)\Delta_{k-i},
\end{aligned}$$

which concludes the proof. □

Lemma 4.6.6. *For Algorithm 11, given the parameters $\gamma_k, a_{i,k}, b_{i,k}$, the following inequality holds*

$$\Phi(x_{k+1}) + \underline{\zeta}\Delta_{k+1}^2 \leq \Phi(x_k) + \textstyle\sum_{i\in\mathcal{S}}\bar{\lambda}_i\Delta_{k-i}^2. \tag{4.6.35}$$

Proof. Define the function

$$\mathcal{L}_k(x) = \gamma_k R(x) + \frac{1}{2}\|x - y_{a,k}\|^2 + \gamma_k\langle x,\, \nabla F(y_{b,k})\rangle.$$

It can be shown that the update of x_{k+1} in (4.4.2) is equivalent to

$$x_{k+1} \in \operatorname{argmin}_{x\in\mathcal{H}}\mathcal{L}_k(x), \tag{4.6.36}$$

which means that $\mathcal{L}_k(x_{k+1}) \leq \mathcal{L}_k(x_k)$, and

$$R(x_{k+1}) + \frac{1}{2\gamma_k}\|x_{k+1} - y_{a,k}\|^2 + \langle x_{k+1},\, \nabla F(y_{b,k})\rangle \leq R(x_k) + \frac{1}{2\gamma_k}\|x_k - y_{a,k}\|^2 + \langle x_k,\, \nabla F(y_{b,k})\rangle,$$

and leads to,

$$\begin{aligned}
R(x_k) &\geq R(x_{k+1}) + \frac{1}{2\gamma_k}\|x_{k+1} - y_{a,k}\|^2 + \langle x_{k+1} - x_k,\, \nabla F(y_{b,k})\rangle - \frac{1}{2\gamma_k}\|x_k - y_{a,k}\|^2 \\
&= R(x_{k+1}) + \langle x_{k+1} - x_k,\, \nabla F(x_k)\rangle + \frac{1}{2\gamma_k}\Delta_{k+1}^2 \\
&\quad + \frac{1}{\gamma_k}\langle x_k - x_{k+1},\, \textstyle\sum_{i\in\mathcal{S}}a_{i,k}(x_{k-i} - x_{k-i-1})\rangle + \langle x_{k+1} - x_k,\, \nabla F(y_{b,k}) - \nabla F(x_k)\rangle.
\end{aligned} \tag{4.6.37}$$

Since F is $(1/\beta)$-Lipschitz, then

$$\langle \nabla F(x_k),\, x_{k+1} - x_k\rangle \geq F(x_{k+1}) - F(x_k) - \frac{1}{2\beta}\Delta_{k+1}^2.$$

Apply Young's inequality, we obtain

$$\begin{aligned}\langle x_k - x_{k+1},\, \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1})\rangle &\geq -\big(\frac{\mu}{2}\Delta_{k+1}^2 + \frac{1}{2\mu}\|\textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1})\|^2\big)\\ &\geq -\big(\frac{\mu}{2}\Delta_{k+1}^2 + \textstyle\sum_{i\in\mathcal{S}} \frac{sa_{i,k}^2}{2\mu}\Delta_{k-i}^2\big),\end{aligned} \tag{4.6.38}$$

where $\mu > 0$. Then similarly, for $\nu > 0$, we have

$$\begin{aligned}\langle x_{k+1} - x_k,\, \nabla F(y_{b,k}) - \nabla F(x_k)\rangle &\geq -\big(\frac{\nu}{2}\Delta_{k+1}^2 + \frac{1}{2\nu}\|\nabla F(y_{b,k}) - \nabla F(x_k)\|^2\big)\\ &\geq -\big(\frac{\nu}{2}\Delta_{k+1}^2 + \textstyle\sum_{i\in\mathcal{S}} \frac{sb_{i,k}^2}{2\nu\beta^2}\Delta_{k-i}^2\big).\end{aligned} \tag{4.6.39}$$

Combining the above 3 inequalities with (4.6.37) leads to

$$\Phi(x_{k+1}) + \zeta_k\Delta_{k+1}^2 \leq \Phi(x_k) + \textstyle\sum_{i\in\mathcal{S}}\big(\frac{sa_{i,k}^2}{2\gamma_k\mu} + \frac{sb_{i,k}^2}{2\nu\beta^2}\big)\Delta_{k-i}^2 = \Phi(x_k) + \sum_{i\in\mathcal{S}}\lambda_{i,k}\Delta_{k-i}^2. \tag{4.6.40}$$

Therefore, we obtain

$$\begin{aligned}\Phi(x_{k+1}) + \underline{\zeta}\Delta_{k+1}^2 \leq \Phi(x_{k+1}) + \zeta_k\Delta_{k+1}^2 &\leq \Phi(x_k) + \textstyle\sum_{i\in\mathcal{S}}\lambda_{i,k}\Delta_{k-i}^2\\ &\leq \Phi(x_k) + \textstyle\sum_{i\in\mathcal{S}}\bar{\lambda}_i\Delta_{k-i}^2,\end{aligned}$$

which concludes the proof. □

Define $\boldsymbol{\mathcal{H}}$ the product space $\boldsymbol{\mathcal{H}} \overset{\text{def}}{=} \underbrace{\mathcal{H}\times\cdots\times\mathcal{H}}_{s \text{ times}}$ and $\boldsymbol{z}_k = (x_k, x_{k-1}, ..., x_{k-s+1}) \in \boldsymbol{\mathcal{H}}$. Then given $\boldsymbol{z}_k$, define the function

$$\Psi(\boldsymbol{z}_k) = \Phi(x_k) + \textstyle\sum_{i\in\mathcal{S}}\sum_{j=i}^{s-1}\bar{\lambda}_j\Delta_{k-i}^2,$$

which is is a KL function if Φ is. Denote $\mathcal{C}_{x_k}, \mathcal{C}_{\boldsymbol{z}_k}$ the set of cluster points of sequences x_k and $\boldsymbol{z}_k$ respectively, and $\mathrm{crit}(\Psi) = \{\boldsymbol{z} = (x, ..., x) \in \boldsymbol{\mathcal{H}} : x \in \mathrm{crit}(\Phi)\}$.

Lemma 4.6.7. *For Algorithm 11, choose $\mu, \nu, \gamma_k, a_{i,k}, b_{i,k}$ such that (4.4.5) holds. If Φ is bounded from below, then*

(i) $\sum_{k\in\mathbb{N}}\Delta_k^2 < +\infty$.

(ii) *The sequence $\Psi(\boldsymbol{z}_k)$ is monotonically decreasing and convergent.*

(iii) *The sequence $\Phi(x_k)$ is convergent.*

Proof. Define

$$\tau = \underline{\zeta} - \textstyle\sum_{i\in\mathcal{S}}\bar{\lambda}_i > 0.$$

From the Lemma 4.6.6, we have

$$\tau\Delta_{k+1}^2 \leq \big(\Phi(x_k) - \Phi(x_{k+1})\big) + \textstyle\sum_{i\in\mathcal{S}}\bar{\lambda}_i(\Delta_{k-i}^2 - \Delta_{k+1}^2).$$

Since we let $x_{1-s} = ... = x_0 = x_1$, for the above inequality, sum over k we get

$$\begin{aligned}\tau\textstyle\sum_{k\in\mathbb{N}}\Delta_{k+1}^2 &\leq \textstyle\sum_{k\in\mathbb{N}}\big(\Phi(x_k) - \Phi(x_{k+1})\big) + \sum_{k\in\mathbb{N}}\sum_{i\in\mathcal{S}}\bar{\lambda}_i(\Delta_{k-i}^2 - \Delta_{k+1}^2)\\ &\leq \Phi(x_0) + \textstyle\sum_{i\in\mathcal{S}}\bar{\lambda}_i\sum_{k\in\mathbb{N}}(\Delta_{k-i}^2 - \Delta_{k+1}^2)\\ &= \Phi(x_0) + \textstyle\sum_{i\in\mathcal{S}}\bar{\lambda}_i\sum_{j=1-i}^{1}\Delta_j^2 = \Phi(x_0),\end{aligned}$$

which means, as $\Phi(x_0)$ is bounded,

$$\textstyle\sum_{k\in\mathbb{N}}\Delta_{k+1}^2 \leq \frac{\Phi(x_0)}{\tau} < +\infty.$$

From Lemma 4.6.6, by pairing terms on both sides of (4.6.35), we get

$$\Psi(\boldsymbol{z}_{k+1}) + \big(\underline{\zeta} - \textstyle\sum_{i\in\mathcal{S}}\bar{\lambda}_i\big)\Delta_{k+1}^2 \leq \Psi(\boldsymbol{z}_k).$$

Since we assume $\underline{\zeta} - \sum_{i\in\mathcal{S}}\bar{\lambda}_i > 0$, hence $\Psi(\boldsymbol{z}_k)$ is monotonically non-increasing. The convergence of $\Phi(x_k)$ is straightforward. □

Lemma 4.6.8. *For Algorithm 11, choose $\mu, \nu, \gamma_k, a_{i,k}, b_{i,k}$ such that (4.4.5) holds. If Φ is bounded from below and $\{x_k\}_{k\in\mathbb{N}}$ is bounded, then x_k converges to a critical point of Φ.*

Proof. Since $\{x_k\}_{k\in\mathbb{N}}$ is bounded, there exists a subsequence $\{x_{k_j}\}_{k\in\mathbb{N}}$ and cluster point $\bar{x}$ such that $x_{k_j} \to \bar{x}$ as $j \to \infty$. Next we show that $\Phi(x_{k_j}) \to \Phi(\bar{x})$ and that $\bar{x}$ is a critical point of Φ.

Since R is lsc, then $\liminf_{j\to\infty} R(x_{k_j}) \geq R(\bar{x})$. From (4.6.36), we have $\mathcal{L}_{k_j-1}(x_{k_j}) \leq \mathcal{L}_{k_j-1}(\bar{x})$,

$$\begin{aligned} R(\bar{x}) &\geq R(x_{k_j}) + \tfrac{1}{2\gamma_{k_j-1}}\|x_{k_j} - y_{a,k_j-1}\|^2 + \langle x_{k_j} - \bar{x}, \nabla F(y_{b,k_j-1})\rangle - \tfrac{1}{2\gamma_{k_j-1}}\|\bar{x} - y_{a,k_j-1}\|^2 \\ &= R(x_{k_j}) + \tfrac{1}{2\gamma_{k_j-1}}(\|x_{k_j} - \bar{x}\|^2 + 2\langle x_{k_j} - \bar{x}, \bar{x} - y_{a,k_j-1}\rangle) + \langle x_{k_j} - \bar{x}, \nabla F(y_{b,k_j-1})\rangle \end{aligned}$$

Since $\Delta_k^2 \to 0$ and $x_{k_j} \to \bar{x}$, then taking the inequality to limit yields $\limsup_{j\to\infty} R(x_{k_j}) \leq R(\bar{x})$. As a result, $\lim_{k\to\infty} R(x_{k_j}) = R(\bar{x})$. Since F is continuous, then $F(x_{k_j}) \to F(\bar{x})$, hence $\Phi(x_{k_j}) \to \Phi(\bar{x})$.

Furthermore, owing to Lemma 4.6.5, $g_{k_j} \in \partial\Phi(x_{k_j})$, and (i) of Lemma 4.6.7 we have $g_{k_j} \to 0$ as $k \to \infty$. As a consequence,

$$g_{k_j} \in \partial\Phi(x_{k_j}),\ (x_{k_j}, g_{k_j}) \to (\bar{x}, 0) \quad \text{and} \quad \Phi(x_{k_j}) \to \Phi(\bar{x}),$$

as $j \to \infty$. Hence $0 \in \partial\Phi(\bar{x})$, *i.e.* $\bar{x}$ is a critical point. □

Now we present the proof of Theorem 4.4.3.

Proof of Theorem 4.4.3. Concluding from above lemmas, we have the following useful results

(**R.1**) Denote $\tau = \zeta - \sum_{i\in\mathcal{S}} \bar{\lambda}_i$, then $\Psi(\boldsymbol{z}_{k+1}) + \tau\Delta_{k+1}^2 \leq \Psi(\boldsymbol{z}_k)$.

(**R.2**) Define

$$\boldsymbol{w}_{k+1} \overset{\text{def}}{=} \begin{pmatrix} g_{k+1} + 2\sum_{i=0}^{s-1} \bar{\lambda}_i(x_{k+1} - x_k) \\ 2\sum_{i=0}^{s-1} \bar{\lambda}_i(x_k - x_{k+1}) + 2\sum_{i=1}^{s-1} \bar{\lambda}_i(x_k - x_{k-1}) \\ \vdots \\ 2\bar{\lambda}_{s-1}(x_{k+2-s} - x_{k+1-s}) \end{pmatrix},$$

then we have $\boldsymbol{w}_{k+1} \in \partial\Psi(\boldsymbol{z}_{k+1})$. Owing to Lemma 4.6.5, there exists a $\sigma > 0$ such that $\|\boldsymbol{w}_{k+1}\| \leq \sigma \sum_{j=k+2-s}^{k+1} \Delta_j$.

(**R.3**) if x_{k_j} is a subsequence such that $x_{k_j} \to \bar{x}$, then $\Psi(\boldsymbol{z}_k) \to \Psi(\bar{\boldsymbol{z}})$ where $\bar{\boldsymbol{z}} = (\bar{x}, ..., \bar{x})$.

(**R.4**) $\mathcal{C}_{\boldsymbol{z}_k} \subseteq \mathrm{crit}(\Psi)$.

(**R.5**) $\lim_{k\to\infty} \mathrm{dist}(\boldsymbol{z}_k, \mathcal{C}_{\boldsymbol{z}_k}) = 0$.

(**R.6**) $\mathcal{C}_{\boldsymbol{z}_k}$ is non-empty, compact and connected.

(**R.7**) Ψ is finite and constant on $\mathcal{C}_{\boldsymbol{z}_k}$.

Next we prove the claims of Theorem 4.4.3.

(i) Consider a critical point of Φ, $\bar{x} \in \mathrm{crit}(\Phi)$, such that $\bar{\boldsymbol{z}} = (\bar{x}, ..., \bar{x}) \in \mathcal{C}_{\boldsymbol{z}_k}$, then owing to (**R.3**), we have $\Psi(\boldsymbol{z}_k) \to \Psi(\bar{\boldsymbol{z}})$.

Suppose there exists K such that $\Psi(z_K) = \Psi(\bar{\boldsymbol{z}})$, then the descent property (**R.1**) implies that $\Psi(\boldsymbol{z}_k) = \Psi(\bar{\boldsymbol{z}})$ holds for all $k \geq K$. Then $\boldsymbol{z}_k$ is constant for $k \geq K$, hence has finite length.

On the other hand, let $\Psi(\boldsymbol{z}_k) > \Psi(\bar{\boldsymbol{z}})$, denote $\psi_k = \Psi(\boldsymbol{z}_k) - \Psi(\bar{\boldsymbol{z}})$. Owing to (**R.6**), (**R.7**) and Definition 4.4.2, the KL property of Ψ means that there exist ϵ, η and a concave function φ, and

$$\mathcal{U} \overset{\text{def}}{=} \{\boldsymbol{u} \in \boldsymbol{\mathcal{H}} : \mathrm{dist}(\boldsymbol{u}, \mathcal{C}_{\boldsymbol{z}_k}) < \epsilon\} \bigcap [\Psi(\bar{\boldsymbol{z}}) < \Psi(\boldsymbol{u}) < \Psi(\bar{\boldsymbol{z}}) + \eta], \tag{4.6.41}$$

sucht ath for all $\boldsymbol{z} \in \mathcal{U}$,

$$\varphi'\big(\Psi(\boldsymbol{z}) - \Psi(\bar{\boldsymbol{z}})\big)\mathrm{dist}\big(0, \partial\Psi(\boldsymbol{z})\big) \geq 1. \tag{4.6.42}$$

Let $k_1 \in \mathbb{N}$ be such that $\Psi(\boldsymbol{z}_k) < \Psi(\bar{\boldsymbol{z}}) + \eta$ holds for all $k \geq k_1$. Owing to (**R.5**), there exists another $k_2 \in \mathbb{N}$ such that $\operatorname{dist}(\boldsymbol{z}_k, \mathcal{C}_{\boldsymbol{z}_k}) < \epsilon$ holds for all $k \geq k_2$. Let $K = \max\{k_1, k_2\}$, then $\boldsymbol{z}_k \in \mathcal{U}$ holds for all $k \geq K$. Then from (4.6.42), we have for $k \geq K$

$$\varphi'(\psi_k)\operatorname{dist}\big(0, \partial\Psi(\boldsymbol{z}_k)\big) \geq 1.$$

Since φ is concave, φ' is decreasing, and $\Psi(\boldsymbol{z}_k)$ is decreasing, we have

$$\varphi(\psi_k) - \varphi(\psi_{k+1}) \geq \varphi'(\psi_k)\big(\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}_{k+1})\big) \geq \frac{\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}_{k+1})}{\operatorname{dist}(0, \partial\Psi(\boldsymbol{z}_k))}.$$

From (**R.1**), since $\operatorname{dist}(0, \partial\Psi(\boldsymbol{z}_k)) \leq \|w_k\|$, then

$$\varphi(\psi_k) - \varphi(\psi_{k+1}) \geq \frac{\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}_{k+1})}{\|w_k\|} \geq \frac{\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}_{k+1})}{\sigma \sum_{j=k+1-s}^{k} \Delta_j}.$$

Moreover, $\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}_{k+1}) \geq \delta\Delta_{k+1}^2$ from (**R.2**), therefore we get

$$\varphi(\psi_k) - \varphi(\psi_{k+1}) \geq \frac{\delta\Delta_{k+1}^2}{\sigma \sum_{j=k+1-s}^{k} \Delta_j},$$

which yields

$$\Delta_{k+1}^2 \leq \big(\tfrac{\sigma}{\delta}(\varphi(\psi_k) - \varphi(\psi_{k+1}))\big) \sum_{j=k+1-s}^{k} \Delta_j. \tag{4.6.43}$$

Taking the square root of both sides and applying Young's inequality with $\kappa > 0$, we further obtain

$$\begin{aligned} 2\Delta_{k+1} &\leq \frac{1}{\kappa} \sum_{j=k+1-s}^{k} \Delta_j + \frac{\kappa\sigma}{\delta}(\varphi(\psi_k) - \varphi(\psi_{k+1})) \\ (\kappa = s) \quad &\leq \frac{1}{s} \sum_{j=k+1-s}^{k} \Delta_j + \frac{s\sigma}{\delta}(\varphi(\psi_k) - \varphi(\psi_{k+1})). \end{aligned} \tag{4.6.44}$$

Summing up both sides over k, and since $x_0 = \ldots = x_{-s}$, we get

$$\ell \overset{\text{def}}{=} \sum\nolimits_{k\in\mathbb{N}} \Delta_k \leq \Delta_1 + \frac{s\sigma}{\delta}\varphi(\psi_1) < +\infty,$$

which concludes the finite length property of x_k.

(ii) Follows from the fact that $\{x_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence, we have $\{x_k\}_{k\in\mathbb{N}}$ is hence convergent. Owing to Lemma 4.6.8, there exists a critical point $x^\star \in \operatorname{crit}(\Phi)$ such that $\lim_{k\to\infty} x_k = x^\star$.

(iii) We now turn to proving local convergence to a global minimmizer. Note that if $x^\star$ is a global minimizer of Φ, then $\boldsymbol{z}^\star$ is a global minimizer of Ψ. Let $r > \rho > 0$ such that $\mathbb{B}_r(\boldsymbol{z}^\star) \subset \mathcal{U}$ and $\eta < \delta(r-\rho)^2$. Suppose that the initial point x_0 is chosen such that following conditions hold,

$$\Psi(\boldsymbol{z}^\star) \leq \Psi(\boldsymbol{z}_0) < \Psi(\boldsymbol{z}^\star) + \eta \tag{4.6.45}$$

$$\|x_0 - x^\star\| + \ell(s-1) + 2\sqrt{\frac{\Psi(\boldsymbol{z}_0) - \Psi(\boldsymbol{z}^\star)}{\delta}} + \frac{\sigma}{\delta}\varphi(\psi_0) < \rho. \tag{4.6.46}$$

The descent property (**R.1**) of Ψ together with (4.6.45) imply that for any $k \in \mathbb{N}$, $\Psi(\boldsymbol{z}^\star) \leq \Psi(\boldsymbol{z}_{k+1}) \leq \Psi(\boldsymbol{z}_k) \leq \Psi(\boldsymbol{z}_0) < \Psi(\boldsymbol{z}^\star) + \eta$, and

$$\|x_{k+1} - x_k\| \leq \sqrt{\frac{\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}_{k+1})}{\delta}} \leq \sqrt{\frac{\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}^\star)}{\delta}}. \tag{4.6.47}$$

Therefore, given any $k \in \mathbb{N}$, if we have $x_k \in \mathbb{B}_\rho(x^\star)$, then

$$\begin{aligned} \|x_{k+1} - x^\star\| &\leq \|x_k - x^\star\| + \|x_{k+1} - x_k\| \leq \|x_k - x^\star\| + \sqrt{\frac{\Psi(\boldsymbol{z}_k) - \Psi(\boldsymbol{z}^\star)}{\delta}} \\ &\leq \rho + (r - \rho) = r, \end{aligned} \tag{4.6.48}$$

which means that $x_{k+1} \in \mathbb{B}_r(x^\star)$.

For any $k \in \mathbb{N}$, define the following partial sum

$$p_k = \sum\nolimits_{j=k+1-s}^{k-1} \sum\nolimits_{i=1}^{j} \Delta_i.$$

Note that $p_k = 0$ for $k = 1$, and $\lim_{k\to\infty} p_k = \ell(s-1)$. Next we prove the following claims through induction: for $k \in \mathbb{N}$

$$x_k \in \mathbb{B}_\rho(x^\star) \tag{4.6.49}$$

$$\sum_{j=1}^{k}\Delta_{j+1} + \Delta_{k+1} \leq \Delta_1 + p_k + \frac{\sigma}{\delta}(\varphi(\psi_1) - \varphi(\psi_{k+1})). \tag{4.6.50}$$

From (4.6.47) we have

$$\|x_1 - x_0\| \leq \sqrt{\frac{\Psi(\boldsymbol{z}_0) - \Psi(\boldsymbol{z}^\star)}{\delta}}. \tag{4.6.51}$$

Applying the triangle inequality we then obtain

$$\|x_1 - x^\star\| \leq \|x_0 - x^\star\| + \|x_1 - x_0\| \leq \|x_0 - x^\star\| + \sqrt{\frac{\Psi(\boldsymbol{z}_0) - \Psi(\boldsymbol{z}^\star)}{\delta}} < \rho,$$

which means $x_1 \in \mathbb{B}_\rho(x^\star)$. Now, taking $\kappa = 1$ in (4.6.44) yields, for any $k \in \mathbb{N}$,

$$2\Delta_{k+1} \leq \sum_{j=k+1-s}^{k}\Delta_j + \frac{\sigma}{\delta}(\varphi(\psi_k) - \varphi(\psi_{k+1})). \tag{4.6.52}$$

Let $k = 1$. Since $x_0 = \ldots = x_{-s}$, we have

$$2\Delta_2 \leq \Delta_1 + \frac{\sigma}{\delta}(\varphi(\psi_1) - \varphi(\psi_2)).$$

Therefore, (4.6.49) and (4.6.50) hold for $k = 1$.
Now assume that they hold for some $k > 1$. Using the triangle inequality and (4.6.50),

$$\begin{aligned}
\|x_{k+1} - x^\star\| &\leq \|x_0 - x^\star\| + \Delta_1 + \sum_{j=1}^{k}\Delta_j \\
&\leq \|x_0 - x^\star\| + 2\Delta_1 + p_k + \frac{\sigma}{\delta}(\varphi(\psi_1) - \varphi(\psi_{k+1})) \\
&\leq \|x_0 - x^\star\| + 2\Delta_1 + (s-1)\ell + \frac{\sigma}{\delta}(\varphi(\psi_1) - \varphi(\psi_{k+1})) \\
(4.6.51) &\leq \|x_0 - x^\star\| + 2\sqrt{\frac{\Psi(\boldsymbol{z}_0) - \Psi(\boldsymbol{z}^\star)}{\delta}} + (s-1)\ell + \frac{\sigma}{\delta}(\varphi(\psi_1) - \varphi(\psi_{k+1})).
\end{aligned}$$

As $\varphi(\psi) \geq 0$ and $\varphi'(\psi) > 0$ for $\psi \in]0, \eta[$, and in view of (4.6.46), we arrive at

$$\|x_{k+1} - x^\star\| \leq \|x_0 - x^\star\| + 2\sqrt{\frac{\Psi(\boldsymbol{z}_0) - \Psi(\boldsymbol{z}^\star)}{\delta}} + (s-1)\ell + \frac{\sigma}{\delta}\varphi(\psi_0) < \rho$$

whence we deduce that (4.6.49) holds at $k+1$. Now, taking (4.6.52) at $k+1$ gives

$$\begin{aligned}
2\Delta_{k+2} &\leq \sum_{j=k+2-s}^{k+1}\Delta_j + \frac{\sigma}{\delta}(\varphi(\psi_{k+1}) - \varphi(\psi_{k+2})) \\
&\leq \Delta_{k+1} + \sum_{j=k+2-s}^{k}\Delta_j + \frac{\sigma}{\delta}(\varphi(\psi_{k+1}) - \varphi(\psi_{(k+2)})).
\end{aligned} \tag{4.6.53}$$

Adding both sides of (4.6.53) and (4.6.50) we get

$$\begin{aligned}
\sum_{j=1}^{k+1}\Delta_{j+1} + \Delta_{k+2} &\leq \Delta_1 + p_k + \sum_{j=k+2-s}^{k}\Delta_j + \frac{\sigma}{\delta}(\varphi(\psi_1) - \varphi(\psi_{k+2})) \\
&= \Delta_1 + p_{k+1} + \frac{\sigma}{\delta}(\varphi(\psi_1) - \varphi(\psi_{k+2})),
\end{aligned}$$

meaning that (4.6.50) holds at $k+1$. This concludes the induction proof.
In summary, the above result shows that if we start close enough from $x^\star$ (so that (4.6.45)-(4.6.46) hold), then the sequence $\{x_k\}_{k\in\mathbb{N}}$ will remain in the neighbourhood $\mathbb{B}_\rho(x^\star)$ and thus converges to a critical point $\bar{x}$ owing to Lemma 4.6.8. Moreover, $\Psi(\boldsymbol{z}_k) \to \Psi(\bar{\boldsymbol{z}}) \geq \Psi(\boldsymbol{z}^\star)$ by virtue of (**R.3**). Now we need to show that $\Psi(\bar{\boldsymbol{z}}) = \Psi(\boldsymbol{z}^\star)$. Suppose that $\Psi(\bar{\boldsymbol{z}}) > \Psi(\boldsymbol{z}^\star)$. As Ψ has the KL property at $\boldsymbol{z}^\star$, we have

$$\varphi'\big(\Psi(\bar{\boldsymbol{z}}) - \Psi(\boldsymbol{z}^\star)\big)\operatorname{dist}\big(0, \partial\Psi(\bar{\boldsymbol{z}})\big) \geq 1.$$

But this is impossible since $\varphi'(s) > 0$ for $s \in]0, \eta[$, and $\operatorname{dist}\big(0, \partial\Psi(\bar{\boldsymbol{z}})\big) = 0$ as $\bar{\boldsymbol{z}}$ is a critical point. Hence we have $\Psi(\bar{\boldsymbol{z}}) = \Psi(\boldsymbol{z}^\star)$, which means $\Phi(\bar{x}) = \Phi(x^\star)$, *i.e.* the cluster point $\bar{x}$ is actually a global minimizer. This concludes the proof. □

Part II

Local Linear Convergence under Partial Smoothness

Chapter 5

Partial Smoothness

Contents

From the numerical experiments presented in Section 4.5, we obtain the following local convergence observations which cannot be explained by metric sub-regularity:

(i) The local linear convergence of the inertial schemes (Theorem 3.3.3 needs the fixed-point operator to be non-expansive, which is not the case for inertial schemes).

(ii) For the multi-step inertial Forward–Backward and FISTA, $\|x_k - x^\star\|$ locally oscillates when the inertial momentums are too large (Figure 4.4 and 4.5).

(iii) For DR and Primal–Dual splitting, when the objective functions are polyhedral, locally Primal–Dual splitting oscillates even without inertia (Figure 4.8), while no oscillations observed for the DR (Figure 4.7).

In order to interpret these observations, new tools are needed and it turns out that partial smoothness is the key notion. In this chapter we present an overview of the theory of partial smoothness, and many results are borrowed for instance from the work of Lewis and his collaborators.

5.1 Partly smooth functions

The notion of "partial smoothness" was first introduced by Lewis in [108]. This concept, as well as that of identifiable surfaces [174], captures the essential features of the geometry of non-smoothness which are along the so-called active/identifiable sub-manifold. For convex functions, a closely related idea is developed in [106]. Loosely speaking, a partly smooth function behaves smoothly as we move along the identifiable submanifold, and transversal to it, the function behaves sharply. Consequently, the behaviour of the function and of its minimizers depend essentially on its restriction to this manifold, hence offering a powerful framework for algorithmic and sensitivity analysis theory.

5.1.1 Definition

Let $\mathcal{M}$ be a C^2-smooth embedded submanifold of $\mathbb{R}^n$ around a point x. To lighten terminology, henceforth we shall state C^2-manifold instead of C^2-smooth embedded submanifold of $\mathbb{R}^n$. The natural embedding of a submanifold $\mathcal{M}$ into $\mathbb{R}^n$ permits to define a Riemannian structure on $\mathcal{M}$, and we simply say $\mathcal{M}$ is a Riemannian manifold. $\mathcal{T}_{\mathcal{M}}(x)$ denotes the tangent space to $\mathcal{M}$ at any point x in $\mathcal{M}$.

Given $R : \mathbb{R}^n \to \overline{\mathbb{R}}$ and a point $x \in \mathrm{dom}(R)$ such that the sub-differential $\partial R(x)$ of R at x is non-empty, define the following subspace

$$T_x \stackrel{\text{def}}{=} \mathrm{par}\big(\partial R(x)\big)^{\perp}. \tag{5.1.1}$$

The definition below is adapted from [108].

Definition 5.1.1 (Partly smooth function). A function $R : \mathbb{R}^n \to \overline{\mathbb{R}}$ is partly smooth at $\bar{x}$ relative to a set $\mathcal{M}$ for $\bar{v} \in \partial R(\bar{x}) \neq \emptyset$ if $\mathcal{M}$ is a C^2 manifold around $\bar{x}$, and:

Smoothness	R restricted to $\mathcal{M}$ is C^2 around $\bar{x}$;
Regularity	R is regular at all $x \in \mathcal{M}$ near $\bar{x}$, with $\partial R(x) \neq \emptyset$, and prox-regular at $\bar{x}$ for $\bar{v}$;
Sharpness	The tangent space $\mathcal{T}_{\mathcal{M}}(\bar{x}) = \mathrm{par}(\partial R(\bar{x}))^{\perp}$;
Continuity	The set-valued mapping ∂R is continuous at x relative to $\mathcal{M}$.

The class of partly smooth functions at $\bar{x}$ relative to $\mathcal{M}$ for $\bar{v}$ is denoted as $\mathrm{PSF}_{\bar{x},\bar{v}}(\mathcal{M})$, when $R \in \Gamma_0(\mathbb{R}^n)$ is convex, we short this notion as $\mathrm{PSF}_{\bar{x}}(\mathcal{M})$. Popular examples of partly smooth functions that are widely used in signal/image processing, machine learning and statistics are discussed in Section 5.2.

Remark 5.1.2.

(i) When $R \in \Gamma_0(\mathbb{R}^n)$, *i.e.* R is convex, the regularity assumption in Definition 5.1.1 is fulfilled automatically [143, Example 2.6].

(ii) Prox-regularity was not included in the original definition of [108], but it is essential to ensure that the partly smooth submanifolds are locally unique [89, Corollary 4.2].

It is obvious that smooth differentiable functions of $\mathbb{R}^n$ are partly smooth relative to $\mathcal{M} = \mathbb{R}^n$. If $R \in \Gamma_0(\mathbb{R}^n)$ is locally polyhedral around x, then R is partly smooth at x relative to $x + T_x$. Moreover, polyhedrality implies that the subdifferential is locally constant around x along $x + T_x$.

Owing to the result of [108, Proposition 2.10], we have the following lemma.

Lemma 5.1.3 (Local normal sharpness). *If $R \in \mathrm{PSF}_x(\mathcal{M})$, then for all $x' \in \mathcal{M}$ near x, there holds $\mathcal{T}_{\mathcal{M}}(x') = T_{x'}$. In particular, if $\mathcal{M}$ is affine or linear, then $T_{x'} = T_x$.*

5.1.2 Riemann gradient and Hessian

We now present the expressions of the Riemannian gradient and Hessian for the case of partly smooth functions relative to a C^2-smooth manifold. This is summarized in the following fact which follows by combining (2.6.3), (2.6.4), Definition 5.1.1, Lemma 5.1.3 and [70, Proposition 17] (or [126, Lemma 2.4]).

Lemma 5.1.4 (Riemannian gradient and Hessian). *If $R \in \mathrm{PSF}_x(\mathcal{M})$, then for any $x' \in \mathcal{M}$ near x*

$$\nabla_{\mathcal{M}} R(x') = \mathcal{P}_{T_{x'}}(\partial R(x')),$$

and this does not depend on the smooth representation of R on $\mathcal{M}$. In turn, for all $h \in T_{x'}$

$$\nabla^2_{\mathcal{M}} R(x')h = \mathcal{P}_{T_{x'}} \nabla^2 \widetilde{R}(x')h + \mathfrak{W}_{x'}\big(h, \mathcal{P}_{T_{x'}^{\perp}} \nabla \widetilde{R}(x')\big),$$

where $\widetilde{R}$ is a smooth extension (representative) of R on $\mathcal{M}$, and $\mathfrak{W}_x(\cdot,\cdot) : T_x \times T_x^{\perp} \to T_x$ is the Weingarten map of $\mathcal{M}$ at x (see Section 2.6 for definitions).

An important implication of Lemma 5.1.4 is that it allows us to linearize the proximity operators along the manifold. Suppose $R \in \Gamma_0(\mathbb{R}^n)$ is partly smooth at x relative to $\mathcal{M}$, and there exists a point $z \in \mathbb{R}^n$ such that $x = \mathrm{prox}_R(z)$. From the Definition 2.1.8 of the proximity operator, we have $z - x \in \partial R(x)$, then project onto the tangent space T_x we get $\mathcal{P}_{T_x}(\partial R(x)) = \nabla_{\mathcal{M}} R(x) = \mathcal{P}_{T_x}(z - x)$.

5.1.3 Identifiability

The following theorem provides a key property of a partly smooth function, namely that the active manifold, relative to which the function is partly smooth, can be identified in finite time under a reasonable non degeneracy condition.

Theorem 5.1.5 (Finite identification). *Let the function $R : \mathbb{R}^n \to \overline{\mathbb{R}}$ be partly smooth at the point $\bar{x}$ relative to the manifold $\mathcal{M}$, with*

$$\bar{v} \in \mathrm{ri}\big(\partial R(\bar{x})\big). \tag{ND}$$

Suppose that there exists a sequence $\{x_k\}_{k\in\mathbb{N}}$ such that $x_k \to \bar{x}$ and $R(x_k) \to R(\bar{x})$, then

$$x_k \in \mathcal{M} \text{ for all large } k,$$

if and only if

$$\mathrm{dist}\big(\bar{v}, \partial R(x_k)\big) \to 0.$$

Proof. [89, Theorem 5.3] applied to the affine function $\langle -\bar{v}, \cdot \rangle + R$. □

Remark 5.1.6. If $R \in \Gamma_0(\mathbb{R}^n)$, then R is sub-differentially continuous, and we have $R(x_k) \to R(\bar{x})$ as long as $x_k \to \bar{x}$.

5.1.4 Calculus rules

The class of partly smooth functions is closed under addition, pre-composition by a linear operator and spectral lifting. In the following, we introduce some basic calculus rules of partial smoothness which are proved in [108, 69].

Consider two Euclidean spaces $\mathcal{X}$ and $\mathcal{Z}$, an open set $\mathcal{W} \subset \mathcal{Z}$ containing a point z, a smooth map $\mathcal{K} : \mathcal{W} \to \mathcal{X}$, and a set $\mathcal{M} \subset \mathcal{X}$. We say $\mathcal{K}$ is transversal to $\mathcal{M}$ at z if $\mathcal{M}$ is a manifold around $\mathcal{K}(z)$, and

$$\mathrm{ran}\big(\nabla \mathcal{K}(z)\big) + \mathcal{T}_{\mathcal{M}}\big(\mathcal{K}(z)\big) = \mathcal{X},$$

or equivalently

$$\ker\big(\nabla \mathcal{K}(z)^*\big) \cap \mathcal{N}_{\mathcal{M}}\big(\mathcal{K}(z)\big) = \{0\}.$$

Theorem 5.1.7 (Chain rule). *Given Euclidean spaces $\mathcal{X}$ and $\mathcal{Z}$, an open set $\mathcal{W} \subset \mathcal{Z}$ containing a point $\bar{z}$. Suppose $\mathcal{K}$ is transversal to $\mathcal{M}$ at $\bar{z}$. Then if the function $R : \mathcal{X} \to \overline{\mathbb{R}}$ is partly smooth at $\mathcal{K}(\bar{z})$ relative to $\mathcal{M}$, then the composition $R(\mathcal{K})$ is partly smooth at $\bar{z}$ relative to $\mathcal{K}^{-1}(\mathcal{M})$.*

Proof. Theorem 4.2 of [108]. □

Theorem 5.1.8 (Separability). *Let m be a positive integer. For each $i = 1, 2, ..., m$, suppose $\mathcal{X}_i$ is a real Euclidean space, that the set $\mathcal{M}_i \subset \mathcal{X}_i$ contains the point $\bar{x}_i$, and that the function $R_i : \mathcal{X}_i \to \overline{\mathbb{R}}$ is partly smooth at $\bar{x}_i$ relative to $\mathcal{M}_i$. Then the function $\boldsymbol{R} : \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_m \to \overline{\mathbb{R}}$ defined by*

$$\boldsymbol{R}(x_1, x_2, ..., x_m) = \textstyle\sum_{i=1}^m R_i(x_i), \;\; x_i \in \mathcal{X}_i, \;\; i = 1, 2, ..., m,$$

is partly smooth at $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_m)$ relative to $\mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \mathcal{M}_m$.

Proof. Proposition 4.5 of [108]. □

Theorem 5.1.9 (Sum rule). *Let m be a positive integer. Consider the sets $\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_m$ in a real Euclidean space $\mathcal{X}$. Suppose that for each $i = 1, ..., m$, the function $R_i : \mathcal{X} \to \overline{\mathbb{R}}$ is partly smooth at $\bar{x}$ relative to $\mathcal{M}_i$. Assume moreover that*

$$\sum_{i=1}^{m} y_i = 0 \text{ and } y_i \in \mathcal{N}_{\mathcal{M}_i}(\bar{x}) \text{ for each } i \implies y_i = 0 \text{ for each } i.$$

Then the function $\sum_i R_i$ is partly smooth at $\bar{x}$ relative to the intersection $\cap_i \mathcal{M}_i$.

Proof. This a corollary of Theorem 5.1.7, see also Corollary 4.6 of [108]. □

Corollary 5.1.10 (Smooth perturbation). *If the function $R : \mathcal{X} \to \overline{\mathbb{R}}$ is partly smooth at the point $\bar{x}$ relative to the set $\mathcal{M} \subset \mathcal{X}$, and the function $F : \mathcal{X} \to \overline{\mathbb{R}}$ is C^2-smooth on an open set containing $\bar{x}$, then the function $R + F$ is partly smooth at $\bar{x}$ relative to $\mathcal{M}$.*

Proof. This a consequence of Theorem 5.1.9. □

A function R is called absolutely permutation-invariant, if it is invariant under all signed permutations of coordinates. For a matrix $x \in \mathbb{R}^{m \times n}$, the operator $\sigma(x)$ returns the singular values of x sorted in descending order.

Theorem 5.1.11 (Spectral lifting). *Let function $R \in \Gamma_0(\mathbb{R}^m)$ and $\bar{x} \in \mathbb{R}^{m \times n}$ be a $m \times n$ real matrix (suppose $m \leq n$). Suppose that R is absolutely permutation-invariant locally around $\sigma(\bar{x})$. Then R is partly smooth at $\sigma(\bar{x})$ relative to $\mathcal{M} \subset \mathbb{R}^m$ if and only if $R \circ \sigma$ is partly smooth at $\bar{x}$ relative to $\sigma^{-1}(\mathcal{M})$.*

Proof. Theorem 5.3 of [69]. □

5.2 Examples

In the following, we present several popular examples of partly smooth functions that are widely used in fields like inverse problems, signal/image processing and machine learning, etc.

5.2.1 The convex case

Example 5.2.1 (ℓ_1-norm). For $x \in \mathbb{R}^n$, the ℓ_1-norm is defined as

$$R(x) = \|x\|_1 \stackrel{\text{def}}{=} \sum_{i=1}^{n} |x_i|,$$

which is polyhedral, hence partly smooth at any x relative to the subspace

$$\mathcal{M} = T_x \stackrel{\text{def}}{=} \big\{u \in \mathbb{R}^n : \operatorname{supp}(u) \subseteq \operatorname{supp}(x)\big\}, \ \operatorname{supp}(x) \stackrel{\text{def}}{=} \{i : x_i \neq 0\}.$$

Its Riemannian gradient at x is $\operatorname{sign}(x_i)$ for $i \in \operatorname{supp}(x)$ and 0 otherwise, where

$$\operatorname{sign}(x_i) = \begin{cases} +1, & x_i > 0, \\ -1, & x_i < 0. \end{cases}$$

Its Riemannian Hessian vanishes as for any polyhedral functions.

Example 5.2.2 ($\ell_{1,2}$-norm). Let the index set $\{1, \ldots, n\}$ be partitioned into a set of m non-overlapping blocks $\mathcal{B}$ such that $\bigcup_{\flat \in \mathcal{B}} \flat_j = \{1, \ldots, n\}$. The $\ell_{1,2}$-norm of x is defined by

$$R(x) = \|x\|_{1,2} \stackrel{\text{def}}{=} \sum_{\flat \in \mathcal{B}} \|x_\flat\|,$$

where $x_\flat = (x_i)_{i \in \flat} \in \mathbb{R}^{|\flat|}$. Though this function is not polyhedral, it is easy to see that it is partly smooth at x relative to the subspace

$$\mathcal{M} = T_x \stackrel{\text{def}}{=} \big\{u \in \mathbb{R}^n : \operatorname{supp}_{\mathcal{B}}(u) \subseteq \mathcal{S}_{\mathcal{B}}\big\}, \ \mathcal{S}_{\mathcal{B}} \stackrel{\text{def}}{=} \bigcup \{\flat : x_\flat \neq 0\}.$$

It is straightforward to show that the Riemannian gradient and Hessian of $\ell_{1,2}$-norm are

$$\left(\nabla_{\mathcal{M}}\|x\|_{1,2}\right)_\flat = \begin{cases} x_\flat/\|x_\flat\| & \text{if } x_\flat \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \nabla^2_{\mathcal{M}}\|x\|_{1,2}(x) = \delta_x Q_{x^\perp},$$

where,

$$\delta_x : T_x \to T_x, v \mapsto \begin{cases} \frac{v_\flat}{\|x_\flat\|} & \text{if } x_\flat \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$Q_{x^\perp} : T_x \to T_x, v \mapsto \begin{cases} v_\flat - \frac{\langle x_\flat, v_\flat \rangle}{\|x_\flat\|^2} x_\flat & \text{if } x_\flat \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Example 5.2.3 (Total Variation). Let $R_0 \in \mathrm{PSF}_{\mathcal{K}x}(\mathcal{M}_0)$, and $\mathcal{K} : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator. Owing to the chain rule Theorem 5.1.7, if $\mathcal{K}$ satisfies the transversality condition (*i.e.* $\mathrm{ran}(K) + \mathcal{T}_{\mathcal{M}_0}(\mathcal{K}x) = \mathbb{R}^m, x \in \mathcal{M}_0$), then $R \overset{\text{def}}{=} R_0\mathcal{K} \in \mathrm{PSF}_x(\mathcal{M})$ where $\mathcal{M} = \left\{u \in \mathbb{R}^n : \mathcal{K}u \in \mathcal{M}_0\right\}$.

Popular examples complying with such formulation include the anisotropic total variation (TV) semi-norm in which case $R_0 = \|\cdot\|_1$ and $\mathcal{K} = D_{\mathrm{DIF}}$ is a finite difference approximation of the derivative [154]. For TV, R is then polyhedral, hence partly smooth at x relative to

$$\mathcal{M} = T_x \overset{\text{def}}{=} \left\{u \in \mathbb{R}^n : \mathrm{supp}(\mathcal{K}u) \subseteq \mathrm{supp}(\mathcal{K}x)\right\}.$$

Its Riemannian gradient reads $\mathcal{P}_{T_x}(\mathcal{K}^T \mathrm{sign}(\mathcal{K}x))$ and its Riemannian Hessian vanishes.

The proximity operator of the anisotropic TV, though not available in closed form, can be obtained efficiently by using either the taut string algorithm [71] or graph cuts [49].

Example 5.2.4 (ℓ_∞-norm). For $x \in \mathbb{R}^n$, the anti-sparsity promoting ℓ_∞-norm is defined as

$$R(x) = \|x\|_\infty \overset{\text{def}}{=} \max_{1 \leq i \leq n} |x_i|.$$

Define the saturation support $\mathcal{I}_x = \{i : |x_i| = \|x\|_\infty\}$. Clearly, R is a polyhedral norm, hence partly smooth at x relative to

$$\mathcal{M} = T_x \overset{\text{def}}{=} \{u \in \mathbb{R}^n : u_{\mathcal{I}_x} \in \mathbb{R}(s_{\mathcal{I}_x})\} \quad \text{where} \quad s_i \overset{\text{def}}{=} \begin{cases} \mathrm{sign}(x_i), & \text{if } i \in \mathcal{I}_x, \\ 0, & \text{otherwise,} \end{cases}$$

which is a subspace.

The Riemannian gradient of $\|\cdot\|_\infty$ at x reads

$$\nabla_{\mathcal{M}}(\|x\|_\infty) = \begin{cases} \frac{s(i)}{|\mathcal{I}_x|}, & i \in \mathcal{I}_x, \\ 0, & i \notin \mathcal{I}_x. \end{cases}$$

The Riemannian Hessian simply vanishes since ℓ_∞-norm is polyhedral.

Example 5.2.5 (Nuclear norm). Given $x \in \mathbb{R}^{m \times n}$ whose rank is r, *i.e.* $\mathrm{rank}(x) = r$. Let $x = U\mathrm{diag}(\sigma(x))V^*$ be a reduced rank-r singular value decomposition (SVD), where $U \in \mathbb{R}^{m \times r}$ and $V \in \mathbb{R}^{n \times r}$ have orthonormal columns, and $\sigma(x) \in (\mathbb{R}_+ \setminus \{0\})^r$ is the vector of singular values $(\sigma_1(x), \cdots, \sigma_r(x))$ in non-increasing order. Low-rank is the spectral extension of vector sparsity to matrix-valued data $x \in \mathbb{R}^{m \times n}$, *i.e.* imposing sparsity on the singular values of x. The nuclear norm is thus defined as

$$R(x) = \|x\|_* \overset{\text{def}}{=} \|\sigma_i(x)\|_1.$$

Putting together Theorem 5.1.11 and Example 5.2.1, it can be shown that the nuclear norm is partly smooth at x relative to the set of fixed-rank matrices

$$\mathcal{M} \overset{\text{def}}{=} \left\{z \in \mathbb{R}^{m \times n} : \mathrm{rank}(z) = r\right\},$$

which is a C^2-manifold around x of dimension $(m+n-r)r$, see [103, Example 8.14].

Moreover, we have

$$T_x = \{UA^* + BV^* : A \in \mathbb{R}^{n\times r}, B \in \mathbb{R}^{m\times r}\} \quad \text{and} \quad \nabla_{\mathcal{M}}\|x\|_* = UV^*. \tag{5.2.1}$$

From [168, Example 21], one can show that for $h \in T_x$,

$$\nabla^2_{\mathcal{M}}\|x\|_*(h) = \mathcal{P}_{T_x}\nabla^2\widetilde{\|x\|}_*(\mathcal{P}_{T_x}h),$$

where

$$\widetilde{\|z\|}_* = \widetilde{\|\sigma(z)\|}_1 = \textstyle\sum_{i=1}^r \sigma_i(z),$$

is a C^2-smooth (and even convex) representation of the nuclear norm on $\mathcal{M}$ near x, obtained owing to the smooth transfer principle [69, Corollary 2.3]. The expression of the (Euclidian) Hessian $\nabla^2\widetilde{\|z\|}_*$ can be obtained in several ways, see [168, Example 21] for details.

5.2.2 The non-convex case

Next we introduce two non-convex examples of partly smooth functions: the ℓ_0 pseudo-norm and the rank function.

Example 5.2.6 (ℓ_0 pseudo-norm). The ℓ_0 pseudo-norm, denoted as $\|x\|_0$ which returns the number of non-zero elements of $x \in \mathbb{R}^n$, is locally constant. Moreover, it is regular on $\mathbb{R}^n$ ([102, Remark 2]) and its subdifferential (see Eq. (2.2.1)) is given by (see [102, Theorem 1])

$$\partial\|x\|_0 = \operatorname{span}\big((e_i)_{i\in\operatorname{supp}(x)^c}\big),$$

where $(e_i)_{i=1,\dots,n}$ is the standard basis, and $\operatorname{supp}(x)$ is the support of x, *i.e.* indexes the non-zero entries of x. The proximity operator of ℓ_0-norm is given by the hard-thresholding,

$$\operatorname{prox}_{\gamma\|x\|_0}(z) = \begin{cases} z & \text{if } |z| > \sqrt{2\gamma}, \\ \operatorname{sign}(z)[0,z] & \text{if } |z| = \sqrt{2\gamma}, \\ 0 & \text{if } |z| < \sqrt{2\gamma}. \end{cases} \tag{5.2.2}$$

It can then be easily verified that the ℓ_0 pseudo-norm is partly smooth at any x relative to the subspace

$$\mathcal{M}_x = T_x = \{z \in \mathbb{R}^n : \operatorname{supp}(z) \subset \operatorname{supp}(x)\}.$$

Note also that the Riemannian gradient and Hessian along T_x of the ℓ_0 pseudo-norm vanish.

Example 5.2.7 (Rank function). The rank function is the spectral extension of ℓ_0 pseudo-norm to matrix-valued data $x \in \mathbb{R}^{m\times n}$ [111]. As the same as the nuclear norm, the rank function is partly smooth relative at x to the set of fixed rank matrices

$$\mathcal{M}_x = \{z \in \mathbb{R}^{m\times n} : \operatorname{rank}(z) = \operatorname{rank}(x)\}.$$

The tangent space of $\mathcal{M}_x$ is as given in (5.2.1). The rank function is also regular, prox-regular. Let $x = U\operatorname{diag}(\sigma(x))V^*$ be the SVD of x, then the subdifferential of $\|x\|_0$ reads

$$\partial\operatorname{rank}(x) = U\partial\big(\|\sigma(x)\|_0\big)V^* = U\operatorname{span}\big((e_i)_{i\in\operatorname{supp}(\sigma(x))^c}\big)V^*,$$

which is a vector space (see [102, Theorem 4 and Proposition 1]). The proximity operator of rank function is obtained by hard-thresholding (5.2.2) to the singular values. Observe that by definition of $\mathcal{M}_x$, the Riemannian gradient and Hessian of the rank function vanish along $\mathcal{M}_x$.

The above 7 examples of partly smooth functions are summarized in Table 5.1.

Table 5.1: Examples of partly smooth functions. For $x \in \mathbb{R}^n$ and some subset of indices $\mathcal{b} \subset \{1, \ldots, n\}$, $x_{\mathcal{b}}$ is the restriction of x to the entries indexed in b. For ℓ_∞-norm, $\mathcal{I}_x = \{i : |x_i| = \|x\|_\infty\}$. $\mathrm{D}_{\mathrm{DIF}}$ stands for the finite differences operator.

Function	Expression	Partial smooth manifold
ℓ_1-norm	$\sum_{i=1}^n \lvert x_i \rvert$	$\mathcal{M} = T_x = \{z \in \mathbb{R}^n : \mathrm{supp}(z) \subseteq \mathrm{supp}(x)\}$
$\ell_{1,2}$-norm	$\sum_{j=1}^m \lVert x_{b_i} \rVert$	$\mathcal{M} = T_x = \{z \in \mathbb{R}^n : \mathrm{supp}(z) \subseteq \mathrm{supp}(x)\}$
ℓ_∞-norm	$\max_{i=1,\ldots,n} \lvert x_i \rvert$	$\mathcal{M} = T_x = \{z \in \mathbb{R}^n : z_{\mathcal{I}_x} \in \mathbb{R}\mathrm{sign}(x_{\mathcal{I}_x})\}$
TV semi-norm	$\lVert \mathrm{D}_{\mathrm{DIF}} x \rVert_1$	$\mathcal{M} = T_x = \{z \in \mathbb{R}^n : \mathrm{supp}(\mathrm{D}_{\mathrm{DIF}} z) \subseteq \mathrm{supp}(\mathrm{D}_{\mathrm{DIF}} x)\}$
Nuclear norm	$\sum_{i=1}^r \sigma(x)$	$\mathcal{M} = \{z \in \mathbb{R}^{m \times n} : \mathrm{rank}(z) = \mathrm{rank}(x)\}$
ℓ_0 pseudo-norm	$\lVert x \rVert_0$	$\mathcal{M} = T_x = \{z \in \mathbb{R}^n : \mathrm{supp}(z) \subseteq \mathrm{supp}(x)\}$
Rank function	$\mathrm{rank}(x)$	$\mathcal{M} = \{z \in \mathbb{R}^{m \times n} : \mathrm{rank}(z) = \mathrm{rank}(x)\}$

Chapter 6

Forward–Backward-type Methods under Partial Smoothness

Main contributions of this chapter

- Finite activity identification (Theorem 6.2.1) and local linear convergence (Theorem 6.3.6) of Forward–Backward-type splitting methods.
- Explain why FISTA locally oscillates and slower than FB (Section 6.4.1).
- New lower complexity bound of first-order methods for certain type of optimization problems (Conjecture 6.3.15).
- Finite activity identification and local linear convergence of multi-step inertial Forward–Backward splitting for non-convex optimization (Section 6.5).

Part of these results appeared in [114].

Contents

In this chapter, we present the local linear convergence analysis for the class of Forward–Backward splitting methods (*e.g.* FB, (multi-step) inertial schemes and sequence convergent FISTA). Our key assumption is that the non-smooth part of the optimization problem is partly smooth, which is reasonable as many popular non-smooth regularization functionals are partly smooth (examples in Section 5.2).

We propose a unified analysis framework, under which we show that the sequence generated by the class of Forward–Backward splitting methods: (i) correctly identifies the active manifolds in a finite number of iterations (finite activity identification), and (ii) then enters a local linear convergence regime, which we characterize precisely in terms of the structure of the underlying active manifolds.

For the 2-step inertial setting, we show that a new lower complexity bound of first-order methods can be established for certain type of functions (Conjecture 6.3.15).

Our local convergence analysis allows us to establish and explain why FISTA (with convergent sequences) locally oscillates and can be slower than FB in terms of the linear rate. Moreover, we provide solutions on how to avoid local oscillations.

Various acceleration techniques are also discussed. For instance, we show that when the finite identification happens, we can either optimize the choices of step-size and inertial parameters, or turn to high-order optimization methods (*e.g.* Riemannian conjugate gradient, Newton method). Moreover, for simple problems such as "polyhedral + quadratic" functions, we show finite termination.

We also build a connection between the metric sub-regularity (Definition 3.3.1) and partial smoothness for the FB method in Section 6.4.4.

Finally in Section 6.5, we generalize the above results to the non-convex case.

6.1 Introduction

6.1.1 Problem statement

Recall the composite "smooth + non-smooth" optimization problem studied in Chapter 4, which we adapt it to the case of n-dimensional Euclidean space $\mathbb{R}^n$,

$$\min_{x\in\mathbb{R}^n} \{\Phi(x) \stackrel{\text{def}}{=} R(x) + F(x)\}, \tag{$\mathcal{P}_{\text{FB}}$}$$

where

(**F.1**) $R \in \Gamma_0(\mathbb{R}^n)$ is proper convex and lower semi-continuous.

(**F.2**) $F \in C^{1,1}(\mathbb{R}^n)$ with gradient ∇F being $(1/\beta)$-Lipschitz continuous.

(**F.3**) $\operatorname{Argmin}(\Phi) \neq \emptyset$, *i.e.* the set of minimizers is non-empty.

We suppose that assumptions (**F.1**)-(**F.3**) hold throughout the chapter.

6.1.2 Forward–Backward-type splitting methods

Specializing the MUSTARD algorithm 7 in Chapter 4 to the case where the metric $\mathcal{V} = \mathrm{Id}$, we obtain the following multi-step inertial Forward–Backward splitting method (MiFB).

Algorithm 12: Multi-step inertial Forward–Backward splitting

Initial: $s \in \mathbb{N}_+$ and $\mathcal{S} \stackrel{\text{def}}{=} \{0, ..., s-1\}$. $x_0 \in \mathbb{R}^n$, $x_{-i} = x_0, i \in \mathcal{S}$. Choose $\bar{\epsilon}, \underline{\epsilon} > 0$ such that $\underline{\epsilon} \leq 2\beta - \bar{\epsilon}$.

repeat

Let $\{a_{i,k}\}_{i\in\mathcal{S}}, \{b_{i,k}\}_{i\in\mathcal{S}} \in]-1, 2]^s$, and $\gamma_k \in [\underline{\epsilon}, 2\beta - \bar{\epsilon}]$:

$$\left\lfloor \begin{aligned} y_{a,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}),\\ y_{b,k} &= x_k + \textstyle\sum_{i\in\mathcal{S}} b_{i,k}(x_{k-i} - x_{k-i-1}),\\ x_{k+1} &= \operatorname{prox}_{\gamma_k R}\big(y_{a,k} - \gamma_k \nabla F(y_{b,k})\big). \end{aligned} \right. \tag{6.1.1}$$

$k = k + 1$;

until *convergence*;

The global convergence results (Theorems 4.2.3, 4.2.5 and 4.2.9) of the MUSTARD algorithm can be adapted to Algorithm 12 straightforwardly, since the only difference is that metric $\mathcal{V}_k \equiv \mathrm{Id}$ for MiFB.

In this chapter, we use the terminology "Forward–Backward-type" (FB-type) splitting methods to represent any variant of Forward–Backward splitting method whose generated sequences are convergent. Those methods include

- The original Forward–Backward splitting method [119, 140].
- 1-step inertial Forward–Backward splitting methods [131, 120, 115].
- Sequence convergent FISTA methods [50, 13].
- The multi-step inertial Forward–Backward splitting method Algorithm 12.

For the rest of the chapter, we use the term "1-MiFB" to denote the MiFB method with inertial step $s = 1$ and "2-MiFB" for MiFB method with $s = 2$.

6.2 Finite activity identification

Throughout the rest of the chapter, $x^\star$ denotes a global minimizer to which the sequence $\{x_k\}_{k\in\mathbb{N}}$ generated by the FB-type method converges, and $\mathcal{M}_{x^\star}$ is the partial smoothness manifold of R at $x^\star$.

Theorem 6.2.1 (Finite activity identification). *Let* (F.1)-(F.3) *hold, and suppose that the FB-type method is used to create a sequence* $\{x_k\}_{k\in\mathbb{N}}$ *that converges to* $x^\star \in \operatorname{Argmin}(\Phi)$ *such that* $R \in \operatorname{PSF}_{x^\star}(\mathcal{M}_{x^\star})$*, and moreover the non-degeneracy condition*

$$-\nabla F(x^\star) \in \operatorname{ri}\big(\partial R(x^\star)\big), \tag{$\mathrm{ND}_{\mathrm{FB}}$}$$

holds. Then, there exists a large enough $K > 0$ *such that for all* $k \geq K$, $x_k \in \mathcal{M}_{x^\star}$.

If moreover,

(i) $\mathcal{M}_{x^\star} = x^\star + T_{x^\star}$ *is an affine subspace, then* $y_{a,k}, y_{b,k} \in \mathcal{M}_{x^\star}$, $\forall k > K + s$.

(ii) R *is locally polyhedral around* $x^\star$*, then* $y_{a,k}, y_{b,k} \in \mathcal{M}_{x^\star} = x^\star + T_{x^\star}$ *for all* $k > K+s$, $\nabla_{\mathcal{M}_{x^\star}} R(x_k) = \nabla_{\mathcal{M}_{x^\star}} R(x^\star)$ *and* $\nabla^2_{\mathcal{M}_{x^\star}} R(x_k) = 0$ *for all* $k \geq K$.

Remark 6.2.2.

(i) If F locally is C^2 around $x^\star$, then owing to Corollary 5.1.10, we have $\Phi \in \operatorname{PSF}_{x^\star}(\mathcal{M}_{x^\star})$.

(ii) Recall that FB-type class of methods we consider contains the original FB, the proposed MiFB, and the sequence convergent FISTA. The MiFB is convergent under the assumptions of Theorem 4.2.3, Theorem 4.2.5 or Theorem 4.2.9. The FISTA is sequence convergent for $a_k = b_k = \frac{k-1}{k+d}$, $d > 2$, and $\gamma_k \equiv \gamma \in]0, \beta]$, see [50, 13]. Thus, the finite identification property holds true for all these instances.

(iii) The non-degeneracy condition ($\mathrm{ND}_{\mathrm{FB}}$) can be viewed as a geometric generalization of the strict complementarity of non-linear programming. Building on the arguments of [90], it is almost a necessary condition for the finite identification of $\mathcal{M}_{x^\star}$. Relaxing this assumption is a challenging problem in general.

See Section 6.7 from page 127 for the proof.

6.2.1 Bounds on the finite identification iteration

In Theorem 6.2.1, we have not provided an estimate $K \in \mathbb{N}_+$ beyond which finite identification occurs. Knowing K has practical interest, for instance, if one wants to switch to higher order acceleration (see Section 6.4.3). It is then legitimate to wonder whether such an estimate of K can be given. In the following, we shall give a bound in some important cases. For the sake of simplicity, we state the result for the case of FB. A similar reasoning can be generalized to the case of any sequence convergent FB-type method.

Proposition 6.2.3. *Suppose that the assumptions of Theorem 6.2.1 hold. Then the following holds.*

(i) *If the iterates are such that* $\partial R(x_k) \subset \operatorname{rbd}(\partial R(x^\star))$ *whenever* $x_k \notin \mathcal{M}_{x^\star}$*, then* $x_k \in \mathcal{M}_{x^\star}$ *for all* $k \geq \dfrac{\|x_0 - x^\star\|^2}{\underline{\epsilon}^2 \operatorname{dist}\big(-\nabla F(x^\star), \operatorname{rbd}(\partial R(x^\star))\big)^2}$.

(ii) *If* R *is separable, i.e.* $R(x) = \sum_{i=1}^m \sigma_{C_i}(x_{\mathcal{b}_i})$*, where* $\forall 1 \leq i \leq m, \mathcal{b}_i \subset \{1, \ldots, n\}$, $\bigcup_{i=1}^m \mathcal{b}_i = \{1, \ldots, n\}$*, and* $\mathcal{b}_i \cap \mathcal{b}_j = \emptyset$, $\forall i \neq j$*, and* $\dim(C_i) = |\mathcal{b}_i|$*, then identification of* $\mathcal{M}_{x^\star}$ *occurs for some* k *larger than* $\dfrac{\|x_0 - x^\star\|^2}{\underline{\epsilon}^2 \sum_{i \in I^c_{x^\star}} \operatorname{dist}(-\nabla F(x^\star_{\mathcal{b}_i}), \operatorname{rbd}(C_i))^2}$*, where* $I_x \overset{\text{def}}{=} \{i : x_{\mathcal{b}_i} \neq 0\}$.

See page 128 for the proof. Note that, as intuitively expected, this bound increases as the non-degeneracy condition ($\mathrm{ND}_{\mathrm{FB}}$) becomes more stringent. However, as it depends on $x^\star$, it is mostly of theoretical interest. The case of the ℓ_1-norm considered in [87] is recovered in the second situation of Proposition 6.2.3 with $C_i \equiv [-\lambda, \lambda]$ for some $\lambda > 0$.

6.2.2 Identifying activity in practice

The bounds on K given in Proposition 6.2.3 give theoretical guides, but are not sufficient to be implemented in practice to detect when activity has been identified. As one may guess, establishing a practically implementable upper-bound on the number of iterations needed for finite identification is a non-trivial problem, even for special cases. Nevertheless, in our implementations, we developed some heuristics to detect the identification that work well for the examples we tested. One possible approach is the dramatic decrease of energy in the transition from sub-linear convergence regime to the linear one. Another strategy uses the fact that for many partly smooth functions R, it holds that $x_k \in \mathcal{M}_{x^\star}$ is equivalent to $\mathcal{M}_{x_k} = \mathcal{M}_{x^\star}$ for x_k nearby $x^\star$. This is true for instance for all examples tested in our numerical experiments. In turn, finite identification of $\mathcal{M}_{x^\star}$ can be translated in terms of stabilization of $\mathcal{M}_{x_k}$ (up to numerical accuracy), which can be easily tested in practice and thus used as an indicator for finite identification. Typical examples are the support for the ℓ_1-norm, block support for the $\ell_{1,2}$-norm, rank for the nuclear norm, *etc.*. This is what we used throughout all numerical examples to be reported in this chapter as well as Chapters 7 and 8, and which worked very well.

6.2.3 Stability to errors

In this part, we investigate the stability of finite identification to errors. Recall the inexact updates of x_{k+1} in (4.2.4), specialized to optimization and let $\varepsilon_k \equiv 0$,

$$x_{k+1} = \mathrm{prox}_{\gamma_k R}\big(y_{a,k} - \gamma_k(\nabla F(y_{b,k}) + \xi_k)\big).$$

Suppose that the convergence of sequence $\{x_k\}_{k\in\mathbb{N}}$ still holds (*e.g.* $\{\xi_k\}_{k\in\mathbb{N}}$ is summable and $\{x_k\}_{k\in\mathbb{N}}$ is bounded). Then, since $\xi_k \to 0$, it can be seen from the proof of Theorem 6.2.1 that the activity identification property holds true for the above inexact iteration.

However, one cannot afford in general having non-zero errors ε_k in the implicit step as in (4.2.4), even if $\{\varepsilon_k\}_{k\in\mathbb{N}}$ summable. The reason behind this is that in the exact case, under condition ($\mathrm{ND}_{\mathrm{FB}}$), the proximal mappings of R and $R + \iota_{\mathcal{M}_{x^\star}}$ locally agree nearby $x^\star$. This property is clearly violated if approximate proximal mappings are involved. Here is a simple example.

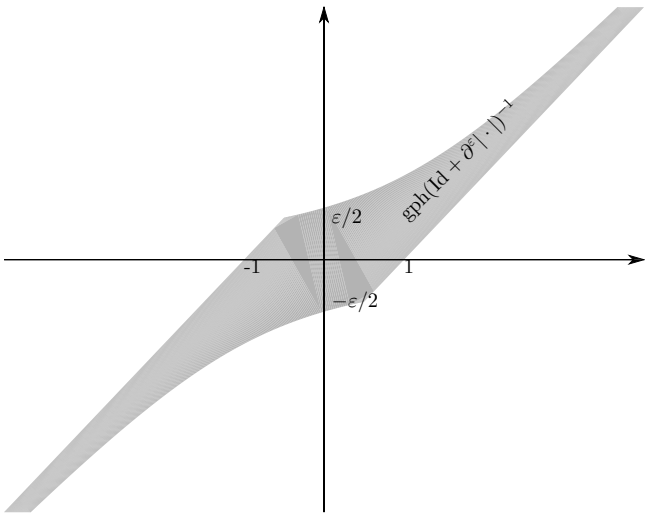


Figure 6.1: Graph of $(\mathrm{Id} + \partial^\varepsilon|\cdot|)^{-1}$.

Example 6.2.4. Let $F : x \in \mathbb{R} \mapsto \frac{1}{2}|\delta - x|^2$, with $\delta \in]-1, 1[$, and $R : x \in \mathbb{R} \mapsto |x|$. It is easy to see that $\Phi \in \Gamma_0(\mathbb{R})$, and it has a unique minimizer $x^\star = \mathrm{prox}_{|\cdot|}(\delta) = \max(1 - 1/|\delta|, 0)\delta = 0$. Moreover, Φ is partly smooth at $x^\star$ relative $\mathcal{M}_{x^\star} = \{0\}$, and $\delta - x^\star = \delta \in \mathrm{ri}\big(\partial R(x^\star)\big) =]-1, 1[$. Consider the

inexact version of the FB algorithm

$$x_{k+1} \in (\mathrm{Id} + \partial^{\varepsilon_k}|\cdot|)^{-1}(\delta), \tag{6.2.1}$$

where we set $\gamma_k \equiv 1$, since ∇F is 1-Lipschitz. From [43, Example 5.2.5], we have

$$\partial^{\varepsilon}|\cdot|(x) = \begin{cases} [1-\varepsilon/x, 1] & \text{if } x > \varepsilon/2 \\ [-1, 1] & \text{if } |x| \leq \varepsilon/2 \\ [-1, -1-\varepsilon/x] & \text{if } x < -\varepsilon/2, \end{cases}$$

whence the graph of $(\mathrm{Id} + \partial^{\varepsilon}|\cdot|)^{-1}$ can be deduced as displayed in Figure 6.1. Thus, depending on ε_k and the choice made in the inclusion (6.2.1), x_k may never vanish for any finite k, *i.e.* $x_k \notin \mathcal{M}_{x^\star}$ for any finite k.

6.3 Local linear convergence

In this section we present the local linear convergence result for FB-type methods. All the proofs in this section are collected in Section 6.7 from page 129. We denote as $T_{x^\star}$ the tangent space of $\mathcal{M}_{x^\star}$ at the minimizer $x^\star$.

Define the following matrices which are all *symmetric*,

$$H \overset{\text{def}}{=} \gamma \mathcal{P}_{T_{x^\star}} \nabla^2 F(x^\star) \mathcal{P}_{T_{x^\star}}, \quad G \overset{\text{def}}{=} \mathrm{Id} - H, \quad U \overset{\text{def}}{=} \gamma \nabla^2_{\mathcal{M}_{x^\star}} \Phi(x^\star) \mathcal{P}_{T_{x^\star}} - H, \tag{6.3.1}$$

where $\nabla^2_{\mathcal{M}_{x^\star}} \Phi$ is the Riemannian Hessian of Φ on the manifold $\mathcal{M}_{x^\star}$ (see Definition 5.1.4).

To establish the local linear convergence analysis, we need the following 3 assumptions or conditions.

I - Restricted injectivity In addition to (**F.2**), in the rest of the chapter, we also assume that F is locally C^2-smooth around $x^\star$, and its Hessian fulfills the following *restricted injectivity* condition,

$$\ker\big(\nabla^2 F(x^\star)\big) \cap T_{x^\star} = \{0\}. \tag{RI}$$

Condition (RI) implies that there exist $\kappa > 0$ such that $\forall h \in T_{x^\star}$,

$$\langle h, \nabla^2 F(x^\star) h \rangle \geq \kappa \|h\|^2. \tag{6.3.2}$$

In turn, it allows to ensure that $\rho(G) < 1$ holds for $\gamma \in [\underline{\epsilon}, 2\beta - \bar{\eta}]$, see Lemma 6.3.4.

It turns out that under conditions ($\mathrm{ND}_{\mathrm{FB}}$)-(RI), we are able show that problem ($\mathcal{P}_{\mathrm{FB}}$) admits a unique minimizer, and local quadratic growth of Φ if R is moreover partly smooth.

Proposition 6.3.1 (Uniqueness of the minimizer). *Under assumptions* (**F.1**)-(**F.3**), *let* $x^\star \in \mathrm{Argmin}(\Phi)$ *be a global minimizer of* ($\mathcal{P}_{\mathrm{FB}}$) *such that* F *is locally* C^2 *around* $x^\star$. *If conditions* ($\mathrm{ND}_{\mathrm{FB}}$) *and* (RI) *are also fulfilled, then*

(i) $x^\star$ *is the unique minimizer of* ($\mathcal{P}_{\mathrm{FB}}$).

(ii) *If moreover* $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star})$, *then* Φ *has at least a quadratic growth near* $x^\star$.

See page 129 for the proof of the proposition.

Remark 6.3.2. In Proposition 6.3.1, partial smoothness of R at $x^\star$ is not needed for the uniqueness claim (i). However, it brings more structure, hence the local quadratic growth property in (ii). Observe that (ii) is equivalent to metric sub-regularity of the subdifferential $\partial\Phi$ at $x^\star$ for 0 [6, Theorem 3.5].

II - Positive semi-definiteness of U Owing to Lemma 6.3.3, we have the following result which guarantees the positive semi-definiteness of the matrix U in (6.3.1).

Lemma 6.3.3. *For problem* ($\mathcal{P}_{\mathrm{FB}}$), *let conditions* (**F.1**)-(**F.3**) *hold and* $x^\star \in \mathrm{Argmin}(\Phi)$ *such that* $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star})$ *and* F *is locally* C^2 *around* $x^\star$. *Then* U *is symmetric positive semi-definite under either of the following circumstances:*

(i) ($\mathrm{ND}_{\mathrm{FB}}$) *holds.*

(ii) $\mathcal{M}_{x^\star}$ *is an affine subspace.*

In turn, $\mathrm{Id}+U$ *is invertible, and* $W \overset{\text{def}}{=} (\mathrm{Id}+U)^{-1}$ *is symmetric positive definite with all eigenvalues in* $]0,1]$.

See page 129 for the proof. The following simple lemma gathers important properties of the matrices appearing in (6.3.1).

Lemma 6.3.4. *For the matrices in* (6.3.1) *and* W,

(i) *Under* (**F.2**) *and* (RI),

(a) H *is symmetric positive definite with eigenvalues in* $[\gamma\kappa, \frac{\gamma}{\beta}]$.

(b) *For* $\gamma \in [\underline{\epsilon}, 2\beta - \bar{\epsilon}]$, G *has eigenvalues in* $[-1 + \frac{\bar{\epsilon}}{\beta}, 1 - \kappa\underline{\epsilon}[\subset]-1, 1[$.

(c) *For* $\gamma \in [\underline{\epsilon}, \beta]$, G *is also symmetric positive semi-definite with eigenvalues in* $[0, 1-\kappa\underline{\epsilon}[\subset [0,1[$.

(ii) *If both the assumptions of Lemma 6.3.3 and* (i) *hold, then* WG *has real eigenvalues lying in* $]-1,1[$. *If moreover* $\gamma \in [\underline{\epsilon}, \beta]$, *then* WG *has eigenvalues lying in* $[0,1[$.

III - Convergent parameters The last condition we need is that the parameters of the FB-type methods are convergent. In terms of the most general multi-step inertial FB method, we need the following condition for Algorithm 12,

$$a_{i,k} \to a_i,\ b_{i,k} \to b_i,\ \forall i \in \mathcal{S} \quad \text{and} \quad \gamma_k \to \gamma \in [\underline{\epsilon}, 2\beta - \bar{\epsilon}]. \tag{6.3.3}$$

Remark 6.3.5. For the sequence convergent FISTA [50], the limit of the inertial parameter $\{a_k\}_{k\in\mathbb{N}}$ is 1.

6.3.1 Local linear convergence

We are now in position to present the local linear convergence of FB-type methods. Define the following notations:

$$\begin{aligned} &M_0 \overset{\text{def}}{=} (a_0 - b_0)W + (1+b_0)WG,\ M_s \overset{\text{def}}{=} -(a_{s-1} - b_{s-1})W - b_{s-1}WG, \\ &M_i \overset{\text{def}}{=} -\big((a_{i-1} - a_i) - (b_{i-1} - b_i)\big)W - (b_{i-1} - b_i)WG,\ i = 1, ..., s-1, \\ &M_{\mathrm{FB}} \overset{\text{def}}{=} \begin{bmatrix} M_0 & \cdots & M_{s-1} & M_s \\ \mathrm{Id} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \mathrm{Id} & 0 \end{bmatrix},\ d_k \overset{\text{def}}{=} \begin{pmatrix} x_k - x^\star \\ \vdots \\ x_{k-s} - x^\star \end{pmatrix}. \end{aligned} \tag{6.3.4}$$

Theorem 6.3.6 (Local linear convergence). *Let* (**F.1**)-(**F.3**) *hold, and assume that an FB-type method is used to create a sequence* $\{x_k\}_{k\in\mathbb{N}}$ *that converges to* $x^\star \in \mathrm{Argmin}(\Phi)$ *such that* $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star})$, F *is locally* C^2 *around* $x^\star$ *and* ($\mathrm{ND}_{\mathrm{FB}}$) *holds. If moreover conditions* (RI) *and* (6.3.3) *are satisfied. Then for all* k *large enough,*

$$d_{k+1} = M_{\mathrm{FB}} d_k + o(\|d_k\|). \tag{6.3.5}$$

If $\rho(M_{\mathrm{FB}}) < 1$, *then given any* $\rho \in]\rho(M_{\mathrm{FB}}), 1[$, *there exists* $K \in \mathbb{N}$ *such that* $\forall k \geq K$,

$$\|x_k - x^\star\| = O(\rho^{k-K}). \tag{6.3.6}$$

The proof of is presented in Section 6.7 from page 129.

Remark 6.3.7.

(i) For all the FB-type methods, except the MiFB method with $s \geq 2$, sufficient conditions ensuring that $\rho(M_{\mathrm{FB}}) < 1$ are discussed in Section 6.3.2.

(ii) When R is locally polyhedral around $x^\star$, the parameters $a_{i,k} \equiv a_i, b_{i,k} \equiv b_i$ and $\gamma_k \equiv \gamma$ are taken constants, then the small o-term in (6.3.5) vanishes.

(iii) For the FB method (*i.e.* $a_{i,k} = b_{i,k} \equiv 0$), (6.3.4) can be further simplified, and the corresponding linearized iteration can be stated in terms of $x_k - x^\star$ directly, which reads

$$x_{k+1} - x^\star = WG(x_k - x^\star) + o(\|x_k - x^\star\|).$$

(iv) The obtained results can be readily extended to the variable metric FB splitting method [62], see Chapter 8 for the study in terms of Primal–Dual splitting methods.

The restricted injectivity condition (RI) plays an important role in our local convergence rate analysis and in general cannot be relaxed. However, for some special cases, such as when R is locally polyhedral, it can be removed, at the price of less sharp rate estimation. This is formalized in the following statement.

Theorem 6.3.8. *Suppose that* (**F.1**)-(**F.3**) *hold, and an FB-type method creates a sequence* $x_k \to x^\star \in \operatorname{Argmin}(\Phi)$ *such that* R *is locally polyhedral around* $x^\star$, F *is* C^2 *near* $x^\star$, *and condition* ($\mathrm{ND}_{\mathrm{FB}}$) *holds. If moreover there exists* $\epsilon > 0$ *and a subspace* V *such that*

$$\ker\big(\mathcal{P}_{T_x} \nabla^2 F(x) \mathcal{P}_{T_x}\big) = V, \quad \forall x \in \mathbb{B}_\epsilon(x^\star) \cap (x^\star + T_{x^\star}),$$

then $\{x_k\}_{k\in\mathbb{N}}$ *converges locally linearly to* $x^\star$.

See page 134 for the proof.

6.3.2 Spectral properties of M_{FB}

In order for (6.3.6) of Theorem 6.3.6 to hold, we need the spectral radius $\rho(M_{\mathrm{FB}})$ to be strictly less than 1. Hence, in the following, we discuss sufficient conditions ensuring that $\rho(M_{\mathrm{FB}}) < 1$. Given the structure of M_{FB}, this is a challenging linear algebra problem, and can only be answered in some special cases: a_i and b_i possibly different but the the function R is locally polyhedral, or R is a general partly smooth function but $a_i = b_i$, moreover for both cases we need to set the inertial steps $s = 1$.

We start with the case that R is locally polyhedral around $x^\star$. When R is locally polyhedral around $x^\star$, U of (6.3.1) vanishes and matrix W defined in Lemma 6.3.3 becomes Id. Then from (6.3.4) we have

$$\begin{aligned} M_0 &\stackrel{\text{def}}{=} (a_0 - b_0)\mathrm{Id} + (1 + b_0)G, \;\; M_s \stackrel{\text{def}}{=} -(a_{s-1} - b_{s-1})\mathrm{Id} - b_{s-1}G, \\ M_i &\stackrel{\text{def}}{=} -\big((a_{i-1} - a_i) - (b_{i-1} - b_i)\big)\mathrm{Id} - (b_{i-1} - b_i)G, \;\; i = 1, ..., s-1. \end{aligned} \tag{6.3.7}$$

Denote η and ρ the eigenvalue of G and M_{FB} respectively. We have the following correspondence between η and ρ.

Theorem 6.3.9. *For* $i \in \{0, ..., s-1\}$, *let* $v_i \in \mathbb{R}^n$. *Suppose that the concatenated column vector* $(v_0; \cdots; v_s)$ *is an eigenvector of* M_{FB} *corresponding to an eigenvalue* ρ. *Then*

(i) *For each* $i \in \{0, ..., s\}$, *there holds*

$$v_i = \rho^{s-i} v_s.$$

(ii) v_s *is an eigenvector of* M *corresponding to an eigenvalue* η, *where* η *and* ρ *satisfy the relation*

$$\begin{aligned} 0 = \rho^{s+1} &- \big((a_0 - b_0) + (1 + b_0)\eta\big)\rho^s \\ &+ \textstyle\sum_{i=1}^{s-1} \big(\big((a_{i-1} - a_i) - (b_{i-1} - b_i)\big) + (b_{i-1} - b_i)\eta\big)\rho^{s-i-1} \\ &+ \big((a_{s-1} - b_{s-1}) + b_{s-1}\eta\big). \end{aligned} \tag{6.3.8}$$

See page 135 for the proof.

Remark 6.3.10. If we choose $b_i = a_i, i \in \{0, ..., s-1\}$, then relation (6.3.8) becomes

$$0 = \rho^{s+1} - (1 + a_0)\eta\rho^s + \sum_{i=1}^{s-1}\big((a_{i-1} - a_i)\eta\big)\rho^{s-i-1} + a_{s-1}\eta. \tag{6.3.9}$$

To express ρ in terms of η, one needs to get analytically the roots of a polynomial of degree $s+1$. When $s = 1$, (6.3.8) is a quadratic function with 3 free parameters η, a_0, b_0, and we can express ρ explicitly as a function of these parameters. However, things become much more challenging when $s \geq 2$. For example, when $s = 2$, studying the properties of ρ involves a 9-order polynomial equation which can be only solved numerically.

6.3.2.1 Inertial memory $s = 1$

To lighten the notation, for $s = 1$, we denote a_0, b_0 as a, b respectively. Then M_{FB} becomes

$$M_{\text{FB}} = \begin{bmatrix} (a-b)\text{Id} + (1+b)G, & -(a-b)\text{Id} - bG \\ \text{Id}, & 0 \end{bmatrix}. \tag{6.3.10}$$

Let $\bar{\eta}, \underline{\eta}$ be the biggest and smallest (signed) eigenvalues of G, *i.e.* $-1 < \underline{\eta} < \bar{\eta} < 1$. From Lemma 6.3.4 and (6.3.2), we have

$$\bar{\eta} = 1 - \kappa\gamma \quad \text{and} \quad \underline{\eta} = 1 - \gamma/\beta. \tag{6.3.11}$$

For the sake of simplicity, we suppose that $\bar{\eta}$ is moreover strictly positive.

Theorem 6.3.11 (Locally polyhedral case). *Let $s = 1$, then (6.3.8) becomes*

$$0 = \rho^2 - \big((a-b) + (1+b)\eta\big)\rho + \big((a-b) + b\eta\big). \tag{6.3.12}$$

Let $a, b \in [0, 1]$, then for M_{FB} in (6.3.10),

(i) *$\rho(M_{\text{FB}}) < 1$ as long as*

$$\frac{2(b-a) - 1}{1 + 2b} < \underline{\eta}.$$

(ii) *Choose γ such that*

$$1 - \gamma/\beta \geq -(3 - \sqrt{\kappa\gamma})(1 - \sqrt{\kappa\gamma}). \tag{6.3.13}$$

Then the value of $\rho(M_{\text{FB}})$ minimizes if we choose a, b such that

$$\begin{aligned} b &\in \left[0, \frac{-(1-\sqrt{\kappa\gamma})^2((1-\sqrt{\kappa\gamma})^2 - 1 - \gamma/\beta) - (1-\sqrt{\kappa\gamma})|(1-\sqrt{\kappa\gamma})^2 - 1 + \gamma/\beta|}{(1-(1-\sqrt{\kappa\gamma})^2)\gamma(1/\beta - \kappa)}\right] \\ a &= (1-\bar{\eta})b + \big(1 - \sqrt{1-\bar{\eta}}\big)^2 = (\kappa\gamma)b + (1-\sqrt{\kappa\gamma})^2, \end{aligned} \tag{6.3.14}$$

and the optimal value of $\rho(M_{\text{FB}})$, denoted by $\rho^\star_{s=1}$, reads

$$\rho^\star_{s=1} = 1 - \sqrt{1-\bar{\eta}} = 1 - \sqrt{\kappa\gamma}. \tag{6.3.15}$$

(iii) *Given $\gamma \in]0, 2\beta[$, the minimal value spectral radius of G, denoted by $\underline{\rho}^\star_G$, is*

$$\underline{\rho}^\star_G = \frac{1-\kappa\beta}{1+\kappa\beta} \quad \text{if} \quad \gamma^\star = \frac{2\beta}{1+\kappa\beta}.$$

For $\rho(M_{\text{FB}})$, denote $\underline{\rho}^\star_{s=1}$ the minimal value $\rho(M_{\text{FB}})$ can reach. Then if we choose γ, a, b as

$$\gamma^\star = \frac{4\beta}{(1+\sqrt{\kappa\beta})^2}, \quad b^\star = 0 \quad \text{and} \quad a^\star = \left(\frac{1-\sqrt{\kappa\beta}}{1+\sqrt{\kappa\beta}}\right)^2, \tag{6.3.16}$$

the value of $\underline{\rho}^\star_{s=1}$ is

$$\underline{\rho}^\star_{s=1} = \frac{1-\sqrt{\kappa\beta}}{1+\sqrt{\kappa\beta}}, \tag{6.3.17}$$

The proof of the theorem starts from page 136.

Remark 6.3.12.

(i) The dependency of $\rho(M_{\text{FB}})$ on $\underline{\eta}, \bar{\eta}$ is rather complicate, Section 6.7.3 from page 142 is dedicated to this dependency.

(ii) The discriminant of the quadratic equation (6.3.12) reads,

$$\Delta_\rho = \big((a-b)+(1+b)\eta\big)^2 - 4\big((a-b)+b\eta\big).$$

We have $\Delta_\rho = 0$ if a, b are chosen according to (6.3.14).

(iii) For the optimal rate $\rho^\star_{s=1}$, if we choose $\gamma = \beta$, then $\rho^\star_{s=1} = 1 - \sqrt{\kappa\beta}$ which agrees with the rate in [133, Theorem 2.2.2] provided by Nesterov.

(iv) Claim (iii) of Theorem 6.3.11 coincides with the optimal convergence rate of the "heavy ball method" [147, Section 3.2.1]. Note that the rate (6.3.17) also matches the lower complexity bounds established by Nesterov for first-order methods to solve the class of problems ($\mathcal{P}_{\text{FB}}$) if F is also κ-strongly convex [133, Theorem 2.1.13].

R is a general partly smooth function For the case of R being an arbitrary partly smooth function, U of (6.3.1) is non-trivial, and the spectral analysis of M_{FB} in (6.3.4) becomes a generalized eigenvalue problem which is much more complex. When choosing $b = a$, from (6.3.4) we have

$$M_{\text{FB}} = \begin{bmatrix} (1+a)WG, & -aWG \\ \text{Id}, & 0 \end{bmatrix}. \tag{6.3.18}$$

Compare with (6.3.10) with $b = a$, the main difference is that we have WG instead of G. Fortunately, the eigenvalues of WG are all real thanks to Lemma 6.3.4.

Denote η and ρ the eigenvalue of WG and M_{FB} respectively, and $\underline{\eta}, \bar{\eta}$ the smallest and largest (signed) eigenvalues of WG, we have the following corollary from Theorem 6.3.11.

Corollary 6.3.13 (General partly smooth case). *For the matrix M_{FB} in (6.3.18), we have*

(i) *$\rho(M_{\text{FB}}) < 1$ as long as*

$$\frac{-1}{1+2a} < \underline{\eta}. \tag{6.3.19}$$

(ii) *If $\underline{\eta}, \bar{\eta}$ are such that $\underline{\eta} \geq -\bar{\eta}/3$, moreover choose a as*

$$a = \frac{(1-\sqrt{1-\bar{\eta}})^2}{\bar{\eta}},$$

then the smallest value of $\rho(M_{\text{FB}})$ is $\rho^\star_{s=1} = 1 - \sqrt{1-\bar{\eta}}$.

See page 140 for the proof.

Remark 6.3.14.

(i) For the case of R being a general partly smooth function, condition $b = a$ does not mean that we should set $b_k = a_k \to a, \forall k \in \mathbb{N}$ along the iterations.

(ii) Condition (6.3.19) holds naturally if we choose $\gamma \in]0, \beta]$.

6.3.2.2 Inertial memory $s = 2$

For the case $s = 2$, (6.3.8) becomes a cubic equation of ρ with 5 free parameters a_0, a_1, b_0, b_1 and η (or γ), which is rather difficult to analyze. Therefore, we choose the inertial parameters $b_i = a_i, i = 0, 1$. Unfortunately, even for this choice, we still can not express everything analytically and only obtain the following conjecture which we prove numerically.

Conjecture 6.3.15 (Spectral radius of M_{FB} for $s = 2$). *Let $s = 2$ and $b_i = a_i, i = 0, 1$, then (6.3.8) becomes*

$$0 = \rho^3 - (1 + a_0)\eta\rho^2 + (a_0 - a_1)\eta\rho + a_1\eta. \tag{6.3.20}$$

Let $a_0, a_1 \in]-1, 2[$, then

(i) *given $\eta \in]-1, 1[$, the optimal choice (i.e. such that the value of $|\rho|$ minimizes) of a_0, a_1 yields*

$$|\rho| = |1 - \sqrt[3]{1-\eta}| \leq |1 - \sqrt{1-\eta}|.$$

(ii) *Denote $\rho^\star_{s=2}$ the optimal value of $\rho(M_{\mathrm{FB}})$ for $s = 2$. Let $\gamma = \beta$, suppose that the eigenvalues of H distribute uniformly in $[\underline{\eta}, \bar{\eta}]$, then if κ is such that $1 - \gamma\kappa < 0.988$, then*

$$\rho^\star_{s=2} \leq \rho^\star_{s=1}.$$

(iii) *Denote $\kappa' > \kappa > 0$ the second smallest eigenvalue of $\mathcal{P}_{T_{x^\star}} \nabla^2 F(x^\star) \mathcal{P}_{T_{x^\star}}$. If $1-\beta\kappa' < 0.494(1-\beta\kappa)$, then*

$$\rho^\star_{s=2} = 1 - \sqrt[3]{1-\bar{\eta}} = 1 - \sqrt[3]{\beta\kappa}. \tag{6.3.21}$$

A numerical proof of the conjecture is given at page 140.

Remark 6.3.16. Given the step-size $\gamma = \beta$, if the condition of (iii) of Conjecture 6.3.15 is satisfied, then for $s \in \{0, 1, 2\}$ the smallest spectral radius of M_{FB} are

$$\rho^\star_{s=0} = 1 - \beta\kappa, \quad \rho^\star_{s=1} = 1 - \sqrt{\beta\kappa} \quad \text{and} \quad \rho^\star_{s=2} = 1 - \sqrt[3]{\beta\kappa},$$

which can be unified as $\rho^\star_s = 1 - \sqrt[s+1]{1-\bar{\eta}}$. One may wonder whether one can generalize this for instance to $s = 3$? Unfortunately, it is most likely not possible. Indeed for $s = 3$, even with $b_i = a_i, i \in \{0, 1, 2\}$, we still have 3 inertial parameters to deal with, moreover, the polynomial (6.3.8) is then of order 4. All these make the spectral analysis for 3-MiFB method almost impossible.

It should be noted that Conjecture 6.3.15 can be generalized to the class of κ-strongly convex and $(1/\beta)$-Lipschitz continuous functions. For such a class of functions, the lower complexity bounds [133, Theorem 2.1.13] for the first-order methods is $\underline{\rho}^\star_{s=1}$ as given in (6.3.17). In the following, we compare the values of $\underline{\rho}^\star_{s=1}$ and $\rho^\star_{s=2}$.

A numerical example, as shown in Figure 6.2 (a), is designed to compare $\underline{\rho}^\star_{s=1}$ and $\rho^\star_{s=2}$, where we consider a strongly convex least square estimation with

$$\kappa = 10^{-3}, \quad \kappa' = 0.64 \quad \text{and} \quad \beta = 1.$$

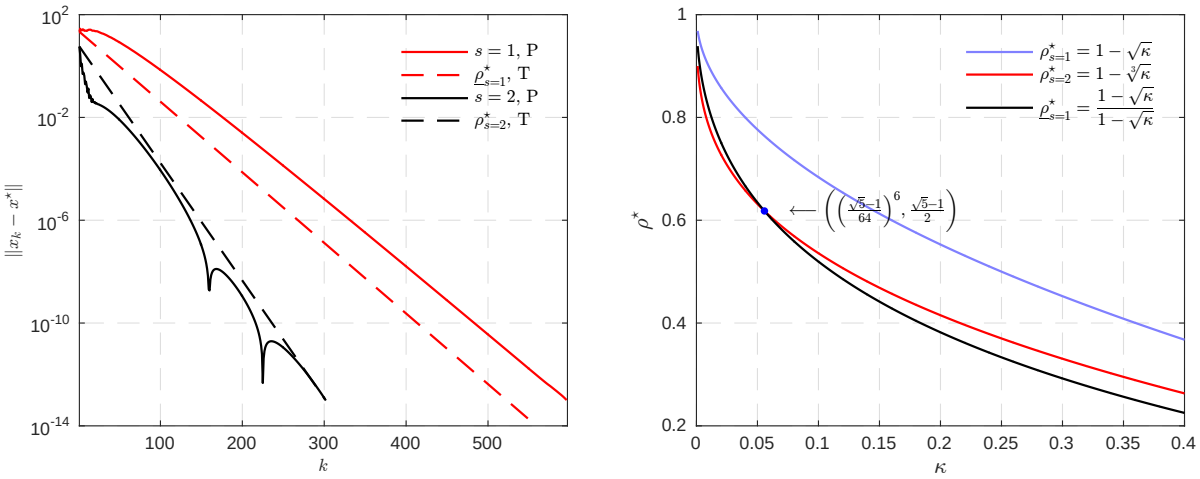
The inertial gradient descent with $s = 1$ and $s = 2$ is applied:

(i) For $s = 1$, *i.e.* heavy ball method, the parameters are chosen according to (6.3.16) and the convergence rate is $\underline{\rho}^\star_{s=1} = 0.9387$. The red solid line in Figure 6.2 (a) denotes the practical observation and the red dashed line stands for theoretical bound $\underline{\rho}^\star_{s=1}$.

(ii) For $s = 2$, we choose $\gamma = \beta$ and the inertial parameters are computed via Conjecture 6.3.15, the obtained convergence rate is $\rho^\star_{s=2} = 0.9 < \underline{\rho}^\star_{s=1}$. The black solid line in Figure 6.2 (a) denotes the practical observation, and the black dashed line stands for theoretical bound $\rho^\star_{s=2}$.

Such observation implies that, for certain class of optimization problems, $\rho^\star_{s=2}$ can serve as the new lower complexity bound of the first-order methods.

It should be noted that this new bound does not contradict Nesterov's result [133, Theorem 2.1.13], since the class of problems that $\rho^\star_{s=2}$ can be applied to is rather limited. Fix $\beta = 1$ and $\gamma = \beta$ for 2-MiFB, then the values of $\rho^\star_{s=2}$ and $\underline{\rho}^\star_{s=1}$ are simply controlled by κ (assume that κ' satisfies $1 - \beta\kappa' < 0.494(1 - \beta\kappa)$). A comparison of $\rho^\star_{s=2}$ and $\underline{\rho}^\star_{s=1}$ is demonstrated in Figure 6.2 (b) for $\kappa \in]0, 0.4]$. It can be observed that $\rho^\star_{s=2} \leq \underline{\rho}^\star_{s=1}$ for $\kappa \in]0, (\frac{\sqrt{5}-1}{2})^6]$.



(a) An experiment on least square estimation

(b) Comparison between $\rho^\star_{s=2}$ and $\underline{\rho}^\star_{s=1}$

Figure 6.2: (a) A numerical example illustrating the conjectured new lower bound. Red lines stand for practical observation and theoretical bound of the heavy ball method, while black lines stand for practical observation and theoretical bound of 2-step inertial gradient descent. (b) Comparison of optimal convergence rate obtained by inertial step $s = 1, 2$ and the lower bound of first-order methods. The parameters are chosen as $\beta = 1$ and $\kappa \in]0, 0.4]$.

6.4 Discussions

In this part, we present some discussions on the obtained local linear convergence result, and mainly focus on the difference between FB and FISTA/1-MiFB.

6.4.1 FB is locally faster than FISTA

For the sake of brevity (the same conclusions hold true in the general case), we consider the situation $b_k = a_k \equiv a \in [0, 1]$ and $\gamma_k \equiv \gamma \in]0, \beta]$ is fixed, in which case $\bar{\eta} \geq \underline{\eta} \geq 0$ (see Lemma 6.3.4(ii)), and thus condition $-1/(1 + 2a) < \underline{\eta}$ is in force.

Moreover, note that $\bar{\eta}$ is the local convergence rate of the FB method, and $\rho(M_{\text{FB}})$ depends solely on $\bar{\eta}$ and the value of a, since γ is fixed. Recall that $\rho(M_{\text{FB}})$ is the best local linear convergence rate can be obtained by FB-type methods under given parameters

Figure 6.3 displays the value of $\rho(M_{\text{FB}})$ as a function of a for fixed $\bar{\eta}$ (*i.e.* fixed γ). We obtain the following observations:

(i) When $a \in [0, \bar{\eta}]$, we have $\rho(M_{\text{FB}}) \leq \bar{\eta}$. This entails that if 1-MiFB is used with such a choice of inertial parameter, it will converges locally lineally faster than FB. For $a \in [\bar{\eta}, 1]$, the situation reverses as $\rho(M_{\text{FB}}) \geq \bar{\eta}$, and 1-MiFB becomes slower than FB.

(ii) In particular, as $a = 1$ for FISTA, we have $\rho(M_{\text{FB}}) = \sqrt{\bar{\eta}} > \bar{\eta}$. In plain words, though FISTA is known to be globally faster (in terms of the objective) than FB, attaining the optimal $O(1/k^2)$ rate, locally, the situation radically changes as FISTA will always ends up being locally slower than FB. A similar observation is made in [165] for the special case of a variant of FISTA used to solve the LASSO problem. This explains in particular why many authors [86, 137] resort to restarting to accelerate local convergence of FISTA, which consists in resetting periodically the scheme to $a = 0$ which is more favourable to FISTA. Our predictions in Figure 6.3 gives clues on when to restart (*i.e.* detect the point in red on the rate curve). We will elaborate more on this in the numerical simulations in Section 6.6.

(iii) $\rho(M_{\text{FB}})$ attains its minimal value at $a = \frac{(1-\sqrt{1-\bar{\eta}})^2}{\bar{\eta}}$, and this is the best convergence rate that can be achieved locally for FB-type methods (see Eq. (6.3.14) for $b = a$).

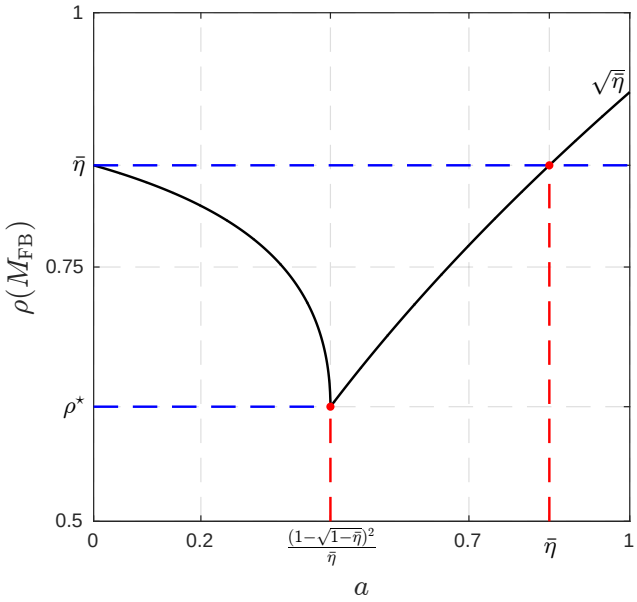


Figure 6.3: Let $b = a$, and assume $\underline{\eta}, \bar{\eta}$ are known and also close enough such that the spectral radius $\rho(M_{\text{FB}})$ is only affected by $\bar{\eta}$, then $\rho(M_{\text{FB}})$ is a function of a.

Remark 6.4.1. For the 2-MiFB, taking $b_{i,k} = a_{i,k} \equiv a_i, i = 0, 1$ for example. If both a_0 and a_1 are positive, then 2-MiFB locally is moreover slower than FB if $a_0 + a_1 \in [\sqrt{\bar{\eta}}, 1[$.

6.4.2 Oscillation of the FISTA method

A typical feature of the FISTA method is that it is not monotone and locally oscillates [23], which makes the local convergence even slower, see Figure 6.4 or 6.6 for example. In fact, the 1-MiFB scheme shares this property as well when the inertial parameters are large.

Such oscillatory behaviour is due to the fact that, for those inertial parameters, the eigenvalue $\rho_{\max}$ such that $|\rho_{\max}| = \rho(M_{\text{FB}})$ is complex. It can then be seen that the oscillation period of $\|x_k - x^\star\|$ is exactly $\frac{\pi}{\theta}$, where θ is the argument of $\rho_{\max}$. For the parameter settings used in Figure 6.3, *i.e.* $b = a$ and $\gamma \in]0, \beta]$, we have

$$\begin{cases} a \in [0, \frac{(1-\sqrt{1-\bar{\eta}})^2}{\bar{\eta}}] : \rho_{\max} \text{ is real}, \\ a \in]\frac{(1-\sqrt{1-\bar{\eta}})^2}{\bar{\eta}}, 1] : \rho_{\max} \text{ is complex}, \end{cases}$$

so that as long as $a > \frac{(1-\sqrt{1-\bar{\eta}})^2}{\bar{\eta}}$, the 1-MiFB method locally oscillates. See Figure 6.6 for an example.

Remark 6.4.2. For 2-MiFB, oscillation is also inevitable if $\sum_i a_i$ is too big. In fact, it is more easier to make 2-MiFB oscillate than 1-MiFB, as we have seen from the numerical experiments in Section 4.5 that $x_{k-1} - x_{k-2}$ has higher momentum than $x_k - x_{k-1}$.

6.4.3 Acceleration

The finite time activity identification property (Theorem 6.2.1) implies that the globally convex but non-smooth problem eventually becomes locally C^2-smooth, but possibly non-convex, constrained on the activity manifold. This opens the door to acceleration, and even finite termination, exploiting the structure of the objective and of the identified manifold. There are several ways to achieve this goal as we explain hereafter.

Optimal first-order method The idea of optimal first-order method is to keep using an inertial schemes, while locally jumping to the optimal inertial parameters and/or step-size provided by Theorem 6.3.11 and Conjecture 6.3.15.

In the mean time, it is worth mentioning that, for all the discussion we have made above, we assume that the local Lipschitz constant of ∇F along $\mathcal{M}_{x^\star}$ is as the same as the global one, *i.e.* $1/\beta$. However, as the dimension of the manifold $\mathcal{M}_{x^\star}$ can be much smaller than $\mathbb{R}^n$, it is possible that the Lipschitz constant of ∇F along $\mathcal{M}_{x^\star}$ could be much smaller than $1/\beta$. Denote $1/\beta_{\mathcal{M}_{x^\star}}$ the Lipschitz continuity of ∇F along $\mathcal{M}_{x^\star}$, then we have $\beta \leq \beta_{\mathcal{M}_{x^\star}}$. In turn, the upper bound of step-size γ locally is enlarged to $2\beta_{\mathcal{M}_{x^\star}}$.

Finite convergence in the polyhedral case Finite termination can be obtained if R is locally polyhedral around $x^\star$, and F is quadratic, *i.e.* problem (3.6.2) with R locally polyhedral around $x^\star$. In this situation, under hypothesis ($\mathrm{ND}_{\mathrm{FB}}$), we have finite identification of $x^\star + T_{x^\star}$. In addition, (RI) is equivalent to injectivity of the linear operator $\mathcal{K}$ on $T_{x^\star}$. Altogether, this allows to show that $x^\star$ can be written explicitly as

$$x^\star = \mathcal{K}^{*,+}_{T_{x_K}} f - \lambda\big(\mathcal{K}^*_{T_{x_K}}\mathcal{K}_{T_{x_K}}\big)^+ \mathcal{P}_{T_{x_K}}\big(\partial R(x_K)\big) - \mathcal{K}\mathcal{P}_{x_K+T_{x_K}}(0),$$

for K sufficiently large.

High-order acceleration: Newton method Once the activity manifold has been identified, one can switch to Newton-type methods for locally minimizing Φ. This can be done either using local parameterizations obtained from $\mathcal{U}$-Lagrangian theory or from Riemannian geometry [106, 126, 156]. One can also use the Riemannian version of the non-linear conjugate gradient method [156]. For these schemes, one can also show respectively quadratic and superlinear convergence since $\nabla^2_{\mathcal{M}_{x^\star}}\Phi(x^\star)$ is positive definite by (ii) of Proposition 6.3.1.

6.4.4 Partial smoothness vs metric sub-regularity

In this section, we establish a connection between partial smoothness and metric sub-regularity for the FB algorithm when R is locally polyhedral around $x^\star$ and F is quadratic (*i.e.* $F = \frac{1}{2}\|\mathcal{K}x - f\|^2$). In turn, this will allow to compare the local linear rate estimate obtained in Section 3.3 and the one of this chapter.

Let the step-size $\gamma_k \equiv \gamma \in]0, 2\beta[$ be a constant, and $a_k = b_k \equiv 0$, *i.e.* apply FB with constant step-size to solve ($\mathcal{P}_{\mathrm{FB}}$). We denote $M'_{\mathrm{FB}} \stackrel{\mathrm{def}}{=} \mathrm{Id} - M_{\mathrm{FB}}$.

Proposition 6.4.3. *Let* (**F.1**)-(**F.3**) *hold, and assume that the FB method is used to create a sequence* $\{x_k\}_{k\in\mathbb{N}}$ *that converges to* $x^\star \in \mathrm{Argmin}(\Phi)$ *such that* R *is locally polyhedral around* $x^\star$, F *is quadratic and conditions* ($\mathrm{ND}_{\mathrm{FB}}$) *and* (RI) *are satisfied. Then* M'_{FB} *is* $\frac{1}{\gamma\kappa}$*-metrically sub-regular at* $x^\star$ *for* 0*, where* κ *is defined in* (6.3.2).

Proof. Owing to Remark 6.3.7(ii), under the stated assumptions, the linearized iteration reads

$$x_{k+1} - x^\star = M_{\mathrm{FB}}(x_k - x^\star),$$

where $M_{\mathrm{FB}} = \mathrm{Id} - H = \mathrm{Id} - \gamma\mathcal{P}_{x^\star}\mathcal{K}^*\mathcal{K}\mathcal{P}_{x^\star}$. M_{FB} is linear and non-expansive. Moreover, in view of Lemma 6.3.4(ii), $M'_{\mathrm{FB}} = \gamma\mathcal{P}_{x^\star}\mathcal{K}^*\mathcal{K}\mathcal{P}_{x^\star}$ has eigenvalues in $[\gamma\kappa, \frac{\gamma}{\beta}]$. It then follows from [22, Theorem 2.5] that M'_{FB} is $\frac{1}{\gamma\kappa}$-linearly regular, hence metrically sub-regular. □

In terms of rate estimations, the one provided by metric sub-regularity is less sharp than that of partial smoothness. For simplicity, let $\gamma \in]0, \beta]$. Then the rate estimation obtained by partial smoothness is $\rho = 1 - \gamma\kappa$ (Lemma 6.3.4), whereas the rate estimate obtained through metric sub-regularity (see (3.3.3)) is

$$\sqrt{1 - (\gamma\kappa)^2} \geq \rho.$$

6.5 Extension to the non-convex case

For the local convergence analysis of Section 6.3, it can be observed that, unlike the global convergence analysis, convexity plays a less important role. For instance, it was only needed to show that the matrix U defined in (6.3.1) is positive semi-definite. This suggests that the local linear convergence claims in Section 6.3 can be extended to the non-convex case, provided that the Riemannian Hessian of R is assumed positive semi-definite at $x^\star$. In the following, we present the extension of the above results to the non-convex case.

6.5.1 Finite activity identification

The finite activity identification of the ncvx-MiFB (Algorithm 11) is rather straightforward.

Corollary 6.5.1 (Finite activity identification). *Suppose that the ncvx-MiFB Algorithm 11 is run under the conditions of Theorem 4.4.3 such that $\{x_k\}_{k\in\mathbb{N}}$ converges to a critical point $x^\star \in \mathrm{crit}(\Phi)$. Assume that $R \in \mathrm{PSF}_{x^\star,-\nabla F(x^\star)}(\mathcal{M}_{x^\star})$ and the non-degeneracy condition ($\mathrm{ND}_{\mathrm{FB}}$) holds. Then, $x_k \in \mathcal{M}_{x^\star}$ for all k large enough.*

If moreover, $\mathcal{M}_{x^\star} = x^\star + T_{x^\star}$ is an affine subspace, then $y_{a,k}, y_{b,k} \in \mathcal{M}_{x^\star}$, $\forall k > K + s$.

Proof. Under the conditions of Theorem 4.4.3, there exists a critical point $x^\star \in \mathrm{crit}(\Phi)$ such that $x_k \to x^\star$, $R(x_k) \to R(x^\star)$ and $\Phi(x_k) \to \Phi(x^\star)$ (see the proof of Lemma 4.6.8).

The convergence property of $\{x_k\}_{k\in\mathbb{N}}$ gives $\|y_{a,k} - x_k\| \to 0$ and $\|y_{b,k} - x^\star\| \to 0$. By (**A.8**) and the proof of Theorem 6.2.1, we get

$$\mathrm{dist}\big(-\nabla F(x^\star), \partial R(x_{k+1})\big) \leq \frac{1}{\underline{\gamma}}\|y_{a,k} - x_{k+1}\| + \frac{1}{\beta}\|y_{b,k} - x^\star\| \to 0.$$

Altogether, this shows that the conditions of Theorem 5.1.5 are fulfilled on $\langle \nabla F(x^\star), \cdot\rangle + R(\cdot)$, and the identification result follows. The identification of $y_{a,k}, y_{b,k}$ is immediate by definition of affine/linear subspace. □

6.5.2 Local linear convergence

Similarly to the result in Section 6.3, we need the following 3 assumptions to be hold.

I' - Restricted injectivity The restricted injectivity (RI) needs to be fulfilled.

II' - Positive semi-definiteness of the Riemannian Hessian Given $\gamma \in]0, \beta[$ and $x^\star \in \mathrm{crit}(\Phi)$, let $\mathcal{M}_{x^\star}$ be a C^2-smooth submanifold and $R \in \mathrm{PSF}_{x^\star,-\nabla F(x^\star)}(\mathcal{M}_{x^\star})$. In addition to (**A.8**), now assume F is locally C^2-smooth around $x^\star$. Denote $T_{x^\star} \overset{\mathrm{def}}{=} \mathcal{T}_{\mathcal{M}_{x^\star}}(x^\star)$ and the following matrices as in (6.3.1),

$$H \overset{\mathrm{def}}{=} \gamma \mathcal{P}_{T_{x^\star}} \nabla^2 F(x^\star) \mathcal{P}_{T_{x^\star}}, \quad U \overset{\mathrm{def}}{=} \gamma \nabla^2_{\mathcal{M}_{x^\star}} \Phi(x^\star) \mathcal{P}_{T_{x^\star}} - H.$$

Since Φ is non-convex, in general there is no result similar to Lemma 6.3.3 which guarantees the positive semi-definiteness of U. Therefore, in the following, we assume that U is positive semi-definite, *i.e.* $\forall h \in T_{x^\star}$,

$$\langle h, Uh\rangle \geq 0. \tag{6.5.1}$$

III' - Convergent parameters Similarly to (6.3.3), we need thee parameters of ncvx-MiFB (Algorithm 11) to be convergent, *i.e.*

$$a_{i,k} \to a_i,\ b_{i,k} \to b_i,\ \forall i \in \mathcal{S} \quad \text{and} \quad \gamma_k \to \gamma \in [\underline{\gamma}, \min\{\bar{\gamma}, \bar{r}\}], \tag{6.5.2}$$

where $\bar{r} < r$, and r is the prox-regularity modulus of R (see Definition 2.2.1). The main difference here is that, γ_k is not only smaller than β but also the prox-regularity modulus of R.

Remark 6.5.2.

(i) Condition (6.5.1) can be satisfied by various non-convex functions, such as polyhedral functions which include the ℓ_0 pseudo-norm (Example 5.2.6) and the rank function (Example 5.2.7).

(ii) It can be shown that conditions (6.5.1) and (RI) together imply that $x^\star$ is a local minimizer of Φ in ($\mathcal{P}_{\text{ncvx}}$), and Φ grows at least quadratically near $x^\star$. The arguments to prove this are essentially adapted from those used to show Proposition 6.3.1.

Corollary 6.5.3 (Local linear convergence). *Suppose that the ncvx-MiFB Algorithm 11 is run under the conditions of Theorem 6.5.1. Moreover, assume that F is locally C^2-smooth around $x^\star$ and conditions (RI), (6.5.1) and (6.5.2) hold. Then for all k large enough,*

$$d_{k+1} = M_{\text{FB}} d_k + o(\|d_k\|), \tag{6.5.3}$$

where M_{FB} is defined in (6.3.4). If $\rho(M_{\text{FB}}) < 1$, then given any $\rho \in]\rho(M_{\text{FB}}), 1[$, there exists $K \in \mathbb{N}$ such that $\forall k \geq K$,

$$\|x_k - x^\star\| = O(\rho^{k-K}).$$

Proof. The main difference of proving this corollary from that of Theorem 6.3.6 is that $\text{prox}_{\gamma_k R}$ is no more firmly non-expansive. However, from prox-regularity of R at $x^\star$ for $-\nabla F(x^\star)$, invoking [153, Proposition 13.37], we have that there exists $\bar{r} > 0$ such that for all $\gamma_k \in]0, \min(\bar{\gamma}, \bar{r})[$, there exists a neighbourhood U of $x^\star - \gamma_k \nabla F(x^\star)$ on which $\text{prox}_{\gamma_k R}$ is single-valued and l-Lipschitz continuous with $l = \bar{r}/(\bar{r} - \gamma_k)$. Since ∇F is continuous and $x_k \to x^\star$, we have $y_{a,k} - \gamma_k \nabla F(y_{b,k}) \to x^\star - \gamma_k \nabla F(x^\star)$. In turn, $y_{a,k} - \gamma_k \nabla F(y_{b,k}) \in U$ for all k sufficiently large. Applying this property to the proof of Theorem 6.3.6 proves the corollary. □

Remark 6.5.4. The spectral analysis of M_{FB} is the same as that of Section 6.3.2, hence locally we can optimize both the step-size and inertial parameters owing to Theorem 6.3.11.

6.6 Numerical experiments

In this section, we present various numerical experiments to support our result. Continuing from the numerical experiments in Section 4.5, consider the problem

$$\min_{x \in \mathbb{R}^n} R(x) + \frac{1}{2}\|\mathcal{K}x - f\|^2. \tag{6.6.1}$$

Now, 6 examples are considered. The first 5 are linear inverse problems, where $\mathcal{K}$ is generated from the standard random Gaussian ensemble, and for each example of R:

ℓ_1**-norm**	$(m, n) = (48, 128)$, $\|x_{\text{ob}}\|_0 = 8$, *i.e.* x_{ob} has 8 non-zero elements.
$\ell_{1,2}$**-norm**	$(m, n) = (48, 128)$, x_{ob} has 3 non-zero blocks of size 4.
ℓ_∞**-norm**	$(m, n) = (63, 64)$, $\lvert\mathcal{I}_{x_{\text{ob}}}\rvert = 8$ where $\mathcal{I}_{x_{\text{ob}}} = \{i : \lvert x_i\rvert = \|x_{\text{ob}}\|_\infty\}$.
Nuclear norm	$(m, n) = (640, 1024)$, $x_{\text{ob}} \in \mathbb{R}^{32\times 32}$ and $\text{rank}(x_{\text{ob}}) = 4$.
Total Variation	$(m, n) = (48, 128)$, $\|D_{\text{DIF}} x_{\text{ob}}\|_0 = 8$.

The last example is the anisotropic TV deconvolution of $2D$ image as we have seen in Section 3.6, but without the box intensity constraint.

Besides verifying our proposed result with the above 6 examples, we will present the following comparisons through (6.6.1) with R being ℓ_1-norm and nuclear norm:

(i) Local optimal rate of MiFB with $s = 1, 2$.

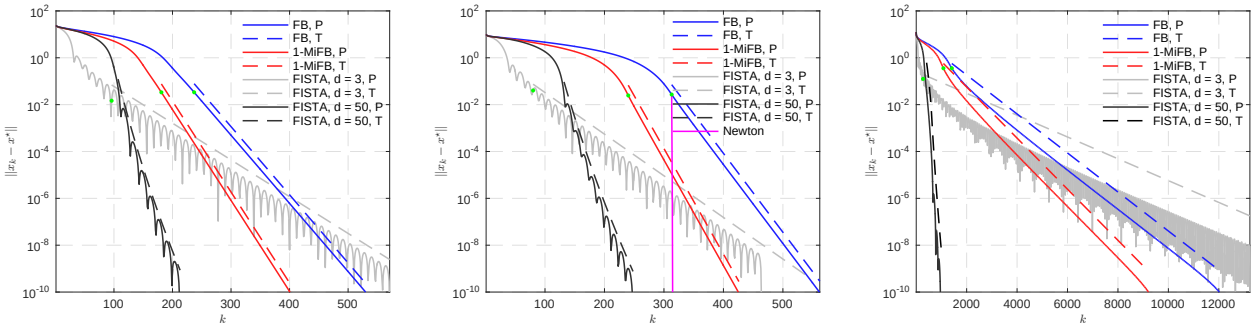
(ii) Oscillation behaviour of MiFB/FISTA, and the difference between the inertial parameters a_k and b_k for 1-MiFB.

We shall skip the comparisons of different choices of step-size γ since it is already presented in Section 4.5.1.

6.6.1 Local linear convergence of FB-type methods

We will verify our local linear convergence result Theorem 6.3.6 for FB, 1-MiFB and the sequence convergent FISTA [50]. The step-size is fixed as $\gamma_k \equiv \beta$ for all 3 methods. For FISTA, two choices of d are considered: $d = 3, 3$ for (4.1.4). For 1-MiFB we choose $b_k = a_k \equiv \sqrt{5} - 2 - 10^{-3}$ such that Theorem 4.2.9 applies and the generated sequence is guaranteed to be convergent.

The convergence profiles of $\|x_k - x^\star\|$ are shown in Figure 6.4. As demonstrated by all the plots, identification and local linear convergence occurs after finite time. The solid lines (denoted as "P") represent the observed profiles, while dashed ones (denoted as "T") stand for the theoretically predicted ones. The positions of the green points (or the starting points of the dashed lines) stand for the iteration at which $\mathcal{M}_{x^\star}$ has been identified.



(a) ℓ_1-norm (b) $\ell_{1,2}$-norm (c) ℓ_∞-norm

(d) TV semi-norm (e) Nuclear norm (f) Anisotropic TV deconvolution

Figure 6.4: Local linear convergence and comparison of the FB-type methods (FB, 1-MiFB and FISTA) in terms of $\|x_k - x^\star\|$. We fix $\gamma_k \equiv \beta$ for all the methods, moreover, for the 1-MiFB, we let $b_k = a_k \equiv \sqrt{5} - 2 - 10^{-3}$, and for the FISTA, $d = 3, 50$ are considered. For each figure, "P" stands for practical observed profiles, while "T" indicates theoretical predictions. The green points indicate the iteration at which $\mathcal{M}_{x^\star}$ has been identified.

Tightness of predicted rates For the ℓ_1, ℓ_∞-norms and TV semi-norm, our predicted rates (Theorem 6.3.11) coincide exactly with the observed ones (same slopes for the dashed and solid lines). This is due to the fact that they are all polyhedral and F is quadratic. Note that for FISTA, which is non-monotone, the prediction coincides with the envelope of the oscillations. For the $\ell_{1,2}$-norm, though it is not polyhedral, our predicted rates are still are very tight, due to the fact that the Riemannian Hessian is taken into account. Then for the nuclear norm, whose active manifold is not anymore a subspace, our estimation becomes slightly less sharp compared to the other examples, though barely visible on the plots. For both the $\ell_{1,2}$-norm and nuclear norm, as the Riemannian Hessian is taken into account, the predicted rates are rather sharp.

For the image deconvolution problem, assumptions (ND_{FB}) and (RI) are checked a posteriori (verified

for this experiment). This together with the fact that the anisotropic TV is polyhedral justifies that the predicted rate is again exact.

Comparison of the methods From the numerical results, we can draw the following remarks:

(i) Overall, FISTA with $d = 50$ (black line) is the fastest while $d = 3$ (gray line) is the slowest. FB and 1-MiFB are sandwiched between them with 1-MiFB being faster.

(ii) For the finite activity identification, however, FISTA $d = 3$ in general shows the fastest identification (see the starting points of the dashed lines), and FB is the slowest.

(iii) Locally, similar to the global convergence, FISTA $d = 50$ has the fastest rate and $d = 3$ is the slowest. Again, FB and 1-MiFB are between them with 1-MiFB being faster.

It can be concluded from the above remarks that, in practice, FISTA [50] with $d = 3$ is not a wise choice if high accuracy solutions are needed. Indeed, under this choice, a_k converges to 1 too fast, and this hampers its local behaviour as the discussions we anticipated in Section 6.4 (see Figure 6.3). In fact, such behaviour of a_k can be avoided by choosing relatively bigger q, and this is exactly what the difference between $d = 3$ and $d = 50$ implies. In our tests, $q \in [50, 100]$ seems to a good trade-off, even bigger d is not recommended since it may lead to a much slower activity identification. A similar observation is also mentioned in [50], where the authors only tried $d = 2, 3, 4$. It should be noted that the original FISTA method [24] has almost the same behaviour as the case $d = 3$.

It should be pointed out that the local rate of FISTA $d = 50$ being faster than FB does not contradict with our claim in Section 6.4 that FB is faster than FISTA locally. The reason is that we are limited by machine accuracy, and bigger value of q delays the speed at which a_k approaches to 1 which actually makes FISTA behaviour similar to the 1-MiFB.

High-order acceleration For the ℓ_1, ℓ_∞-norms and TV semi-norm, since they are polyhedral, finite termination can be obtained once the manifold is identified. For $\ell_{1,2}$-norm which is not polyhedral, we applied the Riemannian Newton method which converges quadratically, leading to a dramatic acceleration as can be seen in Figure 6.4(b). For the nuclear norm, a non-linear conjugate gradient method is applied, leading again to a much faster (super-linear) local convergence.

6.6.2 More comparisons

6.6.2.1 Local optimal rate

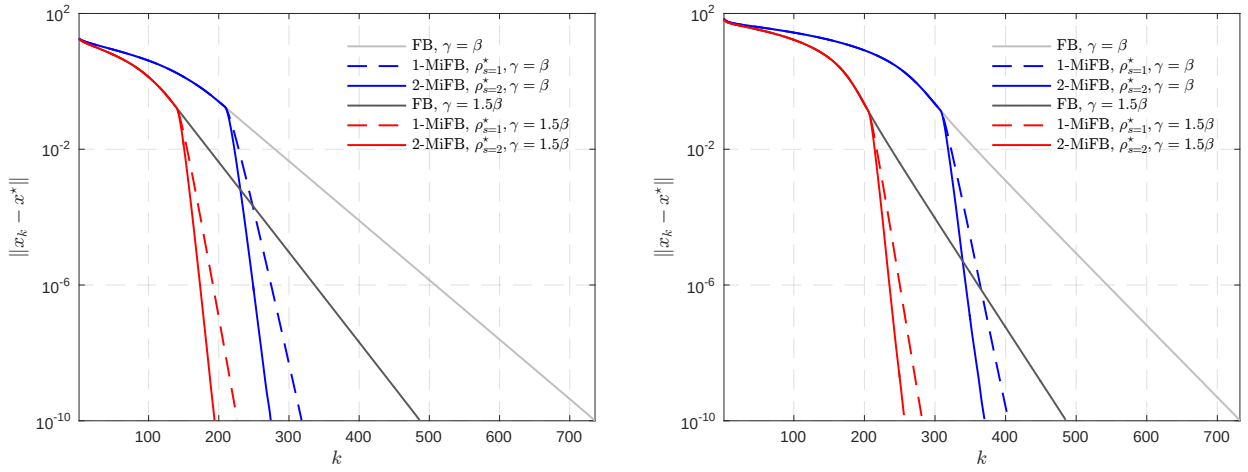
In this part, we provide the comparisons of local optimal rates of MiFB for $s = 1, 2$. Two choices of fixed γ are considered: $\gamma = \beta, 1.5\beta$. For the inertial parameters, we consider the case $b_{i,k} = a_{i,k}, i \in \{0, ..., s-1\}$.

For each choice of γ, we start with the FB scheme, whence the manifold is identified we compute the optimal inertial parameters $\{a_i\}_{i\in\mathcal{S}}$ based on Theorem 6.3.11 and Conjecture 6.3.15, and then continue the iteration with the obtained inertial parameters. The obtained results for ℓ_1-norm and nuclear norm are shown in Figure 6.5, it can be observed that with given γ, the optimal rate obtained by 2-MiFB is faster than 1-MiFB. Such an observation means that the optimal rate established by Nesterov for first-order non-smooth optimization methods indeed can be surpassed as presented in Conjecture 6.3.15.

Moreover, the corresponding optimal values of $a^\star$ for $s = 1$ and (a_0, a_1) for $s = 2$ are shown in Table 6.1 below, from which we draw the following remarks:

(i) For $s = 2$, the optimal choice of inertial parameters is "positive $a_0^\star$ and negative $a_1^\star$". Moreover the value of $a_0^\star$ can be bigger than 1.

(ii) Under given γ, $a^\star$ of $s = 1$ is smaller than $\sum_i a_i^\star$ of $s = 2$.



(a) ℓ_1-norm (b) Nuclear norm

Figure 6.5: Comparison of local optimal rate obtained by MiFB for $s = 1, 2$, two choices of $\gamma = \beta, 1.5\beta$ are considered.

(iii) The bigger the value of γ, the smaller the value of $a^\star$ or $\sum_i a_i^\star$.

Table 6.1: Optimal inertial parameter $a^\star$ for $s = 1$ and (a_0, a_1) for $s = 2$.

	$\gamma = \beta$			$\gamma = 1.5\beta$		
	$s = 1$	$s = 2$		$s = 1$	$s = 2$	
	$a^\star$	$(a_0^\star, a_1^\star)$	$\sum_i a_i$	$a^\star$	$(a_0^\star, a_1^\star)$	$\sum_i a_i$
ℓ_1-norm	0.6683	$(1.0597, -0.2986)$	0.7611	0.6083	$(0.9454, -0.2413)$	0.7041
Nuclear norm	0.6427	$(1.0099, -0.2729)$	0.7370	0.5798	$(0.8932, -0.2170)$	0.6762

6.6.2.2 Oscillations of FISTA/MiFB

As observed from Figure 6.4, FISTA schemes oscillate when a_k is large enough. In order to have a better visualisation of the oscillation of FISTA/MiFB, we choose the LASSO problem (*i.e.* $R = \|\cdot\|_1$ for (4.5.1)) for illustration, set $b_k = a_k = a$ and locally adjust the value of a so that the oscillation period is an integer. The result is shown in Figure 6.6, where the oscillation period of the tested example is 20.

6.6.2.3 Comparison of a_k vs b_k

Now let us assess the influence of inertial parameter choice for 1-MiFB. The step-size γ is fixed as $\gamma_k \equiv \beta$, inertial parameters $b_k \equiv b, a_k \equiv a$ and the online updated rule (4.2.6) is applied. In total, 4 different combinations of (a, b) are considered, which are

$$(0.3, 0.2 \text{ or } 0.6) \quad \text{and} \quad (0.8, 0.2 \text{ or } 0.6).$$

For both examples, if we let $b_k \equiv 0$, then the optimal local choice a_{opt} obtained through (6.3.14) is between 0.3 and 0.8. The obtained plots are depicted in Figure 6.7, from which we summarize the following observations:

(i) The time to activity identification is more dependent on the value of a. Clearly, relatively bigger values of a lead to a faster identification. On the other hand, when $a < a_{\text{opt}}$ (case $a = 0.3$),

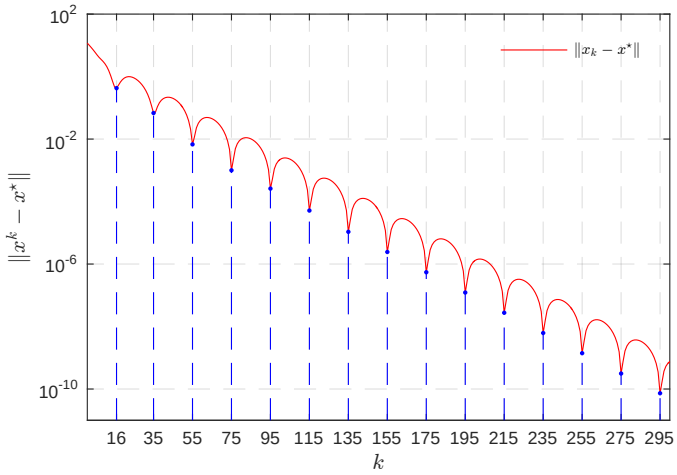
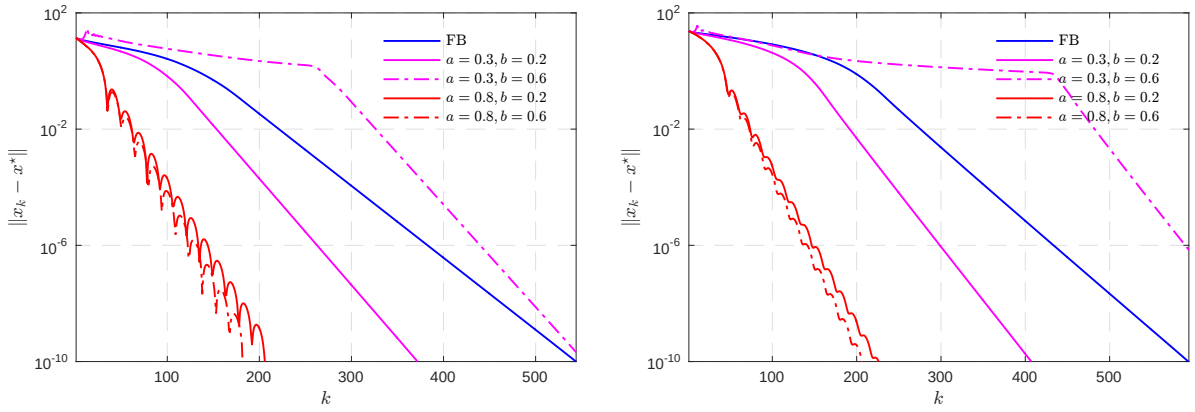


Figure 6.6: Local oscillation of the FISTA/MiFB methods on LASSO problem. Local oscillation of the iFB method, where the oscillation period is 20.

bigger values of b lead to slower identification, while the opposite situation occurs when $a > a_{\text{opt}}$ (case $a = 0.8$).

(ii) The convergence rate also depends more on the choice of a, since with fixed a, the rate difference caused by different values of b is small, see the blue dashed/solid lines, and magenta ones.



(a) ℓ_1-norm (b) Nuclear norm

Figure 6.7: Comparisons on the difference between the inertial parameters a_k and b_k, the step-size γ is fixed as β.

6.6.3 The non-convex case

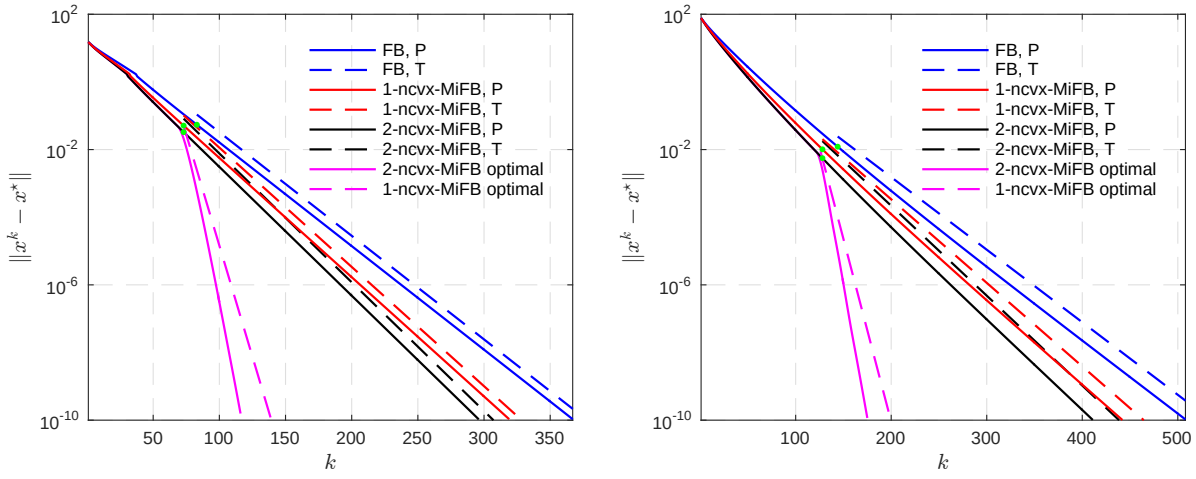
We dedicate the last numerical experiment to verify the result in Section 6.5 for non-convex optimization problems. The same problems and parameter settings as in Section 4.5.4 are considered, *i.e.* problem (6.6.1) with R being ℓ_0 pseudo-norm and the rank function.

Tightness of predicted rates The convergence profiles of $\|x_k - x^\star\|$ are shown in Figure 6.8. As it can be seen from all the plots, finite identification and local linear convergence indeed occur. The positions of the *green dots* indicate the iteration from which x_k identifies the submanifold $\mathcal{M}_{x^\star}$. The solid lines ("P") represents practical observations, while the dashed lines ("T") denotes theoretical predictions.

As the Riemannian Hessians of ℓ_0 and rank function both vanish in all examples, our predicted rates

coincide exactly with the observed ones.

Optimal first-order method To highlight the advantages of multi-steps, we also added an example for the locally optimal choices of inertial parmeters. The results are depicted in Figure 6.8, where the magenta solid line stands for the observation of 2-step inertia, while the magenta dashed line corresponds to the 1-step inertia. This again shows that 2-step inertia can outperform the 1-step one.



(a) ℓ_0 pseudo-norm (b) Rank function

Figure 6.8: Comparison of ncvx-MiFB under different inertial settings. We fix $\gamma_k \equiv 0.5\beta$. For the inertial schemes, the inertial parameters are chosen such that (4.4.5) holds.

6.7 Proofs of main theorems

6.7.1 Proofs of Section 6.2

Proof of Theorem 6.2.1. By assumption, the sequence $\{x_k\}_{k\in\mathbb{N}}$ created by the FB-type method converges to $x^\star \in \operatorname{Argmin}(\Phi)$, and the latter is non-empty by assumption (**F.3**). From the definition of proximity operator, the step of updating x_{k+1} in (6.1.1) is equivalent to

$$y_{a,k} - \gamma_k \nabla F(y_{b,k}) - x_{k+1} \in \gamma_k \partial R(x_{k+1}).$$

Since the sequences $\{x_k\}_{k\in\mathbb{N}}, \{y_{a,k}\}_{k\in\mathbb{N}}$ and $\{y_{b,k}\}_{k\in\mathbb{N}}$ are convergent, hence we have

$$\lim_{k\to+\infty} \|y_{a,k} - x_{k+1}\| = 0 \quad \text{and} \quad \lim_{k\to+\infty} \|y_{b,k} - x^\star\| = 0. \tag{6.7.1}$$

By (**F.2**), we get

$$\begin{aligned}
\operatorname{dist}\big(-\nabla F(x^\star), \partial R(x_{k+1})\big) &\le \frac{1}{\gamma_k}\|y_{a,k} - x_{k+1} - \gamma_k\big(\nabla F(y_{b,k}) - \nabla F(x^\star)\big)\| \\
&\le \frac{1}{\gamma_k}\big(\|y_{a,k} - x_{k+1}\| + \gamma_k\|\nabla F(y_{b,k}) - \nabla F(x^\star)\|\big) \\
&\le \frac{1}{\gamma_k}\|y_{a,k} - x_{k+1}\| + \frac{1}{\beta}\|y_{b,k} - x^\star\|.
\end{aligned}$$

Since $\liminf \gamma_k = \underline{\epsilon} > 0$ and the result of (6.7.1), we obtain $\operatorname{dist}(-\nabla F(x^\star), \partial R(x_{k+1})) \to 0$. Owing to assumption (**F.1**), R is sub-differentially continuous at every point in its domain, and in particular at $x^\star$ for $-\nabla F(x^\star)$, which in turn yields $R(x_k) \to R(x^\star)$. Altogether, this shows that the conditions of Theorem 5.1.5 are fulfilled on $\langle \nabla F(x^\star), \cdot\rangle + R(\cdot)$, and the identification result follows.

(i) When the active manifold $\mathcal{M}_{x^\star}$ is an affine subspace, then $\mathcal{M}_{x^\star} = x^\star + T_{x^\star}$ owing to the normal sharpness property (Lemma 5.1.3) and the claim follows immediately.

(ii) When R is locally polyhedral around $x^\star$, then $\mathcal{M}_{x^\star}$ is an affine subspace and the identification of $y_{a,k}, y_{b,k}$ follows from (i). For the rest, it is sufficient to observe that by polyhedrality, for any $x \in \mathcal{M}_{x^\star}$ near $x^\star$, $\partial R(x) = \partial R(x^\star)$. Therefore, combining Lemma 5.1.3 and Definition 5.1.4, we get the second conclusion. □

Proof of Proposition 6.2.3.

(i) By firm non-expansiveness of $\mathrm{prox}_{\gamma_{k-1}R}$, and non-expansiveness of $\mathrm{Id} - \gamma_{k-1}\nabla F$, we have

$$\begin{aligned}\|x_k - x^\star\|^2 &\leq \|(\mathrm{Id} - \gamma_{k-1}\nabla F)(x_{k-1}) - (\mathrm{Id} - \gamma_{k-1}\nabla F)(x^\star)\|^2 \\ &\quad - \|x_{k-1} - \gamma_{k-1}\nabla F(x_{k-1}) - x_k + \gamma_{k-1}\nabla F(x^\star)\|^2 \\ &\leq \|x_{k-1} - x^\star\|^2 - \underline{\epsilon}^2\|u_k - \nabla F(x^\star)\|^2,\end{aligned}$$

where we denoted $u_k \overset{\text{def}}{=} (x_{k-1} - x_k)/\gamma_{k-1} - \nabla F(x_{k-1})$. By definition, we have $u_k \in \partial R(x_k)$. Suppose that identification has not occurred at k, *i.e.* that $x_k \notin \mathcal{M}_{x^\star}$, and hence $u_k \in \partial R(x_k) \subset \mathrm{rbd}(\partial R(x^\star))$. Therefore, continuing the above inequality, we get

$$\begin{aligned}\|x_k - x^\star\|^2 &\leq \|x_{k-1} - x^\star\|^2 - \underline{\epsilon}^2\mathrm{dist}\big(-\nabla F(x^\star), \partial R(x_k)\big)^2 \\ &\leq \|x_{k-1} - x^\star\|^2 - \underline{\epsilon}^2\mathrm{dist}\big(-\nabla F(x^\star), \mathrm{rbd}(\partial R(x^\star))\big)^2 \\ &\leq \|x_0 - x^\star\|^2 - k\underline{\epsilon}^2\mathrm{dist}\big(-\nabla F(x^\star), \mathrm{rbd}(\partial R(x^\star))\big)^2,\end{aligned}$$

and $\mathrm{dist}(-\nabla F(x^\star), \mathrm{rbd}(\partial R(x^\star))) > 0$ owing to ($\mathrm{ND}_{\mathrm{FB}}$). Taking k as the largest integer such that the right hand is positive, we deduce that the number of iterations where identification has not occurred, does not exceed the given bound, whence our conclusion follows.

(ii) We have $\partial\sigma_{C_i}(x^\star_{\delta_i}) = C_i, \forall i \in I^c_{x^\star}$. In turn, by separability, R is partly smooth at $x^\star$ relative to $\mathcal{M}_{x^\star} = \bigtimes_{i=1}^m \mathcal{M}_{x^\star_{\delta_i}}$, where $\mathcal{M}_{x^\star_{\delta_i}} = 0$ if $i \in I^c_{x^\star}$ and $\mathcal{M}_{x^\star_{\delta_i}} \neq 0$ otherwise. Suppose that at iteration k, $I^c_{x^\star} \cap I_{x_k} \neq \emptyset$. Denote $h_{k-1} = x_{k-1} - \gamma_{k-1}\nabla F(x_{k-1})$, and $h^\star = x^\star - \gamma_{k-1}\nabla F(x^\star)$. Thus for any $i \in I^c_{x^\star} \cap I_{x_k}$, we have

$$\begin{aligned}x_{k,\delta_i} - x^\star_{\delta_i} &= h_{k-1,\delta_i} - \mathcal{P}_{\gamma_{k-1}C_i}(h_{k-1,\delta_i}) \\ &= (h_{k-1,\delta_i} - h^\star_{\delta_i}) - (\mathcal{P}_{\gamma_{k-1}C_i}(h_{k-1,\delta_i}) - \mathcal{P}_{\gamma_{k-1}C_i}(h^\star_{\delta_i}))\end{aligned}$$

where we used Moreau identity in the first equality. Since $i \in I_{x_k} \cap I^c_{x^\star}$, we have $h_{k-1,\delta_i} \notin \gamma_{k-1}C_i$ and $h^\star_{\delta_i} \in \gamma_{k-1}C_i$, or equivalently, that $\mathcal{P}_{\gamma_{k-1}C_i}(h_{k-1,\delta_i}) \in \gamma_{k-1}\mathrm{rbd}(C_i) = \gamma_{k-1}\mathrm{rbd}(\partial\sigma_{C_i}(x^\star_{\delta_i}))$ and $\mathcal{P}_{\gamma_{k-1}C_i}(h^\star_{\delta_i}) = h^\star_{\delta_i}$. Combining this with the fact that the orthogonal projector on $\gamma_{k-1}C_i$ is firmly non-expansive, we obtain

$$\begin{aligned}\|x_{k,\delta_i} - x^\star_{\delta_i}\|^2 &\leq \|h_{k-1,\delta_i} - h^\star_{\delta_i}\|^2 - \|\mathcal{P}_{\gamma_{k-1}C_i}(h_{k-1,\delta_i}) - h^\star_{\delta_i}\|^2 \\ &= \|h_{k-1,\delta_i} - h^\star_{\delta_i}\|^2 - \|\mathcal{P}_{\gamma_{k-1}C_i}(h_{k-1,\delta_i}) + \gamma_{k-1}\nabla F(x^\star)_{\delta_i}\|^2 \\ &\leq \|h_{k-1,\delta_i} - h^\star_{\delta_i}\|^2 - \gamma^2_{k-1}\mathrm{dist}\big(-\nabla F(x^\star)_{\delta_i}, \mathrm{rbd}(C_i)\big)^2 \\ &\leq \|h_{k-1,\delta_i} - h^\star_{\delta_i}\|^2 - \underline{\epsilon}^2\mathrm{dist}\big(-\nabla F(x^\star)_{\delta_i}, \mathrm{rbd}(C_i)\big)^2.\end{aligned}$$

This bound together with non-expansiveness of $\mathrm{prox}_{\gamma_{k-1}C_i}$ and $\mathrm{Id} - \gamma_{k-1}\nabla F$ yield

$$\begin{aligned}\|x_k - x^\star\|^2 &= \textstyle\sum_{i\in I^c_{x^\star}} \|x_{k,\delta_i} - x^\star_{\delta_i}\|^2 + \sum_{j\in I_{x^\star}} \|x_{k,b_j} - x^\star_{b_j}\|^2 \\ &\leq \|h_{k-1} - h^\star\|^2 - \underline{\epsilon}^2 \textstyle\sum_{i\in I^c_{x^\star}} \mathrm{dist}\big(-\nabla F(x^\star)_{\delta_i}, \mathrm{rbd}(C_i)\big)^2 \\ &\leq \|x_{k-1} - x^\star\|^2 - \underline{\epsilon}^2 \textstyle\sum_{i\in I^c_{x^\star}} \mathrm{dist}\big(-\nabla F(x^\star)_{\delta_i}, \mathrm{rbd}(C_i)\big)^2 \\ &\leq \|x_0 - x^\star\|^2 - k\underline{\epsilon}^2 \textstyle\sum_{i\in I^c_{x^\star}} \mathrm{dist}\big(-\nabla F(x^\star)_{\delta_i}, \mathrm{rbd}(C_i)\big)^2,\end{aligned}$$

where the last term in the right hand side is strictly positive by ($\mathrm{ND}_{\mathrm{FB}}$). Taking k as the largest integer such that the right hand side is positive, we deduce that the number of iterations where

$I^c_{x^\star} \cap I_{x_k} \neq \emptyset$ does not exceed the given bound. We then conclude that beyond this bound, there is no i such that $\mathcal{M}_{x_{k,b_i}} \neq 0$ while $\mathcal{M}_{x^\star_{b_i}} = 0$. The proof is complete. □

6.7.2 Proofs of Section 6.3

Proof of Proposition 6.3.1.

(i) Since F is locally C^2 around $x^\star$, there exists $\epsilon > 0$ sufficiently small such that for any $\delta \in \mathbb{B}_\epsilon(0)$, we have

$$\begin{aligned}\Phi(x^\star+\delta) - \Phi(x^\star) &= F(x^\star+\delta) - F(x^\star) - \langle \nabla F(x^\star),\, \delta\rangle + R(x^\star+\delta) - R(x^\star) + \langle \nabla F(x^\star),\, \delta\rangle \\ &= \frac{1}{2}\langle \delta,\, \nabla^2 F(x^\star+t\delta)\delta\rangle + R(x^\star+\delta) - R(x^\star) + \langle \nabla F(x^\star),\, \delta\rangle, \quad t \in]0,1[.\end{aligned}$$

Let $x_t = x^\star + t\delta \in \mathbb{B}_\epsilon(x^\star)$. Since (RI) holds and $\nabla^2 F(x)$ depends continuously on $x \in \mathbb{B}_\epsilon(x^\star)$, we have $\mathcal{P}_{T_{x^\star}} \nabla^2 F(x) \mathcal{P}_{T_{x^\star}} \succeq \kappa \mathrm{Id}$ for any such x. This holds in particular at x_t. We then distinguish two cases.

(a) $\delta \notin \ker(\nabla^2 F(x_t))$. In this case, it is clear that

$$\Phi(x^\star+\delta) - \Phi(x^\star) \geq \frac{1}{2}\langle \delta,\, \nabla^2 F(x_t)\delta\rangle \geq \kappa/2\|\delta\|^2 > 0$$

since F is convex and locally C^2, and R is convex with $-\nabla F(x^\star) \in \partial R(x^\star)$.

(b) $\delta \in \ker(\nabla^2 F(x_t)) \setminus \{0\}$. Since R is a proper closed convex function, it is sub-differentially regular at $x^\star$. Moreover $\partial R(x^\star) \neq \emptyset$ ($-\nabla F(x^\star)$ is in it), and thus the directional derivative $R'(x^\star, \cdot)$ is proper and closed, and it is the support of $\partial R(x^\star)$ [153, Theorem 8.30]. It then follows from the separation theorem [96, Theorem V.2.2.3] that

$$\begin{aligned}&-\nabla F(x^\star) \in \mathrm{ri}(\partial R(x^\star)) \\ \iff\ & R'(x^\star,\delta) > -\langle \nabla F(x^\star),\, \delta\rangle,\ \forall \delta \text{ s.t. } R'(x^\star;\delta) + R'(x^\star;-\delta) > 0.\end{aligned}$$

As $\ker(R'(x^\star;\cdot)) = T_{x^\star}$ [169, Proposition 3(iii) and Lemma 10], and in view of (RI), we get

$$\begin{aligned}&-\nabla F(x^\star) \in \mathrm{ri}(\partial R(x^\star)) \\ \iff\ & R'(x^\star;\delta) > -\langle \nabla F(x^\star),\, \delta\rangle,\ \forall \delta \notin T_{x^\star} \\ \implies\ & R'(x^\star;\delta) > -\langle \nabla F(x^\star),\, \delta\rangle,\ \forall \delta \in \ker(\nabla^2 F(x_t)) \setminus \{0\}.\end{aligned}$$

Combining this with classical properties of the directional derivative of a convex function yields

$$\begin{aligned}\Phi(x^\star+\delta) - \Phi(x^\star) &= R(x^\star+\delta) - R(x^\star) + \langle \nabla F(x^\star),\, \delta\rangle \\ &\geq R'(x^\star;\delta) + \langle \nabla F(x^\star),\, \delta\rangle > 0,\end{aligned}$$

which concludes the first claim.

(ii) Let $\overline{R}$ as defined in the proof of Lemma 6.3.3. If $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star})$, the Riemannian Hessian of Φ reads

$$\nabla^2_{\mathcal{M}_{x^\star}} \Phi(x^\star) = \mathcal{P}_{T_{x^\star}} \nabla F(x^\star) \mathcal{P}_{T_{x^\star}} + \nabla^2_{\mathcal{M}_{x^\star}} \overline{R}(x^\star).$$

In view of Lemma 6.3.3(i), $\nabla^2_{\mathcal{M}_{x^\star}} \overline{R}(x^\star)$ is positive semi-definite on $T_{x^\star}$. On the other hand, hypothesis (RI) entails positive definiteness of $\mathcal{P}_{T_{x^\star}} \nabla F(x^\star) \mathcal{P}_{T_{x^\star}}$. Altogether, this shows that $\nabla^2_{\mathcal{M}_{x^\star}} \Phi(x^\star)$ is positive definite on $T_{x^\star} \setminus \{0\}$. Local quadratic growth of Φ near $x^\star$ then follows by combining [108, Definition 5.4], [126, Theorem 3.4] and [89, Theorem 6.2]. □

Proof of Lemma 6.3.3. By definition of U, $Uh = 0$ for any $h \in T^\perp_{x^\star}$. Thus, in the following we only examine the case $h \in T_{x^\star}$.

(i) Let $\Psi(x) \overset{\text{def}}{=} R(x) + \langle x, \nabla F(x^\star)\rangle$. Owing to Corollary 5.1.10, $\Psi \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star})$. Moreover, from Lemma 5.1.4 and normal sharpness, the Riemannian Hessian of Ψ at $x^\star$ is such that, $\forall h \in T_{x^\star}$,

$$\begin{aligned}\gamma\nabla^2_{\mathcal{M}_{x^\star}}\Psi(x^\star)h &= \gamma\mathcal{P}_{T_{x^\star}}\nabla^2\widetilde{\Psi}(x^\star)h + \gamma\mathfrak{W}_{x^\star}\big(h, \mathcal{P}_{T^\perp_{x^\star}}\nabla\widetilde{\Psi}(x^\star)\big)\\ &= \gamma\mathcal{P}_{T_{x^\star}}\nabla^2\widetilde{R}(x^\star)h + \gamma\mathfrak{W}_{x^\star}\big(h, \mathcal{P}_{T^\perp_{x^\star}}\nabla\widetilde{\Phi}(x^\star)\big)\\ &= \gamma\nabla^2_{\mathcal{M}_{x^\star}}\Phi(x^\star)\mathcal{P}_{T_{x^\star}}h - Hh = Uh,\end{aligned}$$

where $\widetilde{\cdot}$ is the smooth representative of the corresponding function.
Since $-\nabla F(x^\star) \in \mathrm{ri}(\partial R(x^\star))$, we have from [112, Corollary 5.4] that

$$\partial^2 R\big(x^\star|-\nabla F(x^\star)\big)h = \begin{cases}\nabla^2_{\mathcal{M}_{x^\star}}\Psi(x^\star)h + T^\perp_{x^\star}, & h \in T_{x^\star},\\ \emptyset, & h \notin T_{x^\star},\end{cases}$$

where $\partial^2 R(x^\star|-\nabla F(x^\star))$ denotes the Mordukhovich generalized Hessian mapping of function R at $(x^\star, -\nabla F(x^\star)) \in \mathrm{gra}(\partial R)$ [129]. As $R \in \Gamma_0(\mathbb{R}^n)$, ∂R is a maximal monotone operator, and in view of [144, Theorem 2.1] we have that the mapping $\partial^2 R(x^\star|-\nabla F(x^\star))$ is positive semi-definite, whence we conclude that $\forall h \in T_{x^\star}$,

$$0 \le \gamma\langle\partial^2 R\big(x^\star|-\nabla F(x^\star)\big)h,\, h\rangle = \gamma\langle\nabla^2_{\mathcal{M}_{x^\star}}\Psi(x^\star)h,\, h\rangle = \langle Uh,\, h\rangle.$$

(ii) In this case, $U = \gamma\mathcal{P}_{T_{x^\star}}\nabla^2\widetilde{R}(x^\star)\mathcal{P}_{T_{x^\star}}$. Let $x_t = x^\star + th$, $t > 0$, for any scalar t and $h \in T_{x^\star}$. Obviously, $x_t \in x^\star + T_{x^\star} = \mathcal{M}_{x^\star}$, and for t sufficiently small, by Lemma 5.1.3, $T_{x_t} = T_{x^\star}$. Thus, $\forall u \in \partial R(x^\star)$ and $\forall v \in \partial R(x_t)$

$$\begin{aligned}0 \le t^{-2}\langle v-u,\, x_t - x^\star\rangle &= t^{-1}\langle v-u,\, \mathcal{P}_{T_{x^\star}}h\rangle\\ &= t^{-1}\langle\mathcal{P}_{T_{x^\star}}(v-u),\, h\rangle\\ &= t^{-1}\langle\mathcal{P}_{T_{x_t}}v - \mathcal{P}_{T_{x^\star}}u,\, h\rangle\\ \text{(by Lemma 5.1.4)} &= \langle t^{-1}(\nabla_{\mathcal{M}_{x^\star}}R(x_t) - \nabla_{\mathcal{M}_{x^\star}}R(x^\star)),\, h\rangle\\ \text{(by (2.6.3))} &= \langle t^{-1}\mathcal{P}_{T_{x^\star}}(\nabla\widetilde{R}(x^\star + t\mathcal{P}_{T_{x^\star}}h) - \nabla\widetilde{R}(x^\star)),\, h\rangle.\end{aligned}$$

Since $\widetilde{R}$ is C^2, passing to the limit as $t \to 0$ leads to the desired result. □

Proof of Lemma 6.3.4.

(i) (a) is proved using the assumptions and Rademacher theorem. (b) and (c) follow from simple linear algebra arguments.

(ii) From Lemma 6.3.3, we have $WG = W^{1/2}W^{1/2}GW^{1/2}W^{-1/2}$, meaning that WG is similar to $W^{1/2}GW^{1/2}$. The latter is symmetric and obeys

$$\|W^{1/2}GW^{1/2}\| \le \|W^{1/2}\|\|G\|\|W^{1/2}\| < 1,$$

where we used (i):(b) to get the last inequality. Thus $W^{1/2}GW^{1/2}$ has real eigenvalues in $]-1, 1[$, and so does WG by similarity. The last statement follows using (i):(c). □

The proof of Theorem 6.3.6 consists of several steps, first we prove that under the required setting, we can obtain (6.3.5), *i.e.* the linearized fixed-point iteration.

Proposition 6.7.1. *Let* (F.1)-(F.3) *hold, and assume that an FB-type method is used to create a sequence* $\{x_k\}_{k\in\mathbb{N}}$ *that converges to* $x^\star \in \mathrm{Argmin}(\Phi)$ *such that conditions* ($\mathrm{ND}_{\mathrm{FB}}$), (RI) *and* (6.3.3) *hold. Then for all* k *large enough,*

$$d_{k+1} = M_{\mathrm{FB}}d_k + o(\|d_k\|).$$

The term $o(\cdot)$ *vanishes if* R *is polyhedral around* $x^\star$ *and* $(\gamma_k, a_{i,k}, b_{i,k})$ *are chosen constants.*

Define the iteration-dependent versions of the matrices in (6.3.1) and (6.3.4), *i.e.*

$$\begin{aligned}
&H_k \overset{\text{def}}{=} \gamma_k \mathcal{P}_{T_{x^\star}} \nabla^2 F(x^\star) \mathcal{P}_{T_{x^\star}},\ G_k \overset{\text{def}}{=} \mathrm{Id} - H_k,\ U_k \overset{\text{def}}{=} \gamma_k \nabla^2_{\mathcal{M}_{x^\star}} \Phi(x^\star) \mathcal{P}_{T_{x^\star}} - H_k,\\
&M_{0,k} \overset{\text{def}}{=} (a_{0,k} - b_{0,k})P + (1 + b_{0,k})PG,\ M_{s,k} \overset{\text{def}}{=} -(a_{s-1,k} - b_{s-1,k})P - b_{s-1,k}PG,\\
&M_{i,k} \overset{\text{def}}{=} -\big((a_{i-1,k} - a_{i,k}) - (b_{i-1,k} - b_{i,k})\big)P - (b_{i-1,k} - b_{i,k})PG,\ i = 1, ..., s-1,\\
&M_{\mathrm{FB},k} \overset{\text{def}}{=} \begin{bmatrix} M_{0,k} & M_{1,k} & \cdots & M_{s-1,k} & M_{s,k} \\ \mathrm{Id} & 0 & \cdots & 0 & 0 \\ 0 & \mathrm{Id} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathrm{Id} & 0 \end{bmatrix}.
\end{aligned} \tag{6.7.2}$$

After the finite identification of $\mathcal{M}_{x^\star}$, we have $x_k \in \mathcal{M}_{x^\star}$ for x_k close enough to $x^\star$. Let T_{x_k} be their corresponding tangent spaces, and define $\tau_k : T_{x^\star} \to T_{x_k}$ the parallel translation along the unique geodesic joining from x_k to $x^\star$.

Before proving Proposition 6.7.1, we first establish the following intermediate result which provides useful estimates. Define variable $r_k \overset{\text{def}}{=} x_k - x^\star$, and recall that $\mathcal{S} = 0, ..., s-1$.

Proposition 6.7.2. *Under the assumptions of Proposition 6.7.1, we have*

$$\begin{aligned}
&\|y_{a,k} - x^\star\| = O(\|d_k\|),\ \|y_{b,k} - x^\star\| = O(\|d_k\|),\ \|r_{k+1}\| = O(\|d_k\|),\\
&(\tau_{k+1}^{-1} \mathcal{P}_{T_{x_{k+1}}} - \mathcal{P}_{T_{x^\star}})\big(\nabla F(y_{b,k}) - \nabla F(x_{k+1})\big) = o(\|d_k\|).
\end{aligned} \tag{6.7.3}$$

and

$$\|W(U_k - U) r_{k+1}\| = o(\|d_k\|),\ \ \|(M_{\mathrm{FB},k} - M_{\mathrm{FB}}) d_k\| = o(\|d_k\|). \tag{6.7.4}$$

Proof. Denote $\mathcal{C} = \sum_i |a_{i,k}|$ which is bounded, then

$$\begin{aligned}
\|y_{a,k} - x^\star\| &= \|x_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(x_{k-i} - x_{k-i-1}) - x^\star + \sum_{i\in\mathcal{S}} a_{i,k}(x^\star - x^\star)\| \\
&\leq \|x_k - x^\star\| + \textstyle\sum_{i\in\mathcal{S}} |a_{i,k}|(\|x_{k-i} - x^\star\| + \|x_{k-i-1} - x^\star\|) \\
&\leq (\mathcal{C}+1) \textstyle\sum_{i\in\mathcal{S}} \|r_{k-i}\| \leq (\mathcal{C}+1)\sqrt{s+1} \left\| \begin{bmatrix} r_k \\ \vdots \\ r_{k-s} \end{bmatrix} \right\| = (\mathcal{C}+1)\sqrt{s+1}\|d_k\|,
\end{aligned} \tag{6.7.5}$$

hence we get the first and second estimates. Since $\mathrm{prox}_{\gamma_k R}$ firmly non-expansive, and ∇F is $(1/\beta)$-Lipschitz continuous, then

$$\begin{aligned}
\|r_{k+1}\| &= \|\mathrm{prox}_{\gamma_k R}(y_{a,k} - \gamma_k \nabla F(y_{b,k})) - \mathrm{prox}_{\gamma_k R}(x^\star - \gamma_k \nabla F(x^\star))\| \\
&\leq \|(y_{a,k} - x^\star) - \gamma_k(\nabla F(y_{b,k}) - \nabla F(x^\star))\| \\
&\leq (\|y_{a,k} - x^\star\| + \tfrac{\gamma_k}{\beta}\|y_{b,k} - x^\star\|) \\
&\leq 2\sqrt{s+1}\big(1 + \tfrac{\gamma_k}{\beta}\big)\|d_k\| \leq 6\sqrt{s+1}\|d_k\|,
\end{aligned} \tag{6.7.6}$$

which yields the third estimate. Combining Lemma 2.6.2, (6.7.5) and (6.7.6), we get

$$\begin{aligned}
(\tau_{k+1}^{-1} \mathcal{P}_{T_{x_{k+1}}} - \mathcal{P}_{T_{x^\star}})\big(\nabla F(y_{b,k}) - \nabla F(x_{k+1})\big) &= o(\|\nabla F(y_{b,k}) - \nabla F(x_{k+1})\|) \\
&= o(\|y_{b,k} - x^\star\|) + o(\|r_{k+1}\|) = o(\|d_k\|).
\end{aligned}$$

Let's now turn to (6.7.4). First, define the function $\overline{R}(x) \overset{\text{def}}{=} R(x) + \langle x, \nabla F(x^\star)\rangle$. From Corollary 5.1.10, we have $\overline{R} \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star})$. Moreover, from Definition 5.1.4 and local normal sharpness Lemma

5.1.3, the Riemannian Hessian of $\overline{R}$ at $x^\star$ is such that, $\forall h \in T_{x^\star}$,

$$\begin{aligned}\gamma\nabla^2_{\mathcal{M}_{x^\star}}\overline{R}(x^\star)h &= \gamma\mathcal{P}_{T_{x^\star}}\nabla^2\widetilde{\overline{R}}(x^\star)h + \gamma\mathfrak{W}_{x^\star}\big(h, \mathcal{P}_{T^\perp_{x^\star}}\nabla\widetilde{\overline{R}}(x^\star)\big)\\ &= \gamma\mathcal{P}_{T_{x^\star}}\nabla^2\widetilde{R}(x^\star)h + \gamma\mathfrak{W}_{x^\star}\big(h, \mathcal{P}_{T^\perp_{x^\star}}\nabla\widetilde{\Phi}(x^\star)\big)\\ &= \gamma\nabla^2_{\mathcal{M}_{x^\star}}\Phi(x^\star)\mathcal{P}_{T_{x^\star}}h - Hh = Uh,\end{aligned}$$

where $\widetilde{\cdot}$ is the smooth representative of the corresponding function. Then we have

$$\begin{aligned}\lim_{k\to\infty}\frac{\|W(U_k-U)r_{k+1}\|}{\|r_{k+1}\|} &= \lim_{k\to\infty}\frac{\|W(\gamma_k-\gamma)\nabla^2_{\mathcal{M}_{x^\star}}\overline{R}(x^\star)\mathcal{P}_{T_{x^\star}}r_{k+1}\|}{\|r_{k+1}\|}\\ &\le \lim_{k\to\infty}|\gamma_k-\gamma|\|W\|\|\nabla^2_{\mathcal{M}_{x^\star}}\overline{R}(x^\star)\mathcal{P}_{T_{x^\star}}\| = 0,\end{aligned}$$

which entails $\|W(U_k-U)r_{k+1}\| = o(\|r_{k+1}\|) = o(\|d_k\|)$. Similarly, since H is (γ/β)-Lipschitz, we have

$$\lim_{k\to\infty}\frac{\|W(G_k-G)r_k\|}{\|r_k\|} = \lim_{k\to\infty}\frac{\|W(\gamma_k-\gamma)Hr_k\|}{\|r_k\|} \le \lim_{k\to\infty}\frac{\gamma|\gamma_k-\gamma|}{\beta}\|W\| = 0. \tag{6.7.7}$$

Now, let's consider $(M_{\mathrm{FB},k}-M_{\mathrm{FB}})d_k$

$$M_{\mathrm{FB},k}-M_{\mathrm{FB}} = \begin{bmatrix} M_{0,k}-M_0 & M_{1,k}-M_1 & \cdots & M_{s-1,k}-M_{s-1} & M_{s,k}-M_s\\ 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & 0\end{bmatrix}.$$

Take $(M_{0,k}-M_0)r_k$, we have

$$\begin{aligned}(M_{0,k}-M_0)r_k &= \big((a_{0,k}-b_{0,k})W + (1+b_{0,k})WG_k\big)r_k - \big((a_0-b_0)W + (1+b_0)WG\big)r_k\\ &= \big((a_{0,k}-b_{0,k})-(a_0-b_0)\big)Wr_k + (1+b_{0,k})W(G_k-G)r_k + (b_{0,k}-b_0)WGr_k.\end{aligned}$$

Since we assume that $a_{i,k}\to a_i, b_{i,k}\to b_i, i\in\mathcal{S}$ and $\gamma_k\to\gamma$, plus (6.7.7), it can be shown that

$$\begin{aligned}&\lim_{k\to\infty}\frac{\|(M_{0,k}-M_0)r_k\|}{\|r_k\|}\\ &\le \lim_{k\to\infty}|(a_{0,k}-b_{0,k})-(a_0-b_0)|\|W\| + |1+b_{0,k}|\frac{\gamma|\gamma_k-\gamma|}{\beta}\|W\| + |b_{0,k}-b_0|\|W\|\|G\| = 0,\end{aligned}$$

that is $\|(M_{0,k}-M_0)r_k\| = o(\|r_k\|)$. Using the same arguments, we can show that

$$\|(M_{i,k}-M_i)r_{k-i}\| = o(\|r_{k-i}\|),\ i=1,...,s-1 \quad\text{and}\quad \|(M_{s,k}-M_s)r_{k,s}\| = o(\|r_{s,k}\|).$$

Assemble them together, we obtain

$$\|(M_{\mathrm{FB},k}-M_{\mathrm{FB}})d_k\| = o(\|d_k\|),$$

which concludes the proof. □

Proof of Proposition 6.7.1. From the update (6.1.1) and the condition for a critical point $x^\star$ of problem ($\mathcal{P}_{\mathrm{FB}}$), we have

$$\begin{aligned}y_{a,k}-x_{k+1}-\gamma_k\big(\nabla F(y_{b,k})-\nabla F(x_{k+1})\big) &\in \gamma_k\partial\Phi(x_{k+1})\\ 0 &\in \gamma_k\partial\Phi(x^\star).\end{aligned}$$

Projecting into $T_{x_{k+1}}$ and $T_{x^\star}$, respectively, and using Definition 5.1.4, leads to

$$\begin{aligned}\gamma_k\tau_{k+1}^{-1}\nabla_{\mathcal{M}_{x^\star}}\Phi(x_{k+1}) &= \tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}\big(y_{a,k}-x_{k+1}-\gamma_k\big(\nabla F(y_{b,k})-\nabla F(x_{k+1})\big)\big)\\ \gamma_k\nabla_{\mathcal{M}_{x^\star}}\Phi(x^\star) &= 0.\end{aligned}$$

Adding both identities, and subtracting $\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}x^\star$ on both sides, we arrive at

$$\begin{aligned}&\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}r_{k+1}+\gamma_k\big(\tau_{k+1}^{-1}\nabla_{\mathcal{M}_{x^\star}}\Phi(x_{k+1})-\nabla_{\mathcal{M}_{x^\star}}\Phi(x^\star)\big)\\&=\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}(y_{a,k}-x^\star)-\gamma_k\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}\big(\nabla F(y_{b,k})-\nabla F(x_{k+1})\big).\end{aligned}\tag{6.7.8}$$

By virtue of Lemma 2.6.2, we get

$$\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}r_{k+1}=\mathcal{P}_{T_{x^\star}}r_{k+1}+(\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}-\mathcal{P}_{T_{x^\star}})r_{k+1}=\mathcal{P}_{T_{x^\star}}r_{k+1}+o(\|r_{k+1}\|).$$

Using Lemma 2.6.1, we also have

$$r_{k+1}=\mathcal{P}_{T_{x^\star}}r_{k+1}+o(\|r_{k+1}\|),$$

and thus

$$\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}r_{k+1}=r_{k+1}+o(\|r_{k+1}\|)=r_{k+1}+o(\|d_k\|),\tag{6.7.9}$$

where we also used (6.7.3). Similarly

$$\begin{aligned}&\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}(y_{a,k}-x^\star)\\&=\mathcal{P}_{T_{x^\star}}(y_{a,k}-x^\star)+(\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}-\mathcal{P}_{T_{x^\star}})(y_{a,k}-x^\star)\\&=\mathcal{P}_{T_{x^\star}}(y_{a,k}-x^\star)+o(\|y_{a,k}-x^\star\|)\\&=\mathcal{P}_{T_{x^\star}}(y_{a,k}-x^\star)+o(\|d_k\|)\\&=\mathcal{P}_{T_{x^\star}}(x_k-x^\star)+\textstyle\sum_{i\in\mathcal{S}}a_{i,k}\mathcal{P}_{T_{x^\star}}\big((x_{k-i}-x^\star)-(x_{k-i-1}-x^\star)\big)+o(\|d_k\|)\\&=r_k+o(\|r_k\|)+\textstyle\sum_{i\in\mathcal{S}}a_{i,k}\big(r_{k-i}-r_{k-i-1}+o(\|r_{k-i}\|)+o(\|r_{k-i-1}\|)\big)+o(\|d_k\|)\\&=r_k+\textstyle\sum_{i\in\mathcal{S}}a_{i,k}(r_{k-i}-r_{k-i-1})+\sum_{i\in\mathcal{S}\cup\{s\}}o(\|r_{k-i}\|)+o(\|d_k\|)\\&=(y_{a,k}-x^\star)+o(\|d_k\|).\end{aligned}\tag{6.7.10}$$

Moreover owing to Lemma 2.6.3 and (6.7.3),

$$\begin{aligned}\tau^{-1}\nabla_{\mathcal{M}_{x^\star}}\Phi(x_{k+1})-\nabla_{\mathcal{M}_{x^\star}}\Phi(x^\star)&=\nabla^2_{\mathcal{M}_{x^\star}}\Phi(x^\star)\mathcal{P}_{T_{x^\star}}r_{k+1}+o(\|r_{k+1}\|)\\&=\nabla^2_{\mathcal{M}_{x^\star}}\Phi(x^\star)\mathcal{P}_{T_{x^\star}}r_{k+1}+o(\|d_k\|).\end{aligned}\tag{6.7.11}$$

Therefore, inserting (6.7.9), (6.7.10) and (6.7.11) into (6.7.8), we obtain

$$\begin{aligned}&\big(\mathrm{Id}+\gamma_k\nabla^2_{\mathcal{M}_{x^\star}}\Phi(x^\star)\mathcal{P}_{T_{x^\star}}\big)r_{k+1}\\&=(y_{a,k}-x^\star)-\gamma_k\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}\big(\nabla F(y_{b,k})-\nabla F(x_{k+1})\big)+o(\|d_k\|).\end{aligned}\tag{6.7.12}$$

Owing to (6.7.3) and local C^2-smoothness of F, we have

$$\begin{aligned}&\tau_{k+1}^{-1}\mathcal{P}_{T_{x_{k+1}}}\big(\nabla F(y_{b,k})-\nabla F(x_{k+1})\big)\\&=\mathcal{P}_{T_{x^\star}}\big(\nabla F(y_{b,k})-\nabla F(x_{k+1})\big)+o(\|d_k\|)\\&=\mathcal{P}_{T_{x^\star}}\big(\nabla F(y_{b,k})-\nabla F(x^\star)\big)-\mathcal{P}_{T_{x^\star}}\big(\nabla F(x_{k+1})-\nabla F(x^\star)\big)+o(\|d_k\|)\\&=\mathcal{P}_{T_{x^\star}}\nabla^2F(x^\star)(y_{b,k}-x^\star)+o(\|y_{b,k}-x^\star\|)-\mathcal{P}_{T_{x^\star}}\nabla^2F(x^\star)r_{k+1}+o(\|r_{k+1}\|)+o(\|d_k\|)\\&=\mathcal{P}_{T_{x^\star}}\nabla^2F(x^\star)\mathcal{P}_{T_{x^\star}}(y_{b,k}-x^\star)-\mathcal{P}_{T_{x^\star}}\nabla^2F(x^\star)\mathcal{P}_{T_{x^\star}}(x_{k+1}-x^\star)+o(\|d_k\|).\end{aligned}\tag{6.7.13}$$

Injecting (6.7.13) in (6.7.12), we get

$$\begin{aligned}&\big(\mathrm{Id}+\gamma_k\nabla^2_{\mathcal{M}_{x^\star}}\Phi(x^\star)\mathcal{P}_{T_{x^\star}}-\gamma_k\mathcal{P}_{T_{x^\star}}\nabla^2F(x^\star)\mathcal{P}_{T_{x^\star}}\big)r_{k+1}\\&=(\mathrm{Id}+U_k)r_{k+1}=(y_{a,k}-x^\star)-H_k(y_{b,k}-x^\star)+o(\|d_k\|),\end{aligned}\tag{6.7.14}$$

which can be further written as, recall that $H_k = \mathrm{Id} - G_k$,

$$\begin{aligned}
(\mathrm{Id} + U_k)r_{k+1} &= (\mathrm{Id} + U)r_{k+1} + (U_k - U)r_{k+1} \\
&= (y_{a,k} - x^\star) - H_k(y_{b,k} - x^\star) + o(\|d_k\|) \\
&= r_k + \textstyle\sum_{i\in\mathcal{S}} a_{i,k}(r_{k-i} - r_{k-i-1}) - H_k\big(r_k + \sum_{i\in\mathcal{S}} b_{i,k}(r_{k-i} - r_{k-i-1})\big) + o(\|d_k\|) \\
&= (1 + a_{0,k})r_k - \textstyle\sum_{i=1}^{s-1}(a_{i-1,k} - a_{i,k})r_{k-i} - a_{s-1,k}r_{k-s} \\
&\quad - H_k\big((1 + b_{0,k})r_k - \textstyle\sum_{i=1}^{s-1}(b_{i-1,k} - b_{i,k})r_{k-i} - b_{s-1,k}r_{k-s}\big) + o(\|d_k\|) \\
&= (1 + a_{0,k})r_k - \textstyle\sum_{i=1}^{s-1}(a_{i-1,k} - a_{i,k})r_{k-i} - a_{s-1,k}r_{k-s} \\
&\quad - (1 + b_{0,k})H_k r_k + H_k \textstyle\sum_{i=1}^{s-1}(b_{i-1,k} - b_{i,k})r_{k-i} + H_k b_{s-1,k}r_{k-s} + o(\|d_k\|) \\
&= \big((1 + a_{0,k})\mathrm{Id} - (1 + b_{0,k})H_k\big)r_k - (a_{s-1,k}\mathrm{Id} - b_{s-1,k}H_k)r_{k-s} \\
&\quad - \textstyle\sum_{i=1}^{s-1}\big((a_{i-1,k} - a_{i,k})\mathrm{Id} - (b_{i-1,k} - b_{i,k})H_k\big)r_{k-i} + o(\|d_k\|) \\
&= \big((a_{0,k} - b_{0,k})\mathrm{Id} + (1 + b_{0,k})G_k\big)r_k - \big((a_{s-1,k} - b_{s-1,k})\mathrm{Id} + b_{s-1,k}G_k\big)r_{k-s} \\
&\quad - \textstyle\sum_{i=1}^{s-1}\big((a_{i-1,k} - a_{i,k})\mathrm{Id} - (b_{i-1,k} - b_{i,k})\mathrm{Id} + (b_{i-1,k} - b_{i,k})G_k\big)r_{k-i} + o(\|d_k\|).
\end{aligned}$$

Inverting $\mathrm{Id} + U$, we obtain

$$\begin{aligned}
&r_{k+1} + W(U_k - U)r_{k+1} \\
&= \big((a_{0,k} - b_{0,k})W + (1 + b_{0,k})WG_k\big)r_k - \big((a_{s-1,k} - b_{s-1,k})W + b_{s-1,k}WG_k\big)r_{k-s} \\
&\quad - \textstyle\sum_{i=1}^{s-1}\big((a_{i-1,k} - a_{i,k})W - (b_{i-1,k} - b_{i,k})W + (b_{i-1,k} - b_{i,k})WG_k\big)r_{k-i} + o(\|d_k\|) \\
&= M_{0,k}r_k + M_{s,k}r_{k-s} + \textstyle\sum_{i=1}^{s-1} M_{i,k}r_{k-i} + o(\|d_k\|).
\end{aligned}$$

Using the estimates (6.7.4), we get

$$d_{k+1} = \big(M_{\mathrm{FB}} + (M_{\mathrm{FB},k} - M_{\mathrm{FB}})\big)d_k + o(\|d_k\|) = M_{\mathrm{FB}}d_k + o(\|d_k\|). \qquad \square$$

With the above result, we are able to prove the claim (6.3.6), hence Theorem 6.3.6.

Proof of Theorem 6.3.6. Since $\rho(M_{\mathrm{FB}}) < 1$, then M_{FB} is convergent with $\lim_{k\to\infty} M_{\mathrm{FB}}^k = 0$. Define $\psi_k = o(d_k)$, suppose after $K > 0$ iterations, (6.3.5) holds, then for $k \geq K$

$$d_{k+1} = {M_{\mathrm{FB}}}^{k+1-K} d_K + \textstyle\sum_{j=K}^{k} {M_{\mathrm{FB}}}^{k-j}\psi_j. \tag{6.7.15}$$

Since the spectral radius $\rho(M_{\mathrm{FB}}) < 1$, then from the spectral radius formula, given any $\rho \in]\rho(M_{\mathrm{FB}}), 1[$, there exists a constant $\mathcal{C}$ such that, for any $k \in \mathbb{N}$

$$\|M_{\mathrm{FB}}^k\| \leq \|M_{\mathrm{FB}}\|^k \leq \mathcal{C}\rho^k.$$

Therefore, from (6.7.15), we get

$$\begin{aligned}
\|d_{k+1}\| &\leq \|{M_{\mathrm{FB}}}^{k+1-K} d_K + \textstyle\sum_{j=K}^{k} {M_{\mathrm{FB}}}^{k-j}\psi_j\| \\
&\leq \|M_{\mathrm{FB}}\|^{k+1-K}\|d_K\| + \textstyle\sum_{j=K}^{k}\|M_{\mathrm{FB}}\|^{k-j}\|\psi_j\| \\
&\leq \mathcal{C}\rho^{k+1-K}\|d_K\| + \mathcal{C}\textstyle\sum_{j=K}^{k}\rho^{k-j}\|\psi_j\|.
\end{aligned}$$

Together with the fact that $\psi_j = o(\|d_j\|)$ leads to the claimed result. See also the result of [147, Section 2.1.2, Theorem 1]. $\square$

Proof of Theorem 6.3.8. For the sake of simplicity, we prove the theorem for the case $s = 2$, while for other choices of s, the proof extends smoothly. Since R is locally polyhedral, then $\nabla_{\mathcal{M}_{x^\star}}\Phi(x_k)$ is locally constant along $\mathcal{M}_{x^\star} = x^\star + T_{x^\star}$ around $x^\star$. Thus, derive from (6.7.14), for k large enough, we get

$$x_{k+1} - x^\star = (y_{a,k} - x^\star) - E_k(y_{b,k} - x^\star),$$

where we used the mean-value theorem with $E_k = \gamma_k \int_0^1 \nabla^2 F(x^\star + t(y_{b,k} - x^\star))\mathrm{d}t \succeq 0$. Using that E_k is symmetric and $\mathrm{ran}(E_k)^\perp = V$, we have

$$\begin{aligned}&\mathcal{P}_V(x_{k+1} - x^\star)\\ &= (1 + a_{0,k})\mathcal{P}_V(x_k - x^\star) - (a_{0,k} - a_{1,k})\mathcal{P}_V(x_{k-1} - x^\star) - a_{1,k}\mathcal{P}_V(x_{k-2} - x^\star).\end{aligned}$$

If $a_{0,k}, a_{1,k} = 0$, then $\mathcal{P}_V(x_{k+1} - x^\star) = \mathcal{P}_V(x_k - x^\star)$. Thus, in the rest, without loss of generality, we assume that $a_{0,k}, a_{1,k} \neq 0$ for k large enough. The above iteration leads to

$$\begin{pmatrix}\mathcal{P}_V(x_{k+1} - x^\star)\\ \mathcal{P}_V(x_k - x^\star)\\ \mathcal{P}_V(x_{k-1} - x^\star)\end{pmatrix} = \begin{bmatrix}(1 + a_{0,k})\mathrm{Id} & -(a_{0,k} - a_{1,k})\mathrm{Id} & -a_{1,k}\mathrm{Id}\\ \mathrm{Id} & 0 & 0\\ 0 & \mathrm{Id} & 0\end{bmatrix}\begin{pmatrix}\mathcal{P}_V(x_k - x^\star)\\ \mathcal{P}_V(x_{k-1} - x^\star)\\ \mathcal{P}_V(x_{k-2} - x^\star)\end{pmatrix}.$$

It is straightforward to check that

$$N_k \overset{\text{def}}{=} \begin{bmatrix}(1 + a_{0,k})\mathrm{Id} & -(a_{0,k} - a_{1,k})\mathrm{Id} & -a_{1,k}\mathrm{Id}\\ \mathrm{Id} & 0 & 0\\ 0 & \mathrm{Id} & 0\end{bmatrix}$$

is invertible. Iterating the above argument, and owing to the fact that $x_k, y_{a,k}, y_{b,k} \to x^\star$, we get

$$\begin{pmatrix}0\\ 0\end{pmatrix} = \left(\prod_{j=k}^{\infty} N_j\right)\begin{pmatrix}\mathcal{P}_V(x_k - x^\star)\\ \mathcal{P}_V(x_{k-1} - x^\star)\\ \mathcal{P}_V(x_{k-2} - x^\star)\end{pmatrix},$$

and $\prod_{j=k}^{\infty} N_j$ is invertible. Therefore, we obtain that $x_k - x^\star \in V^\perp$, and in turn, $y_{a,k} - x^\star \in V^\perp$ and $y_{b,k} - x^\star \in V^\perp$, for all large enough k. Observe that $V^\perp \subset T_{x^\star}$, it then follows that

$$x_{k+1} - x^\star = y_{a,k} - x^\star - \mathcal{P}_{V^\perp} E_k \mathcal{P}_{V^\perp}(y_{b,k} - x^\star).$$

By definition, $\mathcal{P}_{V^\perp} E_k \mathcal{P}_{V^\perp}$ is symmetric positive definite. Thus, replace H_k by $\mathcal{P}_{V^\perp} E_k \mathcal{P}_{V^\perp}$ in G and M_{FB} accordingly, and apply Lemma 6.3.4, Theorem 6.3.11 and Theorem 6.3.6 leads to the result. □

Proof of Theorem 6.3.9. Owing to the definition of M_{FB}, we have

$$\begin{bmatrix}M_0 & M_1 & \cdots & M_{s-1} & M_s\\ \mathrm{Id} & 0 & \cdots & 0 & 0\\ 0 & \mathrm{Id} & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & \mathrm{Id} & 0\end{bmatrix}\begin{pmatrix}v_0\\ v_1\\ v_2\\ \vdots\\ v_s\end{pmatrix} = \begin{pmatrix}\sum_{i=0}^s M_i v_i\\ v_0\\ v_1\\ \vdots\\ v_{s-1}\end{pmatrix} = \rho\begin{pmatrix}v_0\\ v_1\\ v_2\\ \vdots\\ v_s\end{pmatrix},$$

which implies that

$$v_{s-1} = \rho v_s, v_{s-2} = \rho v_{s-1} = \rho^2 v_s, \cdots, v_0 = \rho v_1 = \rho^s v_s,$$

and

$$\sum_{i=0}^s M_i v_i = \rho v_0 = \rho^{s+1} v_s. \tag{6.7.16}$$

The first claim fo the theorem is proved.

For the matrices $M_0, ..., M_s$, as $Mv_s = \eta v_s$, first we have,

$$\begin{aligned}M_0 v_0 &= \big((a_0 - b_0)\mathrm{Id} + (1 + b_0)G\big)\rho^s v_s = \big((a_0 - b_0) + (1 + b_0)\eta\big)\rho^s v_s,\\ M_s v_s &= \big(-(a_{s-1} - b_{s-1})\mathrm{Id} - b_{s-1}G\big)v_s = \big(-(a_{s-1} - b_{s-1}) - b_{s-1}\eta\big)v_s\end{aligned}$$

Then for $i \in \{1, ..., s-1\}$

$$\begin{aligned}M_i v_i &= \big(-((a_{i-1} - a_i) - (b_{i-1} - b_i))\mathrm{Id} - (b_{i-1} - b_i)G\big)\rho^{s-i} v_s\\ &= \big(-((a_{i-1} - a_i) - (b_{i-1} - b_i)) - (b_{i-1} - b_i)\eta\big)\rho^{s-i} v_s\end{aligned}$$

Therefore, combining the above result with (6.7.16) leads to

$$\begin{aligned}\rho^{s+1}v_s &= \big((a_0-b_0)+(1+b_0)\eta\big)\rho^s v_s + \big(-(a_{s-1}-b_{s-1})-b_{s-1}\eta\big)v_s \\ &\quad + \textstyle\sum_{i=1}^{s-1}\big(-((a_{i-1}-a_i)-(b_{i-1}-b_i))-(b_{i-1}-b_i)\eta\big)\rho^{s-i}v_s \\ &= (a_0-b_0)\rho^s v_s + (1+b_0)\eta\rho^s v_s - (a_{s-1}-b_{s-1})v_s - b_{s-1}\eta v_s \\ &\quad + \textstyle\sum_{i=1}^{s-1}\big(-((a_{i-1}-a_i)-(b_{i-1}-b_i))\rho^{s-i}v_s - (b_{i-1}-b_i)\eta\rho^{s-i}v_s\big) \\ &= \Big((a_0-b_0)\rho^s - (a_{s-1}-b_{s-1}) - \textstyle\sum_{i=1}^{s-1}\big((a_{i-1}-a_i)-(b_{i-1}-b_i)\big)\rho^{s-i}\Big)v_s \\ &\quad + \Big((1+b_0)\rho^s - b_{s-1} - \textstyle\sum_{i=1}^{s-1}(b_{i-1}-b_i)\rho^{s-i}\Big)\eta v_s,\end{aligned}$$

which leads to

$$\eta v_s = \frac{\rho^{s+1} - (a_0-b_0)\rho^s + \sum_{i=1}^{s-1}\big((a_{i-1}-a_i)-(b_{i-1}-b_i)\big)\rho^{s-i} + (a_{s-1}-b_{s-1})}{(1+b_0)\rho^s - \sum_{i=1}^{s-1}(b_{i-1}-b_i)\rho^{s-i} - b_{s-1}}v_s,$$

and simplifies to (6.3.8). □

We prove the first claim of Theorem 6.3.11 through the following lemma.

Lemma 6.7.3. *From the relation (6.3.12). Given $a, b \in \mathbb{R}$, we have $\rho(M_{\text{FB}}) < 1$ if any of the following conditions are satisfied by a, b*

$$\begin{aligned} a &\in \big](1-\bar\eta)b - \tfrac{1+\bar\eta}{2}, -b\underline{\eta} + (b+1)\big[: b \in \big]\tfrac{-(3+\bar\eta)}{2(\bar\eta-\underline{\eta})}, -\tfrac{1}{2}\big] \\ a &\in \big](1-\underline{\eta})b - \tfrac{1+\underline{\eta}}{2}, (1-\underline{\eta})b + 1\big[: b \in \big[-\tfrac{1}{2}, 0\big] \\ a &\in \big](1-\underline{\eta})b - \tfrac{1+\underline{\eta}}{2}, (1-\bar\eta)b + 1\big[: b \in \big[0, \tfrac{3+\underline{\eta}}{2(\bar\eta-\underline{\eta})}\big[.\end{aligned}$$

In particular, if we restrict $a, b \in [0,1]$, then $\rho(M_{\text{FB}}) < 1$ as long as

$$\frac{2(b-a)-1}{1+2b} < \underline{\eta}.$$

Proof. For the quadratic equation (6.3.12) of ρ, given η, a and b, the two roots are

$$\rho_1 = \frac{((a-b)+(1+b)\eta) + \sqrt{\Delta_\rho}}{2}, \quad \rho_2 = \frac{((a-b)+(1+b)\eta) - \sqrt{\Delta_\rho}}{2}. \tag{6.7.17}$$

where Δ_ρ is the discriminant

$$\Delta_\rho = \big((a-b)+(1+b)\eta\big)^2 - 4\big((a-b)+b\eta\big),$$

which is a quadratic function of 3 variables. Moreover, we have

$$\begin{aligned}\Delta_\rho &= \big((a-b)+(1+b)\eta\big)^2 - 4\big((a-b)+b\eta\big) \\ &= \big(a - \big((1-\eta)b + (1-\sqrt{1-\eta})^2\big)\big)\big(a - \big((1-\eta)b + (1+\sqrt{1-\eta})^2\big)\big),\end{aligned}$$

Which means that $\Delta_\rho \le 0$ for

$$a \in \big[(1-\eta)b + (1-\sqrt{1-\eta})^2, (1-\eta)b + (1+\sqrt{1-\eta})^2\big],$$

and $(1-\eta)b + (1+\sqrt{1-\eta})^2 > 1$. Hence we prove the first claim of Lemma 6.7.3.
Now consider the following 3 linear functions of a

$$\begin{aligned} f_1 &= (1-\eta)b - \eta, \\ f_2 &= (1-\eta)b + (1-\sqrt{1-\eta})^2 \begin{cases} \Delta_\rho \le 0 : f_2 \le a \le (1-\eta)b + (1+\sqrt{1-\eta})^2, \\ \Delta_\rho \ge 0 : a \le f_2, \end{cases} \\ f_3 &= (1-\eta)b - \frac{1+\eta}{2} \begin{cases} |\rho| \le 1 : f_3 \le a, \\ |\rho| \ge 1 : a \le f_3. \end{cases}\end{aligned} \tag{6.7.18}$$

Let $\eta \in]-1,1[$, then

$$\begin{cases} \mathcal{f}_1 \geq \mathcal{f}_2 : \eta \in]-1,0], \\ \mathcal{f}_1 \leq \mathcal{f}_2 : \eta \in [0,1[, \end{cases}$$

and $\mathcal{f}_3$ is smaller than both $\mathcal{f}_1, \mathcal{f}_2$ independently of η.

For the rest of the proof, we discuss in case on the conditions of a, b, η such that $|\rho| \leq 1$.

Case $\eta \in]-1,0]$**:** We have $\mathcal{f}_1 \geq \mathcal{f}_2$,

- **Subcase** $a \in [\mathcal{f}_2, (1-\eta)b + (1+\sqrt{1-\eta})^2]$**:** $\rho_{1,2}$ are complex, hence

$$\begin{aligned} |\rho|^2 &= \frac{((a-b)+(1+b)\eta)^2 - \big(((a-b)+(1+b)\eta\big)^2 - 4\big((a-b)+b\eta\big)\big)}{4} \\ &= a - b + b\eta \in \big[(1-\sqrt{1-\eta})^2, (1+\sqrt{1-\eta})^2\big]. \end{aligned} \tag{6.7.19}$$

  Moreover we have,

$$|\rho|^2 \in \big[(1-\sqrt{1-\eta})^2, 1\big] \quad \text{for} \quad a \in [\mathcal{f}_2, 1 + b(1-\eta)]. \tag{6.7.20}$$

- **Subcase** $a \in]-\infty, \mathcal{f}_2]$**:** $\Delta_\rho \geq 0$ and ρ_2 has the bigger absolute value, then

$$\begin{aligned} |\rho_2| < 1 &\iff -\big((a-b)+(1+b)\eta\big) + \sqrt{\Delta_\rho} < 2 \\ &\iff \Delta_\rho < 4 + 4\big((a-b)+(1+b)\eta\big) + \big((a-b)+(1+b)\eta\big)^2 \\ &\iff 0 < 4 + 4\big((a-b)+(1+b)\eta\big) + 4\big((a-b)+b\eta\big) \\ &\iff (1-\eta)b - \frac{1+\eta}{2} < a, \end{aligned} \tag{6.7.21}$$

  which means that $|\rho_2| \leq 1$ for $a \in [\mathcal{f}_3, \mathcal{f}_2]$, and $|\rho_2| \geq 1$ for $a \in]-\infty, \mathcal{f}_3]$.

Case $\eta \in [0,1[$**:** First we have $\mathcal{f}_2 \geq \mathcal{f}_1$, and

- **Subcase** $a \in [\mathcal{f}_2, (1-\eta)b + (1+\sqrt{1-\eta})^2]$**:** $\Delta_\rho \leq 0$ and same situation as (6.7.19), and we get (6.7.20).
- **Subcase** $a \in [\mathcal{f}_1, \mathcal{f}_2]$**:** now $\Delta_\rho \geq 0$, and $(a-b)+(1+b)\eta \geq 0$ as $a \geq \mathcal{f}_2$, therefore we have $\rho_1 \geq |\rho_2|$ and

$$\begin{aligned} \rho_1 < 1 &\iff \big((a-b)+(1+b)\eta\big) + \sqrt{\Delta_\rho} < 2 \\ &\iff \Delta_\rho < 4 - 4\big((a-b)+(1+b)\eta\big) + \big((a-b)+(1+b)\eta\big)^2 \\ &\iff 0 < 4 - 4\big((a-b)+(1+b)\eta\big) + 4\big((a-b)+b\eta\big) \\ &\iff 0 < 4(1-\eta), \end{aligned} \tag{6.7.22}$$

  which holds naturally.

- **Subcase** $a \in]-\infty, \mathcal{f}_1]$**:** We have $|\rho_2| \geq |\rho_1|$, hence (6.7.21) applies and the result follows.

Summarize the above discussions, we obtain that given an $\eta \in]-1,1[$ and $b \in \mathbb{R}$, the modulus $|\rho| < 1$ for

$$a \in \mathcal{I}_\eta \overset{\text{def}}{=} \big](1-\eta)b - \frac{1+\eta}{2}, (1-\eta)b + 1\big[. \tag{6.7.23}$$

Now let's turn to the spectral radius of M_{FB}, since M has n eigenvalues, for each $i \in \{1, ..., n\}$ define $\mathcal{I}_{\eta_i}$ be the interval as in (6.7.23) for the i's eigenvalue η_i of M. Then to let $\rho(M_{\text{FB}}) < 1$, the following condition must be satisfied

$$\mathcal{I} \overset{\text{def}}{=} \bigcap_{i \in \{1,\dots,n\}} \mathcal{I}_{\eta_i} \neq \emptyset.$$

Let's look at first the smallest and biggest eigenvalues $\underline{\eta}$ and $\bar{\eta}$, for which we have

$$\mathcal{I}_{\underline{\eta}} = \big](1-\underline{\eta})b - \frac{1+\underline{\eta}}{2}, (1-\underline{\eta})b + 1\big[\quad \text{and} \quad \mathcal{I}_{\bar{\eta}} = \big](1-\bar{\eta})b - \frac{1+\bar{\eta}}{2}, (1-\bar{\eta})b + 1\big[.$$

If $\underline{\eta} = \bar{\eta}$, then there's not need to discuss. For $\underline{\eta} < \bar{\eta}$, $\mathcal{I}_{\underline{\eta}} \cap \mathcal{I}_{\bar{\eta}} \neq \emptyset$ yields the following restriction on b,

$$b \in \Big[\frac{-(3+\bar{\eta})}{2(\bar{\eta}-\underline{\eta})}, \frac{3+\underline{\eta}}{2(\bar{\eta}-\underline{\eta})}\Big] \Longleftarrow \begin{cases} (1-\underline{\eta})b - \frac{1+\underline{\eta}}{2} \leq (1-\bar{\eta})b + 1, \\ (1-\bar{\eta})b - \frac{1+\bar{\eta}}{2} \leq (1-\underline{\eta})b + 1. \end{cases}$$

Define the following two functions of η,

$$l(\eta) = -\big(b+\tfrac{1}{2}\big)\eta + \big(b-\tfrac{1}{2}\big) \quad \text{and} \quad r(\eta) = -b\eta + (b+1).$$

We have the following cases.

- **Case** $b \in]\frac{-(3+\bar{\eta})}{2(\bar{\eta}-\underline{\eta})}, -\frac{1}{2}]$**:** For this choice of b, we have that both functions $l(\eta), r(\eta)$ are *increasing* over $[-1,1]$, therefore we get that
$$\mathcal{I} = \big]l(\bar{\eta}), r(\underline{\eta})\big[= \big] -\big(b+\tfrac{1}{2}\big)\bar{\eta} + \big(b-\tfrac{1}{2}\big), -b\underline{\eta} + (b+1)\big[.$$
- **Case** $b \in [-\frac{1}{2}, 0]$**:** $l(\eta)$ is now a *decreasing* function of η while $r(\eta)$ still is *increasing*, this means that
$$\mathcal{I} = \mathcal{I}_{\underline{\eta}} = \big](1-\underline{\eta})b - \frac{1+\underline{\eta}}{2}, (1-\underline{\eta})b + 1\big[.$$
- **Case** $b \in [0, \frac{3+\underline{\eta}}{2(\bar{\eta}-\underline{\eta})}[$**:**Now both functions $l(\eta), r(\eta)$ are *decreasing* over $[-1,1]$, and we have
$$\mathcal{I} = \big]l(\underline{\eta}), r(\bar{\eta})\big[= \big](1-\underline{\eta})b - \frac{1+\underline{\eta}}{2}, (1-\bar{\eta})b + 1\big[.$$

Summarize these cases and we conclude the proof. □

See from page 142 for detailed discussion on the relation between $\rho(M_{\text{FB}})$ and $\underline{\eta}, \bar{\eta}$.

Proof of Theorem 6.3.11. Recall in the proof of Lemma 6.7.3, the discriminant Δ_ρ of function (6.3.12) is a quadratic function of a, *i.e.*

$$\begin{aligned} \Delta_\rho &= \big((a-b) + (1+b)\eta\big)^2 - 4\big((a-b) + b\eta\big) \\ &= a^2 + 2\big(\eta + b(\eta-1) - 2\big)a + \big(\eta + b(\eta-1)\big)^2 - 4b(\eta-1). \end{aligned}$$

Fix b, η and let $\Delta_\rho = 0$, we get two *real* roots for a, which are

$$r_{a,1} = (1-\eta)b + \big(1-\sqrt{1-\eta}\big)^2 \begin{cases} \leq 1 : b \in \big[0, \frac{1-(1-\sqrt{1-\eta})^2}{1-\eta}\big] \\ \geq 1 : b \in \big[\frac{1-(1-\sqrt{1-\eta})^2}{1-\eta}, 1\big] \end{cases}$$
$$r_{a,2} = (1-\eta)b + \big(1+\sqrt{1-\eta}\big)^2 > 1.$$

Moreover, we have

$$\frac{1-(1-\sqrt{1-\eta})^2}{1-\eta} \begin{cases} \leq 1 : \eta \leq 0 \\ \geq 1 : \eta \geq 0. \end{cases}$$

Case $a \leq r_{a,1}$**:** We have $\Delta_\rho \geq 0$, hence (6.3.12) admits two *real* roots of ρ. Depends on the sign of $(a-b) + (1+b)\eta$, we have the following subcases.

- **Subcase** $(a-b) + (1+b)\eta \leq 0$**:** This means that $\eta \leq \frac{b-a}{1+b}$, from (6.7.17) we have
$$|\rho| = \frac{-((a-b) + (1+b)\eta) + \sqrt{\Delta_\rho}}{2}.$$
Since
$$\frac{(b-a)(1+b)}{(1+b)^2} \leq \frac{(b-a)(1+b) + 2b}{(1+b)^2}$$
we have $|\rho|$ is *monotone decreasing* for $\eta \leq]-1, \frac{b-a}{1+b}]$.

- **Subcase** $(a-b)+(1+b)\eta \geq 0$**:** For this case, we have

$$|\rho| = \frac{((a-b)+(1+b)\eta)+\sqrt{\Delta_\rho}}{2},$$

and it is easy to see that $|\rho|$ is *monotone increasing* for $\eta \leq]\frac{b-a}{1+b}, 1[$.

The above discussions indicate that when $a \leq r_{a,1}$, the value of $|\rho|$ *decreases* as a *increases*.

Case $a \in [r_{a,1}, r_{a,2}]$**:** For this cases, we have $\Delta_\rho \leq 0$, and (6.3.12) admits two *complex* roots of ρ, then owing to (6.7.19), we have that

$$|\rho| = \sqrt{a-b+b\eta},$$

which is an *monotone increasing* function of a.

To sum up, we have that with given η and chosen $b \in [0,1]$, when a is chosen such that

$$a = (1-\eta)b + (1-\sqrt{1-\eta})^2, \tag{6.7.24}$$

i.e. $\Delta_\rho = 0$. Then $|\rho|$ attains the minimal value, which is

$$|\rho| = |1-\sqrt{1-\eta}| < |\eta|.$$

Figure 6.9 shows the value of $|\rho|$ as the function of a, b. The *yellow dashed line* corresponds to the choices of (a,b) such that $|\rho|$ attains its minimal value $\rho^\star$. It can be verified that given b, the choice of a such that $|\rho| \leq \eta$ is $[(1-\eta)b, (1-\eta)b+\eta^2]$.

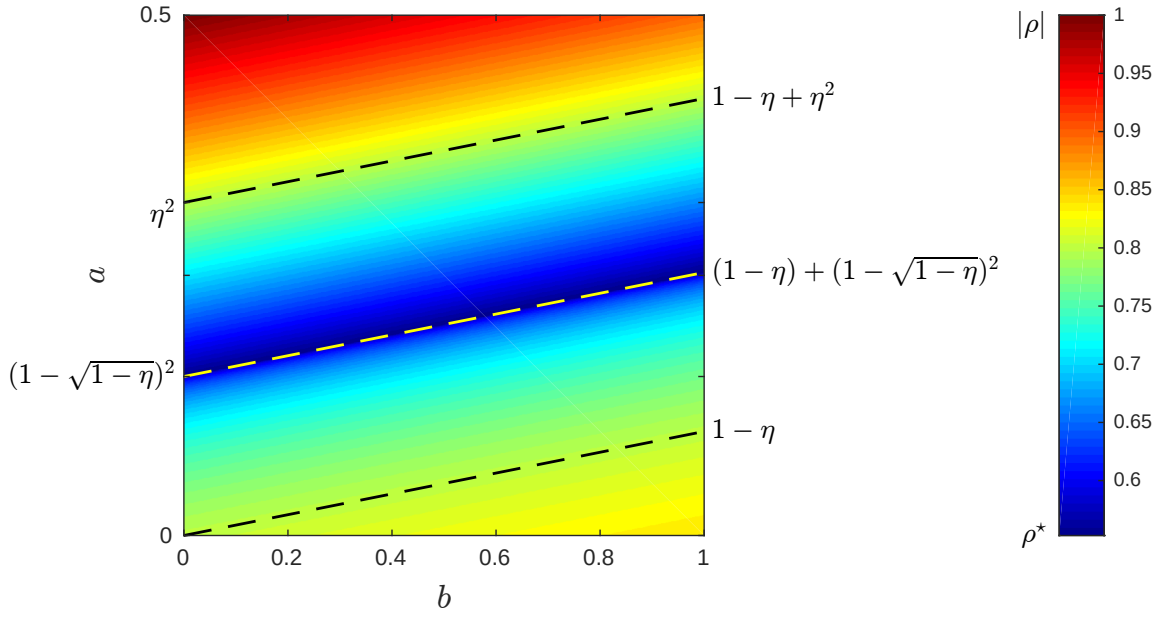


Figure 6.9: Given $\eta \in]-1,1[$, $|\rho|$ is a function of (a,b). The above figure is generated with $\eta = 0.8$.

For the spectral radius of M_{FB}, we need the assistance of Figure 6.12 and the value $\mathcal{g}_1$ defined in (6.7.30). For Figure 6.12, the dark gray areas means that $\rho(M_{\text{FB}})$ is determined by $\bar{\eta}$, while the white area means $\rho(M_{\text{FB}})$ is controlled by $\underline{\eta}$, and the blue line means $\underline{\eta}, \bar{\eta}$ contribute to $\rho(M_{\text{FB}})$ equally. The value $\mathcal{g}_1$ stands for the crossing point of the magenta line ($\mathcal{f}_{2,\bar{\eta}} = (1-\bar{\eta})b + (1-\sqrt{1-\bar{\eta}})^2$) and the blue line, and we have $\mathcal{g}_1 \geq 0$ if (6.7.25) holds, *i.e.*

$$\underline{\eta} \geq -\big(3-\sqrt{1-\bar{\eta}}\big)\big(1-\sqrt{1-\bar{\eta}}\big). \tag{6.7.25}$$

Therefore, apply the result of (6.7.24) we get the condition

$$\begin{aligned} b &\in \left[0, \frac{-(1-\sqrt{1-\bar{\eta}})^2((1-\sqrt{1-\bar{\eta}})^2+\underline{\eta}-2)-(1-\sqrt{1-\bar{\eta}})|(1-\sqrt{1-\bar{\eta}})^2-\underline{\eta}|}{(1-(1-\sqrt{1-\bar{\eta}})^2)(\bar{\eta}-\underline{\eta})}\right] \\ a &= (1-\bar{\eta})b + \big(1-\sqrt{1-\bar{\eta}}\big)^2. \end{aligned} \tag{6.7.26}$$

of a, b such that $\rho(M_{\text{FB}})$ is minimised. Applying Eq. (6.3.11) yields the second claim of Theorem 6.3.11.

Given $\gamma > 0$, let η be an eigenvalue of M, then we have

$$\eta \in [1 - \gamma/\beta, 1 - \gamma\kappa].$$

The spectral radius of M then is

$$\rho(M) = \max\left\{|1 - \gamma/\beta|, |1 - \gamma\kappa|\right\} \begin{cases} 1 - \gamma\kappa : \gamma \in \left[0, \frac{2}{1/\beta + \kappa}\right], \\ \gamma/\beta - 1 : \gamma \in \left[\frac{2}{1/\beta + \kappa}, \frac{2}{\kappa}\right]. \end{cases}$$

which is a piecewise linear function of γ, and minimizes at $\gamma = \frac{2}{1/\beta+\kappa}$.

Adapt condition (6.7.25) to this case we have that if

$$1 - \gamma/\beta \geq -(3 - \sqrt{\gamma\kappa})(1 - \sqrt{\gamma\kappa}) \iff \gamma \in \left]0, \frac{4}{(\sqrt{1/\beta} + \sqrt{\kappa})^2}\right],$$

then $\rho(M_{\text{FB}})$ is determined by $\bar{\eta}$ only. Since $\bar{\eta} = 1 - \gamma\kappa$ is *monotone decreasing* for γ, hence it is straight forward that $\bar{\eta}$ minimises at $\gamma^\star = \frac{4}{(\sqrt{1/\beta}+\sqrt{\kappa})^2}$. Then apply (6.7.26) leads to the optimal choices of a, b, and we conclude the proof. □

Proof of Corollary 6.3.13. The first claim is straightforward from Theorem 6.3.11. Then for then second one. First, in order to make $\rho(M_{\text{FB}})$ be determined by $\bar{\eta}$ for all $b \leq a, a \in [0, 1]$. Then from Figure 6.12, the most left position of $\mathcal{J}_1$ is that it is the crossing point of blue & magenta lines and the diagonal line $a = b$. From (6.7.26) and (6.7.30) we get

$$\frac{(1 - \sqrt{1 - \bar{\eta}})^2}{\bar{\eta}} \leq \frac{-(1 - \sqrt{1 - \bar{\eta}})^2((1 - \sqrt{1 - \bar{\eta}})^2 + \underline{\eta} - 2) - (1 - \sqrt{1 - \bar{\eta}})((1 - \sqrt{1 - \bar{\eta}})^2 - \underline{\eta})}{(1 - (1 - \sqrt{1 - \bar{\eta}})^2)(\bar{\eta} - \underline{\eta})}$$

which leads to $\underline{\eta} \geq -\bar{\eta}/3$. Applying the result of (ii) of Theorem 6.3.11 leads to the claimed result. □

Numerical proof of Conjecture 6.3.15. For $s = 2$ and $b_i = a_i, i = 0, 1$, we have from (6.3.8) that

$$0 = \rho^3 - (1 + a_0)\eta\rho^2 + (a_0 - a_1)\eta\rho + a_1\eta.$$

We have seen from the proof of 6.3.11, that for $s = 1$, given $\eta \in]-1, 1[$, the optimal choice of inertial parameters, given by $a = (1-\eta)b + (1 - \sqrt{1-\eta})^2$ (6.3.14), makes the discriminant Δ_ρ of the quadratic function (6.3.12) equal to 0.

For $s = 2$, it is reasonable to imagine that if such an optimal choice of the inertial parameters exists, it should also make the corresponding discriminant equal to 0. After computing the discriminant Δ_ρ of this cubic equation and let it equal to 0, we get

$$\begin{aligned} 0 = {} & 4\eta^3 a_1^3 + \eta^2\big(18\eta + 6\eta a_0 + \eta^2(1 + a_0)^2 - 27\big)a_1^2 \\ & + \eta^3\big(-18a_0 - 6a_0^2 + 2\eta(1 + a_0)^2(2 + a_0)\big)a_1 + \eta^3 a_0^2\big(\eta(1 + a_0)^2 - 4a_0\big), \end{aligned} \tag{6.7.27}$$

which is another cubic equation of a_1. Compute the discriminant Δ_{a_1} of (6.7.27), we obtain the following equation

$$\begin{aligned} \Delta_{a_1} = {} & 16\eta^{15}(\eta - 1)a_0^9 + 144\eta^{14}(\eta - 1)(\eta - 3)a_0^8 + 576\eta^{13}(-9 + 15\eta - 7\eta^2 + \eta^3)a_0^7 \\ & + \eta^{12}(34992 - 69984\eta + 47088\eta^2 - 13440\eta^3 + 1344\eta^4)a_0^6 \\ & + \eta^{11}(-139968 + 326592\eta - 287712\eta^2 + 125280\eta^3 - 26208\eta^4 + 2016\eta^5)a_0^5 \\ & + \eta^{10}(314928 - 839808\eta^1 + 933120\eta^2 - 570240\eta^3 + 192240\eta^4 - 32256\eta^5 + 2016\eta^6)a_0^4 \\ & + \eta^9(-314928 + 944784\eta^1 - 1364688\eta^2 + 1201392\eta^3 - 622080\eta^4 + 179712\eta^5 - 25536\eta^6 + 1344\eta^7)a_0^3 \\ & + \eta^{10}(314928 - 769824\eta^1 + 746496\eta^2 - 381024\eta^3 + 101520\eta^4 - 12672\eta^5 + 576\eta^6)a_0^2 \\ & + \eta^{11}(-104976 + 198288\eta^1 - 121824\eta^2 + 31968\eta^3 - 3600\eta^4 + 144\eta^5)a_0 \\ & + \eta^{12}(11664 - 15552\eta + 4320\eta^2 - 448\eta^3 + 16\eta^4), \end{aligned} \tag{6.7.28}$$

which is a 9th order equation of a_0, and can be solved only numerically. It turns out from this equation that Δ_{a_1} admits real root(s) for $\eta \in]-1, 1[$, which then contains the optimal value of a_0 that we are looking for. Putting back such a_0 into (6.7.27), we can obtain "the" optimal value of a_1.

Given fixed η, the optimal value of a_0 is denoted as $a_0^\star$, then for each $a_0 \in [0, a_0^\star]$, the optimal value of a_1 obtained by (6.7.27) is plotted in Figure 6.10 (a)[1]. Figure 6.10 (b) shows the value of ρ under each pair of (a_0, a_1) in Figure 6.10 (a). In comparison, $s = 1$ is added.

(i) For the left figure,

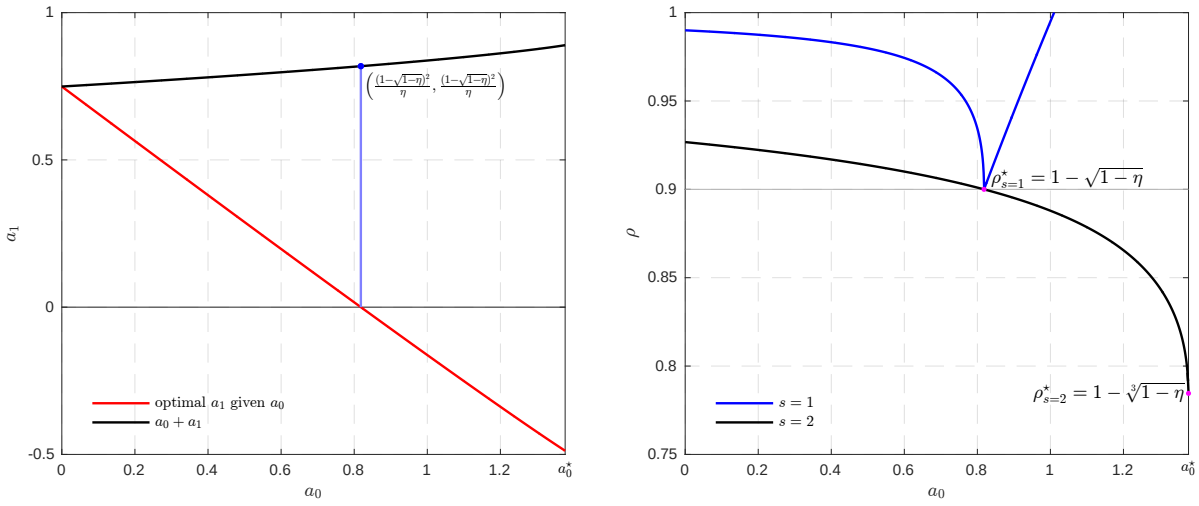
(a) The value $a^\star = (1-\sqrt{1-\eta})^2/\eta$ is the optimal choice for $s = 1$. The red line implies that, for the case $s = 2$, a_1 should be positive when $a_0 \le a^\star$, and negative when $a_0 \ge a^\star$. Moreover, $a_0 + a_1 \le a^\star$ when $a_0 \le a^\star$, and the contrary when $a_0 \ge a^\star$.

(b) For $s = 2$, the largest value that can be attained by a_0 is greater than 1 (this also depends on the value of η). Since a_1 is negative for such choice of a_0, we still have $\sum_i a_i < 1$, see the black line.

(ii) For the right figure,

(a) Given a_0, the optimal value of ρ obtained by $s = 2$ is smaller than that of $s = 1$, except at the point $a_0 = (1 - \sqrt{1-\eta})^2/\eta$ for which $a_1 = 0$ for $s = 2$.

(b) When $s = 1$, the optimal value of ρ can be obtained is $\rho^\star_{s=1} = 1 - \sqrt{1-\eta}$. While for $s = 2$, the optimal value can be obtained is $\rho^\star_{s=2} = 1 - \sqrt[3]{1-\eta}$.



(a) Fixed η, optimal a_1 for given a_0

(b) ρ under given optimal (a_0, a_1)

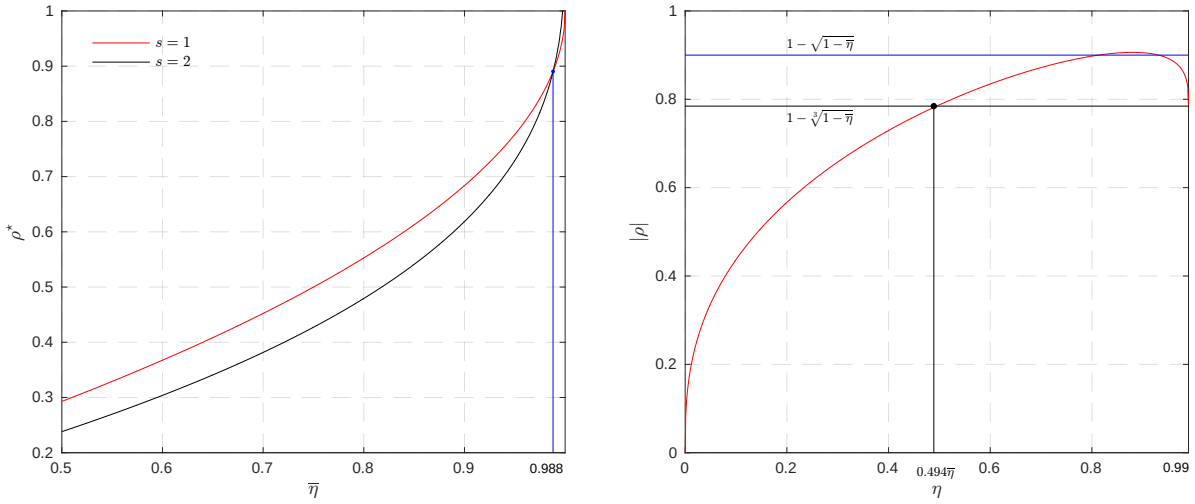
Figure 6.10: Under fixed $\eta = 0.99$, the optimal choice of a_1 for given a_0 and the value ρ under the optimal pair (a_0, a_1).

Let $\bar\eta \in]0, 1[$, and suppose that the eigenvalues of G distributes uniformly in $[0, \bar\eta]$, the value of $\rho^\star_{s=2}$ is then a function of $\bar\eta$. Such a relation is shown in Figure 6.11 (a), and $\rho^\star_{s=1}$ is added in comparison. As we can see, when $\bar\eta < 0.988$, we have $\rho^\star_{s=2} \le \rho^\star_{s=1}$, which means that 2-MiFB is faster than 1-MiFB.

Fix $\bar\eta = 0.99$ as in Figure 6.10, and compute the optimal values of $a_0^\star, a_1^\star$ according to (6.7.28) and (6.7.27) respectively. Assume that the eigenvalues of G distributes uniformly in $[0, \bar\eta]$. Then for each $\eta \in [0, \bar\eta]$, we compute the value of $|\rho|$ (the maximum of the 3 roots) from the following equation

$$0 = \rho^3 - (1 + a_0^\star)\eta\rho^2 + (a_0^\star - a_1^\star)\eta\rho + a_1^\star\eta.$$

[1]With given η, (6.3.20) is a cubic equation of a_1. For the displayed interval of choice of a_0 shown in Figure 6.10, (6.3.20) admits 3 real roots for a_1, for the other two roots which are not plotted, one make the value of $|\rho|$ strictly bigger than 1, the other one is negative and and makes $|\rho|$ bigger than the one obtained by the plotted one.



(a) Fixed η, optimal a_1 for given a_0

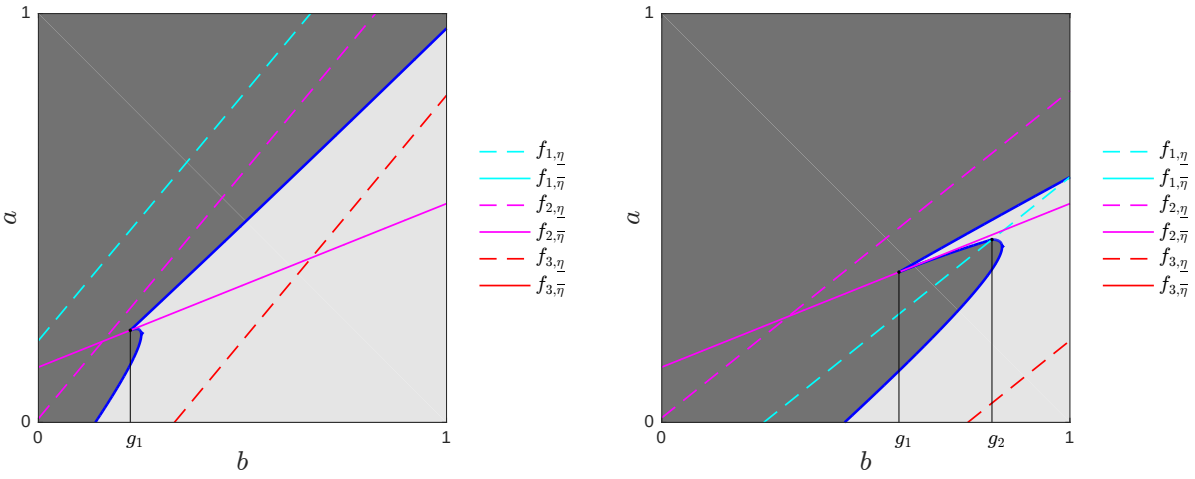
(b) $|\rho|$ under given η and $(a_0^\star, a_1^\star)$

Figure 6.11: Under fixed $\eta = 0.99$, the optimal choice of a_1 for given a_0 and the value ρ under the optimal pair (a_0, a_1).

Figure 6.11 (b) shows graphically the value of $|\rho|$ as a function of $\eta \in [0, \bar{\eta}]$ under given $(a_0^\star, a_1^\star)$. As we can see, $\rho(M_{\text{FB}})$ is determined by some $\eta \in [0.494\bar{\eta}, \bar{\eta}]$. Therefore, in order to let $\rho(M_{\text{FB}})$ determined by $\bar{\eta}$, the spectrum of G should be $[0, 0.494\bar{\eta}] \cup \{\bar{\eta}\}$, which indicates that $1 - \beta\kappa' < 0.494(1 - \beta\kappa)$. □

6.7.3 Spectral radius of M_{FB}

Given a, b, the value of $\rho(M_{\text{FB}})$ depends only on $\underline{\eta}$ and $\bar{\eta}$, Figure 6.12 shows the dependency of $\rho(M_{\text{FB}})$ on $\underline{\eta}, \bar{\eta}$, where the *white* areas of the figures means that $\rho(M_{\text{FB}})$ is determined by $\underline{\eta}$ while the *gray* areas correspond to $\bar{\eta}$, and the solid blue lines means that $|\rho|_{\underline{\eta}} = |\rho|_{\bar{\eta}}$. It also can be observed that, the regions that $\rho(M_{\text{FB}})$ is determined by $\underline{\eta}$ is under the function $f_{2,\underline{\eta}}$, see the *magenta* dashed lines in both figures.



(a) $\underline{\eta} = -0.1, \bar{\eta} = 0.6$

(b) $\underline{\eta} = 0.2, \bar{\eta} = 0.6$

Figure 6.12: Illustration of f_1, f_2 and f_3 for 2 different choice of $\underline{\eta}, \bar{\eta}$. The lines that are invisible (*e.g.* cyan and red solide lines in subfigure (a) and (b)) are all below the horizontal axis.

Given $\underline{\eta}, \bar{\eta}$, define the functions as in (6.7.22),

$$\begin{cases} f_{1,\underline{\eta}} = (1-\underline{\eta})b - \underline{\eta}, \\ f_{2,\underline{\eta}} = (1-\underline{\eta})b + (1-\sqrt{1-\underline{\eta}})^2, \\ f_{3,\underline{\eta}} = (1-\underline{\eta})b - \dfrac{1+\underline{\eta}}{2}, \end{cases} \qquad \begin{cases} f_{1,\bar{\eta}} = (1-\bar{\eta})b - \bar{\eta}, \\ f_{2,\bar{\eta}} = (1-\bar{\eta})b + (1-\sqrt{1-\bar{\eta}})^2, \\ f_{3,\bar{\eta}} = (1-\bar{\eta})b - \dfrac{1+\bar{\eta}}{2}. \end{cases} \tag{6.7.29}$$

For $b \in [0,1]$, it is easy to show that $f_{1,\underline{\eta}} \geq f_{1,\bar{\eta}}$ and $f_{3,\underline{\eta}} \geq f_{3,\bar{\eta}}$.

Area above $a_{2,\bar{\eta}}$ and below $a_{2,\underline{\eta}}$ For this situation, we have $\Delta_{\underline{\eta}} \geq 0, \Delta_{\bar{\eta}} \leq 0$. Moreover since $\bar{\eta} > 0$, tha value of a in this area satisfies $a > f_{2,\bar{\eta}} > f_{1,\bar{\eta}}$, and the modulus of the eigenvalues determined by $\underline{\eta}, \bar{\eta}$ are

$$|\rho|_{\underline{\eta}} = \frac{(a-b) + (1+b)\underline{\eta} + \sqrt{\Delta_{\rho_{\underline{\eta}}}}}{2}, \quad |\rho|_{\bar{\eta}} = \sqrt{a - b + b\bar{\eta}}.$$

Let the two values equal,

$$\begin{aligned} |\rho|_{\underline{\eta}} = |\rho|_{\bar{\eta}} &\iff \big((a-b) + (1+b)\underline{\eta}\big) + \sqrt{\Delta_{\rho_{\underline{\eta}}}} = 2\sqrt{a - b + b\bar{\eta}} \\ &\iff \big((a-b) + (1+b)\underline{\eta}\big)^2 = \big(2\sqrt{a - b + b\bar{\eta}} - \sqrt{\Delta_{\rho_{\underline{\eta}}}}\big)^2 \\ &\iff \mathcal{C}_{1,a}(a), \end{aligned}$$

where $\mathcal{C}_{1,a}(a)$ is a cubic function of a and reads

$$\begin{aligned} 0 = a^3 &+ \big(2\big(b(\underline{\eta}-1) + \underline{\eta} - 2\big) + b(\bar{\eta}-1)\big)a^2 \\ &+ \big(\big(b(\underline{\eta}-1) + \underline{\eta}\big)^2 - 4b(\underline{\eta}-1) + 2b(\bar{\eta}-1)\big(b(\underline{\eta}-1) + \underline{\eta} - 2\big)\big)a \\ &+ b(\bar{\eta}-1)\big(\big(b(\underline{\eta}-1) + \underline{\eta}\big)^2 - 4b(\underline{\eta}-1)\big) - (\bar{\eta}-\underline{\eta})^2 b^2. \end{aligned}$$

In Figure 6.12, the crossing point of the solid magenta and blue lines is determined by the following quadratic equation of b. The crossing point means that, for that b, $a = (1-\bar{\eta})b + (1-\sqrt{1-\bar{\eta}})^2$ is a roots of $\mathcal{C}_{1,a}(a)$, which leads to a quadratic function of b,

$$\begin{aligned} 0 = &(1 - (1-\sqrt{1-\bar{\eta}})^2)(\underline{\eta}-\bar{\eta})^2 b^2 \\ &- 2(1-\sqrt{1-\bar{\eta}})^2(\underline{\eta}-\bar{\eta})\big((1-\sqrt{1-\bar{\eta}})^2 + \underline{\eta} - 2\big)b \\ &- (1-\sqrt{1-\bar{\eta}})^2\big(((1-\sqrt{1-\bar{\eta}})^2 + \underline{\eta})^2 - 4(1-\sqrt{1-\bar{\eta}})^2\big). \end{aligned}$$

We denote the corresponding value of b as g_1, then

$$g_1 = \frac{-(1-\sqrt{1-\bar{\eta}})^2((1-\sqrt{1-\bar{\eta}})^2 + \underline{\eta} - 2) - (1-\sqrt{1-\bar{\eta}})|(1-\sqrt{1-\bar{\eta}})^2 - \underline{\eta}|}{(1-(1-\sqrt{1-\bar{\eta}})^2)(\bar{\eta}-\underline{\eta})}. \tag{6.7.30}$$

We have, if $\underline{\eta} > 0$

$$(1-\sqrt{1-\bar{\eta}})^2 - \underline{\eta} \begin{cases} \geq 0 : \bar{\eta} \in \big[1 - (1-\sqrt{\underline{\eta}})^2, 1\big], \\ \leq 0 : \bar{\eta} \in \big]\underline{\eta}, 1 - (1-\sqrt{\underline{\eta}})^2\big]. \end{cases}$$

$(1-\sqrt{1-\bar{\eta}})^2 - \underline{\eta} \geq 0$ holds automatically if $\underline{\eta} \leq 0$.

Area below both $a_{2,\bar{\eta}}$ and $a_{2,\underline{\eta}}$, above $f_{1,\bar{\eta}}$ For this situation, we have both $\Delta_{\underline{\eta}}, \Delta_{\bar{\eta}} \geq 0$, and it can be shown that this region is always above $f_{1,\bar{\eta}}$, which gives

$$|\rho|_{\underline{\eta}} = \begin{cases} \dfrac{\big((a-b) + (1+b)\underline{\eta}\big) + \sqrt{\Delta_{\rho_{\underline{\eta}}}}}{2} : a \geq f_{1,\underline{\eta}}, \\ \dfrac{-\big((a-b) + (1+b)\underline{\eta}\big) + \sqrt{\Delta_{\rho_{\underline{\eta}}}}}{2} : a \leq f_{1,\underline{\eta}}, \end{cases}, \quad |\rho|_{\bar{\eta}} = \frac{\big((a-b) + (1+b)\bar{\eta}\big) + \sqrt{\Delta_{\rho_{\bar{\eta}}}}}{2}.$$

Consider the crossing point of $f_{1,\underline{\eta}}$ and $f_{2,\bar{\eta}}$, which is denoted by g_2

$$g_2 = \frac{(1-\sqrt{1-\bar{\eta}})^2 + \underline{\eta}}{\bar{\eta} - \underline{\eta}}.$$

The following analysis will depend on the relation between $\mathcal{g}_1$ and $\mathcal{g}_2$.

Case $\mathcal{g}_2 \le \mathcal{g}_1$: We have $|\rho|_{\underline{\eta}} = \frac{-((a-b)+(1+b)\underline{\eta})+\sqrt{\Delta_{\rho_{\underline{\eta}}}}}{2}$, then

$$\begin{aligned}
|\rho|_{\underline{\eta}} = |\rho|_{\bar{\eta}} &\Longleftrightarrow -\big((a-b)+(1+b)\underline{\eta}\big) + \sqrt{\Delta_{\rho_{\underline{\eta}}}} = \big((a-b)+(1+b)\bar{\eta}\big) + \sqrt{\Delta_{\rho_{\bar{\eta}}}} \\
&\Longleftrightarrow \sqrt{\Delta_{\rho_{\underline{\eta}}}} - \sqrt{\Delta_{\rho_{\bar{\eta}}}} = \big((a-b)+(1+b)\bar{\eta}\big) + \big((a-b)+(1+b)\underline{\eta}\big) \\
&\Longleftrightarrow \mathcal{C}_{2,a}(a),
\end{aligned}$$

where $\mathcal{C}_{2,a}(a)$ is a cubic function of a, which reads

$$\begin{aligned}
0 = {} & 4a^3 + 2\big(2c + 3b(c-2)\big)a^2 \\
& + \big(2\big(\underline{\eta}(\bar{\eta}-1) + \bar{\eta}(\underline{\eta}-1)\big)b + 4(\underline{\eta}-1)(\bar{\eta}-1)b^2 + \big(c + 2b(c-2)\big)\big(c + b(c-2)\big)\big)a \\
& + \big(\big(\underline{\eta}(\bar{\eta}-1) + \bar{\eta}(\underline{\eta}-1)\big)b + 2(\underline{\eta}-1)(\bar{\eta}-1)b^2\big)\big(b(c-2)+c\big) + d^2b^2,
\end{aligned}$$

where $c = \bar{\eta} + \underline{\eta}, d = \bar{\eta} - \underline{\eta}$.

Case $\mathcal{g}_2 \ge \mathcal{g}_1$: For this case, there are 2 subcases to be discussed. Let's first check the crossing point of $\mathcal{f}_{1,\underline{\eta}}$ and the curve that $|\rho|_{\underline{\eta}} \ge |\rho|_{\bar{\eta}}$. For that b, $a = (1-\underline{\eta})b - \underline{\eta}$ is a roots of $\mathcal{C}_{2,a}(a)$, therefore,

$$\begin{aligned}
|\rho|_{\underline{\eta}} = |\rho|_{\bar{\eta}} &\Longleftrightarrow -\big((a-b)+(1+b)\underline{\eta}\big) + \sqrt{\Delta_{\rho_{\underline{\eta}}}} = \big((a-b)+(1+b)\bar{\eta}\big) + \sqrt{\Delta_{\rho_{\bar{\eta}}}} \\
&\Longleftrightarrow \sqrt{\underline{\eta}}(b+1) = b,
\end{aligned}$$

and we denote such b as $\mathcal{g}_3$. Now we have 3 different crossing points

$$\begin{aligned}
&\mathcal{g}_1 : \text{crossing point of } \mathcal{f}_{2,\bar{\eta}} \text{ and root of } \mathcal{C}_{1,a}(a), \\
&\mathcal{g}_2 = \frac{(1-\sqrt{1-\bar{\eta}})^2 + \underline{\eta}}{\bar{\eta} - \underline{\eta}}, \text{crossing point of } \mathcal{f}_{1,\underline{\eta}} \text{ and } \mathcal{f}_{2,\bar{\eta}}, \\
&\mathcal{g}_3 = \frac{\sqrt{\underline{\eta}}}{1-\sqrt{\underline{\eta}}}, \text{crossing point of } \mathcal{f}_{1,\underline{\eta}} \text{ and root of } \mathcal{C}_{2,a}(a),
\end{aligned} \tag{6.7.31}$$

and we have $\mathcal{g}_1 \le \mathcal{g}_3 \le \mathcal{g}_2$.

- **Subcase $a \ge \mathcal{f}_{1,\underline{\eta}}$:** For this case, we have $|\rho|_{\underline{\eta}} = \frac{((a-b)+(1+b)\underline{\eta})+\sqrt{\Delta_{\rho_{\underline{\eta}}}}}{2}$, then

$$\begin{aligned}
|\rho|_{\underline{\eta}} = |\rho|_{\bar{\eta}} &\Longleftrightarrow \big((a-b)+(1+b)\underline{\eta}\big) + \sqrt{\Delta_{\rho_{\underline{\eta}}}} = \big((a-b)+(1+b)\bar{\eta}\big) + \sqrt{\Delta_{\rho_{\bar{\eta}}}} \\
&\Longleftrightarrow \sqrt{\Delta_{\rho_{\underline{\eta}}}} - \sqrt{\Delta_{\rho_{\bar{\eta}}}} = \big((a-b)+(1+b)\bar{\eta}\big) - \big((a-b)+(1+b)\underline{\eta}\big) \\
&\Longleftrightarrow a = \frac{b}{1+b}.
\end{aligned}$$

- **Subcase $a \le \mathcal{f}_{1,\underline{\eta}}$:** The result is the same as **Case $\mathcal{g}_2 \le \mathcal{g}_1$**.

Chapter 7

Douglas–Rachford and ADMM under Partial Smoothness

Main contributions of this chapter

- Global convergence of the non-stationary Douglas–Rachford splitting (Proposition 7.2.1).
- Finite activity identification (Theorem 7.3.1) and local linear convergence (Theorem 7.4.10) of Douglas–Rachford/ADMM.
- Finite termination of Douglas–Rachford splitting (Theorem 7.5.1).

A preliminary version of these results has appeared in [118].

Contents

In the context of optimization, the DR and ADMM are two algorithms designed to minimize the sum of two proper lower semi-continuous convex functions whose proximity operators are easy to compute. The goal of this chapter is to understand the local linear convergence behaviour of DR/ADMM when the involved functions are moreover partly smooth. More precisely, when the two functions are partly smooth relative to their respective smooth submanifolds, we show that DR/ADMM (i) identifies these manifolds in finite time. (ii) enters a local linear convergence regime. When both functions are locally polyhedral, we show that the optimal convergence rate is given in terms of the cosine of the Friedrichs angle (Definition 7.4.2) between the tangent spaces of the identified manifolds. Under polyhedrality of both functions, we also provide sufficient conditions for finite convergence of DR.

7.1 Introduction

7.1.1 Problem statement

Consider the following optimization problem which is the sum of two non-smooth functions

$$\min_{x\in\mathbb{R}^n} R(x) + J(x), \tag{$\mathcal{P}_{\mathrm{DR}}$}$$

where we have

(**D.1**) $R, J \in \Gamma_0(\mathbb{R}^n)$, the proper convex and lower semi-continuous functions.

(**D.2**) $\mathrm{ri}(\mathrm{dom}(R)) \cap \mathrm{ri}(\mathrm{dom}(J)) \neq \emptyset$, *i.e.* the domain qualification condition.

(**D.3**) $\mathrm{Argmin}(R + J) \neq \emptyset$, *i.e.* the set of minimizers is nonempty.

From now on, we suppose that assumptions (**D.1**)-(**D.3**) hold throughout the chapter.

7.1.2 The non-stationary DR

Let us recall the DR iteration (2.4.10) introduced in Section 2.4.3.2. By letting $A_1 = \partial R$ and $A_2 = \partial J$, we obtain the following iteration for solving $(\mathcal{P}_{\mathrm{DR}})$

$$\begin{cases} u_{k+1} = \mathrm{prox}_{\gamma R}(2x_k - z_k), \\ z_{k+1} = (1-\lambda_k)z_k + \lambda_k(z_k + u_{k+1} - x_k), \\ x_{k+1} = \mathrm{prox}_{\gamma J}(z_{k+1}), \end{cases} \tag{7.1.1}$$

where $\gamma \in]0, +\infty[$, $\lambda_k \in]0, 2[$.

The iteration scheme (7.1.1) above is stationary, meaning that parameter γ is fixed and does not change along the iterations. The fixed-point formulation of stationary DR with respect to z_k is

$$\begin{gathered} z_{k+1} = \mathcal{F}_{\gamma,\lambda_k}(z_k), \\ \text{where} \quad \mathcal{F}_{\gamma,\lambda} \overset{\text{def}}{=} (1-\lambda)\mathrm{Id} + \lambda\mathcal{F}_\gamma \quad \text{and} \quad \mathcal{F}_\gamma \overset{\text{def}}{=} \frac{1}{2}(\mathcal{R}_{\gamma R}\mathcal{R}_{\gamma J} + \mathrm{Id}). \end{gathered} \tag{7.1.2}$$

Moreover $\mathrm{fix}(\mathcal{F}_\gamma) = \mathrm{fix}(\mathcal{F}_{\gamma,\lambda}) = \mathrm{fix}(\mathcal{F}_{\gamma,\lambda_k})$. Note that in previous chapters $\mathcal{F}_\gamma$ is also denoted as $\mathcal{F}_{\mathrm{DR}}$, for the sake of the upcoming result, we choose to use $\mathcal{F}_\gamma$ here.

Owing to the result in Section 2.4.3.2, the convergence of stationary DR is guaranteed under the assumptions **(D.1)**-**(D.3)**, and the condition that $\sum_{k\in\mathbb{N}} \lambda_k(2-\lambda_k) = +\infty$. That z_k converges to some fixed point $z^\star \in \mathrm{fix}(\mathcal{F}_\gamma) \neq \emptyset$, and that the shadow points x_k and u_k both converge to $x^\star \overset{\text{def}}{=} \mathrm{prox}_{\gamma J}(z^\star) \in \mathrm{Argmin}(R+J)$.

In this chapter, we consider a non-stationary version of (7.1.1), which is summarized in Algorithm 13.

Algorithm 13: Non-stationary Douglas–Rachford splitting

Initial: $k = 0$, $z_0 \in \mathbb{R}^n, x_0 = \mathrm{prox}_{\gamma_0 J}(z_0)$;
repeat
  Let $\gamma_k \in]0, +\infty[$, $\lambda_k \in]0, 2[$:

$$\begin{aligned} u_{k+1} &= \mathrm{prox}_{\gamma_k R}(2x_k - z_k), \\ z_{k+1} &= (1-\lambda_k)z_k + \lambda_k(z_k + u_{k+1} - x_k), \\ x_{k+1} &= \mathrm{prox}_{\gamma_{k+1} J}(z_{k+1}), \end{aligned} \tag{7.1.3}$$

  $k = k + 1$;
until *convergence*;

Remark 7.1.1. By definition, the DR method is not symmetric with respect to the order of the functions J and R, see [21] for a systematic study of the two possible versions in the exact, stationary and unrelaxed case. Nevertheless, all of our statements throughout hold true, with minor adaptations, when the order of J and R is reversed in (7.1.3). Note also that the standard DR only accounts for the sum of two functions. Extension to more than two functions is straightforward through a product space trick, see Section 7.7 for details.

7.2 Global convergence of non-stationary DR

Recall the operators defined in (7.1.2). The non-stationary DR iteration (7.1.3) can also be written

$$z_{k+1} = \mathcal{F}_{\gamma_k,\lambda_k}(z_k) = \mathcal{F}_{\gamma,\lambda_k}(z_k) + (\mathcal{F}_{\gamma_k,\lambda_k} - \mathcal{F}_{\gamma,\lambda_k})(z_k). \tag{7.2.1}$$

In plain words, the non-stationary iteration (7.1.3) is a perturbed version of the stationary one (7.1.1).

Proposition 7.2.1 (Global convergence). *For the non-stationary DR iteration (7.1.3). Suppose that the following conditions are fulfilled*

(**D.4**) *Assumptions* (**D.1**)-(**D.3**) *hold.*

(**D.5**) $\lambda_k \in [0,2]$ *such that* $\sum_{k\in\mathbb{N}} \lambda_k(2-\lambda_k) = +\infty$.

(**D.6**) $(\gamma_k, \gamma) \in]0, +\infty[^2$ *such that*

$$\{\lambda_k|\gamma_k - \gamma|\}_{k\in\mathbb{N}} \in \ell^1_+.$$

Then $\{z_k\}_{k\in\mathbb{N}}$ *converges to a point* $z^\star \in \operatorname{fix}(\mathcal{F}_\gamma)$ *and* $x^\star = \operatorname{prox}_{\gamma J}(z^\star) \in \operatorname{Argmin}(R+J)$. *Moreover, the shadow sequences* $\{x_k\}_{k\in\mathbb{N}}$ *and* $\{u_k\}_{k\in\mathbb{N}}$ *both converge to* $x^\star$ *if* γ_k *is convergent.*

Proposition 7.2.1 is an extension of Theorem 3.5.2 (convergence of non-stationary Krasnosel'skiĭ-Mann iteration) to the non-stationary DR case. See Section 7.9 from page 163 for the proof.

Remark 7.2.2.

(i) The conclusions of Proposition 7.2.1 remain true if x_k and u_k are computed inexactly with additive errors $\varepsilon_{1,k}$ and $\varepsilon_{2,k}$, provided that $\{\lambda_k\|\varepsilon_{1,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$ and $\{\lambda_k\|\varepsilon_{2,k}\|\}_{k\in\mathbb{N}} \in \ell^1_+$.

(ii) The summability assumption (**D.6**) is weaker than imposing it without λ_k. Indeed, following the discussion in [56, Remark 5.7], take $q \in]0,1]$, and let

$$\lambda_k = 1 - \sqrt{1-1/k} \quad \text{and} \quad |\gamma_k - \gamma| = \frac{1+\sqrt{1-1/k}}{k^q},$$

then it can be verified that

$$|\gamma_k - \gamma| \notin \ell^1_+,\ \lambda_k|\gamma_k - \gamma| = \frac{1}{k^{1+q}} \in \ell^1_+ \quad \text{and} \quad \lambda_k(2-\lambda_k) = \frac{1}{k} \notin \ell^1_+.$$

(iii) The assumptions made on the sequence $\{\gamma_k\}_{k\in\mathbb{N}}$ imply that $\gamma_k \to \gamma$ (see Lemma 7.9.1). In fact, if $\inf_{k\in\mathbb{N}} \lambda_k > 0$, we have $\{|\gamma_k - \gamma|\}_{k\in\mathbb{N}} \in \ell^+_1$, entailing $\gamma_k \to \gamma$, and thus the convergence assumption on γ_k is superfluous.

(iv) It is worth mentioning that the Proposition 7.2.1 holds true for any real Hilbert space setting, where a weak convergence can be obtained (Theorem 3.5.2).

7.3 Finite activity identification

With the above global convergence result at hand, we are now ready to state the finite time activity identification property of the non-stationary DR method.

Let $z^\star \in \operatorname{fix}(\mathcal{F}_\gamma)$ and $x^\star = \operatorname{prox}_{\gamma J}(z^\star) \in \operatorname{Argmin}(R+J)$, then at convergence of the DR iteration (7.1.3), we have the following inclusion holds,

$$x^\star - z^\star \in \gamma\partial R(x^\star) \quad \text{and} \quad z^\star - x^\star \in \gamma\partial J(x^\star),$$

which is equivalent to $z^\star \in x^\star + \gamma(-\partial R(x^\star) \cap \partial J(x^\star))$. Our identification result is built upon this inclusion.

Theorem 7.3.1 (Finite activity identification). *For the DR iteration* (7.1.3), *suppose that* (**D.4**)-(**D.6**) *hold and* γ_k *is convergent, entailing that* $(z_k, x_k, u_k) \to (z^\star, x^\star, x^\star)$, *where* $z^\star \in \operatorname{fix}(\mathcal{F}_\gamma)$ *and* $x^\star \in \operatorname{Argmin}(R+J)$. *Assume also that* $\inf_{k\in\mathbb{N}} \gamma_k \geq \underline{\gamma} > 0$. *If* $R \in \operatorname{PSF}_{x^\star}(\mathcal{M}^R_{x^\star})$ *and* $J \in \operatorname{PSF}_{x^\star}(\mathcal{M}^J_{x^\star})$, *and the non-degeneracy condition*

$$x^\star - z^\star \in \gamma\operatorname{ri}\big(\partial R(x^\star)\big) \quad \textit{and} \quad z^\star - x^\star \in \gamma\operatorname{ri}\big(\partial J(x^\star)\big) \tag{ND$_{\text{DR}}$}$$

holds. Then

(i) *There exists* $K \in \mathbb{N}$ *large enough such that for all* $k \geq K$, $(u_k, x_k) \in \mathcal{M}^R_{x^\star} \times \mathcal{M}^J_{x^\star}$.

(ii) *Moreover,*

(a) *if* $\mathcal{M}^J_{x^\star} = x^\star + T^J_{x^\star}$, *then* $\forall k \geq K$, $T^J_{x_k} = T^J_{x^\star}$.

(b) *If* $\mathcal{M}^R_{x^\star} = x^\star + T^R_{x^\star}$, *then* $\forall k \geq K$, $T^R_{u_k} = T^R_{x^\star}$.

(c) *J is locally polyhedral around $x^\star$, then* $\forall k \geq K$, $x_k \in \mathcal{M}^J_{x^\star} = x^\star + T^J_{x^\star}$, $T^J_{x_k} = T^J_{x^\star}$, $\nabla_{\mathcal{M}^J_{x^\star}} J(x_k) = \nabla_{\mathcal{M}^J_{x^\star}} J(x^\star)$, *and* $\nabla^2_{\mathcal{M}^J_{x^\star}} J(x_k) = 0$.

(d) *R is locally polyhedral around $x^\star$, then* $\forall k \geq K$, $u_k \in \mathcal{M}^R_{x^\star} = x^\star + T^R_{x^\star}$, $T^R_{u_k} = T^R_{x^\star}$, $\nabla_{\mathcal{M}^R_{x^\star}} R(u_k) = \nabla_{\mathcal{M}^R_{x^\star}} R(x^\star)$, *and* $\nabla^2_{\mathcal{M}^R_{x^\star}} R(u_k) = 0$.

See Section 7.9 from page 164 for the proof.

Remark 7.3.2.

(i) The non-degeneracy condition is equivalent to assume

$$z^\star \in x^\star + \gamma \mathrm{ri}\big(-\partial R(x^\star) \cap \partial J(x^\star)\big),$$

which means that $z^\star \in \mathrm{ri}(\mathrm{fix}(\mathcal{F}_\gamma))$.

(ii) The theorem remains true if the condition on γ_k is replaced with $\gamma_k \geq \underline{\gamma} > 0$ and $\lambda_k \geq \underline{\lambda} > 0$, (use (**D.6**) in the proof).

(iii) Similarly to Example 6.2.4, in general, we have no identification guarantees for x_k and u_k if the proximity operators are computed with errors, even if the latter are summable, in which case one can still prove global convergence (see Remark 7.2.2).

Bounds on the finite identification iteration Similarly to the identification properties of the FB-type methods (Proposition 6.2.3), in this part we provide two estimations for the identification number K. Denote $\tau_k \overset{\text{def}}{=} \lambda_k(2-\lambda_k)$.

Proposition 7.3.3. *Suppose that the assumptions of Theorem 7.3.1 hold, that* $\inf_{k\in\mathbb{N}} \tau_k \geq \underline{\tau} > 0$, *and moreover, the iterates are such that* $\partial J(x_k) \subset \mathrm{rbd}(\partial J(x^\star))$ *whenever* $x_k \notin \mathcal{M}^J_{x^\star}$ *and* $\partial R(u_k) \subset \mathrm{rbd}(\partial R(x^\star))$ *whenever* $u_k \notin \mathcal{M}^R_{x^\star}$. *Then, $\mathcal{M}^J_{x^\star}$ and $\mathcal{M}^R_{x^\star}$ will be identified for some k obeying*

$$k \geq \frac{\|z_0 - z^\star\|^2 + O(\sum_{k\in\mathbb{N}} \lambda_k|\gamma_k - \gamma|)}{\underline{\gamma}^2 \underline{\tau} \mathrm{dist}(0, \mathrm{rbd}(\partial J(x^\star) + \partial R(x^\star)))^2}. \tag{7.3.1}$$

See Section 7.9 from page 165 for the proof.

Observe that the assumption on τ_k automatically implies (**D.5**). As one intuitively expects, this upper-bound (7.3.1) increases as ($\mathrm{ND}_{\mathrm{DR}}$) becomes more stringent.

Example 7.3.4 (Indicators of polyhedral sets). We will discuss the case of J, and the same reasoning applies to R. Consider J as the indicator function of a polyhedral set C^J, *i.e.*

$$J(x) = \iota_{C^J}(x), \quad \text{where} \quad C^J = \{x \in \mathbb{R}^n : \langle c^J_i, x\rangle \leq d^J_i, i = 1, \ldots, m\}.$$

Define $I^J_x \overset{\text{def}}{=} \{i : \langle c_i, x\rangle = d_i\}$ the active set at x. The normal cone to C^J at $x \in C^J$ is polyhedral and given by [153, Theorem 6.46]

$$\partial J(x) = N_{C^J}(x) = \mathrm{cone}((c^J_i)_{i\in I^J_x})).$$

It is immediate then to show that J is partly smooth at $x \in C^J$ relative to the affine subspace $\mathcal{M}^J_x = x + T^J_x$, where, $T^J_x = \mathrm{span}((c^J_i)_{i\in I^J_x})^\perp$. Let $\mathcal{F}^J_x$ be the face of C^J containing x. From [152, Theorem 18.8], one can deduce that

$$\mathcal{F}^J_x = \mathcal{M}^J_x \cap C^J. \tag{7.3.2}$$

We then have

$$\mathcal{M}^J_{x^\star} \subsetneq \mathcal{M}^J_x \underset{(7.3.2)}{\iff} \mathcal{F}^J_{x^\star} \subsetneq \mathcal{F}^J_x \tag{7.3.3}$$

$$\underset{[100,\ \text{Proposition 3.4}]}{\Longrightarrow} N_{C^J}(x) \text{ is a face of (other than) } N_{C^J}(x^\star) \tag{7.3.4}$$

$$\underset{[152,\ \text{Corollary 18.1.3}]}{\Longrightarrow} \partial J(x) \subset \mathrm{rbd}(\partial J(x^\star)). \tag{7.3.5}$$

Suppose that $\mathcal{M}_{x^\star}^J$ has not been identified yet. Then, since $x_k = \mathcal{P}_{C^J}(z_k) = \mathcal{P}_{\mathcal{F}_{x_k}^J \setminus \mathcal{F}_{x^\star}^J}(z_k)$, and thanks to (7.3.3), this is equivalent to

$$\text{either} \quad \mathcal{F}_{x^\star}^J \subsetneq \mathcal{F}_{x_k}^J \quad \text{or} \quad \mathcal{F}_{x_k}^J \cap \mathcal{F}_{x^\star}^J = \emptyset.$$

It then follows from (7.3.4) and Proposition 7.3.3 that the number of iterations where $\mathcal{F}_{x^\star}^J \subsetneq \mathcal{F}_{x_k}^J$ and $\mathcal{F}_{x^\star}^R \subsetneq \mathcal{F}_{x_k}^R$ cannot exceed the bound in (7.3.1), and thus identification happens indeed for some large enough k obeying (7.3.1).

7.4 Local linear convergence

Building upon the identification results from the previous section, we now turn to the local behaviour of the DR iteration (7.2.1) under partial smoothness. Similarly to the analysis of the FB-type methods 6.3, the proof strategy is first to show that DR iteration locally linearises along the identified manifolds, and then converges linearly owing to the linearization.

Angles between subspaces Before presenting the main result, we introduce the angles between subspaces. Let T_1, T_2 be two subspaces, and without loss of generality, let $1 \leq p \stackrel{\text{def}}{=} \dim(T_1) \leq q \stackrel{\text{def}}{=} \dim(T_2) \leq n-1$.

Definition 7.4.1 (Principal angles). The principal angles $\theta_k \in [0, \frac{\pi}{2}]$, $k = 1, \ldots, p$ between subspaces T_1 and T_2 are defined by, with $u_0 = v_0 \stackrel{\text{def}}{=} 0$

$$\cos(\theta_k) \stackrel{\text{def}}{=} \langle u_k, v_k \rangle = \max \langle u, v \rangle \ \text{ s.t. } \ u \in T_1, v \in T_2, \|u\| = 1, \|v\| = 1, \\ \langle u, u_i \rangle = \langle v, v_i \rangle = 0, \ i = 0, \cdots, k-1.$$

The principal angles θ_k are unique and satisfy $0 \leq \theta_1 \leq \theta_2 \leq \cdots \leq \theta_p \leq \pi/2$.

Definition 7.4.2 (Friedrichs angle). The Friedrichs angle $\theta_F \in]0, \frac{\pi}{2}]$ between T_1 and T_2 is

$$\cos\left(\theta_F(T_1, T_2)\right) \stackrel{\text{def}}{=} \max \langle u, v \rangle \ \text{ s.t. } \ u \in T_1 \cap (T_1 \cap T_2)^\perp, \|u\| = 1, \ v \in T_2 \cap (T_1 \cap T_2)^\perp, \|v\| = 1.$$

The following lemma shows the relation between the Friedrichs and principal angles. The proof can be found in [18, Proposition 3.3].

Lemma 7.4.3 (Principal angles and Friedrichs angle). *The Friedrichs angle is exactly θ_{d+1} where $d \stackrel{\text{def}}{=} \dim(T_1 \cap T_2)$. Moreover, $\theta_F(T_1, T_2) > 0$.*

Remark 7.4.4. One approach to obtain the principal angles is through the singular value decomposition. For instance, let $X \in \mathbb{R}^{n \times p}$ and $Y \in \mathbb{R}^{n \times q}$ form orthonormal bases for the subspaces T_1 and T_2 respectively. Let $U \Sigma V^T$ be the SVD of $X^T Y \in \mathbb{R}^{p \times q}$, then $\cos(\theta_k) = \sigma_k$, $k = 1, 2, \ldots, p$ and σ_k corresponds to the k^{th} largest singular value in Σ.

7.4.1 Local linearized iteration

Let $z^\star \in \operatorname{fix}(\mathcal{F}_{\gamma,\lambda})$ and $x^\star = \operatorname{prox}_{\gamma J}(z^\star) \in \operatorname{Argmin}(R + J)$. Define the following two functions

$$\overline{R}(x) \stackrel{\text{def}}{=} \gamma R(x) - \langle x, x^\star - z^\star \rangle, \quad \overline{J}(x) \stackrel{\text{def}}{=} \gamma J(x) - \langle x, z^\star - x^\star \rangle. \tag{7.4.1}$$

We start with the following key lemma.

Lemma 7.4.5. *Suppose that $R \in \operatorname{PSF}_{x^\star}(\mathcal{M}_{x^\star}^R)$ and $J \in \operatorname{PSF}_{x^\star}(\mathcal{M}_{x^\star}^J)$. Define the two matrices*

$$H_{\overline{R}} \stackrel{\text{def}}{=} \mathcal{P}_{T_{x^\star}^R} \nabla^2_{\mathcal{M}_{x^\star}^R} \overline{R}(x^\star) \mathcal{P}_{T_{x^\star}^R} \quad \text{and} \quad H_{\overline{J}} \stackrel{\text{def}}{=} \mathcal{P}_{T_{x^\star}^J} \nabla^2_{\mathcal{M}_{x^\star}^J} \overline{J}(x^\star) \mathcal{P}_{T_{x^\star}^J}. \tag{7.4.2}$$

$H_{\overline{R}}$ and $H_{\overline{J}}$ are symmetric positive semi-definite under either of the following circumstances:

(i) ($\mathrm{ND}_{\mathrm{DR}}$) *holds.*

(ii) $\mathcal{M}^R_{x^\star}$ *and* $\mathcal{M}^J_{x^\star}$ *are affine subspaces.*

In turn, the matrices

$$W_{\overline{R}} \overset{\text{def}}{=} (\mathrm{Id} + H_{\overline{R}})^{-1} \quad \textit{and} \quad W_{\overline{J}} \overset{\text{def}}{=} (\mathrm{Id} + H_{\overline{J}})^{-1} \tag{7.4.3}$$

are both firmly non-expansive.

Proof. Here we prove the case for J since the same arguments apply to R just as well. Claims (i) and (ii) follow from Lemma 6.3.3 since $J \in \mathrm{PSF}_{x^\star}(\mathcal{M}^J_{x^\star})$. Consequently, $W_{\overline{J}}$ is symmetric positive definite with eigenvalues in $]0,1]$. Thus by virtue of [16, Corollary 4.3(ii)], it is firmly non-expansive. □

Now define $M_{\overline{R}} \overset{\text{def}}{=} \mathcal{P}_{T^R_{x^\star}} W_{\overline{R}} \mathcal{P}_{T^R_{x^\star}}$ and $M_{\overline{J}} \overset{\text{def}}{=} \mathcal{P}_{T^J_{x^\star}} W_{\overline{J}} \mathcal{P}_{T^J_{x^\star}}$, and the matrices

$$\begin{aligned} M_{\mathrm{DR}} &\overset{\text{def}}{=} \mathrm{Id} + 2M_{\overline{R}}M_{\overline{J}} - M_{\overline{R}} - M_{\overline{J}} = \frac{1}{2}\mathrm{Id} + \frac{1}{2}(2M_{\overline{R}} - \mathrm{Id})(2M_{\overline{J}} - \mathrm{Id}), \\ M_\lambda &\overset{\text{def}}{=} (1-\lambda)\mathrm{Id} + \lambda M_{\mathrm{DR}}, \quad \lambda \in]0,2[. \end{aligned} \tag{7.4.4}$$

We have the following locally linearized version of (7.1.3).

Proposition 7.4.6 (Locally linearized DR iteration). *Suppose that the DR iteration (7.1.3) is run under the assumptions of Theorem 7.3.1. Assume also that $\lambda_k \to \lambda \in]0,2[$. Then M_{DR} is* firmly non-expansive *and M_λ is $\frac{\lambda}{2}$-averaged. Moreover, for all k large enough, we have*

$$z_{k+1} - z^\star = M_\lambda(z_k - z^\star) + o(\|z_k - z^\star\|) + O(\lambda_k|\gamma_k - \gamma|). \tag{7.4.5}$$

The term $O(\lambda_k|\gamma_k - \gamma|)$ disappears if (γ_k, λ_k) are chosen constant in $]0, +\infty[\times]0,2[$. If moreover R and J are locally polyhedral around $x^\star$, the term $o(\|z_k - z^\star\|)$ vanishes.

See Section 7.9 from page 166 for the proof.

Remark 7.4.7. If $\lambda_k|\gamma_k - \gamma| = o(\|z_k - z^\star\|)$, then (7.4.5) becomes

$$z_{k+1} - z^\star = M_\lambda(z_k - z^\star) + o(\|z_k - z^\star\|).$$

However, this is of little practical interest as $z^\star$ is unknown.

We now derive a characterization of the spectral properties of M_λ, which in turn allows to study the linear convergence rates of its powers. Recall the notion of convergent matrices from Definition 2.5.1. To lighten the notation in the following, we define the subspaces $S^J_{x^\star} \overset{\text{def}}{=} (T^J_{x^\star})^\perp$ and $S^R_{x^\star} \overset{\text{def}}{=} (T^R_{x^\star})^\perp$.

Lemma 7.4.8. *Suppose that $\lambda \in]0,2[$, then,*

(i) *M_λ is convergent with*

$$M^\infty_{\mathrm{DR}} = \mathcal{P}_{\ker(M_{\overline{R}}(\mathrm{Id}-M_{\overline{J}})+(\mathrm{Id}-M_{\overline{R}})M_{\overline{J}})}.$$

In particular, $M^\infty_{\mathrm{DR}} = 0$ if holds

$$T^J_{x^\star} \cap T^R_{x^\star} = \{0\}, \ \ \mathrm{span}(\mathrm{Id} - M_{\overline{J}}) \cap S^R_{x^\star} = \{0\} \quad \textit{and} \quad \mathrm{span}(\mathrm{Id} - M_{\overline{R}}) \cap T^R_{x^\star} = \{0\}.$$

(ii) *If, moreover, R and J are locally polyhedral around $x^\star$, then M_λ is* normal *and converges linearly to $\mathcal{P}_{(T^J_{x^\star}\cap T^R_{x^\star})\oplus(S^J_{x^\star}\cap S^R_{x^\star})}$ with the optimal rate (Definition 2.5.1)*

$$\sqrt{(1-\lambda)^2 + \lambda(2-\lambda)\cos^2\left(\theta_F(T^J_{x^\star}, T^R_{x^\star})\right)} < 1.$$

In particular, if $T^J_{x^\star} \cap T^R_{x^\star} = S^J_{x^\star} \cap S^R_{x^\star} = \{0\}$, then M_λ converges linearly to 0 with the optimal rate

$$\sqrt{(1-\lambda)^2 + \lambda(2-\lambda)\cos^2\left(\theta_1(T^J_{x^\star}, T^R_{x^\star})\right)} < 1.$$

See Section 7.9 from page 168 for the proof.

Combining Proposition 7.4.6 and Lemma 7.4.8, we have the following equivalent characterization of the locally linearized iteration.

Corollary 7.4.9. *Suppose that the DR iteration (7.1.3) is run under the assumptions of Theorem 7.3.1. Assume also that $\lambda_k \to \lambda \in]0,2[$. Then the following holds.*

(i) *(7.4.5) is equivalent to*

$$\begin{aligned}&(\mathrm{Id} - M_{\mathrm{DR}}^{\infty})(z_{k+1} - z^{\star})\\ &= (M_{\lambda} - M_{\mathrm{DR}}^{\infty})(\mathrm{Id} - M_{\mathrm{DR}}^{\infty})(z_k - z^{\star}) + (\mathrm{Id} - M_{\mathrm{DR}}^{\infty})o(\|z_k - z^{\star}\|) + O(\lambda_k|\gamma_k - \gamma|).\end{aligned} \tag{7.4.6}$$

(ii) *If R and J are locally polyhedral around $x^{\star}$ and (γ_k, λ_k) are constant in $]0, +\infty[\times]0,2[$, then*

$$z_{k+1} - z^{\star} = (M_{\lambda} - M_{\mathrm{DR}}^{\infty})(z_k - z^{\star}). \tag{7.4.7}$$

See Section 7.9 from page 169 for the proof.

7.4.2 Local linear convergence

We are now in position to present the local linear convergence of DR method.

Theorem 7.4.10 (Local linear convergence). *Suppose that the DR iteration (7.1.3) is run under the assumptions of Proposition 7.4.6. Recall M_{DR}^{∞} from Lemma 7.4.8. The following holds:*

(i) *Given any $\rho \in]\rho(M_{\lambda} - M_{\mathrm{DR}}^{\infty}), 1[$, there exists $K \in \mathbb{N}$ large enough such that for all $k \geq K$, if $\lambda_k|\gamma_k - \gamma| = O(\eta^k)$ for $0 \leq \eta < \rho$, then*

$$\|(\mathrm{Id} - M_{\mathrm{DR}}^{\infty})(z_k - z^{\star})\| = O(\rho^{k-K}). \tag{7.4.8}$$

(ii) *If R and J are locally polyhedral around $x^{\star}$ and $(\gamma_k, \lambda_k) \equiv (\gamma, \lambda) \in]0, +\infty[\times]0,2[$, then there exists $K \in \mathbb{N}$ large enough such that for all $k \geq K$,*

$$\|z_k - z^{\star}\| \leq \rho^{k-K}\|z_K - z^{\star}\|, \tag{7.4.9}$$

where the value of ρ is

$$\rho = \sqrt{(1-\lambda)^2 + \lambda(2-\lambda)\cos^2\big(\theta_F(T_{x^{\star}}^{J}, T_{x^{\star}}^{R})\big)} \in [0,1[.$$

The proof is given at page 169 in Section 7.9 for the proof.

Remark 7.4.11.

(i) If $M_{\mathrm{DR}}^{\infty} = 0$ in (7.4.8) or the situation of (7.4.9), we also have local linear convergence of x_k and v_k to $x^{\star}$ by non-expansiveness of the proximity operator.

(ii) The condition on $\lambda_k|\gamma_k - \gamma|$ in Theorem 7.4.10(i) amounts to saying that γ_k should converge fast enough to γ. Otherwise, the local convergence rate would be dominated by that of $\lambda_k|\gamma_k - \gamma|$. Especially when $\lambda_k|\gamma_k - \gamma|$ converges sub-linearly to 0, then the local convergence rate will eventually become sub-linear. See Figure 7.5 in the numerical experiments section.

(iii) For (ii) of Theorem 7.4.10, it can be observed that the best rate is obtained for $\lambda = 1$. This has been also pointed out in [73] for basis pursuit. This assertion is however only valid for the local convergence behaviour and does not mean in general that the DR will be globally faster for $\lambda_k \equiv 1$. See the discussions below for more details.

(iv) Observe also that the local linear convergence rate does not depend on γ when both R and J are locally polyhedral around $x^{\star}$. This means that the choice of γ_k only affects the number of iterations needed for finite identification.

However, for general partly smooth functions, γ_k influences both the identification time and the local linear convergence rate, since M_λ depends on it through the matrices $W_{\overline{R}}$ and $W_{\overline{J}}$ (γ weights the Riemannian Hessians of $\overline{R}$ and $\overline{J}$, see (7.4.1)-(7.4.3). See Figure 7.4 for a numerical comparison.

7.4.3 The effect of relaxation

Here we investigate the effect of the relaxation parameter λ. Let $\eta = \eta_{\text{Re}} + \mathrm{i}\eta_{\text{Im}}$ (here $\mathrm{i}^2 = -1$, η_{Re} and η_{Im} are the *real* and *imagery* part of η respectively) be the eigenvalue of $M_{\text{DR}} - M_{\text{DR}}^\infty$, then the eigenvalue of $M_\lambda - M_{\text{DR}}^\infty$ if a function of η and λ,

$$\rho_{\eta,\lambda} = (1-\lambda) + \lambda(\eta_{\text{Re}} + \mathrm{i}\eta_{\text{Im}}) = 1 + \lambda(\eta_{\text{Re}} - 1) + \mathrm{i}\lambda\eta_{\text{Im}}$$

whose modulus is

$$|\rho_{\eta,\lambda}| = \sqrt{(1 + \lambda(\eta_{\text{Re}} - 1))^2 + \lambda^2\eta_{\text{Im}}^2}. \tag{7.4.10}$$

Eigenvalue η is real When η is real, *i.e.* $\eta_{\text{Im}} = 0$, then ρ is real.

Lemma 7.4.12. *Let $M_{\text{DR}} \in \mathbb{R}^{n\times n}$ be a convergent square matrix whose eigenvalues are all real, denote $\bar{\eta}, \underline{\eta}$ the biggest and smallest (signed) eigenvalues of $M_{\text{DR}} - M_{\text{DR}}^\infty$ respectively. Then the spectral radius $\rho(M_{\text{DR}} - M_{\text{DR}}^\infty)$ attains the minimum at*

$$\lambda = \frac{2}{2 - (\bar{\eta} + \underline{\eta})}.$$

See page 170 for the proof. The above result implies if all the eigenvalues of M_{DR} are real, then *over-relaxation* (*i.e.* $\lambda > 1$) makes M_λ converge faster to M_{DR}^∞ than M_{DR}.

Eigenvalue η is complex The situation becomes much more complicated when η is complex. That is, if all the eigenvalues of $M_{\text{DR}} - M_{\text{DR}}^\infty$ are complex, then in general it is very difficult to obtain result similar to Lemma 7.4.12.

However, there are extra properties we can apply when both R, J are locally polyhedral around $x^\star$. Let θ be a principal angle between $T_{x^\star}^J$ and $T_{x^\star}^R$, then owing to the result of [73], the eigenvalue of $M_{\text{DR}} - M_{\text{DR}}^\infty$ corresponding to θ reads

$$\eta_\theta = \cos^2(\theta) + \mathrm{i}\cos(\theta)\sin(\theta). \tag{7.4.11}$$

Putting such η_θ back in (7.4.10), we get

$$|\rho_{\eta_\theta,\lambda}|^2 = (\lambda - 1)^2\sin^2(\theta) + \cos^2(\theta) \geq \cos^2(\theta),$$

which implies that non-relaxation (*i.e.* $\lambda = 1$) yields the best local convergence rate.

7.4.4 Partial smoothness vs metric sub-regularity

In this section, we discuss the connection between partial smoothness and metric sub-regularity for DR when both R and J are locally polyhedral around $x^\star$. We also assume that the iteration (7.1.3) is stationary (*i.e.* $\gamma_k \equiv \gamma$) and non-relaxed (*i.e.* $\lambda_k \equiv 1$).

Proposition 7.4.13. *Suppose that the DR iteration (7.1.3) is run under the assumptions of Proposition 7.4.6 with $\gamma_k \equiv \gamma$ and $\lambda_k \equiv 1$. If R and J are locally polyhedral around $x^\star$, then $M'_{\text{DR}} \overset{\text{def}}{=} \text{Id} - M_{\text{DR}}$ is $\frac{1}{\sin(\theta_F(T_{x^\star}^J, T_{x^\star}^R))}$-metrically sub-regular at $z^\star$ for 0.*

Proof. Under the assume conditions, we have the locally

$$M_{\text{DR}} = \mathcal{P}_{T_{x^\star}^J}\mathcal{P}_{T_{x^\star}^R} + \mathcal{P}_{S_{x^\star}^J}\mathcal{P}_{S_{x^\star}^R},$$

which is linear and non-expansive, hence metrically subregular owing to [22, Example 2.5]. Now consider the eigenvalues of M_{DR}. Without loss of generality, let $p = \dim(T_{x^\star}^J) \leq q = \dim(T_{x^\star}^R)$, and denote $\{\theta_i\}_{i=1,\dots,p} \in [0, \pi/2[$ the principal angles between $T_{x^\star}^J$ and $T_{x^\star}^R$ sorted in ascending order. Owing to (7.4.11) (see also [18, Proposition 3.3 and Theorem 3.10]), the eigenvalues of M'_{DR} lie in

$$\begin{cases} \{\sin^2(\theta_i) + \mathrm{i}\cos(\theta_i)\sin(\theta_i) : i = s+1, \dots, p\} \cup \{0\} & \text{if } p = q, \\ \{\sin^2(\theta_i) + \mathrm{i}\cos(\theta_i)\sin(\theta_i) : i = s+1, \dots, p\} \cup \{0\} \cup \{1\} & \text{if } p < q. \end{cases}$$

Thus, for any $i = s+1, \dots, p$, each eigenvalue has modulus $\sin(\theta_i) \in [0, 1[$. Therefore, since M'_{DR} is normal, given any $z \in \ker(M'_{\mathrm{DR}})^\perp$, we have

$$\|M'_{\mathrm{DR}} z\| \geq \sin(\theta_{s+1})\|z\| = \sin(\theta_F(T_{x^\star}^J, T_{x^\star}^R))\|z\|.$$

In view of [22, Theorem 2.5], this entails that M'_{DR} is $\frac{1}{\sin(\theta_F(T_{x^\star}^J, T_{x^\star}^R))}$-metrically sub-regular. □

Unlike the case of Forward–Backward splitting (Section 6.4.4), for the DR splitting with polyhedral functions, partial smoothness and metric sub-regularity yield the same estimate of the convergence rate. Indeed, from (3.3.3) the rate estimate obtained through metric sub-regularity reads

$$\sqrt{1 - \sin^2(\theta_F(T_{x^\star}^J, T_{x^\star}^R))} = \cos\left(\theta_F(T_{x^\star}^J, T_{x^\star}^R)\right),$$

which is exactly the result of Theorem 7.4.10(ii) for $\lambda = 1$.

7.5 Finite termination

In this section we characterize the situations where finite convergence of DR can be obtained.

Theorem 7.5.1 (Finite termination of DR). *Assume that the unrelaxed and stationary DR iteration is used (i.e., $\gamma_k \equiv \gamma \in]0, +\infty[$ and $\lambda_k \equiv 1$), such that $(z_k, x_k, u_k) \to (z^\star, x^\star, x^\star)$, where R and J are locally polyhedral around $x^\star$. Suppose that either J or R is locally C^2 at $x^\star$. Then the DR sequences $\{z_k, x_k, u_k\}_{k\in\mathbb{N}}$ converge in finite number of steps to $(z^\star, x^\star, x^\star)$.*

Proof. We will prove the statement when J is locally C^2 at $x^\star$, and the same reasoning holds if the assumption is on R. Local C^2-smoothness of J at $x^\star$ entails that $\partial J(x^\star) = \{\nabla J(x^\star)\}$ and J is partly smooth at $x^\star$ relative to $\mathcal{M}_{x^\star}^J = \mathbb{R}^n$. Moreover, the non-degeneracy condition ($\mathrm{ND}_{\mathrm{DR}}$) is in force. It then follows from Proposition 7.4.6 and (i) of Lemma 7.4.8 that there exists $K \in \mathbb{N}$ large enough such that

$$\forall k \geq K, \quad z_{k+1} - z^\star = \mathcal{P}_{T_{x^\star}^R}(z_k - z^\star) \Rightarrow \forall k \geq K+1, \quad z_k - z^\star \in T_{x^\star}^R,$$

whence we conclude that

$$\forall k \geq K+1, \quad z_k = z_{k+1} = \cdots = z^\star.$$ □

DR is known (see, *e.g.*, Section 4.3 or [78, Theorem 6]) to be a special case of the exact proximal point algorithm (PPA) with constant step-size $\gamma_k \equiv 1$. This suggests that many results related to PPA can be carried over to DR. For instance, finite convergence of PPA has been studied in [151, 121] under different conditions. However, [78, Theorem 9] gave a negative result which suggests that these previous conditions which are sufficient for finite termination of PPA are difficult or impossible to carry over to DR even for the polyhedral case. The authors in [19] consider the unrelaxed and stationary DR for solving the convex feasibility problem

$$\text{Find a point in } \mathcal{S}_1 \cap \mathcal{S}_2,$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are nonempty closed convex sets in $\mathbb{R}^n$, $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$, $\mathcal{S}_1$ is an affine subspace and $\mathcal{S}_2$ is a polyhedron. They established finite convergence under Slater's condition

$$\mathcal{S}_1 \cap \mathrm{int}(\mathcal{S}_2) \neq \emptyset.$$

They also provided examples where this condition holds while the conditions of [151, 121] for finite convergence do not apply.

Specializing our result to $R = \iota_{\mathcal{S}_1}$ and $J = \iota_{\mathcal{S}_2}$, then under Slater's condition, if $x^\star \in \mathcal{S}_1 \cap \operatorname{int}(\mathcal{S}_2)$, we have that R is partly smooth at any $x \in \mathcal{S}_1$ relative to $\mathcal{S}_1$ with $T_{x^\star}^R = \operatorname{par}(\mathcal{S}_1)$ (*i.e.* a translate of $\mathcal{S}_1$ to the origin), and $\partial J(x^\star) = \mathcal{N}_{\mathcal{S}_2}(x^\star) = \{0\}$, and we recover the result of [19]. In fact, [19, Theorem 3.7] shows that the cluster point $x^\star$ is always an interior point regardless of the starting point of DR. The careful reader may have noticed that in the current setting, thanks to Example 7.3.4, the estimate in (7.3.3) gives a bound on the number of iterations for finite convergence.

7.6 Alternating direction method of multipliers

For problem ($\mathcal{P}_{\text{DR}}$), let's now compose J with a linear operator L, *i.e.*

$$\min_{x\in\mathbb{R}^n} R(x) + J(Lx), \tag{$\mathcal{P}_{\text{ADMM}}$}$$

where now

(**D.7**) $R \in \Gamma_0(\mathbb{R}^n)$, $J \in \Gamma_0(\mathbb{R}^m)$ and $L : \mathbb{R}^n \to \mathbb{R}^m$ is an *injective* linear operator.

(**D.8**) $\operatorname{ri}(\operatorname{dom}(R) \cap \operatorname{dom}(J)) \neq \emptyset$.

(**D.9**) $\operatorname{Argmin}(R + JL) \neq \emptyset$, *i.e.* the set of minimizers is non-empty.

The main difficulty of applying DR to solve ($\mathcal{P}_{\text{ADMM}}$) is that the proximity operator of the composition JL in general can not be solved explicitly. The alternating direction method of multipliers [82] is an efficient way to deal with such a difficulty.

The stationary version (*i.e.* constant step-size) of ADMM is described in Algorithm 14. For convenience, here with a slight abuse of notations, we use x_k, u_k and z_k to denote the sequences generated by ADMM.

Algorithm 14: Alternating Direction method of Multipliers

Initial: $\gamma > 0$. $k = 0$, $x_0, y_0 \in \mathbb{R}^m$;

repeat

$$\begin{aligned} u_{k+1} &= \operatorname{argmin}_{u\in\mathbb{R}^n} R(u) + \langle Lu,\, y_k\rangle + \tfrac{\gamma}{2}\|Lu - x_k\|^2, \\ x_{k+1} &= \operatorname{argmin}_{x\in\mathbb{R}^m} J(x) - \langle x,\, y_k\rangle + \tfrac{\gamma}{2}\|x - Lu_{k+1}\|^2, \\ y_{k+1} &= y_k + \gamma(Lu_{k+1} - x_{k+1}), \end{aligned} \tag{7.6.1}$$

until *convergence*;

It is shown in [82] (see also [78]), that ADMM amounts to applying DR to the Fenchel-Rockafellar dual problem of ($\mathcal{P}_{\text{ADMM}}$). In the following, we recall in short the derivation of transforming ADMM to DR. First, the dual form of ($\mathcal{P}_{\text{ADMM}}$) is

$$\max_{y\in\mathbb{R}^m} -\big(R^*(-L^*y) + J^*(y)\big), \tag{$\mathcal{D}_{\text{ADMM}}$}$$

where R^*, J^* denote the Fenchel conjugate of R and J respectively (Definition 2.1.14). Define $z_{k+1} = y_k + \gamma L u_{k+1}$, then following the iteration from (7.6.1) we get

$$\begin{cases} u_{k+1} = \operatorname{argmin}_{u\in\mathbb{R}^n} R(u) + \frac{\gamma}{2}\|Lu - \frac{1}{\gamma}(z_k - 2y_k)\|^2, \\ z_{k+1} = y_k + \gamma L u_{k+1}, \\ x_{k+1} = \operatorname{argmin}_{x\in\mathbb{R}^m} J(x) + \frac{\gamma}{2}\|x - \frac{1}{\gamma}z_{k+1}\|^2, \\ y_{k+1} = z_{k+1} - \gamma x_{k+1}. \end{cases} \tag{7.6.2}$$

Apply the Moreau's identity to R, J respectively [152], then we obtain

$$\begin{cases} z_{k+1} = \operatorname{argmin}_{z\in\mathbb{R}^m} R^*(-L^*z) + \frac{1}{2\gamma}\|z - (2y_k - z_k)\|^2 + (z_k - y_k), \\ y_{k+1} = \operatorname{argmin}_{y\in\mathbb{R}^m} J^*(y) + \frac{1}{2\gamma}\|y - z_{k+1}\|^2, \end{cases} \tag{7.6.3}$$

which is clearly applying the non-relaxed and stationary DR method (*i.e.* $\lambda_k \equiv 1, \gamma \equiv \gamma$) to the dual problem ($\mathcal{D}_{\text{ADMM}}$).

When L injective (*i.e.* has full column rank), the convergence of all the sequences in (7.6.2) are guaranteed [78], that is

$$u_k \to u^\star, \;\; x_k \to x^\star = Lu^\star, \;\; z_k \to z^\star, \;\; y_k \to y^\star = z^\star - \gamma L u^\star, \tag{7.6.4}$$

where $u^\star$ is a global minimizer of ($\mathcal{P}_{\text{ADMM}}$), $x^\star$ is a dual solution of ($\mathcal{D}_{\text{ADMM}}$), $z^\star$ is a fixed point of (7.6.3), and $y^\star$ is the Lagrangian multiplier.

Local linear convergence of ADMM By virtue of the chain rule of partial smoothness (Theorem 5.1.7), if R^* is a partly smooth function, then so is the composition $R^*(-L^*)$ since L is injective. As a consequence, besides conditions (**D.7**)-(**D.9**), if we assume that R^* and J^* are moreover partly smooth functions, then based on the result of Section 7.3 and 7.4, we can obtain the local linear convergence of the ADMM through the dual iteration (7.6.3). As a matter of fact, we can establish the result directly on the primal ADMM iteration (7.6.1), which is what we present now.

Let $u^\star$ be a global minimizer of ($\mathcal{P}_{\text{ADMM}}$), $x^\star$ be a dual solution of ($\mathcal{D}_{\text{ADMM}}$) such that $Lu^\star = x^\star$, and $y^\star$ be the Lagrangian multiplier. Then at convergence, we have the following inclusion from (7.6.2),

$$y^\star \in \partial J(x^\star) \quad \text{and} \quad -L^*y^\star \in \partial R(u^\star).$$

Similarly to the analysis procedure of DR, we present first the finite identification of ADMM.

Corollary 7.6.1 (Finite activity identification). *For the ADMM method (7.6.1), suppose that (**D.7**)-(**D.9**) hold such that the generated sequences satisfy $(u_k, x_k, y_k) \to (u^\star, x^\star, y^\star)$ where $u^\star \in \operatorname{Argmin}(R + JL)$ and $x^\star = Lu^\star$. If $J \in \operatorname{PSF}_{x^\star}(\mathcal{M}^J_{x^\star}), R \in \operatorname{PSF}_{u^\star}(\mathcal{M}^R_{u^\star})$, and moreover the non-degeneracy condition*

$$y^\star \in \operatorname{ri}\big(\partial J(x^\star)\big) \quad \text{and} \quad -L^*y^\star \in \operatorname{ri}\big(\partial R(u^\star)\big), \tag{ND$_{\text{ADMM}}$}$$

holds. Then,

(i) *For all k sufficiently large, $(u_k, x_k) \in \mathcal{M}^R_{u^\star} \times \mathcal{M}^J_{x^\star}$.*

(ii) *Moreover,*

(a) *if $\mathcal{M}^R_{u^\star} = u^\star + T^R_{u^\star}$, then $\forall k \geq K$, $T^R_{u_k} = T^R_{u^\star}$;*

(b) *if $\mathcal{M}^J_{x^\star} = x^\star + T^J_{x^\star}$, then $\forall k \geq K$, $T^J_{x_k} = T^J_{x^\star}$;*

(c) *if R is locally polyhedral around $u^\star$, then $\forall k \geq K$, $u_k \in \mathcal{M}^R_{u^\star} = u^\star + T^R_{u^\star}$, $T^R_{u_k} = T^R_{u^\star}$, $\nabla_{\mathcal{M}^R_{u^\star}} R(u_k) = \nabla_{\mathcal{M}^R_{u^\star}} R(u^\star)$, and $\nabla^2_{\mathcal{M}^R_{u^\star}} R(u_k) = 0$;*

(d) *if J is locally polyhedral around $x^\star$, then $\forall k \geq K$, $x_k \in \mathcal{M}^J_{x^\star} = x^\star + T^J_{x^\star}$, $T^J_{x_k} = T^J_{x^\star}$, $\nabla_{\mathcal{M}^J_{x^\star}} J(x_k) = \nabla_{\mathcal{M}^J_{x^\star}} J(x^\star)$, and $\nabla^2_{\mathcal{M}^J_{x^\star}} J(x_k) = 0$.*

Next we present the local linear convergence of ADMM. Define the following function similar to those in (7.6.5),

$$\overline{\overline{R}}(u) \overset{\text{def}}{=} \frac{1}{\gamma}\big(R(u) - \langle u, -L^*y^\star\rangle\big), \;\; \overline{\overline{J}}(x) \overset{\text{def}}{=} \frac{1}{\gamma}\big(J(x) - \langle x, y^\star\rangle\big). \tag{7.6.5}$$

Owing to Lemma 6.3.3, with condition (ND$_{\text{ADMM}}$) holding, their Riemannian Hessian are positive semidefinite. Hence, define the following matrices

$$\begin{aligned} H_{\overline{\overline{R}}} &\overset{\text{def}}{=} \mathcal{P}_{T^R_{u^\star}} \nabla^2_{\mathcal{M}^R_{u^\star}} \overline{\overline{R}}(u^\star) \mathcal{P}_{T^R_{u^\star}} && \text{and} \quad H_{\overline{\overline{J}}} \overset{\text{def}}{=} \mathcal{P}_{T^J_{x^\star}} \nabla^2_{\mathcal{M}^J_{x^\star}} \overline{\overline{J}}(x^\star) \mathcal{P}_{T^J_{x^\star}}, \\ W_{\overline{\overline{R}}} &\overset{\text{def}}{=} L_R\big(\mathrm{Id} + (L_R^T L_R)^{-1} H_{\overline{\overline{R}}}\big)^{-1} (L_R^T L_R)^{-1} L_R^T && \text{and} \quad W_{\overline{\overline{J}}} \overset{\text{def}}{=} \big(\mathrm{Id} + H_{\overline{\overline{J}}}\big)^{-1}, \end{aligned} \tag{7.6.6}$$

where $L_R \overset{\text{def}}{=} L\mathcal{P}_{T^R_{u^\star}}$. Define $T^{R,L}_{u^\star}$ the subspace corresponding to the projection operator $\mathcal{P}_{T^{R,L}_{u^\star}} = L_R(L_R^T L_R)^{-1} L_R^T$.

Now define $M_{\overline{R}} \overset{\text{def}}{=} \mathcal{P}_{T^R_{u^\star}} W_{\overline{R}} \mathcal{P}_{T^R_{u^\star}}$ and $M_{\overline{J}} \overset{\text{def}}{=} \mathcal{P}_{T^J_{x^\star}} W_{\overline{J}} \mathcal{P}_{T^J_{x^\star}}$, and the matrix

$$M_{\text{ADMM}} = \text{Id} + 2M_{\overline{R}}M_{\overline{J}} - M_{\overline{R}} - M_{\overline{J}} = \frac{1}{2}(2M_{\overline{R}} - \text{Id})(2M_{\overline{J}} - \text{Id}) + \frac{1}{2}\text{Id}. \tag{7.6.7}$$

We have the following local linear convergence rate result for ADMM.

Corollary 7.6.2. *Let* (**D.7**)-(**D.9**) *and conditions in Theorem 7.6.1 hold.*

(i) *Matrix M_{ADMM} is firmly non-expansive, hence convergent.*

(ii) *For all k large enough, the ADMM iteration (7.6.1) can be written as the following fixed-point iteration*

$$z_{k+1} - z^\star = M_{\text{ADMM}}(z_k - z^\star) + o(\|z_k - z^\star\|). \tag{7.6.8}$$

(iii) *For the sequence $\{z_k\}_{k\in\mathbb{N}}$, we have the following result*

(a) *given any $\rho \in]\rho(M_{\text{ADMM}} - M^\infty_{\text{ADMM}}), 1[$, there is K large enough such that for all $k \geq K$,*

$$\|(\text{Id} - M^\infty_{\text{ADMM}})(z_k - z^\star)\| = O(\rho^{k-K}). \tag{7.6.9}$$

(b) *If R and J are locally polyhedral around $u^\star$ and $x^\star$ respectively, then there exists $K \in \mathbb{N}$ large enough such that for all $k \geq K$,*

$$\|z_k - z^\star\| \leq \rho^{k-K}\|z_K - z^\star\|, \tag{7.6.10}$$

where the value of ρ is

$$\rho = \cos\left(\theta_F(T^J_{x^\star}, T^{R,L}_{u^\star})\right) \in [0, 1[,$$

(iv) *Moreover, if $M^\infty_{\text{ADMM}} = 0$, given any $\rho \in]\rho(M_{\text{ADMM}}), 1[$, there is K large enough such that for all $k \geq K$,*

$$\|x_k - x^\star\| = O(\rho^k),\ \|y_k - y^\star\| = O(\rho^k),\ \|L(u_{k+1} - u^\star)\| = O(\rho^k). \tag{7.6.11}$$

If moreover R, J are locally polyhedral around $u^\star, x^\star$ respectively, then (7.6.11) holds with

$$\rho = \cos\left(\theta_F(T^J_{x^\star}, T^{R,L}_{u^\star})\right) \in [0, 1[,$$

which is the optimal convergence rate.

See page 171 for the proof.

7.7 Sum of more than 2 functions

We now want to tackle the problem of solving

$$\min_{x\in\mathbb{R}^n} \sum_{i=1}^m J_i(x), \tag{$\boldsymbol{\mathcal{P}_m}$}$$

where

(**D′.1**) $J_i \in \Gamma_0(\mathbb{R}^n)$, $\forall i = 1, \cdots, m$.

(**D′.2**) $\bigcap_{1\leq i\leq m} \text{ri}(\text{dom}(J_i)) \neq \emptyset$.

(**D′.3**) $\text{Argmin}(\sum_{i=1}^m J_i) \neq \emptyset$.

In fact, problem ($\boldsymbol{\mathcal{P}_m}$) can be equivalently reformulated as ($\mathcal{P}_{\text{DR}}$) in a product space, see *e.g.* [59, 148]. Recall the product space of GFB method defined in Section 3.4.1, and specialize it to the case of Euclidean space. Let $\boldsymbol{\mathcal{H}} = \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{m \text{ times}}$ endowed with the scalar inner-product and norm

$$\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{H}}, \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^m \langle x_i, y_i \rangle, \|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^m \|x_i\|^2}.$$

Let $\boldsymbol{\mathcal{S}} = \{\boldsymbol{x} = (x_i)_i \in \boldsymbol{\mathcal{H}} : x_1 = \cdots = x_m\}$ and its orthogonal complement $\boldsymbol{\mathcal{S}}^\perp = \{\boldsymbol{x} = (x_i)_i \in \boldsymbol{\mathcal{H}} : \sum_{i=1}^m x_i = 0\}$. Now define the canonical isometry,

$$\boldsymbol{C} : \mathbb{R}^n \to \boldsymbol{\mathcal{S}},\ x \mapsto (x, \cdots, x),$$

then we have $\mathcal{P}_{\boldsymbol{\mathcal{S}}}(\boldsymbol{z}) = \boldsymbol{C}(\frac{1}{m}\sum_{i=1}^m z_i)$.

Problem ($\boldsymbol{\mathcal{P}}_m$) is now equivalent to

$$\min_{\boldsymbol{x}\in\boldsymbol{\mathcal{H}}} \boldsymbol{J}(\boldsymbol{x}) + \boldsymbol{R}(\boldsymbol{x}), \quad \text{where} \quad \boldsymbol{J}(\boldsymbol{x}) = \sum_{i=1}^m J_i(x_i) \quad \text{and} \quad \boldsymbol{R}(\boldsymbol{x}) = \iota_{\boldsymbol{\mathcal{S}}}(\boldsymbol{x}), \tag{$\boldsymbol{\mathcal{P}}_{\mathrm{DR}}$}$$

which has the same structure on $\boldsymbol{\mathcal{H}}$ as ($\mathcal{P}_{\mathrm{DR}}$) on $\mathbb{R}^n$.

Obviously, $\boldsymbol{J}$ is separable and therefore,

$$\mathrm{prox}_{\gamma\boldsymbol{J}}(\boldsymbol{x}) = \big(\mathrm{prox}_{\gamma J_i}(x_i)\big)_i.$$

Let $\boldsymbol{x}^\star = \boldsymbol{C}(x^\star)$. Clearly, $\boldsymbol{R}$ is polyhedral, hence partly smooth relative to $\boldsymbol{\mathcal{S}}$ with $\boldsymbol{T}^{\boldsymbol{R}}_{\boldsymbol{x}^\star} = \boldsymbol{\mathcal{S}}$. Suppose that $J_i \in \mathrm{PSF}_{x^\star}(\mathcal{M}^{J_i}_{x^\star})$ for each i. Denote $\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star} = \bigtimes_i T^{J_i}_{x^\star}$ and $\boldsymbol{S}^{\boldsymbol{J}}_{\boldsymbol{x}^\star} = \bigtimes_i (T^{J_i}_{x^\star})^\perp$. Similarly to (7.4.2), define

$$\boldsymbol{H}_{\overline{\boldsymbol{J}}} \overset{\text{def}}{=} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \nabla^2 \overline{\boldsymbol{J}}(x^\star) \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \quad \text{and} \quad \boldsymbol{W}_{\overline{\boldsymbol{J}}} \overset{\text{def}}{=} (\mathbf{Id} + \boldsymbol{H}_{\overline{\boldsymbol{J}}})^{-1},$$

where $\overline{\boldsymbol{J}}(\boldsymbol{x}) \overset{\text{def}}{=} \gamma \sum_{i=1}^m J_i(x_i) - \langle \boldsymbol{x}, \boldsymbol{z}^\star - \boldsymbol{x}^\star \rangle$, and $\mathbf{Id}$ is the identity operatror on $\boldsymbol{\mathcal{H}}$. Since $\boldsymbol{R}$ is polyhedral, we have $\boldsymbol{W}_{\overline{\boldsymbol{R}}} = \mathbf{Id}$. Now we can provide the product space form of (7.4.4), which reads

$$\begin{aligned}
\boldsymbol{M}_{\mathrm{DR}} &= \mathbf{Id} + 2\mathcal{P}_{\boldsymbol{T}^{\boldsymbol{R}}_{\boldsymbol{x}^\star}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \boldsymbol{W}_{\overline{\boldsymbol{J}}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} - \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{R}}_{\boldsymbol{x}^\star}} - \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \boldsymbol{W}_{\overline{\boldsymbol{J}}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \\
&= \frac{1}{2}\mathbf{Id} + \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{R}}_{\boldsymbol{x}^\star}} (2\mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \boldsymbol{W}_{\overline{\boldsymbol{J}}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} - \mathbf{Id}) - \frac{1}{2}(2\mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \boldsymbol{W}_{\overline{\boldsymbol{J}}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} - \mathbf{Id}) \\
&= \frac{1}{2}\mathbf{Id} + \frac{1}{2}(2\mathcal{P}_{\boldsymbol{T}^{\boldsymbol{R}}_{\boldsymbol{x}^\star}} - \mathbf{Id})(2\mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \boldsymbol{W}_{\overline{\boldsymbol{J}}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} - \mathbf{Id}),
\end{aligned} \tag{7.7.1}$$

and $\boldsymbol{M}_\lambda \overset{\text{def}}{=} (1-\lambda)\mathbf{Id} + \lambda \boldsymbol{M}_{\mathrm{DR}}$. Owing to Lemma 7.4.8, we have

$$\boldsymbol{M}^\infty_{\mathrm{DR}} = \mathcal{P}_{\ker(\mathcal{P}_{\boldsymbol{T}^{\boldsymbol{R}}_{\boldsymbol{x}^\star}}(\mathbf{Id} - \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \boldsymbol{W}_{\overline{\boldsymbol{J}}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}}) + (\mathbf{Id} - \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{R}}_{\boldsymbol{x}^\star}})\mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}} \boldsymbol{W}_{\overline{\boldsymbol{J}}} \mathcal{P}_{\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}})},$$

and when all J_i's are locally polyhedral nearby $x^\star$, $\boldsymbol{M}^\infty_{\mathrm{DR}}$ specializes to

$$\boldsymbol{M}^\infty_{\mathrm{DR}} = \mathcal{P}_{(\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star} \cap \boldsymbol{\mathcal{S}}) \oplus (\boldsymbol{S}^{\boldsymbol{J}}_{\boldsymbol{x}^\star} \cap \boldsymbol{\mathcal{S}}^\perp)}.$$

Corollary 7.7.1. *Suppose that conditions* (D′.1)-(D′.3) *and* (D.5)-(D.6) *hold. Consider the sequence* $\{\boldsymbol{z}_k, \boldsymbol{x}_k, \boldsymbol{u}_k\}_{k\in\mathbb{N}}$ *provided by the non-stationary DR method* (7.1.3) *applying to solve* ($\boldsymbol{\mathcal{P}}_{\mathrm{DR}}$). *One has the following results:*

(i) $(\boldsymbol{z}_k, \boldsymbol{x}_k, \boldsymbol{u}_k)$ *converges to* $(\boldsymbol{z}^\star, \boldsymbol{x}^\star, \boldsymbol{x}^\star)$, *where* $\boldsymbol{x}^\star = \boldsymbol{C}(x^\star)$ *and* $x^\star$ *is a global minimizer of* ($\boldsymbol{\mathcal{P}}_m$).

(ii) *Assume, moreover, that* $\gamma_k \geq \underline{\gamma} > 0$ *and* $\gamma_k \to \gamma$, $J_i \in \mathrm{PSF}_{x^\star}(\mathcal{M}^{J_i}_{x^\star})$ *and*

$$\boldsymbol{z}^\star \in \boldsymbol{x}^\star + \gamma \mathrm{ri}\Big(\bigtimes_i \partial J_i(x^\star)\Big) \cap \boldsymbol{\mathcal{S}}^\perp. \tag{$\mathbf{ND}_{\mathrm{DR}}$}$$

Then,

(a) *for all* k *large enough,* $\boldsymbol{x}_k \in \bigtimes_i \mathcal{M}^{J_i}_{x^\star}$.

(b) *in addition, if* $\lambda_k \to \lambda \in]0, 2[$, *then given any* $\rho \in]\rho(\boldsymbol{M}_\lambda - \boldsymbol{M}^\infty_{\mathrm{DR}}), 1[$, *there exists* $K \in \mathbb{N}$ *large enough such that for all* $k \geq K$, *if* $\lambda_k|\gamma_k - \gamma| = O(\eta^k)$ *where* $0 \leq \eta < \rho$, *then*

$$\|(\mathbf{Id} - \boldsymbol{M}^\infty_{\mathrm{DR}})(\boldsymbol{z}_k - \boldsymbol{z}^\star)\| = O(\rho^{k-K}).$$

In particular, if all J_i*'s are locally polyhedral around* $x^\star$ *and* $(\gamma_k, \lambda_k) \equiv (\gamma, \lambda) \in]0, +\infty[\times]0, 2[$, *then* $\boldsymbol{z}_k$ *(resp.* $x_k \overset{\text{def}}{=} \frac{1}{m}\sum_{i=1}^m \boldsymbol{x}_{k,i}$*) converges locally linearly to* $\boldsymbol{z}^\star$ *(resp.* $x^\star$*) at the optimal rate* $\rho = \sqrt{(1-\lambda)^2 + \lambda(2-\lambda)\cos^2(\theta_F(\boldsymbol{T}^{\boldsymbol{J}}_{\boldsymbol{x}^\star}, \boldsymbol{\mathcal{S}}))} \in [0, 1[$.

Proof.

(i) Apply Proposition 7.2.1 to ($\boldsymbol{\mathcal{P}}_{\mathrm{DR}}$).

(ii) (a) By the separability rule, we have $\boldsymbol{J} \in \mathrm{PSF}_{\boldsymbol{x}^\star}(\bigtimes_i \mathcal{M}_{x^\star}^{J_i})$, see [108, Proposition 4.5]. We have also $\partial \boldsymbol{R}(\boldsymbol{x}^\star) = N_{\boldsymbol{\mathcal{S}}}(\boldsymbol{x}^\star) = \boldsymbol{\mathcal{S}}^\perp$. Then ($\mathbf{ND}_{\mathrm{DR}}$) is simply a specialization of condition ($\mathrm{ND}_{\mathrm{DR}}$) to problem ($\boldsymbol{\mathcal{P}}_{\mathrm{DR}}$). The claim then follows from Theorem 7.3.1.

(b) This is a direct consequence of Theorem 7.4.10. For the local linear convergence of x_k to $x^\star$ in the last part, observe that

$$\begin{aligned}\|x_k - x^\star\|^2 &= \|\tfrac{1}{m}\textstyle\sum_{i=1}^m \boldsymbol{x}_{k,i} - \tfrac{1}{m}\sum_{i=1}^m \boldsymbol{x}_i^\star\|^2 \\ &\leq \tfrac{1}{m}\textstyle\sum_{i=1}^m \|\boldsymbol{x}_{k,i} - \boldsymbol{x}_i^\star\|^2 \\ &= \tfrac{1}{m}\textstyle\sum_{i=1}^m \|\mathrm{prox}_{\gamma J_i}(\boldsymbol{z}_{k,i}) - \mathrm{prox}_{\gamma J_i}(\boldsymbol{z}_i^\star)\|^2 \\ &\leq \tfrac{1}{m}\textstyle\sum_{i=1}^m \|\boldsymbol{z}_{k,i} - \boldsymbol{z}_i^\star\|^2 = \tfrac{1}{m}\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|^2.\end{aligned}$$

□

We also have the following corollary of Theorem 7.5.1.

Corollary 7.7.2. *Assume that the unrelaxed stationary DR iteration is used (i.e., $\gamma_k \equiv \gamma \in]0, +\infty[$ and $\lambda_k \equiv 1$), such that $(\boldsymbol{z}_k, \boldsymbol{x}_k, \boldsymbol{u}_k) \to (\boldsymbol{z}^\star, \boldsymbol{C}(x^\star), \boldsymbol{C}(x^\star))$, where for each i function J_i is locally polyhedral nearby $x^\star$ and is differentiable at $x^\star$. Then the sequences $\{\boldsymbol{z}_k, \boldsymbol{x}_k, \boldsymbol{u}_k, \frac{1}{m}\sum_{i=1}^m \boldsymbol{x}_{i,k}\}_{k\in\mathbb{N}}$ converge in finite number of steps to $(\boldsymbol{z}^\star, \boldsymbol{C}(x^\star), \boldsymbol{C}(x^\star), x^\star)$.*

7.8 Numerical experiments

7.8.1 Linear inverse problem

We first consider the 4 linear inverse problems (4.5.2) of Section 4.5. The function R is either ℓ_1, ℓ_∞, $\ell_{1,2}$-norms or the nuclear norm, and $J = \iota_{\mathcal{S}}(\cdot)$ is the indicator function of the constraint $\mathcal{S} \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : \mathcal{K}x_{\mathrm{ob}} = \mathcal{K}x\} = x_{\mathrm{ob}} + \ker(\mathcal{K})$. It is immediate to see that J is indeed polyhedral and partly smooth at any $x \in \mathcal{S}$ relative to $\mathcal{S}$.

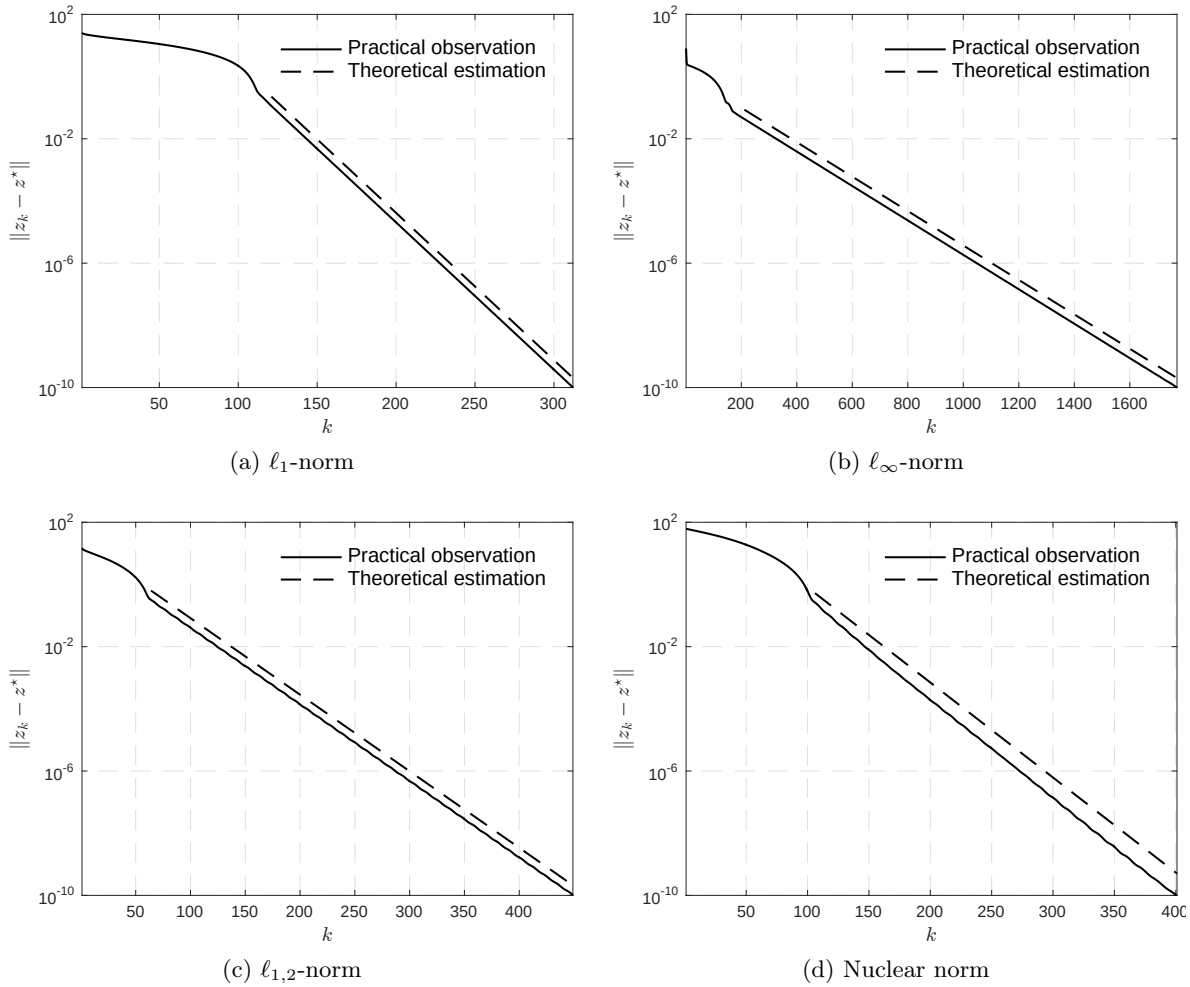
The linear operator $\mathcal{K}$ is generated from the standard random Gaussian ensemble. For each example of R, the setting of x_{ob} is (*i.e.* the same as the ones in Section 4.5.1):

ℓ_1**-norm** $(m, n) = (48, 128)$, $\|x_{\mathrm{ob}}\|_0 = 8$, *i.e.* x_{ob} has 8 non-zero elements.
$\ell_{1,2}$**-norm** $(m, n) = (48, 128)$, x_{ob} has 3 non-zero blocks of size 4.
ℓ_∞**-norm** $(m, n) = (63, 64)$, $|\mathcal{I}_{x_{\mathrm{ob}}}| = 8$ where $\mathcal{I}_{x_{\mathrm{ob}}} = \{i : |x_i| = \|x_{\mathrm{ob}}\|_\infty\}$.
Nuclear norm $(m, n) = (640, 1024)$, $x_{\mathrm{ob}} \in \mathbb{R}^{32\times 32}$ and $\mathrm{rank}(x_{\mathrm{ob}}) = 4$.

For each setting, the number of measurements is sufficiently large so that one can prove that the minimizer $x^\star$ is unique, and in particular that $\ker(\mathcal{K}) \cap T_{x^\star}^R = \{0\}$ (with high probability), see *e.g.* [170]. We also checked that $\ker(\mathcal{K})^\perp \cap S_{x^\star}^R = \{0\}$, which is equivalent to the uniqueness of the fixed point and also implies that $M_{\mathrm{DR}}^\infty = 0$ (see (i) of Lemma 7.4.8). Thus ($\mathrm{ND}_{\mathrm{DR}}$) is fulfilled, and Theorem 7.4.10 applies. DR is run in its stationary version (*i.e.* constant γ).

Figure 7.1 displays the profile of $\|z_k - z^\star\|$ as a function of k, and the starting point of the dashed line is the iteration number at which the active partial smoothness manifold of R is identified (recall that $\mathcal{M}_{x^\star}^J = \mathcal{S}$ which is trivially identified from the first iteration). One can see that for ℓ_1- and ℓ_∞-norms, the observed linear convergence coincides with the optimal rate predicted by (ii) of Theorem 7.4.10. For the case of $\ell_{1,2}$-norm and nuclear norm, though not optimal, our estimates are very tight.

Noise removal In the following two examples, we suppose that we observe $f = x_{\mathrm{ob}} + w$, where x_{ob} is a piecewise-constant vector, and w is an unknown noise supposed to be either uniform or sparse. The goal is to recover x_{ob} from f using the prior information on x_{ob} (*i.e.* piecewise-smooth) and w



(a) ℓ_1-norm (b) ℓ_∞-norm

(c) $\ell_{1,2}$-norm (d) Nuclear norm

Figure 7.1: Observed (solid) and predicted (dashed) convergence profiles of DR (2.4.10) in terms of $\|z_k - z^\star\|$. (a) ℓ_1-norm. (b) ℓ_∞-norm. (c) $\ell_{1,2}$-norm. (d) Nuclear norm. The starting point of the dashed line is the iteration at which the active manifold of R is identified.

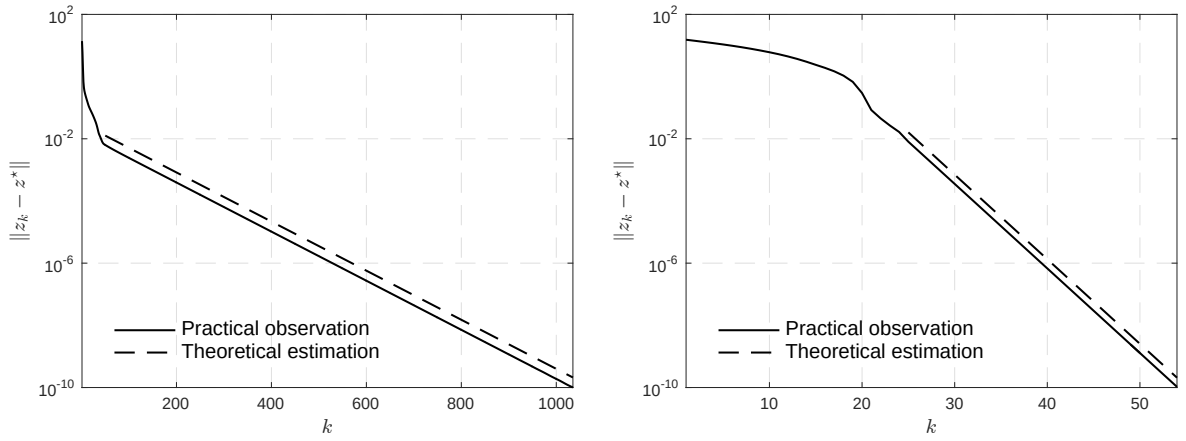
(uniform or sparse). To achieve this goal, a popular and natural approach in the signal processing literature is to solve

$$\min_{x\in\mathbb{R}^n} \|\mathrm{D}_{\mathrm{DIF}}(x)\|_1 \quad \text{subject to} \quad \|f - x\|_p \leq \tau, \tag{7.8.1}$$

where $p = +\infty$ for uniform noise, and $p = 1$ for sparse noise, and $\tau > 0$ is a parameter to be set by the user to adapt to the noise level. Identifying $R = \|\mathrm{D}_{\mathrm{DIF}}(\cdot)\|_1$ and $J = \iota_{\|f-\cdot\|_p\leq\tau}$, one recognises that for $p \in \{1, +\infty\}$, R and J are indeed polyhedral and their proximity operators are simple to compute [67]. For both examples, we set $n = 128$ and x_{ob} is such that $\mathrm{D}_{\mathrm{DIF}}(x_{\mathrm{ob}})$ has 8 nonzero entries. For $p = +\infty$, ϵ is generated uniformly in $[-1, 1]$, and for $p = 1$ ϵ is sparse with 16 nonzero entries. DR is run in its stationary version. The corresponding convergence profiles are depicted in Figure 7.2(a)-(b). Condition ($\mathrm{ND}_{\mathrm{DR}}$) is checked a posteriori, and it is satisfied for the considered examples. Owing to polyhedrality, our predictions are again optimal.

7.8.2 Finite convergence

We now numerically illustrate the finite convergence of DR. For the remainder of this subsection, we set $n = 2$, and solve ($\mathcal{P}_{\mathrm{DR}}$) with $R = \|\cdot\|_1$ and $J = \iota_C$, $C = \{x \in \mathbb{R}^2 : \|x - (3/4 \quad 3/4)^T\|_1 \leq 1/2\}$. The set of minimizers is the segment $[(1/4 \quad 3/4)^T, (3/4 \quad 1/4)^T]$, and R is differentiable at any minimizer with gradient $(1 \quad 1)^T$. The set of fixed points is thus $[(1/4 \quad 3/4)^T, (3/4 \quad 1/4)^T] - \gamma$. Figure 7.3(a)

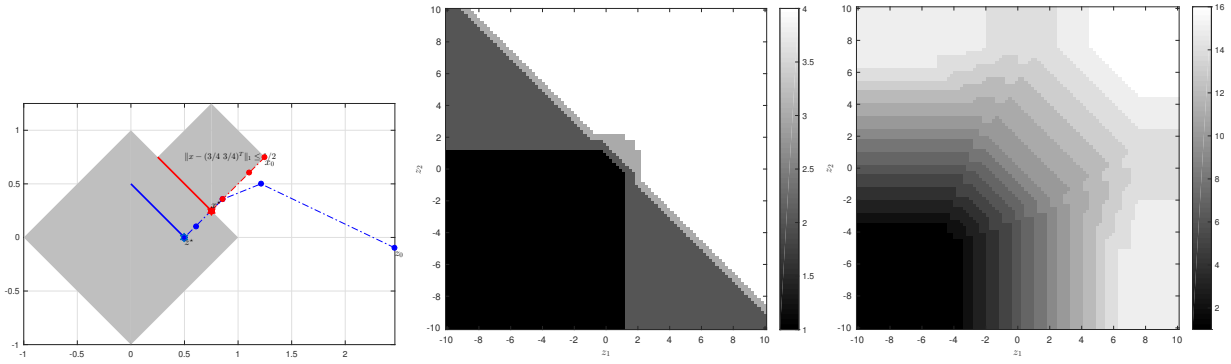


(a) Uniform noise removal $p = +\infty$ (b) Sparse noise removal $p = 1$

Figure 7.2: Observed (solid) and predicted (dashed) convergence profiles of DR (2.4.10) in terms of $\|z_k - z^\star\|$. (a) Uniform noise removal by solving (7.8.1) with $p = +\infty$, (c) Sparse noise removal by solving (7.8.1) with $p = 1$. The starting point of the dashed line is the iteration at which the manifolds $\mathcal{M}^R_{x^\star}$ and $\mathcal{M}^J_{x^\star}$ are identified.

shows the trajectory of the sequence $\{z_k\}_{k\in\mathbb{N}}$ and the shadow sequence $\{x_k\}_{k\in\mathbb{N}}$ which both converge finitely as predicted by Theorem 7.5.1 (DR is used with $\gamma = 0.25$).

For each starting point $z_0 \in [-10, 10]^2$, we run the DR algorithm until $z_{k+1} = z_k$ (up to machine precision), with $\gamma = 0.25$ and $\gamma = 5$. Figure 7.3(b)-(c) show the number of iterations to finite convergence, where $\gamma = 0.25$ for (b) and $\gamma = 5$ for (c). This confirms that DR indeed converges in finitely many iterations regardless of the starting point and choice of γ, though more iterations are needed for higher γ in this example (see next subsection for further discussion on the choice of γ).



(a) (b) $\gamma = 0.25$ (c) $\gamma = 5$

Figure 7.3: (a) Trajectories of $\{z_k\}_{k\in\mathbb{N}}$ and $\{x_k\}_{k\in\mathbb{N}}$. The red segment is the set of minimizers and the blue one is the set of fixed points. (b)-(c) Number of iterations needed for the finite convergence of z_k to $z^\star$. DR is run with $\gamma = 0.25$ for (b) and $\gamma = 5$ for (c).

7.8.3 Choice of γ

Impact of γ on identification We compare the impacts of different values of γ in the DR algorithm. Linear inverse problem (4.5.2) with R being $\ell_1, \ell_{1,2}$ and nuclear norms is are chosen for experiments.

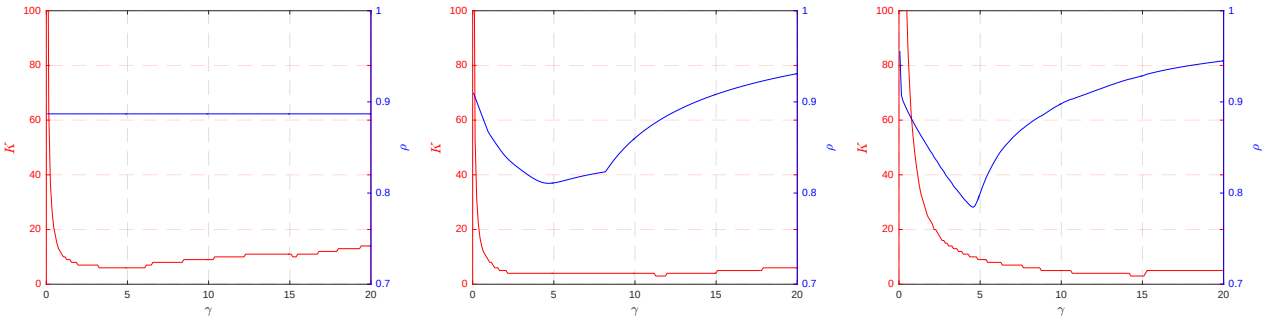
The results are shown in Figure 7.4, where K denotes the number of iterations needed to identify $\mathcal{M}^R_{x^\star}$ and ρ denotes the local linear convergence rate. We summarize our observations as follows:

(i) For all examples, the choice of γ affects the iteration K at which activity identification occurs.

Indeed, K typically decreases monotonically and then either stabilizes or slightly increases. This is in agreement with the bound in (7.3.1).

(ii) When R is ℓ_1-norm, which is polyhedral, the local linear convergence rate is insensitive to γ as anticipated by (ii) of Theorem 7.4.10. For the other two norms, the local rate depends on γ (see (i) of Theorem 7.4.10), and this rate can be optimized for the parameter γ.

(iii) In general, there is no correspondence between the optimal choice of γ for identification and the one for local convergence rate.



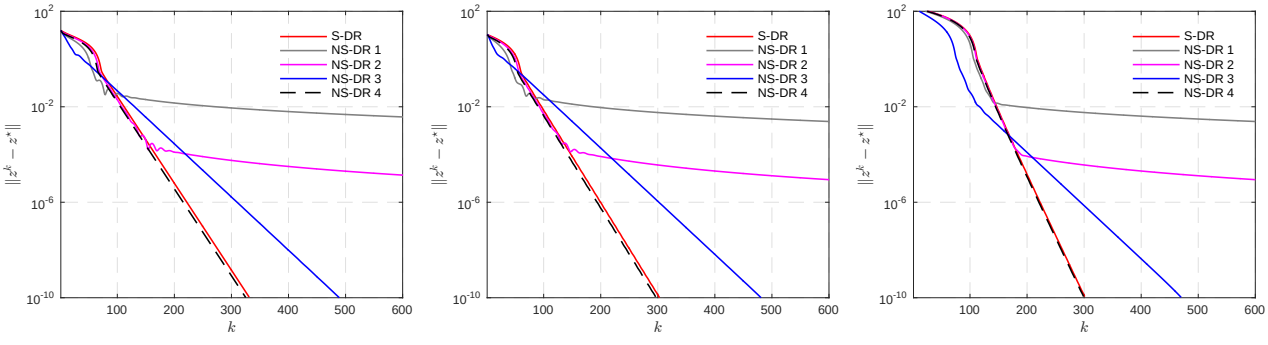
(a) ℓ_1-norm (b) $\ell_{1,2}$-norm (c) Nuclear norm

Figure 7.4: Number of iterations (K) needed for identification and local linear convergence rate (ρ) as a function of γ when solving problem (4.5.2) with different functions R. (a) ℓ_1-norm. (b) $\ell_{1,2}$-norm. (c) Nuclear norm.

Stationary vs non-stationary DR We now investigate numerically the convergence behaviour of the non-stationary version of DR and compare it to the stationary one. We fix $\lambda_k \equiv 1$, *i.e.* the iteration is unrelaxed. The stationary DR algorithm is run with some $\gamma > 0$. For the non-stationary one, four choices of γ_k are considered:

$$\text{Case 1: } \gamma_k = \gamma + \frac{1}{k^{1.1}}, \quad \text{Case 2: } \gamma_k = \gamma + \frac{1}{k^2}, \quad \text{Case 3: } \gamma_k = \gamma + 0.95^k, \quad \text{Case 4: } \gamma_k = \gamma + 0.5^k.$$

Obviously, we have $\{|\gamma_k - \gamma|\}_{k\in\mathbb{N}} \in \ell_+^1$ for all the four cases. Problem (4.5.2) is considered again with R being $\ell_1, \ell_{1,2}$ and the nuclear norms. The comparison results are displayed in Figure 7.5. Table 7.1 shows the number of iteration K needed for the identification of $\mathcal{M}_{x^\star}^R$.



(a) ℓ_1-norm (b) $\ell_{1,2}$-norm (c) Nuclear norm

Figure 7.5: Comparison between stationary ("S-DR") and non-stationary DR ("NS-DR X", X stands for Case X) when solving (4.5.2) with different functions J. (a) ℓ_1-norm. (b) $\ell_{1,2}$-norm. (c) Nuclear norm.

For the stationary iteration, the local convergence rate of the 3 examples are,

$$\ell_1\text{-norm: } \rho = 0.9196, \quad \ell_{1,2}\text{-norm: } \rho = 0.9153, \quad \text{Nuclear norm: } \rho = 0.8904.$$

Table 7.1: Number of iterations K needed for the identification of $\mathcal{M}^R_{x^\star}$ for each tested case. "NS-DR X" stands for the non-stationary DR with choice of γ_k as in Case X.

	S-DR	NS-DR 1	NS-DR 2	NS-DR 3	NS-DR 4
ℓ_1-norm	81	64	77	244	75
$\ell_{1,2}$-norm	62	46	58	233	56
Nuclear norm	114	107	112	77	112

We can make the following observations from the comparison:

(i) The local convergence behaviour of the non-stationary iteration is no better than the stationary one which is in agreement with our analysis.

(ii) As argued in (ii) of Remark 7.4.11, the convergence rate is eventually controlled by the error $|\gamma_k - \gamma|$, except for "Case 4", since 0.5 is strictly smaller than the local linear rate of the stationary version (*i.e.* $|\gamma_k - \gamma| = o(\|z_k - z^\star\|)$).

(iii) The non-stationary DR seems to generally lead to faster identification. But his is not a systematic behaviour as observed for instance for Case 3, where slower identification is obtained for $\ell_1, \ell_{1,2}$-norms.

7.9 Proofs of main theorems

7.9.1 Proofs of Section 7.2

The following lemma is needed in the proof of Proposition 7.2.1.

Lemma 7.9.1. *Suppose that conditions* (**D.5**) *and* (**D.6**) *hold, and that* γ_k *is convergent. Then*

$$\lim_{k\to+\infty} \gamma_k = \gamma.$$

Proof. Since γ_k is convergent, it has a unique cluster point, say $\lim_{k\to+\infty}\gamma_k = \gamma'$. It is then sufficient to show that $\gamma' = \gamma$. Suppose that $\gamma' \neq \gamma$. Fix some $\epsilon \in]0, |\gamma' - \gamma|[$. Thus, there exist an index $K > 0$ such that for all $k \geq K$,

$$|\gamma_k - \gamma'| < \epsilon/2.$$

Therefore

$$|\gamma_k - \gamma| \geq |\gamma' - \gamma| - |\gamma_k - \gamma'| > \epsilon/2.$$

It then follows that

$$\lambda_k(2-\lambda_k)\epsilon \leq 2\lambda_k\epsilon \leq 4\lambda_k|\gamma_k - \gamma|.$$

Denote $\overline{\tau} \stackrel{\text{def}}{=} \sup_{k\mathbb{N}} \lambda_k(2-\lambda_k)$ which is obviously positive and bounded since $\lambda_k \in [0,2]$. Summing both sides for $k \geq K$ we get

$$\begin{aligned}\epsilon \sum\nolimits_{k\in\mathbb{N}} \lambda_k(2-\lambda_k) - K\overline{\tau} \leq \epsilon \sum\nolimits_{k=K}^{+\infty} \lambda_k(2-\lambda_k) &\leq 4\sum\nolimits_{k=K}^{\infty} \lambda_k|\gamma_k - \gamma| \\ &\leq 4\sum\nolimits_{k\in\mathbb{N}} \lambda_k|\gamma_k - \gamma|,\end{aligned}$$

which, in view of (**D.6**), implies

$$\sum\nolimits_{k\in\mathbb{N}} \lambda_k(2-\lambda_k) \leq \epsilon^{-1}(\lambda_k|\gamma_k - \gamma| + K\overline{\tau}) < +\infty,$$

which is a contradiction with (**D.5**). □

Proof of Proposition 7.2.1. To prove our claim, we only need to check the conditions listed in Theorem 3.5.2.

(i) As (**D.3**) assumes the set of minimizers of ($\mathcal{P}_{\mathrm{DR}}$) is nonempty, so is the set $\mathrm{fix}(\mathcal{F}_\gamma)$, since the former is nothing but $\mathrm{prox}_{\gamma J}(\mathrm{fix}(\mathcal{F}_\gamma))$ [16, Proposition 25.1(ii)].

(ii) Since $\mathcal{F}_{\gamma_k}$ is firmly non-expansive by Lemma 2.4.10, $\mathcal{F}_{\gamma_k,\lambda_k}$ is $\frac{\lambda_k}{2}$-averaged, hence non-expansive, owing to Proposition 3.1.2.

(iii) Let $\rho \in [0,+\infty[$ and $z \in \mathbb{R}^n$ such that $\|z\| \leq \rho$, Then we have

$$\begin{aligned}(\mathcal{F}_{\gamma_k} - \mathcal{F}_\gamma)(z) &= \tfrac{\mathcal{R}_{\gamma_k R}\mathcal{R}_{\gamma_k J}}{2}(z) - \tfrac{\mathcal{R}_{\gamma R}\mathcal{R}_{\gamma J}}{2}(z)\\ &= \Big(\tfrac{\mathcal{R}_{\gamma_k R}\mathcal{R}_{\gamma_k J}}{2}(z) - \tfrac{\mathcal{R}_{\gamma_k R}\mathcal{R}_{\gamma J}}{2}(z)\Big) - \Big(\tfrac{\mathcal{R}_{\gamma R}\mathcal{R}_{\gamma J}}{2}(z) - \tfrac{\mathcal{R}_{\gamma_k R}\mathcal{R}_{\gamma J}}{2}(z)\Big)\\ &= \Big(\tfrac{\mathcal{R}_{\gamma_k R}\mathcal{R}_{\gamma_k J}}{2}(z) - \tfrac{\mathcal{R}_{\gamma_k R}\mathcal{R}_{\gamma J}}{2}(z)\Big) - \big(\mathrm{prox}_{\gamma R}\big(\mathcal{R}_{\gamma J}(z)\big) - \mathrm{prox}_{\gamma_k R}\big(\mathcal{R}_{\gamma J}(z)\big)\big)\,.\end{aligned}$$

Thus, by virtue of (ii) of Lemma 2.3.8, we have

$$\|(\mathcal{F}_{\gamma_k} - \mathcal{F}_\gamma)(z)\| \leq \|\mathrm{prox}_{\gamma_k J}(z) - \mathrm{prox}_{\gamma J}(z)\| + \|\mathrm{prox}_{\gamma_k R}(\mathcal{R}_{\gamma J}(z)) - \mathrm{prox}_{\gamma R}(\mathcal{R}_{\gamma J}(z))\|.$$

Let's bound the first term. From the resolvent equation (2.4.1) and Lemma 2.3.8, we have

$$\begin{aligned}\|\mathrm{prox}_{\gamma_k J}(z) - \mathrm{prox}_{\gamma J}(z)\| &= \|\mathrm{prox}_{\gamma_k J}(z) - \mathrm{prox}_{\gamma_k J}\big(\tfrac{\gamma_k}{\gamma}z + \big(1-\tfrac{\gamma_k}{\gamma}\big)\mathrm{prox}_{\gamma J}(z)\big)\|\\ &\leq \frac{|\gamma_k - \gamma|}{\gamma}\|(\mathrm{Id} - \mathrm{prox}_{\gamma J})(z)\| \leq \frac{|\gamma_k - \gamma|}{\gamma}(\rho + \|\mathrm{prox}_{\gamma J}(0)\|). \qquad (7.9.1)\end{aligned}$$

With similar arguments, we also obtain

$$\|\mathrm{prox}_{\gamma_k R}(\mathcal{R}_{\gamma J}(z)) - \mathrm{prox}_{\gamma R}(\mathcal{R}_{\gamma J}(z))\| \leq \frac{|\gamma_k - \gamma|}{\gamma}(\rho + \|\mathrm{prox}_{\gamma R}(0)\| + 2\|\mathrm{prox}_{\gamma J}(0))\|). \qquad (7.9.2)$$

Combining (7.9.1) and (7.9.2) leads to

$$\|(\mathcal{F}_{\gamma_k} - \mathcal{F}_\gamma)(z)\| \leq \frac{|\gamma_k - \gamma|}{\gamma}(2\rho + \|\mathrm{prox}_{\gamma R}(0)\| + 3\|\mathrm{prox}_{\gamma J}(0)\|), \qquad (7.9.3)$$

whence we get

$$\|(\mathcal{F}_{\gamma_k,\lambda_k} - \mathcal{F}_{\gamma,\lambda_k})(z)\| = \lambda_k\|(\mathcal{F}_{\gamma_k} - \mathcal{F}_\gamma)(z)\| \leq \lambda_k\frac{|\gamma_k - \gamma|}{\gamma}(2\rho + \|\mathrm{prox}_{\gamma R}(0)\| + 3\|\mathrm{prox}_{\gamma J}(0)\|).$$

Therefore, from (**D.6**), we deduce that

$$\{\sup_{\|z\|\leq\rho}\|(\mathcal{F}_{\gamma_k,\lambda_k} - \mathcal{F}_{\gamma,\lambda_k})(z)\|\}_{k\in\mathbb{N}} \in \ell^1_+.$$

In other words, the non-stationary iteration (7.2.1) is a perturbed version of the stationary one (7.1.2) with an error term which is summable thanks to (**D.6**). The claim on the convergence of $z^\star$ follows by applying [57, Corollary 5.2]. Moreover, $x^\star \overset{\text{def}}{=} \mathrm{prox}_{\gamma J}(z^\star)$ is a solution of ($\mathcal{P}_{\mathrm{DR}}$). In turn, using non-expansiveness of $\mathrm{prox}_{\gamma_k J}$ and (7.9.1), we have

$$\|x_k - x^\star\| \leq \|z_k - z^\star\| + \frac{|\gamma_k - \gamma|}{\gamma}(\|z^\star\| + \|\mathrm{prox}_{\gamma J}(0)\|),$$

and thus the right hand side goes to zero as $k \to +\infty$ as we are in finite dimension and since $\gamma_k \to \gamma$ owing to Lemma 7.9.1. This entails that the shadow sequence $\{x_k\}_{k\in\mathbb{N}}$ also converges to $x^\star$. With similar arguments, we can also show that $\{u_k\}_{k\in\mathbb{N}}$ converges to $x^\star$ (using for instance (7.9.2) and non-expansiveness of $\mathrm{prox}_{\gamma_k R}$). □

7.9.2 Proofs of Section 7.3

Proof of Theorem 7.3.1. By Proposition 7.2.1, we have the convergence of all the sequences generated by (7.1.3), that is

$$z_k \to z^\star \in \mathrm{fix}(\mathcal{F}_\gamma),\ \ x_k, u_k, \to x^\star = \mathrm{prox}_{\gamma J}(z^\star) \in \mathrm{Argmin}(R+J).$$

The non-degeneracy condition ($\mathrm{ND}_{\mathrm{DR}}$) is equivalent to

$$\frac{x^\star - z^\star}{\gamma} \in \mathrm{ri}\big(\partial R(x^\star)\big) \quad \text{and} \quad \frac{z^\star - x^\star}{\gamma} \in \mathrm{ri}\big(\partial J(x^\star)\big). \qquad (7.9.4)$$

(i) The update of x_{k+1} and u_{k+1} in iteration (7.1.3) is equivalent to the monotone inclusions

$$\frac{2x_k - z_k - u_{k+1}}{\gamma_k} \in \partial R(u_{k+1}) \quad \text{and} \quad \frac{z_k - x_k}{\gamma_k} \in \partial J(x_k).$$

It then follows that

$$\begin{aligned}
&\operatorname{dist}\big(\tfrac{x^\star - z^\star}{\gamma}, \partial R(u_{k+1})\big) \\
&\leq \|\tfrac{x^\star - z^\star}{\gamma} - \tfrac{2x_k - z_k - u_{k+1}}{\gamma_k}\| \\
&= \|\tfrac{(\gamma_k - \gamma)(x^\star - z^\star)}{\gamma\gamma_k} + \tfrac{x^\star - z^\star}{\gamma_k} - \tfrac{2x_k - z_k - u_{k+1}}{\gamma_k}\| \\
&\leq \tfrac{|\gamma_k - \gamma|}{\gamma\underline{\gamma}}\|(\mathrm{Id} - \operatorname{prox}_{\gamma J})(z^\star)\| + \tfrac{1}{\underline{\gamma}}\|(z_k - z^\star) - 2(x_k - x^\star) + (u_{k+1} - x^\star)\| \\
&\leq \tfrac{|\gamma_k - \gamma|}{\gamma\underline{\gamma}}(\|z^\star\| + \operatorname{prox}_{\gamma J}(0)) + \tfrac{1}{\underline{\gamma}}(\|z_k - z^\star\| + 2\|x_k - x^\star\| + \|u_{k+1} - x^\star\|),
\end{aligned}$$

and the right hand side converges to 0 in view of Proposition 7.2.1 and Lemma 7.9.1. Similarly, we have

$$\begin{aligned}
\operatorname{dist}\big(\tfrac{z^\star - x^\star}{\gamma}, \partial J(x_k)\big) &\leq \|\tfrac{z^\star - x^\star}{\gamma} - \tfrac{z_k - x_k}{\gamma_k}\| \\
&= \|\tfrac{(\gamma_k - \gamma)(z^\star - x^\star)}{\gamma\gamma_k} + \tfrac{z^\star - x^\star}{\gamma_k} - \tfrac{z_k - x_k}{\gamma_k}\| \\
&\leq \tfrac{|\gamma_k - \gamma|}{\gamma\underline{\gamma}}(\|z^\star\| + \operatorname{prox}_{\gamma J}(0)) + \tfrac{1}{\underline{\gamma}}(\|z_k - z^\star\| + \|x_k - x^\star\|) \to 0.
\end{aligned}$$

By assumption, $R, J \in \Gamma_0(\mathbb{R}^n)$, hence are subdifferentially continuous at every point in their respective domains [153, Example 13.30], and in particular at $x^\star$. It then follows that $R(u_k) \to R(x^\star)$ and $J(x_k) \to J(x^\star)$. Altogether, this shows that the conditions of Theorem 5.1.5 are fulfilled for R and J, and the finite identification claim follows.

(ii) (a) In this case, $\mathcal{M}^J_{x^\star}$ is an affine subspace, i.e. $\mathcal{M}^J_{x^\star} = x^\star + T^J_{x^\star}$. Since J is partly smooth at $x^\star$ relative to $\mathcal{M}^J_{x^\star}$, the sharpness property holds at all nearby points in $\mathcal{M}^J_{x^\star}$ owing to Lemma 5.1.3. Thus for k large enough, i.e. x_k sufficiently close to $x^\star$ on $\mathcal{M}^J_{x^\star}$, we have indeed $\mathcal{T}_{x_k}(\mathcal{M}^J_{x^\star}) = T^J_{x^\star} = T^J_{x_k}$ as claimed.

(b) Similarly to (ii)(a).

(c) It is immediate to verify that a locally polyhedral function around $x^\star$ is indeed partly smooth relative to the affine subspace $x^\star + T^J_{x^\star}$, and thus, the first claim follows from (ii)(a). For the rest, it is sufficient to observe that by polyhedrality, for any $x \in \mathcal{M}^J_{x^\star}$ near $x^\star$, $\partial J(x) = \partial J(x^\star)$. Therefore, combining local normal sharpness Lemma 5.1.3 and Lemma 6.3.3 yields the second conclusion.

(d) Similarly to (ii)(c). □

Proof of Proposition 7.3.3. From (7.2.1), we have

$$z_{k+1} = \mathcal{F}_{\gamma,\lambda_k}(z_k) + e_k$$

where $\{\|e_k\|\}_{k\in\mathbb{N}} = \{O(\lambda_k|\gamma_k - \gamma|)\}_{k\in\mathbb{N}} \in \ell^1_+$ (see the proof of Proposition 7.2.1). Since $\mathcal{F}_{\gamma_k}$ is firmly non-expansive by Lemma 2.4.10, $\mathcal{F}_{\gamma,\lambda_k}$ is $\frac{\lambda_k}{2}$-averaged non-expansive owing to Proposition 3.1.2. Thus owing to Lemma 3.1.5, we have

$$\begin{aligned}
\|z_k - z^\star\|^2 &\leq \|\mathcal{F}_{\gamma,\lambda_k}(z_{k-1}) - \mathcal{F}_{\gamma,\lambda_k}(z^\star)\|^2 + \mathcal{C}\|e_{k-1}\| \\
&\leq \|\mathcal{F}_{\gamma,\lambda_k}(z_{k-1}) - \mathcal{F}_{\gamma,\lambda_k}(z^\star)\|^2 - \tfrac{2-\lambda_{k-1}}{\lambda_{k-1}}\|z_k - z_{k-1}\|^2 + \mathcal{C}\|e_{k-1}\| \\
&\leq \|z_{k-1} - z^\star\|^2 - \tau_{k-1}\|u_k - x_{k-1}\|^2 + \mathcal{C}\|e_{k-1}\|,
\end{aligned}$$

where $\mathcal{C} < +\infty$ by boundedness of z_k and e_k. Let $g_k = (z_{k-1} - x_{k-1})/\gamma_{k-1}$ and $h_k = (2x_{k-1} - z_{k-1} - u_k)/\gamma_{k-1}$. By definition, we have $(g_k, h_k) \in \partial J(x_{k-1}) \times \partial R(u_k)$. Suppose that neither $\mathcal{M}^J_{x^\star}$ nor $\mathcal{M}^R_{x^\star}$

have been identified at iteration k. That is $x_{k-1} \notin \mathcal{M}_{x^\star}^J$ and $u_k \notin \mathcal{M}_{x^\star}^R$, and by assumption, $g_k \in \mathrm{rbd}(\partial J(x^\star))$ and $h_k \in \mathrm{rbd}(\partial R(x^\star))$, which implies that $g_k + h_k = (u_k - x_{k-1})/\gamma_{k-1} \in \mathrm{rbd}(\partial J(x^\star)) + \mathrm{rbd}(\partial R(x^\star))$. Thus, the above inequality becomes

$$\begin{aligned}\|z_k - z^\star\|^2 &\leq \|z_{k-1} - z^\star\|^2 - \gamma_{k-1}^2\tau_{k-1}\mathrm{dist}\big(0, \mathrm{rbd}(\partial J(x^\star)) + \mathrm{rbd}(\partial R(x^\star))\big)^2 + \mathcal{C}\|e_{k-1}\| \\ &\leq \|z_{k-1} - z^\star\|^2 - \gamma_{k-1}^2\tau_{k-1}\mathrm{dist}\big(0, \mathrm{rbd}(\partial J(x^\star) + \partial R(x^\star))\big)^2 + \mathcal{C}\|e_{k-1}\| \\ &\leq \|z_0 - z^\star\|^2 - k\underline{\gamma}^2\underline{\tau}\mathrm{dist}\big(0, \mathrm{rbd}(\partial J(x^\star) + \partial R(x^\star))\big)^2 + O\big(\textstyle\sum_{k\in\mathbb{N}}\lambda_k|\gamma_k - \gamma|\big),\end{aligned}$$

and $\mathrm{dist}\big(0, \mathrm{rbd}(\partial J(x^\star) + \partial R(x^\star))\big) > 0$ owing to condition ($\mathrm{ND}_{\mathrm{DR}}$). Taking k as the largest integer such that the bound in the right-hand side is positive, we then deduce that the number of iterations where both $\mathcal{M}_{x^\star}^J$ and $\mathcal{M}_{x^\star}^R$ have not been identified yet does not exceed the claimed bound (7.3.1). Thus finite identification necessarily occurs at some k larger than this bound. □

7.9.3 Proofs of Section 7.4

Proof of Proposition 7.4.6. Since $W_{\overline{R}}$ and $W_{\overline{J}}$ are both firmly non-expansive owing to Lemma 7.4.5, it follows from [16, Example 4.7] that the composite matrices $M_{\overline{R}}$ and $M_{\overline{J}}$ are firmly non-expansive. As a result, M_{DR} is firmly non-expansive Lemma 2.4.10, and equivalently that M_λ is $\frac{\lambda}{2}$-averaged by Proposition 3.1.2.

Under the assumptions of Theorem 7.3.1, the identification theorem of DR, there exists $K \in \mathbb{N}$ large enough such that for all $k \geq K$, $(x_k, u_k) \in \mathcal{M}_{x^\star}^J \times \mathcal{M}_{x^\star}^R$. Denote $T_{x_k}^J$ and $T_{x^\star}^J$ be the tangent spaces corresponding to x_k and $x^\star \in \mathcal{M}_{x^\star}^J$, and similarly $T_{x_k}^R$ and $T_{x^\star}^R$ the tangent spaces corresponding to u_k and $x^\star \in \mathcal{M}_{x^\star}^R$. Denote $\tau_k^J : T_{x_k}^J \to T_{x^\star}^J$ (resp. $\tau_k^R : T_{u_k}^R \to T_{x^\star}^R$) the parallel translation along the unique geodesic on $\mathcal{M}_{x^\star}^J$ (resp. $\mathcal{M}_{x^\star}^R$) joining x_k to $x^\star$ (resp. u_k to $x^\star$).

From (7.1.3), for x_k, we have

$$\begin{cases} x_k = \mathrm{prox}_{\gamma_k J}(z_k), \\ x^\star = \mathrm{prox}_{\gamma J}(z^\star), \end{cases} \iff \begin{cases} z_k - x_k \in \gamma_k \partial J(x_k), \\ z^\star - x^\star \in \gamma \partial J(x^\star). \end{cases}$$

Projecting on the corresponding tangent spaces, using Lemma 6.3.3, and applying the parallel translation operator τ_k^J leads to

$$\begin{aligned}\gamma_k\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J} J(x_k) &= \tau_k^J\mathcal{P}_{T_{x_k}^J}(z_k - x_k) = \mathcal{P}_{T_{x^\star}^J}(z_k - x_k) + \big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(z_k - x_k), \\ \gamma\nabla_{\mathcal{M}_{x^\star}^J} J(x^\star) &= \mathcal{P}_{T_{x^\star}^J}(z^\star - x^\star).\end{aligned}$$

We then obtain

$$\begin{aligned}&\gamma_k\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J} J(x_k) - \gamma\nabla_{\mathcal{M}_{x^\star}^J} J(x^\star) \\ &= \gamma\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J} J(x_k) - \gamma\nabla_{\mathcal{M}_{x^\star}^J} J(x^\star) + (\gamma_k - \gamma)\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J} J(x_k) \\ &= \mathcal{P}_{T_{x^\star}^J}\big((z_k - z^\star) - (x_k - x^\star)\big) \\ &\quad + \underbrace{\big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(z_k - x_k - z^\star + x^\star)}_{\textbf{Term 1}} + \underbrace{\big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(z^\star - x^\star)}_{\textbf{Term 2}}.\end{aligned} \tag{7.9.5}$$

For $(\gamma_k - \gamma)\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J} J(x_k)$, since the Riemannian gradient $\nabla_{\mathcal{M}_{x^\star}^J} J(x_k)$ is single-valued and bounded on bounded sets, we have

$$\|(\gamma_k - \gamma)\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J} J(x_k)\| = O(|\gamma_k - \gamma|). \tag{7.9.6}$$

Combining (7.9.1) and Lemma 2.6.2, we have for **Term 1**

$$\big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(z_k - x_k - z^\star + x^\star) = o(\|z_k - z^\star\|) + o(|\gamma_k - \gamma|). \tag{7.9.7}$$

As far as **Term 2** is concerned, with (7.4.1), (7.9.1) and the Riemannian Taylor expansion Lemma 2.6.3, we have

$$\begin{aligned}
&\gamma\tau_k^J\nabla_{\mathcal{M}^J_{x^\star}}J(x_k)-\gamma\nabla_{\mathcal{M}^J_{x^\star}}J(x^\star)-\big(\tau_k^J\mathcal{P}_{T^J_{x_k}}-\mathcal{P}_{T^J_{x^\star}}\big)(z^\star-x^\star)\\
&=\tau_k^J\big(\gamma\nabla_{\mathcal{M}^J_{x^\star}}J(x_k)-\mathcal{P}_{T^J_{x_k}}(z^\star-x^\star)\big)-\big(\gamma\nabla_{\mathcal{M}^J_{x^\star}}J(x^\star)-\mathcal{P}_{T^J_{x^\star}}(z^\star-x^\star)\big)\\
&=\tau_k^J\nabla_{\mathcal{M}^J_{x^\star}}\overline{J}(x_k)-\nabla_{\mathcal{M}^J_{x^\star}}\overline{J}(x^\star)=\mathcal{P}_{T^J_{x^\star}}\nabla^2_{\mathcal{M}^J_{x^\star}}\overline{J}(x^\star)\mathcal{P}_{T^J_{x^\star}}(x_k-x^\star)+o(\|x_k-x^\star\|)\\
&=\mathcal{P}_{T^J_{x^\star}}\nabla^2_{\mathcal{M}^J_{x^\star}}\overline{J}(x^\star)\mathcal{P}_{T^J_{x^\star}}(x_k-x^\star)+o(\|z_k-z^\star\|)+o(|\gamma_k-\gamma|).
\end{aligned}\tag{7.9.8}$$

Therefore, inserting (7.9.6), (7.9.7) and (7.9.8) into (7.9.5), we obtain

$$\begin{aligned}
&H_{\overline{J}}(x_k-x^\star)=\mathcal{P}_{T^J_{x^\star}}(z_k-z^\star)-\mathcal{P}_{T^J_{x^\star}}(x_k-x^\star)+o(\|z_k-z^\star\|)+O(|\gamma_k-\gamma|)\\
\Longrightarrow\quad&(\mathrm{Id}+H_{\overline{J}})\mathcal{P}_{T^J_{x^\star}}(x_k-x^\star)=\mathcal{P}_{T^J_{x^\star}}(z_k-z^\star)+o(\|z_k-z^\star\|)+O(|\gamma_k-\gamma|)\\
\Longrightarrow\quad&\mathcal{P}_{T^J_{x^\star}}(x_k-x^\star)=W_{\overline{J}}\mathcal{P}_{T^J_{x^\star}}(z_k-z^\star)+o(\|z_k-z^\star\|)+O(|\gamma_k-\gamma|)\\
\Longrightarrow\quad&\mathcal{P}_{T^J_{x^\star}}(x_k-x^\star)=\mathcal{P}_{T^J_{x^\star}}W_{\overline{J}}\mathcal{P}_{T^J_{x^\star}}(z_k-z^\star)+o(\|z_k-z^\star\|)+O(|\gamma_k-\gamma|)\\
\Longrightarrow\quad&x_k-x^\star=M_{\overline{J}}(z_k-z^\star)+o(\|z_k-z^\star\|)+O(|\gamma_k-\gamma|),
\end{aligned}\tag{7.9.9}$$

where we used the fact that $x_k-x^\star=\mathcal{P}_{T^J_{x^\star}}(x_k-x^\star)+o(\|x_k-x^\star\|)$ Lemma 2.6.1.

Similarly for u_{k+1}, we have

$$\begin{cases}u_{k+1}=\mathrm{prox}_{\gamma_k R}(2x_k-z_k),\\ x^\star=\mathrm{prox}_{\gamma R}(2x^\star-z^\star),\end{cases}\iff\begin{cases}2x_k-z_k-u_{k+1}\in\gamma\partial R(u_{k+1}),\\ 2x^\star-z^\star-x^\star\in\gamma\partial R(x^\star).\end{cases}$$

Upon projecting onto the corresponding tangent spaces and applying the parallel translation τ^R_{k+1}, we get

$$\begin{aligned}
\gamma_k\tau^R_{k+1}\nabla_{\mathcal{M}^R_{x^\star}}R(u_{k+1})&=\tau^R_{k+1}\mathcal{P}_{T^R_{u_{k+1}}}(2x_k-z_k-u_{k+1})\\
&=\mathcal{P}_{T^R_{x^\star}}(2x_k-z_k-u_{k+1})+\big(\tau^R_{k+1}\mathcal{P}_{T^R_{u_{k+1}}}-\mathcal{P}_{T^R_{x^\star}}\big)(2x_k-z_k-u_{k+1}),\\
\gamma\nabla_{\mathcal{M}^R_{x^\star}}R(x^\star)&=\mathcal{P}_{T^R_{x^\star}}(2x^\star-z^\star-x^\star).
\end{aligned}$$

Subtracting both equations, we obtain

$$\begin{aligned}
&\gamma\tau^R_{k+1}\nabla_{\mathcal{M}^R_{x^\star}}R(u_{k+1})-\gamma\nabla_{\mathcal{M}^R_{x^\star}}R(x^\star)\\
&=\gamma\tau^R_{k+1}\nabla_{\mathcal{M}^R_{x^\star}}R(u_{k+1})-\gamma\nabla_{\mathcal{M}^R_{x^\star}}R(x^\star)+(\gamma_k-\gamma)\tau^R_{k+1}\nabla_{\mathcal{M}^R_{x^\star}}R(u_{k+1})\\
&=\mathcal{P}_{T^R_{x^\star}}\big((2x_k-z_k-u_{k+1})-(2x^\star-z^\star-x^\star)\big)+\underbrace{\big(\tau^R_{k+1}\mathcal{P}_{T^R_{u_{k+1}}}-\mathcal{P}_{T^R_{x^\star}}\big)(x^\star-z^\star)}_{\textbf{Term 4}}\\
&\quad+\underbrace{\big(\tau^R_{k+1}\mathcal{P}_{T^R_{u_{k+1}}}-\mathcal{P}_{T^R_{x^\star}}\big)\big((2x_k-z_k-u_{k+1})-(2x^\star-z^\star-x^\star)\big)}_{\textbf{Term 3}}.
\end{aligned}\tag{7.9.10}$$

As for (7.9.6), we have

$$\|(\gamma_k-\gamma)\tau^R_{k+1}\nabla_{\mathcal{M}^R_{x^\star}}R(u_{k+1})\|=O(|\gamma_k-\gamma|).\tag{7.9.11}$$

With similar arguments to those used for **Term 1**, we have **Term 3** $=o(\|z_k-z^\star\|)+o(|\gamma_k-\gamma|)$. Moreover, similarly to (7.9.8), we have for **Term 4**,

$$\begin{aligned}
&\gamma\tau^R_{k+1}\nabla_{\mathcal{M}^R_{x^\star}}R(u_{k+1})-\gamma\nabla_{\mathcal{M}^R_{x^\star}}R(x^\star)-\big(\tau^R_{k+1}\mathcal{P}_{T^R_{u_{k+1}}}-\mathcal{P}_{T^R_{x^\star}}\big)(x^\star-z^\star)\\
&=\mathcal{P}_{T^R_{x^\star}}\nabla^2_{\mathcal{M}^R_{x^\star}}\overline{R}(x^\star)\mathcal{P}_{T^R_{x^\star}}(u_{k+1}-x^\star)+o(\|z_k-x^\star\|)+o(|\gamma_k-\gamma|).
\end{aligned}\tag{7.9.12}$$

Then for (7.9.10) we have,

$$\begin{aligned}
& H_{\overline{R}}(u_{k+1}-x^\star) = 2\mathcal{P}_{T^R_{x^\star}}(x_k-x^\star) - \mathcal{P}_{T^R_{x^\star}}(z_k-z^\star) - \mathcal{P}_{T^R_{x^\star}}(u_{k+1}-x^\star) \\
& \qquad\qquad + o(\|z_k-z^\star\|) + O(|\gamma_k-\gamma|) \\
\Longrightarrow \quad & (\mathrm{Id}+H_{\overline{R}})\mathcal{P}_{T^R_{x^\star}}(u_{k+1}-x^\star) = 2\mathcal{P}_{T^R_{x^\star}}(x_k-x^\star) - \mathcal{P}_{T^R_{x^\star}}(z_k-z^\star) \\
& \qquad\qquad + o(\|z_k-z^\star\|) + O(|\gamma_k-\gamma|) \\
\Longrightarrow \quad & \mathcal{P}_{T^R_{x^\star}}(u_{k+1}-x^\star) = 2M_{\overline{R}}M_{\overline{J}}(z_k-z^\star) - M_{\overline{R}}(z_k-z^\star) + o(\|z_k-z^\star\|) + O(|\gamma_k-\gamma|) \\
\Longrightarrow \quad & u_{k+1}-x^\star = 2M_{\overline{R}}M_{\overline{J}}(z_k-z^\star) - M_{\overline{R}}(z_k-z^\star) + o(\|z_k-z^\star\|) + O(|\gamma_k-\gamma|),
\end{aligned} \tag{7.9.13}$$

where $u_{k+1}-x^\star = \mathcal{P}_{T^R_{x^\star}}(u_{k+1}-x^\star) + o(\|u_{k+1}-x^\star\|)$ is applied again Lemma 2.6.1.

Summing up (7.9.9) and (7.9.13), we get

$$\begin{aligned}
& (z_k+u_{k+1}-x_k)-(z^\star+x^\star-x^\star) \\
&= (z_k-z^\star)+(u_{k+1}-x^\star)-(x_k-x^\star) \\
&= (\mathrm{Id}+2M_{\overline{R}}M_{\overline{J}}-M_{\overline{R}}-M_{\overline{J}})(z_k-z^\star) + o(\|z_k-z^\star\|) + O(|\gamma_k-\gamma|) \\
&= M_{\mathrm{DR}}(z_k-z^\star) + o(\|z_k-z^\star\|) + O(|\gamma_k-\gamma|).
\end{aligned}$$

Hence for the relaxed DR iteration, we have

$$\begin{aligned}
z_{k+1}-z^\star &= (1-\lambda_k)(z_k-z^\star) + \lambda_k\big((z_k+u_{k+1}-x_k)-(z^\star+x^\star-x^\star)\big) \\
&= (1-\lambda_k)(z_k-z^\star) + \lambda_k M_{\mathrm{DR}}(z_k-z^\star) + o(\|z_k-z^\star\|) + O(\lambda_k|\gamma_k-\gamma|) \\
&= M_\lambda(z_k-z^\star) - (\lambda_k-\lambda)(\mathrm{Id}-M_{\mathrm{DR}})(z_k-z^\star) + o(\|z_k-z^\star\|) + O(\lambda_k|\gamma_k-\gamma|)
\end{aligned}$$

Since $\mathrm{Id}-M_{\mathrm{DR}}$ is also (firmly) non-expansive (Lemma 2.3.8) and $\lambda_k \to \lambda \in]0,2[$, we thus get

$$\begin{aligned}
\lim_{k\to\infty} \frac{\|(\lambda_k-\lambda)(\mathrm{Id}-M_{\mathrm{DR}})(z_k-z^\star)\|}{\|z_k-z^\star\|} &= \lim_{k\to\infty} \frac{|\lambda_k-\lambda|\|(\mathrm{Id}-M_{\mathrm{DR}})(z_k-z^\star)\|}{\|z_k-z^\star\|} \\
&\leq \lim_{k\to\infty} |\lambda_k-\lambda| = 0,
\end{aligned}$$

which means that

$$z_{k+1}-z^\star = M_\lambda(z_k-z^\star) + o(\|z_k-z^\star\|) + O(\lambda_k|\gamma_k-\gamma|),$$

and the claimed result is obtained. □

Proof of Lemma 7.4.8.

(i) Since M_{DR} is firmly non-expansive and M_λ is $\frac{\lambda}{2}$-averaged by Proposition 7.4.6, we deduce from [16, Proposition 5.15] that M_{DR} and M_λ are convergent, and their limit is $M_\lambda^\infty = \mathcal{P}_{\mathrm{fix}(M_\lambda)} = \mathcal{P}_{\mathrm{fix}(M_{\mathrm{DR}})} = M_{\mathrm{DR}}^\infty$ [18, Corollary 2.7(ii)]. Moreover, $M_\lambda^k - M_{\mathrm{DR}}^\infty = (M_\lambda - M_{\mathrm{DR}}^\infty)^k$, $\forall k \in \mathbb{N}$, and $\rho(M_\lambda - M_{\mathrm{DR}}^\infty) < 1$ by [18, Theorem 2.12]. It is also immediate to see that

$$\mathrm{fix}(M_{\mathrm{DR}}) = \ker\big(M_{\overline{R}}(\mathrm{Id}-M_{\overline{J}}) + (\mathrm{Id}-M_{\overline{R}})M_{\overline{J}}\big).$$

Observe that

$$\begin{aligned}
\mathrm{span}(M_{\overline{J}}) \subseteq T^J_{x^\star} \quad &\text{and} \quad \mathrm{span}(M_{\overline{R}}) \subseteq T^R_{x^\star}, \\
\ker(\mathrm{Id}-M_{\overline{R}}) \subseteq T^R_{x^\star} \quad &\text{and} \quad \ker(M_{\overline{R}}) = S^R_{x^\star}, \\
\mathrm{span}\big((\mathrm{Id}-M_{\overline{R}})M_{\overline{J}}\big) \subseteq \mathrm{span}(\mathrm{Id}-M_{\overline{R}}) \quad &\text{and} \quad \mathrm{span}\big(M_{\overline{R}}(\mathrm{Id}-M_{\overline{J}})\big) \subseteq T^R_{x^\star},
\end{aligned}$$

where we used the fact that $W_{\overline{G}}$ and $W_{\overline{J}}$ are positive definite. Therefore, $M_\lambda^\infty = 0$, if and only if, $\mathrm{fix}(M_{\mathrm{DR}}^\infty) = \{0\}$, and for this to hold true, it is sufficient that

$$\begin{aligned}
\mathrm{span}(M_{\overline{J}}) \cap \ker(\mathrm{Id}-M_{\overline{R}}) &\subseteq T^J_{x^\star} \cap T^R_{x^\star} = \{0\}, \\
\mathrm{span}(\mathrm{Id}-M_{\overline{J}}) \cap \ker(M_{\overline{R}}) &= \mathrm{span}(\mathrm{Id}-M_{\overline{J}}) \cap S^R_{x^\star} = \{0\}, \\
\mathrm{span}\big((\mathrm{Id}-M_{\overline{R}})M_{\overline{J}}\big) \cap \mathrm{span}\big(M_{\overline{R}}(\mathrm{Id}-M_{\overline{J}})\big) &\subseteq \mathrm{span}(\mathrm{Id}-M_{\overline{R}}) \cap T^R_{x^\star} = \{0\}.
\end{aligned}$$

(ii) The proof is classical using the spectral radius formula (2.5.1).

(iii) In this case, we have $W_{\overline{R}} = W_{\overline{J}} = \mathrm{Id}$. In turn, $M_{\overline{R}} = \mathcal{P}_{T^R_{x^\star}}$ and $M_{\overline{J}} = \mathcal{P}_{T^J_{x^\star}}$, which yields

$$M_{\mathrm{DR}} = \mathrm{Id} + 2\mathcal{P}_{T^R_{x^\star}}\mathcal{P}_{T^J_{x^\star}} - \mathcal{P}_{T^R_{x^\star}} - \mathcal{P}_{T^J_{x^\star}} = \mathcal{P}_{T^R_{x^\star}}\mathcal{P}_{T^J_{x^\star}} + \mathcal{P}_{S^R_{x^\star}}\mathcal{P}_{S^J_{x^\star}},$$

which is normal, and so is M_λ. From [17, Proposition 3.6(i)], we get that $\mathrm{fix}(M) = (T^J_{x^\star} \cap T^R_{x^\star}) \oplus (S^J_{x^\star} \cap S^R_{x^\star})$. Thus, combining normality, statement (i) and [18, Theorem 2.16] we get that

$$\|M_\lambda^{k+1-K} - M_{\mathrm{DR}}^\infty\| = \|M_\lambda - M_{\mathrm{DR}}^\infty\|^{k+1-K},$$

and $\|M_\lambda - M_{\mathrm{DR}}^\infty\|$ is the optimal convergence rate of M_λ. Combining together [18, Proposition 3.3] and arguments similar to those of the proof of [17, Theorem 3.10(ii)] (see also [18, Theorem 4.1(ii)]), we get indeed that

$$\|M_\lambda - M_{\mathrm{DR}}^\infty\| = \sqrt{(1-\lambda)^2 + \lambda(2-\lambda)\cos^2\big(\theta_F(T^J_{x^\star}, T^R_{x^\star})\big)}.$$

The special case is immediate. This concludes the proof. □

Proof of Corollary 7.4.9. Define $\psi_k \overset{\mathrm{def}}{=} o(\|z_k - z^\star\|)$ and $\phi_k \overset{\mathrm{def}}{=} O(\lambda_k|\gamma_k - \gamma|)$.

(i) Let $K \in \mathbb{N}$ sufficiently large such that the locally linearized iteration (7.4.5) holds. Then we have for $k \geq K$

$$\begin{aligned} z_{k+1} - z^\star &= M_\lambda(z_k - z^\star) + \psi_k + \phi_k \\ &= M_\lambda\big(M_\lambda(z_{k-1} - z^\star) + \psi_{k-1} + \phi_{k-1}\big) + \psi_k + \phi_k \\ &= M_\lambda^{k+1-K}(z_K - z^\star) + \textstyle\sum_{j=K}^k M_\lambda^{k-j}(\psi_j + \phi_j). \end{aligned} \tag{7.9.14}$$

Since $z_k \to z^\star$ from Theorem 7.2.1 and M_λ is convergent to M_{DR}^∞ by Lemma 7.4.8(i), taking the limit as $k \to \infty$, we have for all finite $p \geq K$,

$$\lim_{k\to\infty} \textstyle\sum_{j=p}^k M_\lambda^{k-j}(\psi_j + \phi_j) = -M_{\mathrm{DR}}^\infty(z_p - z^\star). \tag{7.9.15}$$

Using (7.9.15) in (7.9.14), we get

$$\begin{aligned} z_{k+1} - z^\star &= (M_\lambda - M_{\mathrm{DR}}^\infty)(z_k - z^\star) + \psi_k + \phi_k - \lim_{l\to\infty} \textstyle\sum_{j=k}^l M_\lambda^{l-j}(\psi_j + \phi_j) \\ &= (M_\lambda - M_{\mathrm{DR}}^\infty)(z_k - z^\star) + \psi_k + \phi_k - \lim_{l\to\infty} \textstyle\sum_{j=k+1}^l M_\lambda^{l-j}(\psi_j + \phi_j) - M_{\mathrm{DR}}^\infty(\psi_k + \phi_k) \\ &= (M_\lambda - M_{\mathrm{DR}}^\infty)(z_k - z^\star) + (\mathrm{Id} - M_{\mathrm{DR}}^\infty)(\psi_j + \phi_j) + M_{\mathrm{DR}}^\infty(z_{k+1} - z^\star). \end{aligned}$$

It is also immediate to see from Lemma 7.4.8(i) that $\|\mathrm{Id} - M_{\mathrm{DR}}^\infty\| \leq 1$ and

$$(M_\lambda - M_{\mathrm{DR}}^\infty)(\mathrm{Id} - M_{\mathrm{DR}}^\infty) = M_\lambda - M_{\mathrm{DR}}^\infty.$$

Rearranging the terms gives the claimed equivalence.

(ii) Under polyhedrality and constant parameters, we have from Proposition 7.4.6 that both ϕ_k and ψ_k vanish. In this case, (7.9.15) reads

$$z_k - z^\star \in \ker(M_{\mathrm{DR}}^\infty), \qquad \forall k \geq K,$$

and therefore (7.4.5) obviously becomes (7.4.7). □

Proof of Theorem 7.4.10. Recall in the Proof of Corollary 7.4.9 that $\psi_k \overset{\mathrm{def}}{=} o(\|z_k - z^\star\|)$ and $\phi_k \overset{\mathrm{def}}{=} O(\lambda_k|\gamma_k - \gamma|)$.

(i) Let $K \in \mathbb{N}$ sufficiently large such that (7.4.6) holds. We then have from Corollary 7.4.9(i)

$$\begin{aligned} (\mathrm{Id} - M_{\mathrm{DR}}^\infty)(z_{k+1} - z^\star) = (M_\lambda - M_{\mathrm{DR}}^\infty)^{k+1-K}(\mathrm{Id} - M_{\mathrm{DR}}^\infty)(z_K - z^\star) \\ + \textstyle\sum_{j=K}^k (M_\lambda - M_{\mathrm{DR}}^\infty)^{k-j}\big((\mathrm{Id} - M_{\mathrm{DR}}^\infty)\psi_j + \phi_j\big). \end{aligned}$$

Since $\rho(M_\lambda - M_{\mathrm{DR}}^\infty) < 1$ by Lemma 7.4.8(i), from the spectral radius formula, we know that for every $\rho \in]\rho(M_\lambda - M_{\mathrm{DR}}^\infty), 1[$, there is a constant $\mathcal{C}$ such that

$$\|(M_\lambda - M_{\mathrm{DR}}^\infty)^j\| \leq \mathcal{C}\rho^j$$

for all integers j. We thus get

$$\begin{aligned}
&\|(\mathrm{Id} - M_{\mathrm{DR}}^\infty)(z_{k+1} - z^\star)\| \\
&\leq \mathcal{C}\rho^{k+1-K}\|z_K - z^\star\| + \mathcal{C}\sum_{j=K}^k \rho^{k-j}\phi_j + \mathcal{C}\sum_{j=K}^k \rho^{k-j}\|(\mathrm{Id} - M_{\mathrm{DR}}^\infty)\psi_j\| \\
&= \mathcal{C}\rho^{k+1-K}\big(\|z_K - z^\star\| + \rho^{K-1}\sum_{j=K}^k \frac{\phi_j}{\rho^j}\big) + \mathcal{C}\sum_{j=K}^k \rho^{k-j}\|(\mathrm{Id} - M_{\mathrm{DR}}^\infty)\psi_j\|,
\end{aligned} \tag{7.9.16}$$

By assumption, $\phi_j = \mathcal{C}'\eta^j$, for some constant $\mathcal{C}' \geq 0$ and $\eta < \rho$, and we have

$$\rho^{K-1}\sum_{j=K}^k \frac{\phi_j}{\rho^j} \leq \mathcal{C}'\rho^{K-1}\sum_{j=K}^\infty (\eta/\rho)^j = \frac{\mathcal{C}'\eta^K}{\rho - \eta} < +\infty.$$

Setting $\mathcal{C}'' = \mathcal{C}\left(\|z_K - z^\star\| + \frac{\mathcal{C}'\eta^K}{\rho - \eta}\right) < +\infty$, we obtain

$$\|(\mathrm{Id} - M_{\mathrm{DR}}^\infty)(z_{k+1} - z^\star)\| \leq \mathcal{C}''\rho^{k+1-K} + \mathcal{C}\sum_{j=K}^k \rho^{k-j}\|(\mathrm{Id} - M_{\mathrm{DR}}^\infty)\psi_j\|.$$

This, together with the fact that $\|(\mathrm{Id} - M_{\mathrm{DR}}^\infty)\psi_j\| = o(\|(\mathrm{Id} - M_{\mathrm{DR}}^\infty)(z_j - z^\star)\|)$ yields the claimed result.

(ii) From Corollary 7.4.9(ii), we have

$$z_k - z^\star = (M_\lambda - M_{\mathrm{DR}}^\infty)^{k+1-K}(z_K - z^\star).$$

Moreover, by virtue of Lemma 7.4.8(ii), M_λ is normal and converges linearly to

$$M_{\mathrm{DR}}^\infty = \mathcal{P}_{(T_{x^\star}^J \cap T_{x^\star}^G)\oplus(S_{x^\star}^J \cap S_{x^\star}^G)}$$

at the optimal rate

$$\rho = \|M_\lambda - M_{\mathrm{DR}}^\infty\| = \sqrt{(1-\lambda)^2 + \lambda(2-\lambda)\cos^2\left(\theta_F(T_{x^\star}^J, T_{x^\star}^G)\right)}.$$

Combining all this then entails

$$\begin{aligned}
\|z_{k+1} - z^\star\| &\leq \|(M_\lambda - M_{\mathrm{DR}}^\infty)^{k+1-K}\|(z_K - z^\star) \\
&= \|M_\lambda - M_{\mathrm{DR}}^\infty\|^{k+1-K}\|z_K - z^\star\| \\
&= \rho^{k+1-K}\|z_K - z^\star\|,
\end{aligned}$$

and we conclude the proof. □

Proof of Lemma 7.4.12. When η is real, *i.e.* $\eta_{\mathrm{Im}} = 0$, then (7.4.10) can be simplified as

$$|\rho_{\eta,\lambda}| = \begin{cases} (\eta - 1)\lambda + 1 : \lambda \in \left[0, \frac{1}{1-\eta}\right], \\ (1-\eta)\lambda - 1 : \lambda \in \left[\frac{1}{1-\eta}, +\infty\right]. \end{cases}$$

Next we focus on the case $\lambda \in [0, \frac{1}{1-\eta}]$ for which $|\rho_{\eta,\lambda}|$ is a *decreasing* function of λ. Let $\lambda > 0$, and define the function $f(\eta) = (\eta - 1)\lambda + 1$. Then given any eigenvalue η of M_{DR}, we have $\eta \in [\underline{\eta}, \bar{\eta}]$, and moreover

$$f(\underline{\eta}) \leq f(\eta) \leq f(\bar{\eta}).$$

Therefore, we have $\rho(M_\lambda - M_{\mathrm{DR}}^\infty) = \max\{|f(\underline{\eta})|, |f(\bar{\eta})|\}$, and

$$\max\{|f(\underline{\eta})|, |f(\bar{\eta})|\} = \begin{cases} \text{monotone decreasing} : \lambda \in \left[0, \frac{2}{2 - (\bar{\eta} + \underline{\eta})}\right], \\ \text{monotone increasing} : \lambda \in \left[\frac{2}{2 - (\bar{\eta} + \underline{\eta})}, +\infty\right]. \end{cases}$$

Clearly, the minimum of $\rho(M_\lambda - M_{\mathrm{DR}}^\infty)$ is attained at $\lambda = \frac{2}{2-(\bar{\eta}+\underline{\eta})}$. □

7.9.4 Proofs of Section 7.6

Proof of Corollary 7.6.1. At convergence, owing to (7.6.4), we have

$$y^\star = z^\star - \gamma L u^\star = z^\star - \gamma x^\star, \quad -L^* y^\star = L^*(z^\star - 2y^\star - \gamma L u^\star).$$

Then from the update of u_{k+1}, x_{k+1} in (7.6.2), we have the following monotone inclusions

$$\begin{aligned} L^*(z_k - 2y_k - \gamma L u_{k+1}) &\in \partial R(u_{k+1}) \qquad \text{and} \quad z_k - \gamma x_k \in \partial J(x_k), \\ L^*(z^\star - 2y^\star - \gamma L u^\star) &\in \partial R(u^\star) \qquad \text{and} \quad z^\star - \gamma x^\star \in \partial J(x^\star), \end{aligned}$$

Since L is bounded injective linear operator, it then follows that

$$\begin{aligned} \operatorname{dist}\big(-L^* y^\star, \partial R(u_{k+1})\big) &\leq \|L^*(z_k - 2y_k - \gamma L u_{k+1}) - L^*(z^\star - 2y^\star - \gamma L u^\star)\| \\ &\leq \|L\|\|(z_k - z^\star) - 2(y_k - y^\star) - \gamma L(u_{k+1} - u^\star)\| \\ &\leq \|L\|\big(\|z_k - z^\star\| + 2\|y_k - y^\star\| + \gamma\|L\|\|u_{k+1} - u^\star\|\big) \to 0. \end{aligned}$$

and

$$\operatorname{dist}\big(y^\star, \partial J(x_k)\big) \leq \|(z_k - \gamma x_k) - (z^\star - \gamma x^\star)\| \leq \|z_k - z^\star\| + \gamma\|x_k - x^\star\| \to 0.$$

$G \in \Gamma_0(\mathbb{R}^n)$ and $JL \in \Gamma_0(\mathbb{R}^m)$, then by the sub-differentially continuous property of them we have $R(u_k) \to R(u^\star)$ and $J(x_k) \to J(x^\star)$. Hence the conditions of Theorem 5.1.5 are fulfilled for R and J, and the finite identification claim follows.

The rest of the proof follows the proof of Theorem 7.3.1. □

Proof of Corollary 7.6.2.

(i) Similarly to (7.4.3), we have $W_{\overline{\overline{J}}}$ is firmly non-expansive. Then for $W_{\overline{\overline{R}}}$, since L is injective, so is L_R, then both $L_R^T L_R$ and $(L_R^T L_R)^{-1}$ are *symmetric positive definite*. Therefore, we have the following similarity result for $W_{\overline{\overline{R}}}$,

$$\begin{aligned} W_{\overline{\overline{R}}} &= L_R\Big((L_R^T L_R)^{-\frac{1}{2}}\big(\mathrm{Id} + (L_R^T L_R)^{-\frac{1}{2}} H_{\overline{\overline{R}}}(L_R^T L_R)^{-\frac{1}{2}}\big)(L_R^T L_R)^{\frac{1}{2}}\Big)^{-1}(L_R^T L_R)^{-1} L_R \\ &= L_R(L_R^T L_R)^{-\frac{1}{2}}\big(\mathrm{Id} + (L_R^T L_R)^{-\frac{1}{2}} H_{\overline{\overline{R}}}(L_R^T L_R)^{-\frac{1}{2}}\big)^{-1}(L_R^T L_R)^{\frac{1}{2}}(L_R^T L_R)^{-1} L_R \\ &= L_R(L_R^T L_R)^{-\frac{1}{2}}\big(\mathrm{Id} + (L_R^T L_R)^{-\frac{1}{2}} H_{\overline{\overline{R}}}(L_R^T L_R)^{-\frac{1}{2}}\big)^{-1}(L_R^T L_R)^{-\frac{1}{2}} L_R. \end{aligned} \tag{7.9.17}$$

Since $(L_R^T L_R)^{-\frac{1}{2}} H_{\overline{\overline{R}}}(L_R^T L_R)^{-\frac{1}{2}}$ is symmetric positive definite, then the matrix

$$(\mathrm{Id} + (L_R^T L_R)^{-\frac{1}{2}} H_{\overline{\overline{R}}}(L_R^T L_R)^{-\frac{1}{2}})^{-1}$$

is firmly non-expansive. It is easy to show that matrix $\|L_R(L_R^T L_R)^{-\frac{1}{2}}\| \leq 1$, then owing to [16, Corollary 4.6], $W_{\overline{\overline{R}}}$ is firmly non-expansive.

(ii) Under the assumptions of Corollary 7.6.1, the identification theorem of ADMM, there exists $K \in \mathbb{N}$ large enough such that for all $k \geq K$, $(x_k, u_k) \in \mathcal{M}_{x^\star}^J \times \mathcal{M}_{u^\star}^R$. Denote $T_{x_k}^J$ and $T_{x^\star}^J$ be the tangent spaces corresponding to x_k and $x^\star \in \mathcal{M}_{x^\star}^J$, and similarly $T_{u_k}^R$ and $T_{u^\star}^R$ the tangent spaces corresponding to u_k and $u^\star \in \mathcal{M}_{u^\star}^R$. Denote $\tau_k^J : T_{x_k}^J \to T_{x^\star}^J$ (resp. $\tau_k^R : T_{u_k}^R \to T_{u^\star}^R$) the parallel translation along the unique geodesic on $\mathcal{M}_{x^\star}^J$ (resp. $\mathcal{M}_{u^\star}^R$) joining x_k to $x^\star$ (resp. u_k to $x^\star$). For convenience, define $\gamma' = 1/\gamma$. From (7.6.2), for x_k, we have

$$\begin{cases} x_k = \operatorname{prox}_{\gamma' J}(\gamma' z_k), \\ x^\star = \operatorname{prox}_{\gamma' J}(\gamma' z^\star), \end{cases} \iff \begin{cases} \gamma' z_k - x_k \in \gamma' \partial J(x_k), \\ \gamma' z^\star - x^\star \in \gamma' \partial J(x^\star). \end{cases}$$

Projecting on the corresponding tangent spaces, using Lemma 6.3.3, and applying the parallel

translation operator τ_k^J leads to

$$\begin{aligned}\gamma'\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J}J(x_k) &= \tau_k^J\mathcal{P}_{T_{x_k}^J}(\gamma' z_k - x_k)\\ &= \mathcal{P}_{T_{x^\star}^J}(\gamma' z_k - x_k) + \big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(\gamma' z_k - x_k),\\ \gamma'\nabla_{\mathcal{M}_{x^\star}^J}J(x^\star) &= \mathcal{P}_{T_{x^\star}^J}(\gamma' z^\star - x^\star).\end{aligned}$$

We then obtain

$$\begin{aligned}&\gamma'\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J}J(x_k) - \gamma'\nabla_{\mathcal{M}_{x^\star}^J}J(x^\star)\\ &= \mathcal{P}_{T_{x^\star}^J}\big((\gamma' z_k - \gamma' z^\star) - (x_k - x^\star)\big)\\ &\quad + \underbrace{\big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(\gamma' z_k - x_k - \gamma' z^\star + x^\star)}_{\textbf{Term 1}} + \underbrace{\big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(\gamma' z^\star - x^\star)}_{\textbf{Term 2}}.\end{aligned} \tag{7.9.18}$$

Combining (7.9.1) and Lemma 2.6.2, we have for **Term 1**

$$(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J})(\gamma' z_k - x_k - \gamma' z^\star + x^\star) = o(\gamma'\|z_k - z^\star\|). \tag{7.9.19}$$

As far as **Term 2** is concerned, with (7.6.5), (7.9.1) and the Riemannian Taylor expansion Lemma 2.6.3, and recall that $\gamma' y^\star = \gamma' z^\star - x^\star$, we have

$$\begin{aligned}&\gamma'\tau_k^J\nabla_{\mathcal{M}_{x^\star}^J}J(x_k) - \gamma'\nabla_{\mathcal{M}_{x^\star}^J}J(x^\star) - \big(\tau_k^J\mathcal{P}_{T_{x_k}^J} - \mathcal{P}_{T_{x^\star}^J}\big)(\gamma' z^\star - x^\star)\\ &= \tau_k^J\big(\gamma'\nabla_{\mathcal{M}_{x^\star}^J}J(x_k) - \mathcal{P}_{T_{x_k}^J}(\gamma' z^\star - x^\star)\big) - \big(\gamma'\nabla_{\mathcal{M}_{x^\star}^J}J(x^\star) - \mathcal{P}_{T_{x^\star}^J}(\gamma' z^\star - x^\star)\big)\\ &= \tau_k^J\nabla_{\mathcal{M}_{x^\star}^J}\overline{\overline{J}}(x_k) - \nabla_{\mathcal{M}_{x^\star}^J}\overline{\overline{J}}(x^\star) = \mathcal{P}_{T_{x^\star}^J}\nabla^2_{\mathcal{M}_{x^\star}^J}\overline{\overline{J}}(x^\star)\mathcal{P}_{T_{x^\star}^J}(x_k - x^\star) + o(\|x_k - x^\star\|)\\ &= \mathcal{P}_{T_{x^\star}^J}\nabla^2_{\mathcal{M}_{x^\star}^J}\overline{\overline{J}}(x^\star)\mathcal{P}_{T_{x^\star}^J}(x_k - x^\star) + o(\gamma'\|z_k - z^\star\|).\end{aligned} \tag{7.9.20}$$

Therefore, inserting (7.9.19) and (7.9.20) into (7.9.18), we obtain

$$\begin{aligned}&H_{\overline{\overline{J}}}(x_k - x^\star) = \gamma'\mathcal{P}_{T_{x^\star}^J}(z_k - z^\star) - \mathcal{P}_{T_{x^\star}^J}(x_k - x^\star) + o(\gamma'\|z_k - z^\star\|)\\ \Longrightarrow\quad &(\mathrm{Id} + H_{\overline{\overline{J}}})\mathcal{P}_{T_{x^\star}^J}(x_k - x^\star) = \gamma'\mathcal{P}_{T_{x^\star}^J}(z_k - z^\star) + o(\gamma'\|z_k - z^\star\|)\\ \Longrightarrow\quad &\mathcal{P}_{T_{x^\star}^J}(x_k - x^\star) = \gamma' W_{\overline{\overline{J}}}\mathcal{P}_{T_{x^\star}^J}(z_k - z^\star) + o(\gamma'\|z_k - z^\star\|)\\ \Longrightarrow\quad &\mathcal{P}_{T_{x^\star}^J}(x_k - x^\star) = \gamma'\mathcal{P}_{T_{x^\star}^J}W_{\overline{\overline{J}}}\mathcal{P}_{T_{x^\star}^J}(z_k - z^\star) + o(\gamma'\|z_k - z^\star\|)\\ \Longrightarrow\quad &x_k - x^\star = \gamma' M_{\overline{\overline{J}}}(z_k - z^\star) + o(\gamma'\|z_k - z^\star\|),\end{aligned} \tag{7.9.21}$$

where we used the fact that $x_k - x^\star = \mathcal{P}_{T_{x^\star}^J}(x_k - x^\star) + o(\|x_k - x^\star\|)$ Lemma 2.6.1. Similarly for u_{k+1}, we have

$$\begin{cases} L^*(z_k - 2y_k - \gamma L u_{k+1}) \in \partial R(u_{k+1}),\\ L^*(z^\star - 2y^\star - \gamma L u^\star) \in \partial R(u^\star),\end{cases}$$

Upon projecting onto the corresponding tangent spaces and applying the parallel translation τ_{k+1}^R, we get

$$\begin{aligned}\tau_{k+1}^R\nabla_{\mathcal{M}_{u^\star}^R}R(u_{k+1}) &= \tau_{k+1}^R\mathcal{P}_{T_{u_{k+1}}^R}\big(L^*(z_k - 2y_k - \gamma L u_{k+1})\big)\\ &= \mathcal{P}_{T_{u^\star}^R}\big(L^*(z_k - 2y_k - \gamma L u_{k+1})\big)\\ &\quad + \big(\tau_{k+1}^R\mathcal{P}_{T_{u_{k+1}}^R} - \mathcal{P}_{T_{u^\star}^R}\big)\big(L^*(z_k - 2y_k - \gamma L u_{k+1})\big),\\ \nabla_{\mathcal{M}_{u^\star}^R}R(u^\star) &= \mathcal{P}_{T_{u^\star}^R}\big(L^*(z^\star - 2y^\star - \gamma L u^\star)\big).\end{aligned}$$

Subtracting both equations, we obtain

$$\begin{aligned}
&\tau_{k+1}^R \nabla_{\mathcal{M}_{u^\star}^R} R(u_{k+1}) - \nabla_{\mathcal{M}_{u^\star}^R} R(u^\star) \\
&= \mathcal{P}_{T_{u^\star}^R}\big(L^*(z_k - 2y_k - \gamma L u_{k+1}) - L^*(z^\star - 2y^\star - \gamma L u^\star)\big) \\
&\quad + \underbrace{\big(\tau_{k+1}^R \mathcal{P}_{T_{u_{k+1}}^R} - \mathcal{P}_{T_{u^\star}^R}\big)\big(L^*(z_k - 2y_k - \gamma L u_{k+1}) - L^*(z^\star - 2y^\star - \gamma L u^\star)\big)}_{\textbf{Term 3}} \\
&\quad + \underbrace{\big(\tau_{k+1}^R \mathcal{P}_{T_{u_{k+1}}^R} - \mathcal{P}_{T_{u^\star}^R}\big)\big(L^*(z^\star - 2y^\star - \gamma L u^\star)\big)}_{\textbf{Term 4}}.
\end{aligned} \tag{7.9.22}$$

Owing to (7.6.3), we have

$$\begin{aligned}
\|(y_k + \gamma L u_{k+1}) - (y^\star + \gamma L u^\star)\| = \|z_{k+1} - z^\star\| \le \|z_k - z^\star\|, \\
\|(z_k - y_k) - (z^\star - y^\star)\| = \gamma\|x_k - x^\star\| \le \gamma\|\gamma'(z_k - z^\star)\| = \|z_k - z^\star\|.
\end{aligned}$$

Hence, with similar arguments to those used for **Term 1**, we have **Term 3** $= o(\|z_k - z^\star\|)$. Moreover, as $z^\star - 2y^\star - \gamma L u^\star = -y^\star$, then similarly to (7.9.20), we have for **Term 4**,

$$\begin{aligned}
&\tau_{k+1}^R \nabla_{\mathcal{M}_{u^\star}^R} R(u_{k+1}) - \nabla_{\mathcal{M}_{u^\star}^R} R(u^\star) + \big(\tau_{k+1}^R \mathcal{P}_{T_{u_{k+1}}^R} - \mathcal{P}_{T_{u^\star}^R}\big) L^T y^\star \\
&= \gamma \mathcal{P}_{T_{u^\star}^R} \nabla^2_{\mathcal{M}_{u^\star}^R} \overline{R}(u^\star) \mathcal{P}_{T_{u^\star}^R}(u_{k+1} - u^\star) + o(\|z_k - z^\star\|).
\end{aligned} \tag{7.9.23}$$

Then for (7.9.22) we have,

$$\begin{aligned}
H_{\overline{R}}(u_{k+1} - u^\star) &= \gamma' \mathcal{P}_{T_{u^\star}^R} L^*\big((z_k - 2y_k - \gamma L u_{k+1}) - (z^\star - 2y^\star - \gamma L u^\star)\big) + o(\gamma'\|z_k - z^\star\|) \\
&= \gamma' L_R^T\big((z_k - 2y_k - \gamma L u_{k+1}) - (z^\star - 2y^\star - \gamma L u^\star)\big) + o(\gamma'\|z_k - z^\star\|) \\
&= \gamma' L_R^T(z_k - z^\star) - \gamma' L_R^T(2y_k - 2y^\star) - L_R^T L\big(u_{k+1} - u^\star\big) + o(\gamma'\|z_k - z^\star\|) \\
&= \gamma' L_R^T(z_k - z^\star) - \gamma' L_R^T(2y_k - 2y^\star) - L_R^T L_R(u_{k+1} - u^\star) + o(\gamma'\|z_k - z^\star\|),
\end{aligned}$$

which leads to, (recall that $y_k = z_k - \gamma x_k$)

$$\begin{aligned}
&\big(\mathrm{Id} + (L_R^T L_R)^{-1} H_{\overline{R}}\big)(u_{k+1} - u^\star) \\
&= \gamma'(L_R^T L_R)^{-1} L_R^T(z_k - z^\star) - \gamma'(L_R^T L_R)^{-1} L_R^T(2y_k - 2y^\star) + o(\gamma'\|z_k - z^\star\|) \\
&= -\gamma'(L_R^T L_R)^{-1} L_R^T(z_k - z^\star) + 2(L_R^T L_R)^{-1} L_R^T(x_k - x^\star) + o(\gamma'\|z_k - z^\star\|),
\end{aligned}$$

since $V_{\overline{R}} = (\mathrm{Id} + (L_R^T L_R)^{-1} H_{\overline{R}})^{-1}$, then the above equation can be further reformulated as

$$\begin{aligned}
&\gamma L_R(u_{k+1} - u^\star) \\
&= -L_R V_{\overline{R}}(L_R^T L_R)^{-1} L_R^T(z_k - z^\star) + 2\gamma L_R V_{\overline{R}}(L_R^T L_R)^{-1} L_R^T(x_k - x^\star) + o(\|z_k - z^\star\|) \\
&= -W_{\overline{R}}(z_k - z^\star) + 2\gamma W_{\overline{R}}(x_k - x^\star) + o(\|z_k - z^\star\|),
\end{aligned} \tag{7.9.24}$$

Summing up (7.9.21) and (7.9.24), we get

$$\begin{aligned}
z_{k+1} - z^\star &= (z_k - \gamma x_k + \gamma L u_{k+1}) - (z^\star - \gamma x^\star + \gamma L u^\star) \\
&= (z_k - z^\star) - \gamma(x_k - x^\star) + \gamma L(u_{k+1} - u^\star) \\
&= (z_k - z^\star) - M_{\overline{J}}(z_k - z^\star) + o(\|z_k - z^\star\|) \\
&\qquad - W_{\overline{R}}(z_k - z^\star) + 2\gamma W_{\overline{R}}(x_k - x^\star) + o(\|z_k - z^\star\|) \\
&= (z_k - z^\star) - M_{\overline{J}}(z_k - z^\star) + o(\|z_k - z^\star\|) - W_{\overline{R}}(z_k - z^\star) + 2W_{\overline{R}} M_{\overline{J}}(z_k - z^\star) \\
&= \big(\mathrm{Id} - M_{\overline{J}} - W_{\overline{R}} + 2W_{\overline{R}} M_{\overline{J}}\big)(z_k - z^\star) - M_{\overline{J}}(z_k - z^\star) + o(\|z_k - z^\star\|) \\
&= M_{\mathrm{ADMM}}(z_k - z^\star) + o(\|z_k - z^\star\|).
\end{aligned}$$

(iii) The convergence rate of sequence $\{z_k\}_{k\in\mathbb{N}}$ following the proof of Theorem 7.4.10. In the following, we simply derive the form of M_{ADMM} when both R and J are locally polyhedral around $u^\star$ and

$x^\star$ respectively. For this case, $H_{\overline{\overline{R}}}$ and $H_{\overline{\overline{J}}}$ vanish and then $W_{\overline{\overline{R}}}, W_{\overline{\overline{J}}}$ become

$$W_{\overline{\overline{R}}} \overset{\text{def}}{=} L_R(L_R^T L_R)^{-1} L_R^T \qquad \text{and} \quad W_{\overline{\overline{J}}} \overset{\text{def}}{=} \mathcal{P}_{T^J_{x^\star}} x^\star,$$

where $W_{\overline{\overline{R}}}$ now is the projection operator opto the subspace $T^{R,L}_{u^\star}$. As a result, we have

$$M_{\text{ADMM}} = \mathcal{P}_{T^{R,L}_{u^\star}} \mathcal{P}_{T^J_{x^\star}} + (\text{Id} - \mathcal{P}_{T^{R,L}_{u^\star}}) \mathcal{P}_{S^J_{x^\star}}.$$

The optimal convergence result then follows Theorem 7.4.10.

(iv) When $M^\infty_{\text{ADMM}} = 0$ for R, J are locally polyhedral around $u^\star, x^\star$ respectively, we have

$$\|z_{k+1} - z^\star\| \leq O(\rho^{k-K})$$

hold for any $\rho \in]\rho(M_{\text{ADMM}}), 1[$.
We have from (7.6.2) that $x_k = \text{prox}_{J/\gamma}(z_k/\gamma)$, then owing to the non-expansiveness of $\text{prox}_{J/\gamma}$ (Lemma 2.3.8), we have

$$\|x_{k+1} - x^\star\| = \|\text{prox}_{J/\gamma}(z_{k+1}/\gamma) - \text{prox}_{J/\gamma}(z^\star/\gamma)\| \leq \frac{1}{\gamma}\|z_{k+1} - z^\star\|.$$

Then as $y_{k+1} = \text{prox}_{\gamma J^*}(z_{k+1})$ and $\gamma L u_{k+1} = z_{k+1} - y_k$, we have

$$\begin{aligned} \|y_{k+1} - y^\star\| &= \|\text{prox}_{\gamma J^*}(z_{k+1}) - \text{prox}_{\gamma J^*}(z^\star)\| \leq \|z_{k+1} - z^\star\|, \\ \|\gamma L(u_{k+1} - u^\star)\| &= \|(z_{k+1} - y_k) - (z^\star - y^\star)\| \leq \|z_{k+1} - z^\star\| + \|y_k - y^\star\|, \end{aligned} \tag{7.9.25}$$

which leads to the claimed result. □

Chapter 8

Primal–Dual Splitting Methods under Partial Smoothness

Main contributions of this chapter

- Finite activity identification (Theorem 8.2.1) and local linear convergence (Theorem 8.3.6) for the class of Primal–Dual splitting methods.

Contents

In this chapter, we consider a versatile class of Primal-Dual splitting methods for minimizing composite non-smooth optimization problems. Under the assumption that the non-smooth components of the problem are partly smooth relative to smooth manifolds, we present a local convergence analysis framework for Primal–Dual splitting methods. Similarly to the previous two chapters, we first show that (i) the sequences generated by Primal–Dual splitting methods identify the smooth manifolds in finite number of iterations. (ii) then enter a local linear convergence regime, which is characterized in terms of the structure of the underlying active smooth manifolds, the involved functions and linear operators. We also discuss connections between Primal–Dual splitting on the one hand, and the FB and DR/ADMM on the other hand. Moreover, practical acceleration techniques for the class of Primal–Dual splitting methods are discussed.

8.1 Introduction

8.1.1 Problem statement

Let us recall the optimization problem considered in Chapter 6, and add an infimal convolution to ($\mathcal{P}_{\text{opt}}$), and consider the following problem, given in its *primal* form,

$$\min_{x\in\mathbb{R}^n} R(x) + F(x) + (J \mathbin{\dot{\vee}} G)(Lx), \tag{$\mathcal{P}_{\text{Primal}}$}$$

where

(**P.1**) $R, F \in \Gamma_0(\mathbb{R}^n)$, and ∇F is $(1/\beta_{_F})$-Lipschitz continuous for some $\beta_{_F} > 0$.

(**P.2**) $J, G \in \Gamma_0(\mathbb{R}^m)$, G is differentiable and $\beta_{_G}$-strongly convex for $\beta_{_G} > 0$, and $(J \mathbin{\dot{\vee}} G)$ is the infimal convolution of J and G (see Definition 2.1.5).

(**P.3**) $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping.

The main difficulties of solving such problem are that the objective function is non-smooth, and the presence of a linear operator and an infimal convolution. To decouple the composition and make the problem easier to solve, it is natural to consider the *saddle-point problem* associated to ($\mathcal{P}_{\text{Primal}}$),

$$\min_{x\in\mathbb{R}^n}\max_{v\in\mathbb{R}^m} R(x)+F(x)+\langle Lx,\, v\rangle-\big(J^*(v)+G^*(v)\big), \tag{$\mathcal{P}_{\text{SP}}$}$$

where J^*, G^* are the Legendre-Fenchel conjugates of J and G respectively. If we fully dualize ($\mathcal{P}_{\text{Primal}}$), then we obtain its *Fenchel-Rockafellar dual form* (Definition 2.1.18 and Theorem 2.1.19),

$$\min_{v\in\mathbb{R}^m} R^*(v)+G^*(v)+(J^* \mathbin{\triangledown} F^*)(-L^*v). \tag{$\mathcal{P}_{\text{Dual}}$}$$

Denote by $\mathcal{X}$ and $\mathcal{V}$ the sets of solutions of problem ($\mathcal{P}_{\text{Primal}}$) and ($\mathcal{P}_{\text{Dual}}$) respectively, and assume that

(**P.4**) The set of minimizers of ($\mathcal{P}_{\text{Primal}}$) and ($\mathcal{P}_{\text{Dual}}$), *i.e.* $\mathcal{X}$ and $\mathcal{V}$, are both non-empty.

More complex formulations of ($\mathcal{P}_{\text{Primal}}$) (*e.g.* multiple infimal convolutions) will be discussed in Section 8.6.

The connection between the optimization problem ($\mathcal{P}_{\text{Primal}}$) and the monotone inclusion problem (2.4.12) introduced in Section 2.4 can be obtained through the optimality condition of ($\mathcal{P}_{\text{Primal}}$). Let $(x^\star, v^\star)\in\mathcal{X}\times\mathcal{V}$ be a saddle-point of ($\mathcal{P}_{\text{SP}}$), then the corresponding first-order optimality condition is,

$$0\in\begin{bmatrix}\partial R & 0\\ 0 & \partial J^*\end{bmatrix}\begin{pmatrix}x^\star\\ v^\star\end{pmatrix}+\begin{bmatrix}\nabla F & L^*\\ -L & \nabla G^*\end{bmatrix}\begin{pmatrix}x^\star\\ v^\star\end{pmatrix}, \tag{8.1.1}$$

which turns out to be a special case of (2.4.13).

8.1.2 The class of Primal–Dual splitting methods

An efficient class of methods for solving ($\mathcal{P}_{\text{Primal}}$) is the Primal–Dual splitting methods, including [7, 51, 92, 173, 66, 54, 62] to name a few. The very first Arrow–Hurwicz Primal–Dual splitting method dates back to the late 1950s [7], where Arrow *et al.* study ($\mathcal{P}_{\text{Primal}}$) with $F=0, G=0$. This method is generalized by Chambolle and Pock in [51], where they propose the following iterative scheme

$$\left|\begin{aligned} x_{k+1} &= \operatorname{prox}_{\gamma_R R}(x_k-\gamma_R L^* v_k),\\ \bar{x}_{k+1} &= x_{k+1}+\theta(x_{k+1}-x_k),\\ v_{k+1} &= \operatorname{prox}_{\gamma_J J^*}(v_k+\gamma_J L\bar{x}_{k+1}).\end{aligned}\right. \tag{8.1.2}$$

Iteration (8.1.2) recovers the algorithm in [7] when $\theta=0$, the choice of θ is later further generalized to $[-1,1]$ in [92] by casting (8.1.2) as a PPA [151] in an appropriate metric.

Recently, several Primal–Dual splitting methods were proposed to solve ($\mathcal{P}_{\text{Primal}}$) for the scenarios where F and/or G are *not* 0. In [66], the authors consider the problem with only $G=0$, and build the connections between Primal–Dual splitting and Forward–Backward splitting [119]. Such a relation is further investigated in [62], where the authors propose a *variable metric Forward–Backward splitting method*, and design several Primal–Dual splitting methods for different monotone inclusion problems, see also [173] for another example. All these works are later formalized in [58], where the Forward–Backward structure of Primal–Dual splitting methods is proposed. Such a framework also covers [51] since PPA is a special case of Forward–Backward splitting.

All these Primal–Dual splitting methods without *relaxation* are summarized in Algorithm 15.

Algorithm 15: A Primal–Dual splitting method

Initial: Choose $\gamma_R, \gamma_J > 0$ and $\theta \in [-1, +\infty[$. For $k = 0$, $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$;

repeat

$$\left\lfloor \begin{aligned} x_{k+1} &= \operatorname{prox}_{\gamma_R R}(x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k), \\ \bar{x}_{k+1} &= x_{k+1} + \theta(x_{k+1} - x_k), \\ v_{k+1} &= \operatorname{prox}_{\gamma_J J^*}(v_k - \gamma_J \nabla G^*(v_k) + \gamma_J L \bar{x}_{k+1}), \end{aligned} \right. \tag{8.1.3}$$

$k = k + 1$;

until *convergence*;

Remark 8.1.1.

(i) It worth mentioning that Algorithm 15 is somehow new to the literature since here the choice of θ is $[-1, +\infty[$ which is larger than the condition $\theta \in [-1, 1]$ imposed in [92]. It recovers the Primal–Dual splitting methods proposed in [173, 62] when $\theta = 1$, and the one in [66] when moreover $G = 0$. When $F = 0, G = 0$, then it reduces to the Primal–Dual splitting methods in [7, 51, 92].

(ii) It can be also observed that that Algorithm 15 also covers the Forward–Backward splitting [119], Douglas–Rachford splitting [76] and preconditioned alternating direction method of multipliers (preconditioned ADMM) [79] as special cases. Therefore, in Section 8.5, we will build connections with the result presented in Chapter 6 and 7 for FB-type methods and Douglas–Rachford/ADMM.

8.1.3 Global convergence

We have seen in Section 2.4.3.3 the Forward–Backward structure of Algorithm 15, and the condition needed for the convergence.

Define the product space $\boldsymbol{\mathcal{K}} = \mathbb{R}^n \times \mathbb{R}^m$, and the following variable and operators

$$\boldsymbol{z}_k \overset{\text{def}}{=} \begin{pmatrix} x_k \\ v_k \end{pmatrix}, \ \boldsymbol{A} \overset{\text{def}}{=} \begin{bmatrix} \partial R & L^* \\ -L & \partial J^* \end{bmatrix}, \ \boldsymbol{B} \overset{\text{def}}{=} \begin{bmatrix} \nabla F & 0 \\ 0 & \nabla G^* \end{bmatrix}, \ \boldsymbol{\mathcal{V}} \overset{\text{def}}{=} \begin{bmatrix} \mathrm{Id}_n/\gamma_R & -L^* \\ -L & \mathrm{Id}_m/\gamma_J \end{bmatrix}, \tag{8.1.4}$$

where $\mathrm{Id}_n, \mathrm{Id}_m$ denote the identity operators on $\mathbb{R}^n, \mathbb{R}^m$ respectively. Owing to the result of Section 2.4.3.3, Algorithm 15 can be written as the following Forward–Backward splitting under metric $\boldsymbol{\mathcal{V}}$,

$$\boldsymbol{z}_{k+1} = (\boldsymbol{\mathcal{V}} + \boldsymbol{A})^{-1}(\boldsymbol{\mathcal{V}} - \boldsymbol{B})\boldsymbol{z}_k = (\mathbf{Id} + \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{A})^{-1}(\mathbf{Id} - \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{B})\boldsymbol{z}_k. \tag{8.1.5}$$

Proposition 8.1.2 (Convergence of Algorithm 15). *For Algorithm 15, suppose that* (P.1)-(P.4) *hold. Let* $\theta = 1$ *and choose* γ_R, γ_J *such that*

$$2\min\{\beta_F, \beta_G\} \min\left\{\frac{1}{\gamma_J}, \frac{1}{\gamma_R}\right\}\left(1 - \sqrt{\gamma_J \gamma_R \|L\|^2}\right) > 1, \tag{8.1.6}$$

then there exists a saddle point $(x^\star, v^\star) \in \mathcal{X} \times \mathcal{V}$ *such that* $(x_k, v_k) \to (x^\star, v^\star)$.

Proof. See [173, 62]. □

Remark 8.1.3.

(i) It is obvious from (8.1.6) that $\gamma_J \gamma_R \|L\|^2 < 1$, which is also the condition imposed in [51, 66]. The values of γ_J, γ_R can be varying along the iteration, which results in a variable metric Forward–Backward splitting (4.1.6). The convergence is guaranteed if the variable metric satisfies assumption (i) of Theorem 4.2.5 (see also [62]). However, for the sake of brevity, we do not pursue this discussion further in this chapter.

(ii) Proposition 8.1.2 addresses the convergence of Algorithm 15 only for the case $\theta = 1$. For the choices $\theta \in [-1, 1[\cup]1, +\infty[$, so far the corresponding convergence of the iteration cannot be obtained directly, and a *correction step* as proposed in [92] is needed so that the iteration is a *contraction*. Meanwhile, such a correction step will lead to a new iterative scheme, not simply (8.1.3) itself, see [92] for more details. Since the main focus of this chapter is to study the *local* convergence property, we shall skip the discussion of the global convergence of Algorithm 15 and mainly focus on the case $\theta = 1$. Nevertheless, as we will see later, locally $\theta > 1$ gives *faster* convergence rate than $\theta \in [-1, 1]$ does, which points out a future research direction to design new Primal–Dual splitting methods.

(iii) We could also handle the multi-step inertial version just as well as considered in Section 4.3.3, the extension of the MUSTARD algorithm 7 to (8.1.3) with $\theta = 1$.

8.2 Finite activity identification

In this section, we present the local convergence analysis of the Primal–Dual splitting methods. Following the steps of the previous two chapters, we first present the finite activity identification of the sequences, then show that the fixed-point iteration of Primal–Dual splitting methods locally can be linearized along the identified smooth manifolds, and finally characterize the local linear regime.

Theorem 8.2.1 (Finite activity identification). *For Algorithm 15, let* (**P.1**)-(**P.4**) *and condition* (8.1.6) *hold such that* $(x_k, v_k) \to (x^\star, v^\star)$. *Assume moreover that* $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}^R_{x^\star})$ *and* $J^* \in \mathrm{PSF}_{v^\star}(\mathcal{M}^{J^*}_{v^\star})$, *and the non-degeneracy condition*

$$\begin{aligned} -L^*v^\star - \nabla F(x^\star) &\in \mathrm{ri}\big(\partial R(x^\star)\big), \\ Lx^\star - \nabla G^*(v^\star) &\in \mathrm{ri}\big(\partial J^*(v^\star)\big), \end{aligned} \tag{ND$_{\mathrm{PD}}$}$$

holds. Then,

(i) *there* $\exists K > 0$ *such that for all* $k \geq K$,

$$(x_k, v_k) \in \mathcal{M}^R_{x^\star} \times \mathcal{M}^{J^*}_{v^\star}.$$

(ii) *Moreover,*

(a) *if* $\mathcal{M}^R_{x^\star} = x^\star + T^R_{x^\star}$, *then* $\forall k \geq K$ *there holds* $T^R_{x_k} = T^R_{x^\star}$, *and* $\bar{x}_k \in \mathcal{M}^R_{x^\star}$ *for* $k > K$.

(b) *If* $\mathcal{M}^{J^*}_{v^\star} = v^\star + T^{J^*}_{v^\star}$, *then* $\forall k \geq K$, $T^{J^*}_{v_k} = T^{J^*}_{v^\star}$.

(c) R *is locally polyhedral around* $x^\star$, *then* $\forall k \geq K$, $x_k \in \mathcal{M}^R_{x^\star} = x^\star + T^R_{x^\star}$, $T^R_{x_k} = T^R_{x^\star}$, $\nabla_{\mathcal{M}^R_{x^\star}} R(x_k) = \nabla_{\mathcal{M}^R_{x^\star}} R(x^\star)$, *and* $\nabla^2_{\mathcal{M}^R_{x^\star}} R(x_k) = 0$.

(d) J^* *is locally polyhedral around* $v^\star$, *then* $\forall k \geq K$, $v_k \in \mathcal{M}^{J^*}_{v^\star} = v^\star + T^{J^*}_{v^\star}$, $T^{J^*}_{v_k} = T^{J^*}_{v^\star}$, $\nabla_{\mathcal{M}^{J^*}_{v^\star}} J^*(v_k) = \nabla_{\mathcal{M}^{J^*}_{v^\star}} J^*(x^\star)$, *and* $\nabla^2_{\mathcal{M}^{J^*}_{v^\star}} J^*(v_k) = 0$.

See Section 8.8 for the proof.

Remark 8.2.2.

(i) In general, we have no identification guarantees for x_k and v_k if the proximity operators are computed with errors, even if the latter are summable, in which case one can still prove global convergence.

(ii) Theorem 8.2.1 only states the *existence* of K after which the identification of the sequences happen, and no bounds are available. In Chapters 6 and 7, *lower* bounds of K for the Forward–Backward and Douglas–Rachford splitting methods are provided, and similar lower bounds can be obtained here for the Primal–Dual splitting methods. Since lower bound is not that useful in practice, we choose to skip the detailed discussion in this chapter.

8.3 Local linear convergence

Relying on the identification result, now we are able to show that the fixed-point iteration (8.1.3) can be linearized along the manifold $\mathcal{M}_{x^\star}^R \times \mathcal{M}_{v^\star}^{J^*}$, then the convergence rate of the iteration essentially boils down the spectral properties of the matrix obtained in the linearized fixed-point iteration.

8.3.1 Locally linearized iteration

Given a saddle point $(x^\star, v^\star) \in \mathcal{X} \times \mathcal{V}$, define the following two functions

$$\overline{R}(x) \overset{\text{def}}{=} R(x) + \langle x,\, L^* v^\star + \nabla F(x^\star)\rangle, \quad \overline{J^*}(y) \overset{\text{def}}{=} J^*(v) - \langle v,\, Lx^\star - \nabla G^*(v^\star)\rangle, \tag{8.3.1}$$

we have the following lemma.

Lemma 8.3.1. *Let $(x^\star, v^\star)$ be a saddle point of ($\mathcal{P}_{\text{SP}}$) such that $R \in \text{PSF}_{x^\star}(\mathcal{M}_{x^\star}^R)$, $J^* \in \text{PSF}_{x^\star}(\mathcal{M}_{v^\star}^{J^*})$. For the Riemannian Hessian of $\overline{R}$ and $\overline{J^*}$,*

$$H_{\overline{R}} \overset{\text{def}}{=} \gamma_R \mathcal{P}_{T_{x^\star}^R} \nabla^2_{\mathcal{M}_{x^\star}^R} \overline{R}(x^\star) \mathcal{P}_{T_{x^\star}^R} \quad \text{and} \quad H_{\overline{J^*}} \overset{\text{def}}{=} \gamma_J \mathcal{P}_{T_{v^\star}^{J^*}} \nabla^2_{\mathcal{M}_{v^\star}^{J^*}} \overline{J^*}(v^\star) \mathcal{P}_{T_{v^\star}^{J^*}}. \tag{8.3.2}$$

They are symmetric positive semi-definite *under either of the following circumstances:*

(i) (ND_{PD}) *holds.*

(ii) $\mathcal{M}_{x^\star}^R$ *and* $\mathcal{M}_{v^\star}^{J^*}$ *are* affine subspaces.

Define,

$$W_{\overline{R}} \overset{\text{def}}{=} (\text{Id}_n + H_{\overline{R}})^{-1} \quad \text{and} \quad W_{\overline{J^*}} \overset{\text{def}}{=} (\text{Id}_m + H_{\overline{J^*}})^{-1}, \tag{8.3.3}$$

then both $W_{\overline{R}}$ and $W_{\overline{J^*}}$ are firmly non-expansive.

Proof. See Lemma 7.4.5. □

Throughout the rest of the chapter, we assume that F and G^* are actually locally C^2 around $(x^\star, v^\star)$. Define the restricted Hessian of F and G^* respectively,

$$H_F \overset{\text{def}}{=} \mathcal{P}_{T_{x^\star}^R} \nabla^2 F(x^\star) \mathcal{P}_{T_{x^\star}^R} \quad \text{and} \quad H_{G^*} \overset{\text{def}}{=} \mathcal{P}_{T_{v^\star}^{J^*}} \nabla^2 G^*(v^\star) \mathcal{P}_{T_{v^\star}^{J^*}}. \tag{8.3.4}$$

Denote $\overline{H}_F \overset{\text{def}}{=} \text{Id}_n - \gamma_R H_F$, $\overline{H}_{G^*} \overset{\text{def}}{=} \text{Id}_m - \gamma_J H_{G^*}$ and $\overline{L} \overset{\text{def}}{=} \mathcal{P}_{T_{v^\star}^{J^*}} L \mathcal{P}_{T_{x^\star}^R}$,

$$M_{\text{PD}} \overset{\text{def}}{=} \begin{bmatrix} W_{\overline{R}} \overline{H}_F & -\gamma_R W_{\overline{R}} \overline{L}^* \\ \gamma_J(1+\theta) W_{\overline{J^*}} \overline{L} W_{\overline{R}} \overline{H}_F - \theta \gamma_J W_{\overline{J^*}} \overline{L} & W_{\overline{J^*}} \overline{H}_{G^*} - \gamma_J \gamma_R (1+\theta) W_{\overline{J^*}} \overline{L} W_{\overline{R}} \overline{L}^* \end{bmatrix}. \tag{8.3.5}$$

Proposition 8.3.2 (Local linearized iteration). *Suppose that the Algorithm 15 is run with the conditions in Theorem 8.2.1 hold. Then for all k large enough, the fixed-point iteration (8.1.3) can be written as*

$$\boldsymbol{z}_{k+1} - \boldsymbol{z}^\star = M_{\text{PD}}(\boldsymbol{z}_k - \boldsymbol{z}^\star) + o(\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|). \tag{8.3.6}$$

See Section 8.8 from page 192 for the proof.

Remark 8.3.3.

(i) For the case of varying γ_J, γ_R, according to the result of [115], (8.3.6) remains hold if they are converging to some constants such that condition (8.1.6) still stands.

(ii) For the fixed-point operator in (8.3.5), let $\overline{H}_{G^*} = \text{Id}_m$, then it recovers the linearized iteration of the Primal–Dual splitting method proposed in [66]. If we further let $\overline{H}_F = \text{Id}_n$, then it recovers the linearized version of the method in [51].

Now we study the spectral properties of M_{PD}. Let $p \overset{\text{def}}{=} \dim(T^R_{x^\star}), q \overset{\text{def}}{=} \dim(T^{J^*}_{v^\star})$ be the dimensions of $T^R_{x^\star}$ and $T^{J^*}_{v^\star}$ respectively, define $S^R_{x^\star} = (T^R_{x^\star})^\perp$ and $S^{J^*}_{v^\star} = (T^{J^*}_{v^\star})^\perp$. Assume that $q \geq p$ (alternative situations are discussed in Remark 8.3.5). Let $\overline{L} = X\Sigma_{\overline{L}}Y^*$ the singular value decomposition of $\overline{L}$, denote the rank of $\overline{L}$ as $l \overset{\text{def}}{=} \text{rank}(\overline{L})$. Clearly, we have $l \leq p$.

Lemma 8.3.4. *For the matrix operator M_{PD} of (8.3.5).*

(i) *If $\theta = 1$, then M_{PD} is convergent with the convergent matrix as*

$$M^\infty_{\text{PD}} \overset{\text{def}}{=} \lim_{k\to\infty} M^k_{\text{PD}}. \tag{8.3.7}$$

(ii) *If $F = 0, G = 0$, and R, J^* are locally polyhedral around $(x^\star, v^\star)$. Then given any $\theta \in]0, 1]$, M_{PD} is convergent with*

$$M^\infty_{\text{PD}} = \begin{bmatrix} Y & \\ & X \end{bmatrix} \begin{bmatrix} 0_l & & & \\ & \text{Id}_{n-l} & & \\ & & 0_l & \\ & & & \text{Id}_{m-l} \end{bmatrix} \begin{bmatrix} Y^* & \\ & X^* \end{bmatrix}. \tag{8.3.8}$$

Moreover, all the eigenvalues of $M_{\text{PD}} - M^\infty_{\text{PD}}$ are complex *and the spectral radius is*

$$\rho(M_{\text{PD}} - M^\infty_{\text{PD}}) = \sqrt{1 - \theta\gamma_R\gamma_J\sigma^2_{\min}} < 1, \tag{8.3.9}$$

where $\sigma_{\min}$ is the smallest non-zero singular value *of $\overline{L}$.*

If $q = p$ and $l = p$, then

$$M^\infty_{\text{PD}} = \begin{bmatrix} \mathcal{P}_{S^R_{x^\star}} & \\ & \mathcal{P}_{S^{J^*}_{v^\star}} \end{bmatrix}, \tag{8.3.10}$$

which is a projection operator onto the normal spaces $S^R_{x^\star} \times S^{J^*}_{v^\star}$.

See Section 8.8 from page 194 for the proof.

Remark 8.3.5. Let us briefly discuss other possible cases of (8.3.8) when $F = 0, G = 0$ and R, J^* are locally polyhedral around $(x^\star, v^\star)$.

(i) When $L = \text{Id}$, then $\overline{L} = \mathcal{P}_{T^{J^*}_{v^\star}}\mathcal{P}_{T^R_{x^\star}}$ and $\sigma_{\min}$ stands for the *cosine* of the *biggest principal angle* (yet *smaller* than $\pi/2$) between $T^R_{x^\star}$ and $T^{J^*}_{v^\star}$.

(ii) The spectral radius formula (8.3.9) means that for the case $\theta = 0$, *i.e.* the Arrow–Hurwicz scheme [7]. When R, J^* are locally polyhedral, let $\Sigma_{\overline{L}} = (\sigma_j)_{\{j=1,\dots,l\}}$ be the singular values of $\overline{L}$, then the eigenvalues of $M_{\text{PD}} - M^\infty_{\text{PD}}$ are

$$\rho_j = \frac{\left(2 - \gamma_R\gamma_J\sigma_j^2\right) \pm \sqrt{\gamma_R\gamma_J\sigma_j^2(\gamma_R\gamma_J\sigma_j^2 - 4)}}{2}, \ j \in \{1, \dots, l\}, \tag{8.3.11}$$

which apparently is *complex* ($\gamma_R\gamma_J\sigma_j^2 \leq \gamma_R\gamma_J\|L\|^2 < 1$). Moreover,

$$|\rho_j| = \frac{1}{2}\sqrt{\left(2 - \gamma_R\gamma_J\sigma_j^2\right)^2 - \gamma_R\gamma_J\sigma_j^2(\gamma_R\gamma_J\sigma_j^2 - 4)} = 1.$$

This implies that M_{PD} has *multiple* eigenvalues with absolute values *all* equal to 1, then owing to the result of [18], we have M_{PD} is not convergent.

Furthermore, for $\theta \in [-1, 0[$, we have $1 - \theta\gamma_R\gamma_J\sigma^2_{\min} > 1$ meaning that M_{PD} is not convergent, this implies that the *correction step* proposed in [92] is *necessary* for $\theta \in [-1, 0]$. Discussion on $\theta > 1$ is left to Section 8.4.

(iii) If $l \overset{\text{def}}{=} \operatorname{rank}(L) \leq \dim(T^R_{x^\star})$, then

$$M^{\infty}_{\mathrm{PD}} = \begin{bmatrix} \mathcal{P}_{S^R_{x^\star}} & \\ & \mathcal{P}_{S^{J^*}_{v^\star}} \end{bmatrix} + \begin{bmatrix} Y & \\ & X \end{bmatrix} \begin{bmatrix} 0_l & & & & \\ & \mathrm{Id}_{p-l} & & & \\ & & 0_{n-p+l} & & \\ & & & \mathrm{Id}_{q-l} & \\ & & & & 0_{m-q} \end{bmatrix} \begin{bmatrix} Y^* & \\ & X^* \end{bmatrix}.$$

8.3.2 Local linear convergence

Finally, we are able to present the local linear convergence of Primal–Dual splitting methods.

Theorem 8.3.6 (Local linear convergence). *For Algorithm 15, suppose it is run under the conditions of Theorem 8.2.1. Then*

(i) *given any* $\rho \in]\rho(M_{\mathrm{PD}} - M^{\infty}_{\mathrm{PD}}), 1[$, *there exist a* K *large enough such that for all* $k \geq K$,

$$\|(\mathbf{Id} - M^{\infty}_{\mathrm{PD}})(\boldsymbol{z}_k - \boldsymbol{z}^\star)\| = O(\rho^{k-K}). \tag{8.3.12}$$

(ii) *If moreover,* R, J^* *are locally polyhedral around* $(x^\star, v^\star)$, *there exist a* K *large enough such that for all* $k \geq K$, *we have directly*

$$\|\boldsymbol{z}_k - \boldsymbol{z}^\star\| = O(\rho^{k-K}), \tag{8.3.13}$$

for $\rho \in [\rho(M_{\mathrm{PD}} - M^{\infty}_{\mathrm{PD}}), 1[$.

Proof. Direct result of Corollary 7.4.9 and Theorem 7.4.10. □

Remark 8.3.7.

(i) The obtained result can be extended to the case of varying parameters, *i.e.* the value of (γ_J, γ_R) change along the iteration which results in $(\gamma_{J,k}, \gamma_{R,k})$. Then based on the result of [62] for global convergence and Chapter 6 and 7 for the local linear convergence, $(\gamma_{J,k}, \gamma_{R,k})$ must converge to some constants, for instance $(\gamma_{J,k}, \gamma_{R,k}) \to (\gamma_J, \gamma_R)$.

Moreover, the local linear convergence *depends on* how fast they converge, meaning that if $(\gamma_{J,k}, \gamma_{R,k})$ converge sub-linearly to (γ_J, γ_R), then the local convergence rate will eventually become sub-linear.

(ii) When $F = 0, G = 0$ and both R and J^* are *locally polyhedral* around the saddle point, then the convergence rate of the Primal–Dual splitting method is controlled by θ and $\gamma_J\gamma_R$ as shown in (8.3.9); see the upcoming section for a detailed discussion.

For general situations (*i.e.* F, G are nontrivial and R, J^* are general partly smooth functions), the factors that contribute to the local convergence rate are much more complicated, these include the Riemannian Hessians of the involved functions and the curvatures of the underlying active manifolds.

(iii) For the relation between partial smoothness and metric sub-regularity, the result for Primal–Dual splitting is very similar to those of Forward–Backward (Section 6.4.4) and Douglas–Rachford (Section 7.4.4). Therefore, we choose not to pursue the discussion in this chapter.

8.4 Discussions

In this part, we present several discussions on the local linear convergence result, including the acceleration, what will happen when $\theta \geq 1$ and the oscillation of the Primal–Dual splitting method.

For this discussion, to make the result more immediate, we focus on the case where $F = 0, G = 0$, *i.e.* the Primal–Dual splitting method (8.1.2) in [51], and moreover R, J^* are locally polyhedral around the saddle-point. Under such setting, letting $\theta \in [0, 1]$, the matrix defined in (8.3.5) becomes

$$M_{\mathrm{PD}} \overset{\mathrm{def}}{=} \begin{bmatrix} \mathrm{Id}_n & -\gamma_R \overline{L}^* \\ \gamma_J \overline{L} & \mathrm{Id}_m - (1+\theta)\gamma_J\gamma_R \overline{L}\overline{L}^* \end{bmatrix}. \tag{8.4.1}$$

8.4.1 Choice of θ

Owing to Lemma 2.5.6, the matrix M_{PD} in (8.4.1) is convergent for $\theta \in]0, 1]$, see Eq. (8.3.8), with the spectral radius

$$\rho(M_{\mathrm{PD}} - M_{\mathrm{PD}}^{\infty}) = \sqrt{1 - \theta\gamma_R\gamma_J\sigma_{\min}^2} < 1, \tag{8.4.2}$$

with $\sigma_{\min}$ being the *smallest non-zero singular value* of $\overline{L}$. In general, given a saddle point $(x^\star, v^\star)$, $\sigma_{\min}$ is fixed, hence the spectral radius $\rho(M_{\mathrm{PD}} - M_{\mathrm{PD}}^{\infty})$ is simply controlled by θ and the product $\gamma_J\gamma_R$. To have a faster rate, it is obvious that we need to make $\theta\gamma_J\gamma_R$ as big as possible. Recall in the global convergence of Primal–Dual splitting method or the result from [51], that

$$\gamma_J\gamma_R\|L\|^2 < 1.$$

Denote $\sigma_{\max}$ the *biggest singular value* of $\overline{L}$. It is then straightforward that $\gamma_J\gamma_R\sigma_{\max}^2 \leq \gamma_J\gamma_R\|L\|^2 < 1$ and moreover

$$\begin{aligned} \rho(M_{\mathrm{PD}} - M_{\mathrm{PD}}^{\infty}) &= \sqrt{1 - \theta\gamma_R\gamma_J\sigma_{\min}^2} \\ &> \sqrt{1 - \theta(\sigma_{\min}/\|L\|)^2} \geq \sqrt{1 - \theta(\sigma_{\min}/\sigma_{\max})^2}. \end{aligned} \tag{8.4.3}$$

If we define $\mathrm{cnd} \overset{\mathrm{def}}{=} \sigma_{\max}/\sigma_{\min}$ the condition number of L, then we have

$$\rho(M_{\mathrm{PD}} - M_{\mathrm{PD}}^{\infty}) > \sqrt{1 - \theta(1/\mathrm{cnd})^2}.$$

To this end, it is clear that $\theta = 1$ gives the best convergence rate. Next let us look at what happens if $\theta > 1$ (assuming that the corresponding iteration converges globally). The spectral radius formula (8.4.2) implies that *bigger* value of θ yields *smaller* $\rho(M_{\mathrm{PD}} - M_{\mathrm{PD}}^{\infty})$. Therefore, *locally* we should choose θ as big as possible. Next we investigate the upper bound of θ if there is one.

Following Remark 8.3.5, let $\Sigma_{\overline{L}} = (\sigma_j)_{\{j=1,\dots,l\}}$ be the singular values of $\overline{L}$, let ρ_j be the eigenvalue of $M_{\mathrm{PD}} - M_{\mathrm{PD}}^{\infty}$, we have known that ρ_j is complex and

$$\rho_j = \frac{\big(2 - (1+\theta)\gamma_R\gamma_J\sigma_j^2\big) \pm \sqrt{(1+\theta)^2\gamma_R^2\gamma_J^2\sigma_j^4 - 4\gamma_R\gamma_J\sigma_j^2}}{2}, \quad |\rho_j| = \sqrt{1 - \theta\gamma_R\gamma_J\sigma_j^2}.$$

Now let $\theta > 1$, then in order to let $|\rho_j|$ make sense for all $j \in \{1, \dots, l\}$, there must holds

$$1 - \theta\gamma_R\gamma_J\sigma_{\max}^2 \geq 0 \Longleftrightarrow \theta \leq \frac{1}{\gamma_R\gamma_J\sigma_{\max}^2},$$

which means that θ indeed is upper bounded.

Unfortunately, since $\overline{L} = \mathcal{P}_{T_{x^\star}^R} L \mathcal{P}_{T_{v^\star}^{J^*}}$, the upper bound can be only obtained if we had the saddle point. However, we can use *back-tracking* or the *Armijo-Goldstein-rule* to find the proper θ. See Section 8.7.1 for an illustration of online searching of θ. It should be noted that such updating rule can also be applied to γ_J, γ_R since we have $\|\overline{L}\| \leq \|L\|$, hence locally bigger values of them should be applicable.

8.4.2 Oscillations

We have seen in Chapter 6 the local oscillation of FB-type methods, when the inertia momentum are too high. When solving certain type of problems (*i.e.* $F = 0, G = 0$ and R, J^* are locally polyhedral

around the saddle point $(x^\star, v^\star)$), the Primal–Dual splitting method oscillates too. As revealed in the proof of Lemma 2.5.6, all the eigenvalues of $M_{\text{PD}} - M_{\text{PD}}^\infty$ in (8.4.1) are *complex*. This means that locally the sequences generated by (8.1.2) will oscillate.

For $\sigma_{\min}$, the *smallest non-zero singular* of $\overline{L}$, one of its corresponding eigenvalues of M_{PD} reads

$$\rho_{\sigma_{\min}} = \frac{(2-(1+\theta)\gamma_J\gamma_R\sigma_{\min}^2) + \sqrt{(1+\theta)^2\gamma_R^2\gamma_J^2\sigma_{\min}^4 - 4\gamma_J\gamma_R\sigma_{\min}^2}}{2},$$

and $(1+\theta)^2\gamma_R^2\gamma_J^2\sigma_{\min}^4 - 4\gamma_J\gamma_R\sigma_{\min}^2 < 0$. Denote ω the argument of $\rho_{\sigma_{\min}}$, then

$$\cos(\omega) = \frac{2-(1+\theta)\gamma_J\gamma_R\sigma_{\min}^2}{\sqrt{1-\theta\gamma_J\gamma_R\sigma_{\min}^2}}. \tag{8.4.4}$$

The oscillation period of the sequence $\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|$ is then exactly $\frac{\pi}{\omega}$. See Figure 8.6 for an illustration.

Remark 8.4.1.

(i) Complex eigenvalues is only a *sufficient* condition for local oscillation behaviour. When the involved functions are all polyhedral, the eigenvalues of the local linearized operator of DR are also complex, however, the iterates of DR do not oscillate.

(ii) The mechanisms of oscillation between Primal–Dual splitting and the FB-type methods are quite different. The oscillation of FISTA is caused by the *inertial momentum* being too large, while the oscillation of Primal–Dual splitting is due to the polyhedrality of the functions. Furthermore, if $F \neq 0$ and/or $G \neq 0$, then the Primal–Dual splitting method *does not* oscillate.

8.5 Relations with FB and DR/ADMM

In this part, we discuss several special cases of the obtained result, and the relations with the previous chapters.

8.5.1 Forward–Backward splitting

For problem ($\mathcal{P}_{\text{Primal}}$), when $J = 0, G = 0$, Algorithm 15 reduces to, denote $\gamma = \gamma_R$,

$$x_{k+1} = \text{prox}_{\gamma R}(x_k - \gamma\nabla F(x_k)),\ \gamma \in]0, 2\beta[, \tag{8.5.1}$$

which is the non-relaxed FB iteration [119] with constant step-size.

Let $x^\star \in \text{Argmin}(R+F)$ be a global minimizer such that $\{x_k\}_{k\in\mathbb{N}}$ of (8.5.1) converges to $x^\star$ and conditions of Theorem 6.2.1 are satisfied. Then under the notions of Section 8.3, define $M_{\text{FB}} = W_{\overline{R}}(\text{Id}_n - \gamma H_F)$, we have for all k large enough

$$x_{k+1} - x^\star = M_{\text{FB}}(x_k - x^\star) + o(\|x_k - x^\star\|).$$

From Theorem 8.3.6, we obtain the following result for the FB algorithm.

Corollary 8.5.1. *For problem ($\mathcal{P}_{\text{FB}}$), suppose that* (**F.1**)-(**F.3**) *hold, and the FB iteration* (8.5.1) *creates a sequence $x_k \to x^\star \in \text{Argmin}(\Phi)$ such that $R \in \text{PSF}_{x^\star}(\mathcal{M}_{x^\star})$, F is C^2 near $x^\star$, and condition* (ND_{FB}) *holds. Then*

(i) *given any $\rho \in]\rho(M_{\text{FB}} - M_{\text{FB}}^\infty), 1[$, there exist a K large enough such that for all $k \geq K$,*

$$\|(\text{Id} - M_{\text{FB}}^\infty)(x_k - x^\star)\| = O(\rho^{k-K}). \tag{8.5.2}$$

(ii) *If moreover, R are locally polyhedral around $x^\star$, there exist a K large enough such that for all $k \geq K$, we have directly*

$$\|x_k - x^\star\| = O(\rho^{k-K}), \tag{8.5.3}$$

for $\rho \in [\rho(M_{\text{FB}} - M_{\text{FB}}^\infty), 1[$.

Proof. Owing to Lemma 2.4.7, M_{FB} is $\frac{2\beta}{4\beta-\gamma}$-averaged non-expansive, hence convergent. The convergence rates in (8.5.2) and (8.5.3) are straightforward from Theorem 8.3.6. □

In view of this result, one my wonder if it was useful to develop the results of Chapter 6 since they seem to be covered to some extent by the current chapter. The answer is affirmative for several reasons. First, though Corollary 8.5.1 relaxes the restricted injectivity condition (RI), it establishes linear convergence of $\|(\mathrm{Id} - M_{\mathrm{FB}}^{\infty})(x_k - x^\star)\|$ rather than $\|x_k - x^\star\|$. Second, the rate estimate is not explicit unlike Theorem 6.3.6. In Chapter 6, we considered a much larger class than the sole FB, which is not the case in this chapter. Finally, in that chapter, we developed a similar analysis for the non-convex case which cannot be afforded through the primal-dual setting of this chapter where convexity is crucial.

8.5.2 Douglas–Rachford/ADMM

Let $F = 0, G = 0$ and $L = \mathrm{Id}$, the problem ($\mathcal{P}_{\mathrm{Primal}}$) then reads

$$\min_{x\in\mathbb{R}^n} R(x) + J(x).$$

For the Primal–Dual splitting scheme (8.1.2), let $\theta = 1$ and change the update order, then we obtain the following iteration

$$\left|\begin{array}{l} v_{k+1} = \mathrm{prox}_{\gamma_J J^*}(v_k + \gamma_J \bar{x}_k) \\ x_{k+1} = \mathrm{prox}_{\gamma_R R}(x_k - \gamma_R v_{k+1}) \\ \bar{x}_{k+1} = 2x_{k+1} - x_k. \end{array}\right. \tag{8.5.4}$$

Apply the Moreau's identity to $\mathrm{prox}_{\gamma_J J^*}$, let $\gamma_J = 1/\gamma_R$ and define $z_{k+1} = x_k - \gamma_R v_{k+1}$, iteration (8.5.4) becomes

$$\left|\begin{array}{l} u_{k+1} = \mathrm{prox}_{\gamma_R J}(2x_k - z_k) \\ z_{k+1} = z_k + u_{k+1} - x_k \\ x_{k+1} = \mathrm{prox}_{\gamma_R R}(z_{k+1}), \end{array}\right. \tag{8.5.5}$$

which is the non-relaxed Douglas–Rachford splitting, and we have $u_k, x_k \to x^\star = \mathrm{prox}_{\gamma_J J}(w^\star)$ where $w^\star$ is a fixed point of the iteration. See also the discussions in [51, Section 4.2].

Specializing the derivation of (8.3.5) to (8.5.4) and (8.5.5), we obtain the following two linearized fixed-point operator of (8.5.4) and (8.5.5) respectively

$$M_1 = \begin{bmatrix} \mathrm{Id}_n & -\gamma_R \mathcal{P}_{T_{x^\star}^R} \mathcal{P}_{T_{v^\star}^{J^*}} \\ \gamma_J \mathcal{P}_{T_{v^\star}^{J^*}} \mathcal{P}_{T_{x^\star}^R} & \mathrm{Id}_n - 2\gamma_J\gamma_R \mathcal{P}_{T_{v^\star}^{J^*}} \mathcal{P}_{T_{x^\star}^R} \mathcal{P}_{T_{v^\star}^{J^*}} \end{bmatrix},$$
$$M_2 = \begin{bmatrix} \mathrm{Id}_n & -\gamma_R \mathcal{P}_{T_{x^\star}^R} \mathcal{P}_{T_{v^\star}^{J^*}} \\ \frac{1}{\gamma_R} \mathcal{P}_{T_{v^\star}^{J^*}} \mathcal{P}_{T_{x^\star}^R} & \mathrm{Id}_n - 2\mathcal{P}_{T_{v^\star}^{J^*}} \mathcal{P}_{T_{x^\star}^R} \mathcal{P}_{T_{v^\star}^{J^*}} \end{bmatrix}.$$

Owing to (ii) of Lemma 8.3.4, M_1, M_2 are convergent. Let θ be the *largest principal angle (yet smaller than $\pi/2$)* between $T_{x^\star}^R$ and $T_{v^\star}^{J^*}$, then we have the spectral radius of $M_1 - M_1^\infty$ reads ((i) of Remark 8.3.5),

$$\begin{aligned} \rho(M_1 - M_1^\infty) &= \sqrt{1 - \gamma_J\gamma_R \cos^2(\theta)} \\ &\geq \sqrt{1 - \cos^2(\theta)} = \sin(\theta) = \cos(\pi/2 - \theta). \end{aligned} \tag{8.5.6}$$

Suppose that the set of saddle-point is a *singleton*, *i.e.* $(x^\star, v^\star)$ is unique, and moreover that R and J are polyhedral. Thus we have that if J^* is locally polyhedral near $v^\star$ along $v^\star + T_{v^\star}^{J^*}$, then J is locally polyhedral near $x^\star$ around $x^\star + T_{x^\star}^J$, and moreover there holds $T_{x^\star}^J = (T_{v^\star}^{J^*})^\perp$. As a result, the principal

angles between $T^R_{x^\star}, T^{J^*}_{v^\star}$ and the ones between $T^R_{x^\star}, T^J_{x^\star}$ are *complementary*, which means that $\pi/2-\theta$ is the *Friedrichs angle* between tangent spaces $T^R_{x^\star}, T^J_{x^\star}$. Thus, following (8.5.6), we have

$$\rho(M_1 - M_1^\infty) = \sqrt{1-\gamma_J\gamma_R\cos^2(\theta)} \geq \cos(\pi/2-\theta) = \rho(M_2 - M_2^\infty). \tag{8.5.7}$$

This means that locally, the convergence rate of Primal-Dual splitting (8.5.4) is slower than that of DR in (8.5.5).

We emphasise the fact that such connection can be drawn only for the polyehdral case, which justifies the different analysis carried in Chapter 7. In addition, in that chapter, we were able to characterize situations where finite convergence provably occurs (Theorem 7.5.1), while this is not (yet) the case for Primal–Dual splitting even for R and J^* being locally polyhedral around $(x^\star, v^\star)$ and $F = G = 0$ but L is non-trivial.

8.6 Sum of more than 2 infimal convolutions

In this section, we consider problem ($\mathcal{P}_{\text{Primal}}$) with more than one infimal convolution. Le $m \geq 1$ be a positive integer. Consider the problem of minimizing

$$\min_{x\in\mathbb{R}^n} R(x) + F(x) + \textstyle\sum_{i=1}^m (J_i \mathbin{\triangledown} G_i)(L_i x), \tag{$\boldsymbol{\mathcal{P}}^m_{\text{Primal}}$}$$

where (**P.1**) holds for R and F, and for every $i = 1, ..., m$,

(**P'.2**) $J_i, G_i \in \Gamma_0(\mathbb{R}^{m_i})$, with G_i being differentiable and $\beta_{G,i}$-strongly convex for $\beta_{G,i} > 0$.

(**P'.3**) $L_i : \mathbb{R}^n \to \mathbb{R}^{m_i}$ is a non-zero linear operator.

The dual problem of ($\boldsymbol{\mathcal{P}}^m_{\text{Primal}}$) reads,

$$\min_{v_1\in\mathbb{R}^{m_1},\dots,v_m\in\mathbb{R}^{m_m}} (J^* \mathbin{\triangledown} F^*)\Big(-\textstyle\sum_{i=1}^m L_i^T v_i\Big) + \textstyle\sum_{i=1}^m \big(J_i^*(v_i) + G_i^*(v_i)\big). \tag{$\boldsymbol{\mathcal{P}}^m_{\text{Dual}}$}$$

Denote by $\mathcal{X}$ and $\boldsymbol{\mathcal{V}} = \mathcal{V}_1 \times \cdots \times \mathcal{V}_m$ the sets of solutions of problem ($\boldsymbol{\mathcal{P}}^m_{\text{Primal}}$) and ($\boldsymbol{\mathcal{P}}^m_{\text{Dual}}$) respectively, and assume

(**P'.4**) The set of minimizers of ($\boldsymbol{\mathcal{P}}^m_{\text{Primal}}$) and ($\boldsymbol{\mathcal{P}}^m_{\text{Dual}}$), *i.e.* $\mathcal{X}$ and $\boldsymbol{\mathcal{V}}$, are both non-empty.

By combining the primal and dual problems, the saddle-point problem reads

$$\min_{x\in\mathbb{R}^n}\max_{v_1\in\mathbb{R}^{m_1},\dots,v_m\in\mathbb{R}^{m_m}} R(x) + F(x) + \textstyle\sum_{i=1}^m \big(\langle L_i x,\, v_i\rangle - \big(J_i^*(v_i) + G_i^*(v_i)\big)\big). \tag{$\boldsymbol{\mathcal{P}}^m_{\text{SP}}$}$$

Problem ($\boldsymbol{\mathcal{P}}^m_{\text{Primal}}$) is considered in [62], and a Primal–Dual splitting algorithm is proposed there which is the extension of Algorithm 15, see below Algorithm 16 for details. A similar problem is consider in [173], where the author considers the *weighted* sums of the infimal convolutions. The main difference between [173] and [62] is that when casting the corresponding algorithm into the Forward–Backward splitting form, they have different *metrics*.

Product space The following result is taken from [62]. Define the *product space* $\boldsymbol{\mathcal{K}} = \mathbb{R}^n \times \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_m}$, and let $\mathbf{Id}$ be the identity operator on $\boldsymbol{\mathcal{K}}$. Define the following operators

$$\boldsymbol{A} \overset{\text{def}}{=} \begin{bmatrix} \partial R & L_1^* & \cdots & L_m^* \\ -L_1 & \partial J_1 & & \\ \vdots & & \ddots & \\ -L_m & & & \partial J_m \end{bmatrix}, \boldsymbol{B} \overset{\text{def}}{=} \begin{bmatrix} \nabla F & & & \\ & \nabla G_1^* & & \\ & & \ddots & \\ & & & \nabla G_m^* \end{bmatrix}, \boldsymbol{\mathcal{V}} \overset{\text{def}}{=} \begin{bmatrix} \frac{\mathrm{Id}_n}{\gamma_R} & -L_1^* & \cdots & -L_m^* \\ -L_1 & \frac{\mathrm{Id}_{m_1}}{\gamma_{J_1}} & & \\ \vdots & & \ddots & \\ -L_m & & & \frac{\mathrm{Id}_{m_m}}{\gamma_{J_m}} \end{bmatrix}. \tag{8.6.2}$$

Then $\boldsymbol{A}$ is maximal monotone, $\boldsymbol{B}$ is $\min\{\beta_F, \beta_{G_1}, ..., \beta_{G_m}\}$-cocoercive, and $\boldsymbol{\mathcal{V}}$ is symmetric and ν-positive definite with $\nu = (1 - \sqrt{\gamma_R \sum_i \gamma_{J_i}\|L_i\|^2}) \min\{\frac{1}{\gamma_R}, \frac{1}{\gamma_{J_1}}, ..., \frac{1}{\gamma_{J_m}}\}$. Define $\boldsymbol{z}_k = (x_k, v_{1,k}, \cdots, v_{m,k})^T$, then it can be shown that (8.6.1) is equivalent to

$$\boldsymbol{z}_{k+1} = (\boldsymbol{\mathcal{V}} + \boldsymbol{A})^{-1}(\boldsymbol{\mathcal{V}} - \boldsymbol{B})\boldsymbol{z}_k = (\mathbf{Id} + \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{A})^{-1}(\mathbf{Id} - \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{B})\boldsymbol{z}_k. \tag{8.6.3}$$

Algorithm 16: A Primal–Dual splitting method

Initial: Choose $\gamma_R, (\gamma_{J_i})_i > 0$. For $k = 0$, $x_0 \in \mathbb{R}^n$, $v_{i,0} \in \mathbb{R}^{m_i}$, $i \in \{1, ..., m\}$;

repeat

$$\left\lfloor \begin{array}{l} x_{k+1} = \operatorname{prox}_{\gamma_R R}\big(x_k - \gamma_R \nabla F(x_k) - \gamma_R \sum_i L_i^T v_{i,k}\big) \\ \bar{x}_{k+1} = 2x_{k+1} - x_k \\ \text{For } i = 1, ..., m \\ \quad \left\lfloor v_{i,k+1} = \operatorname{prox}_{\gamma_{J_i} J_i^*}\big(v_{i,k} - \gamma_{J_i} \nabla G_i^*(v_{i,k}) + \gamma_{J_i} L_i \bar{x}_{k+1}\big), \right. \end{array} \right. \tag{8.6.1}$$

$k = k + 1$;

until *convergence*;

Local convergence analysis Let $(x^\star, v_1^\star, ..., v_m^\star)$ be a saddle point of $(\boldsymbol{\mathcal{P}}_{\mathrm{SP}}^m)$, define the following functions

$$\overline{J_i^*}(v) \stackrel{\text{def}}{=} J_i^*(v) - \langle v,\, L_i x^\star - \nabla G_i^*(v_i^\star)\rangle,\ \ i \in \{1, ..., m\}, \tag{8.6.4}$$

and the Riemannian Hessian of $\overline{J_i^*}$,

$$H_{\overline{J_i^*}} \stackrel{\text{def}}{=} \mathcal{P}_{T_{v_i^\star}^{J_i^*}} \nabla^2_{\mathcal{M}_{v_i^\star}^{J_i^*}} \overline{J_i^*}(v_i^\star) \mathcal{P}_{T_{v_i^\star}^{J_i^*}} \quad \text{and} \quad W_{\overline{J_i^*}} \stackrel{\text{def}}{=} (\mathrm{Id}_{m_i} + H_{\overline{J_i^*}})^{-1}, i \in \{1, ..., m\}. \tag{8.6.5}$$

Owing to Lemma 8.3.1, we have that $W_{\overline{J_i^*}}, i \in \{1, ..., m\}$ are *firmly non-expansive* if (ND_m) holds. Now suppose that G_i^* locally is C^2 around $v_i^\star$, and define the restricted Hessian $H_{G_i^*} \stackrel{\text{def}}{=} \mathcal{P}_{T_{v_i^\star}^{J_i^*}} \nabla^2 G_i^*(v_i^\star) \mathcal{P}_{T_{v_i^\star}^{J_i^*}}$.

Define $\overline{H}_{G_i^*} \stackrel{\text{def}}{=} \mathrm{Id}_{m_i} - \gamma_{J_i^*} H_{G_i^*}$, $\overline{L}_i \stackrel{\text{def}}{=} \mathcal{P}_{T_{v_i^\star}^{J_i^*}} L_i \mathcal{P}_{T_{x^\star}^R}$, and the matrix

$$M_{\mathrm{PD}} \stackrel{\text{def}}{=} \begin{bmatrix} W_{\overline{R}} \overline{H}_F & -\gamma_R W_{\overline{R}} \overline{L}_1^T & \cdots & -\gamma_R W_{\overline{R}} \overline{L}_m^T \\ \gamma_{J_1^*} W_{J_1^*} \overline{L}_1 (2 W_{\overline{R}} \overline{H}_F - \mathrm{Id}_n) & W_{J_1^*}(\overline{H}_{G_1^*} - 2\gamma_{J_1^*} \gamma_R \overline{L}_1 W_{\overline{R}} \overline{L}_1^T) & & \\ \vdots & & \ddots & \\ \gamma_{J_m^*} W_{J_m^*} \overline{L}_m (2 W_{\overline{R}} \overline{H}_F - \mathrm{Id}_n) & \cdots & & W_{J_m^*}(\overline{H}_{G_m^*} - 2\gamma_{J_m^*} \gamma_R \overline{L}_m W_{\overline{R}} \overline{L}_m^T) \end{bmatrix}. \tag{8.6.6}$$

Using the same strategy of the proof of Lemma 2.5.6, one can show that M_{PD} is convergent, which again is denoted as M_{PD}^∞, and $\rho(M_{\mathrm{PD}} - M_{\mathrm{PD}}^\infty) < 1$.

Corollary 8.6.1. *For the Primal–Dual splitting algorithm 16, choose $\gamma_R, (\gamma_{J_i})_i > 0$ such that*

$$2 \min\{\beta_F, \beta_{G_1}, ..., \beta_{G_m}\} \min\left\{\frac{1}{\gamma_R}, \frac{1}{\gamma_{J_1}}, ..., \frac{1}{\gamma_{J_m}}\right\}\Big(1 - \sqrt{\gamma_R \textstyle\sum_i \gamma_{J_i} \|L_i\|^2}\Big) > 1. \tag{8.6.7}$$

Then there exists a saddle point $(x^\star, v_1^\star, ..., v_m^\star) \in \mathcal{X} \times \boldsymbol{\mathcal{V}}$, such that $(x_k, v_{1,k}, ..., v_{m,k}) \to (x^\star, v_1^\star, ..., v_m^\star)$. If moreover $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star}^R)$ and $J_i^* \in \mathrm{PSF}_{v_i^\star}(\mathcal{M}_{v_i^\star}^{J_i^*})$, $i \in \{1, ..., m\}$, and there holds

$$\begin{array}{c} -\sum_i L_i^T v_i^\star - \nabla F(x^\star) \in \mathrm{ri}\big(\partial R(x^\star)\big) \\ L_i x^\star - \nabla G_i^*(v^\star) \in \mathrm{ri}\big(\partial J_i^*(v^\star)\big),\ \forall i \in \{1, ..., m\}. \end{array} \tag{ND_m}$$

Then,

(i) *there $\exists K > 0$ such that for all $k \geq K$,*

$$(x_k, v_{1,k}, ..., v_{m,k}) \in \mathcal{M}_{x^\star}^R \times \mathcal{M}_{v_1^\star}^{J_1^*} \times \cdots \times \mathcal{M}_{v_m^\star}^{J_m^*}.$$

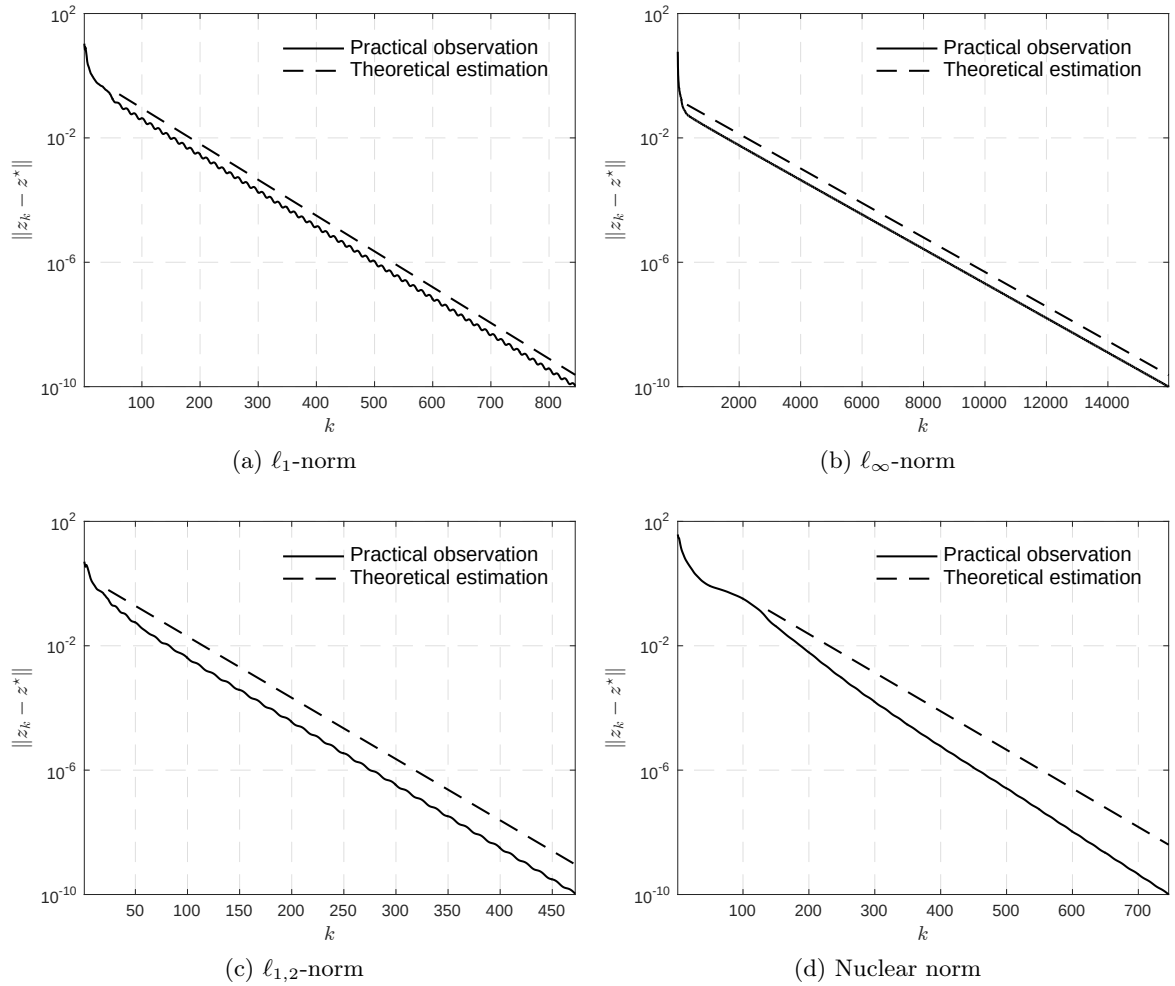
(ii) *Given any $\rho \in]\rho(M_{\mathrm{PD}} - M_{\mathrm{PD}}^\infty), 1[$, there exist a K large enough such that for all $k \geq K$,*

$$\|(\mathbf{Id} - M_{\mathrm{PD}}^\infty)(\boldsymbol{z}_k - \boldsymbol{z}^\star)\| = O(\rho^{k-K}). \tag{8.6.8}$$

If moreover, $R, J_1^*, ..., J_m^*$ are locally polyhedral around $(x^\star, v_1^\star, ..., v_m^\star)$, then we have directly have $\|\boldsymbol{z}_k - \boldsymbol{z}^\star\| = O(\rho^{k-K})$.

8.7 Numerical experiments

Linear inverse problems Consider again the linear inverse problem (4.5.3) of Section 4.5, and apply the Chambolle-Pock Primal–Dual splitting method. Figure 8.1 displays the profile of $\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|$ as a function of k, and the starting point of the *dashed* line is the iteration number at which the active partial smoothness manifold of $\mathcal{M}^R_{x^\star}$ is identified (recall that $\mathcal{M}^{J^*}_{v^\star} = \{0\}$ which is trivially identified from the first iteration). One can see that for the ℓ_1 and ℓ_∞ norms, Theorem 8.3.6 applies and our estimates are exact. For the case of $\ell_{1,2}$-norm and nuclear norm, though not optimal, our estimates are very tight.



(a) ℓ_1-norm (b) ℓ_∞-norm

(c) $\ell_{1,2}$-norm (d) Nuclear norm

Figure 8.1: Observed (solid) and predicted (dashed) convergence profiles of Primal–Dual splitting (8.1.2) in terms of $\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|$. (a) ℓ_1-norm. (b) ℓ_∞-norm. (c) $\ell_{1,2}$-norm. (d) Nuclear norm. The starting point of the *dashed* line is the iteration at which the active manifold of R is identified.

Noise removal Now consider the noise removal problem (7.8.1) in Section 7.8. Let $J = \iota_{\|f-\cdot\|_p \leq \tau}$ and $R = \|\cdot\|_1$, then (7.8.1) can be written as

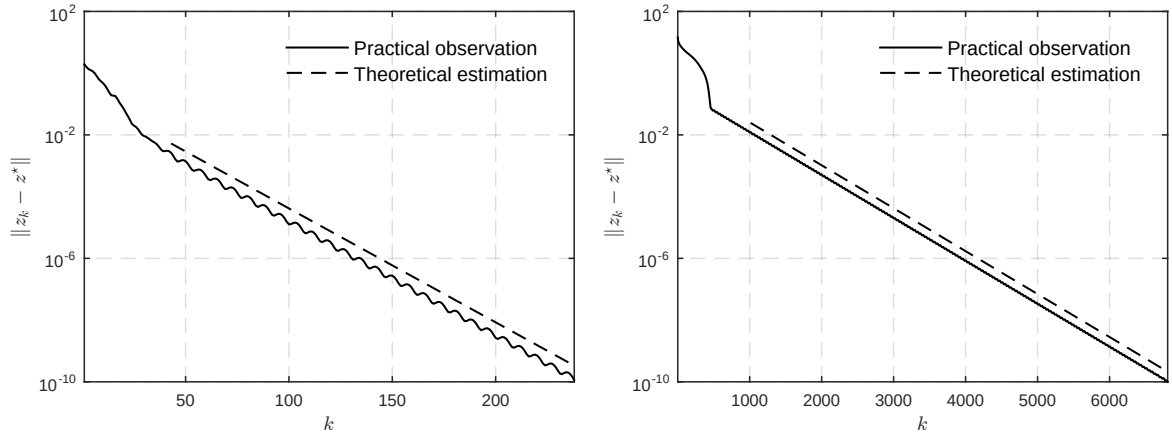
$$\min_{x \in \mathbb{R}^n} \iota_{\|f-\cdot\|_p \leq \tau} + \|\mathrm{D}_{\mathrm{DIF}}\, x\|_1,$$

whose saddle point problem reads,

$$\min_{x \in \mathbb{R}^n} \max_{v \in \mathbb{R}^m} \iota_{\|f-\cdot\|_p \leq \tau} + \langle \mathrm{D}_{\mathrm{DIF}}\, x,\, v \rangle - \iota_{\|\cdot\|_\infty \leq 1}(v).$$

Clearly, both two indicator functions are polyhedral, and their proximal operator are simple to compute. The corresponding local convergence profiles are depicted in Figure 8.2(a)-(b). Owing to polyhedrality,

our rate predictions are again exact.



(a) Sparse noise removal $p = 1$ (b) Uniform noise removal $p = \infty$

Figure 8.2: Observed (solid) and predicted (dashed) convergence profiles of Primal–Dual splitting (8.1.3) in terms of $\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|$. (a) Sparse noise removal. (b) Uniform noise removal.

8.7.1 Choices of θ and γ_J, γ_R

In this part, we present a comparison on different choices of θ and γ_J, γ_R to see their influences on the finite identification and local linear convergence rate. Two examples are consider for these comparisons, problem (4.5.3) with R being ℓ_1-norm and $\ell_{1,2}$-norm.

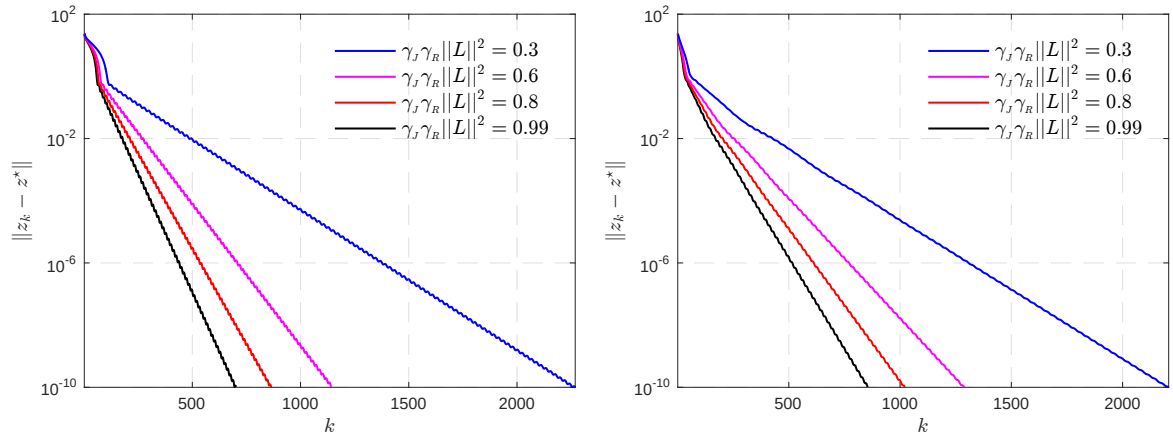
Fixed θ We consider first the case of fixing θ, and changing the value of $\gamma_J\gamma_R\|L\|^2$. 4 different cases are considered, which are

$$\gamma_J\gamma_R\|L\|^2 \in \{0.3, 0.6, 0.8, 0.99\},$$

and we fix $\theta = 1$, moreover we set $\gamma_J = \gamma_R$. The comparison result is shown in Figure 8.3, and we have the following observations

(i) The *smaller* the value of $\gamma_J\gamma_R\|L\|^2$, the *slower* the iteration converges.

(ii) *Bigger* value of γ_J leads to *faster* identification (since J^* is globally C^2-smooth, so only the identification of R for this case).



(a) ℓ_1-norm (b) $\ell_{1,2}$-norm

Figure 8.3: Comparison of the choice of γ_J, γ_R when θ is fixed.

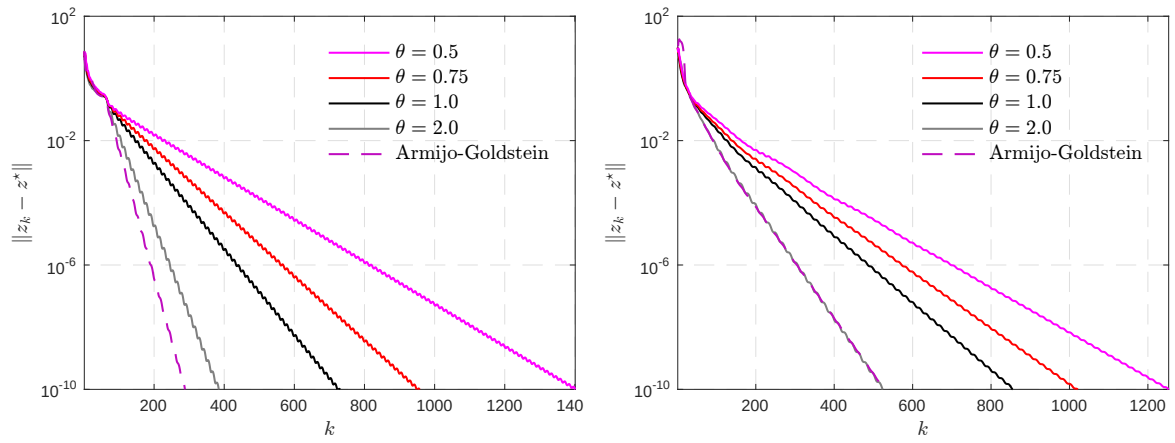
Fixed $\gamma_J\gamma_R\|L\|^2$ Now we turn to the opposite direction, fix the value of $\gamma_J\gamma_R\|L\|^2$ and then change θ. In the test, we fixed $\gamma_J\gamma_R\|L\|^2 = 0.9$ and $\gamma_J = \gamma_R$, 5 different choices of θ are considered, which are

$$\theta \in \{0.5, 0.75, 1.0, 2.0\},$$

plus one with Armijo-Goldstein-rule for adaptive update θ. Although there's no convergence guarantee for $\theta = 2.0$, in the tests it converges and we choose to put it here as an illustration of the effects of θ. The result is shown in Figure 8.4, and we have the following observations

(i) Similarly to the previous one, the *smaller* the value of θ, the *slower* the iteration converges. Also, the Armijo-Goldstein-rule is the *fastest* one of all.

(ii) Interestingly, the value of θ has no impacts to the identification of the iteration.



(a) ℓ_1-norm (b) $\ell_{1,2}$-norm

Figure 8.4: Comparison of the choice of θ when γ_J, γ_R are fixed.

Fixed θ and $\gamma_J\gamma_R$ For the above comparisons, we fix $\gamma_J = \gamma_R$, so for this comparison, we compare the different choices of them. We fix $\theta = 1$ and $\gamma_J\gamma_R\|L\|^2 = 0.99$, then we choose

$$\gamma_J \in \{0.25, 0.5, 1, 2\} \quad \text{and} \quad \gamma_R = \frac{0.99}{\gamma_J\|L\|^2}.$$

Figure 8.5 shows the comparison result, we can also make two observations:

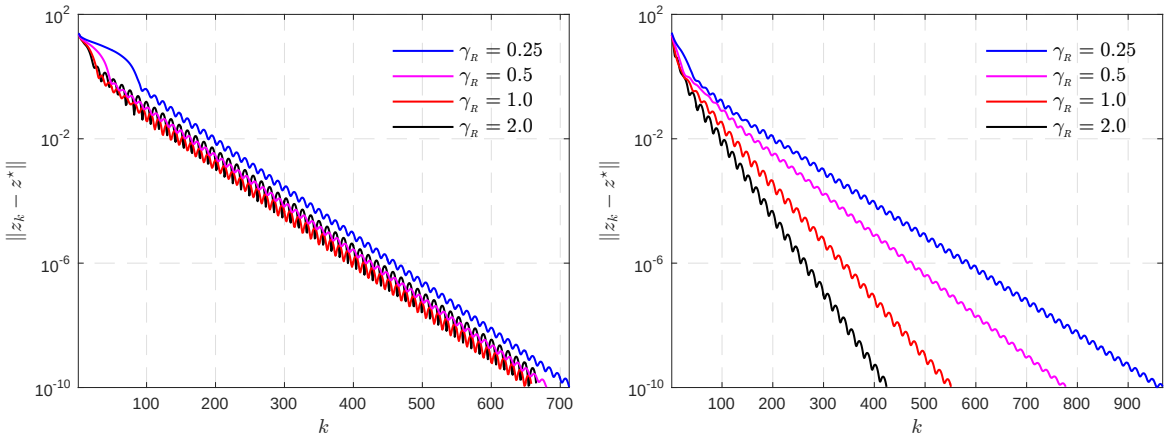
(i) For the ℓ_1-norm, since both functions are polyhedral, local convergence rate are the same for all choices of γ_R, see (8.3.9) for the expression of the rate. The only difference is the identification speed, $\gamma_R = 0.25$ gives the slowest identification, however it uses almost the same number of iterations reaching the given accuracy.

(ii) For the $\ell_{1,2}$-norm, on the other hand, the choice of γ_R affects both the identification and local convergence rate. It can be observed that *bigger* γ_R leads to *faster* local rate, however, it does not mean that the bigger the better. In fact, too big value will slow down the convergence.

To summarize the above comparison, in practice, it is better to chose θ and $\gamma_J\gamma_R$ as big as possible, moreover, the choices of γ_J and γ_R should be determined based on the properties of the functions at hand (*i.e.* polyhedral or others).

8.7.2 Oscillation of the method

We decidate the last part of the numerical experiment to the oscillation of the Primal–Dual splitting method when dealing with polyhedral functions. As we have seen from the above experiments, oscillation of $\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|$ happens for all examples where the functions are polyhedral, even for the



(a) ℓ_1-norm (b) $\ell_{1,2}$-norm

Figure 8.5: Comparison of fixed θ and $\gamma_J\gamma_R$, but varying γ_J.

non-polyhedral $\ell_{1,2}$-norm (for the ℓ_∞-norm, due to the fact that the oscillation period is too small compared to the number of iteration, hence it is not visible).

Now to verify our discussion in Section 8.4, we consider problem (4.5.3) with J being the ℓ_1-norm for this illustration, and the result is shown in Figure 8.6. As revealed in (8.4.4), the *argument* of the leading eigenvalue of $\mathcal{F} - M_{\mathrm{PD}}^{\infty}$ is controlled by $\theta\gamma_J\gamma_R$, so is the oscillation period. Therefore, the value $\gamma_J\gamma_R$ is tuned such that the oscillation period is an *integer*, and $\pi/\omega = 12$ for the example we tested. Figure 8.6 shows graphically the observed oscillation, apparently the oscillation pattern coincides well with the theoretical estimation.

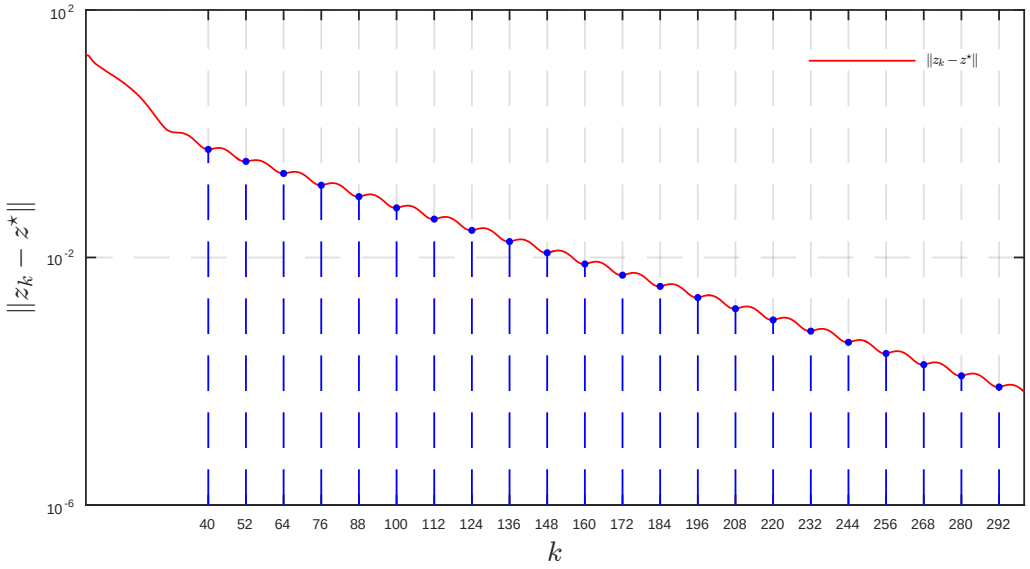


Figure 8.6: Oscillation of the Primal–Dual splitting when solving polyhedral functions.

8.8 Proofs of main theorems

Proof of Theorem 8.2.1.

(i) From the iteration scheme (8.1.3), for the updating of x_{k+1}, we have for x_{k+1},

$$x_{k+1} = \mathrm{prox}_{\gamma_J R}\big(x_k - \gamma_R \nabla F(x_k)) - \gamma_R L^* v_k\big)$$
$$\iff \frac{1}{\gamma_R}\big(x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1}\big) \in \partial R(x_{k+1}),$$

then, as $\mathrm{Id} - \gamma_R \nabla F$ is non-expansive,

$$\begin{aligned}
&\operatorname{dist}\big(-L^*v^\star - \nabla F(x^\star), \partial R(x_{k+1})\big) \\
&\leq \| - L^*v^\star - \nabla F(x^\star) - \tfrac{1}{\gamma_R}(x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1})\| \\
&= \| - L^*v^\star - \nabla F(x^\star) - \tfrac{1}{\gamma_R}(x_k - x^\star - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1} + x^\star)\| \\
&\leq \tfrac{1}{\gamma_R}\|x^\star - \gamma_R \nabla F(x^\star) - (x_k - \gamma_R \nabla F(x_k)) - (x_{k+1} - x^\star)\| + \|L\|\|v_k - v^\star\| \\
&\leq \tfrac{1}{\gamma_R}(\|x_k - x^\star\| + \|x_{k+1} - x^\star\|) + \|L\|\|v_k - v^\star\| \to 0.
\end{aligned}$$

Similarly for v_{k+1}

$$\begin{aligned}
& v_{k+1} = \operatorname{prox}_{\gamma_R J^*}\big(v_k - \gamma_J \nabla G^*(v_k) + \gamma_J L \bar{x}_{k+1}\big) \\
\iff & \tfrac{1}{\gamma_J}\big(v_k - \gamma_J \nabla G^*(v_k) + \gamma_J L \bar{x}_{k+1} - v_{k+1}\big) \in \partial J^*(v_{k+1}),
\end{aligned}$$

then we have

$$\begin{aligned}
&\operatorname{dist}\big(Lx^\star - \nabla G^*(v^\star), \partial J^*(v_{k+1})\big) \\
&\leq \|Lx^\star - \nabla G^*(v^\star) - \tfrac{1}{\gamma_J}(v_k - \gamma_J \nabla G^*(v_k) + \gamma_J L \bar{x}_{k+1} - v_{k+1})\| \\
&\leq \tfrac{1}{\gamma_J}(\|v_k - v^\star\| + \|v_{k+1} - v^\star\|) + \|L\|\big((1+\theta)\|x_{k+1} - x^\star\| + \theta\|x_k - x^\star\|\big) \to 0.
\end{aligned}$$

By assumption, $J^* \in \Gamma_0(\mathbb{R}^m), R \in \Gamma_0(\mathbb{R}^n)$, hence they are sub-differentially continuous at every point in their respective domains [153, Example 13.30], and in particular at $v^\star$ and $x^\star$. It then follows that $J^*(v_k) \to J^*(v^\star)$ and $R(x_k) \to R(x^\star)$. Altogether with the non-degeneracy condition ($\mathrm{ND}_{\mathrm{PD}}$), shows that the conditions of Theorem 5.1.5 are fulfilled for J^* and R, and the finite identification claim follows.

(ii) See proof of Theorem 7.3.1. □

Proof of Proposition 8.3.2. From the update of x_k, we have

$$\begin{aligned}
x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1} &\in \gamma_R \partial R(x_{k+1}), \\
-\gamma_R \nabla F(x^\star) - \gamma_R L^* v^\star &\in \gamma_R \partial R(x^\star).
\end{aligned}$$

Denote τ_k^R the parallel translation from $T^R_{x_k}$ to $T^R_{x^\star}$. Then project on to corresponding tangent spaces and apply parallel translation,

$$\begin{aligned}
\gamma_R \tau_k^R \nabla_{\mathcal{M}^R_{x^\star}} R(x_{k+1}) &= \tau_k^R \mathcal{P}_{T^R_{x_{k+1}}}(x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1}) \\
&= \mathcal{P}_{T^R_{x^\star}}(x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1}) \\
&\quad + (\tau_k^R \mathcal{P}_{T^R_{x_{k+1}}} - \mathcal{P}_{T^R_{x^\star}})(x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1}), \\
\gamma_R \nabla_{\mathcal{M}^R_{x^\star}} R(x^\star) &= \mathcal{P}_{T^R_{x^\star}}(-\gamma_R \nabla F(x^\star) - \gamma_R L^* v^\star),
\end{aligned}$$

which leads to

$$\begin{aligned}
&\gamma_R \tau_k^R \nabla_{\mathcal{M}^R_{x^\star}} R(x_{k+1}) - \gamma_R \nabla_{\mathcal{M}^R_{x^\star}} R(x^\star) \\
&= \mathcal{P}_{T^R_{x^\star}}\big((x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1}) - (x^\star - \gamma_R \nabla F(x^\star) - \gamma_R L^* v^\star - x^\star)\big) \\
&\quad \underbrace{+ (\tau_k^R \mathcal{P}_{T^R_{x_{k+1}}} - \mathcal{P}_{T^R_{x^\star}})(-\gamma_R \nabla F(x^\star) - \gamma_R L^* v^\star)}_{\textbf{Term 1}} \\
&\quad \underbrace{+ (\tau_k^R \mathcal{P}_{T^R_{x_{k+1}}} - \mathcal{P}_{T^R_{x^\star}})\big((x_k - \gamma_R \nabla F(x_k) - \gamma_R L^* v_k - x_{k+1}) + (\gamma_R \nabla F(x^\star) + \gamma_R L^* v^\star)\big)}_{\textbf{Term 2}}.
\end{aligned} \tag{8.8.1}$$

Moving **Term 1** to the other side leads to

$$\begin{aligned}
&\gamma_R\tau_k^R\nabla_{\mathcal{M}_{x^\star}^R}R(x_{k+1})-\gamma_R\nabla_{\mathcal{M}_{x^\star}^R}R(x^\star)-(\tau_k^R\mathcal{P}_{T_{x_{k+1}}^R}-\mathcal{P}_{T_{x^\star}^R})(-\gamma_R\nabla F(x^\star)-\gamma_R L^*v^\star)\\
&=\gamma_R\tau_k^R\big(\nabla_{\mathcal{M}_{x^\star}^R}R(x_{k+1})+(L^*v^\star+\nabla F(x^\star))\big)-\gamma_R\big(\nabla_{\mathcal{M}_{x^\star}^R}R(x^\star)+(L^*v^\star+\nabla F(x^\star))\big)\\
&=\gamma_R\mathcal{P}_{T_{x^\star}^R}\nabla^2_{\mathcal{M}_{x^\star}^R}\overline{R}(x^\star)\mathcal{P}_{T_{x^\star}^R}(x_{k+1}-x^\star)+o(\|x_{k+1}-x^\star\|),
\end{aligned}$$

where Lemma 6.3.3 is applied. Since $x_{k+1}=\mathrm{prox}_{\gamma_R R}(x_k-\gamma_R\nabla F(x_k)-\gamma_R L^*v_k)$, $\mathrm{prox}_{\gamma_R R}$ is firmly non-expansive and $\mathrm{Id}_n-\gamma_R\nabla F$ is non-expansive (under the parameter setting), then

$$\begin{aligned}
&\|(x_k-\gamma_R\nabla F(x_k)-\gamma_R L^*v_k-x_{k+1})-(x^\star-\gamma_R\nabla F(x^\star)-\gamma_R L^*v^\star-x^\star)\|\\
&\le\|(\mathrm{Id}_n-\gamma_R\nabla F)(x_k)-(\mathrm{Id}_n-\gamma_R\nabla F)(x^\star)\|+\gamma_R\|L^*v_k-L^*v^\star\|\\
&\le\|x_k-x^\star\|+\gamma_R\|L\|\|v_k-v^\star\|.
\end{aligned}\tag{8.8.2}$$

Therefore, for **Term 2**, owing to Lemma 2.6.2, we have

$$\begin{aligned}
&(\tau_k^R\mathcal{P}_{T_{x_{k+1}}^R}-\mathcal{P}_{T_{x^\star}^R})\big((x_k-\gamma_R\nabla F(x_k)-\gamma_R L^*v_k-x_{k+1})-(x^\star-\gamma_R\nabla F(x^\star)-\gamma_R L^*v^\star-x^\star)\big)\\
&=o(\|x_k-x^\star\|+\gamma_R\|L\|\|v_k-v^\star\|).
\end{aligned}$$

Therefore, from (8.8.1), and apply Lemma 2.6.1 to $(x_{k+1}-x^\star)$ and $(x_k-x^\star)$, we get

$$\begin{aligned}
&(\mathrm{Id}_n+H_{\overline{R}})(x_{k+1}-x^\star)+o(\|x_{k+1}-x^\star\|)\\
&=(x_k-x^\star)-\gamma_R\mathcal{P}_{T_{x^\star}^R}(\nabla F(x_k)-\nabla F(x^\star))-\gamma_R\mathcal{P}_{T_{x^\star}^R}L^*(v_k-v^\star)+o(\|x_k-x^\star\|+\gamma_R\|L\|\|v_k-v^\star\|),
\end{aligned}$$

then apply Taylor expansion to ∇F, and apply Lemma 2.6.1 to $(v_k-v^\star)$,

$$\begin{aligned}
&(\mathrm{Id}_n+H_{\overline{R}})(x_{k+1}-x^\star)\\
&=(\mathrm{Id}_n-\gamma_R H_F)(x_k-x^\star)-\gamma_R\overline{L}^*(v_k-v^\star)+o(\|x_k-x^\star\|+\gamma_R\|L\|\|v_k-v^\star\|).
\end{aligned}\tag{8.8.3}$$

Then inverse $(\mathrm{Id}_n+H_{\overline{R}})$ and apply Lemma 2.6.1, we get

$$x_{k+1}-x^\star=W_{\overline{R}}\overline{H}_F(x_k-x^\star)-\gamma_R W_{\overline{R}}\overline{L}^*(v_k-v^\star)+o(\|x_k-x^\star\|+\gamma_R\|L\|\|v_k-v^\star\|).\tag{8.8.4}$$

Now from the update of v_{k+1}

$$\begin{aligned}
v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1}&\in\gamma_J\partial J^*(v_{k+1}),\\
v^\star-\gamma_J\nabla G^*(v^\star)+\gamma_J Lx^\star-v^\star&\in\gamma_J\partial J^*(v^\star).
\end{aligned}$$

Denote $\tau_{k+1}^{J^*}$ the parallel translation from $T_{v_{k+1}}^{R^*}$ to $T_{v^\star}^{R^*}$, then

$$\begin{aligned}
\gamma_J\tau_{k+1}^{J^*}\nabla_{\mathcal{M}_{v^\star}^{J^*}}J^*(v_{k+1})&=\tau_{k+1}^{J^*}\mathcal{P}_{T_{v_{k+1}}^{J^*}}(v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1})\\
&=\mathcal{P}_{T_{v^\star}^{J^*}}(v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1})\\
&\quad+(\tau_{k+1}^{J^*}\mathcal{P}_{T_{v_{k+1}}^{J^*}}-\mathcal{P}_{T_{v^\star}^{J^*}})(v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1}),\\
\gamma_J\nabla_{\mathcal{M}_{v^\star}^{J^*}}J^*(v^\star)&=\mathcal{P}_{T_{v^\star}^{J^*}}(v^\star-\gamma_J\nabla G^*(v^\star)+\gamma_J Lx^\star-v^\star)
\end{aligned}$$

which leads to

$$\begin{aligned}
&\gamma_J\tau_{k+1}^{J^*}\nabla_{\mathcal{M}_{v^\star}^{J^*}}J^*(v_{k+1})-\gamma_J\nabla_{\mathcal{M}_{v^\star}^{J^*}}J^*(v^\star)\\
&=\mathcal{P}_{T_{v^\star}^{J^*}}\big((v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1})-(v^\star-\gamma_J\nabla G^*(v^\star)+\gamma_J Lx^\star-v^\star)\big)\\
&\quad+(\tau_{k+1}^{J^*}\mathcal{P}_{T_{v_{k+1}}^{J^*}}-\mathcal{P}_{T_{v^\star}^{J^*}})(v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1})\\
&=\mathcal{P}_{T_{v^\star}^{J^*}}\big((v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1})-(v^\star-\gamma_J\nabla G^*(v^\star)+\gamma_J Lx^\star-v^\star)\big)\\
&\quad\underbrace{+(\tau_{k+1}^{J^*}\mathcal{P}_{T_{v_{k+1}}^{J^*}}-\mathcal{P}_{T_{v^\star}^{J^*}})(\gamma_J Lx^\star-\gamma_J\nabla G^*(v^\star))}_{\textbf{Term 3}}\\
&\quad\underbrace{+(\tau_{k+1}^{J^*}\mathcal{P}_{T_{v_{k+1}}^{J^*}}-\mathcal{P}_{T_{v^\star}^{J^*}})\big((v_k-\gamma_J\nabla G^*(v_k)+\gamma_J L\bar{x}_{k+1}-v_{k+1})+\gamma_J(\nabla G^*(v^\star)-Lx^\star)\big)}_{\textbf{Term 4}}.
\end{aligned}\tag{8.8.5}$$

Similarly to the previous analysis, for **Term 3**, move to the lefthand side of the inequality and apply Lemma 6.3.3,

$$\begin{aligned}
&\gamma_J \tau_{k+1}^{J^*} \nabla_{\mathcal{M}_{v^\star}^{J^*}} J^*(v_{k+1}) - \gamma_J \nabla_{\mathcal{M}_{v^\star}^{J^*}} J^*(v^\star) - (\tau_{k+1}^{J^*} \mathcal{P}_{T_{v_{k+1}}^{J^*}} - \mathcal{P}_{T_{v^\star}^{J^*}})(\gamma_J L x^\star - \gamma_J \nabla G^*(v^\star)) \\
&= \gamma_J \tau_{k+1}^{J^*} \big(\nabla_{\mathcal{M}_{v^\star}^{J^*}} J^*(v_{k+1}) - (L x^\star - \nabla G^*(v^\star))\big) - \gamma_J \big(\nabla_{\mathcal{M}_{v^\star}^{J^*}} J^*(v^\star) - (L x^\star - \nabla G^*(v^\star))\big) \\
&= \gamma_J \mathcal{P}_{T_{v^\star}^{J^*}} \nabla^2_{\mathcal{M}_{v^\star}^{J^*}} \overline{R}^*(v^\star) \mathcal{P}_{T_{v^\star}^{J^*}} (v_{k+1} - v^\star) + o(\|v_{k+1} - v^\star\|).
\end{aligned}$$

Since $\theta \leq 1$, we have

$$\begin{aligned}
\|\bar{x}_{k+1} - x^\star\| &\leq (1+\theta)\|x_{k+1} - x^\star\| + \theta \|x_k - x^\star\| \\
&\leq 2(\|x_k - x^\star\| + \gamma_R \|L\| \|v_k - v^\star\|) + \|x_k - x^\star\| \\
&= 3\|x_k - x^\star\| + 2\gamma_R \|L\| \|v_k - v^\star\|.
\end{aligned}$$

Then for **Term 4**, since $\gamma_J \gamma_R \|L^2\| < 1$, $\operatorname{prox}_{\gamma_J J^*}$ is firmly non-expansive and $\mathrm{Id}_m - \gamma_J \nabla G^*$ is non-expansive, we have

$$\begin{aligned}
&(\tau_{k+1}^{J^*} \mathcal{P}_{T_{v_{k+1}}^{J^*}} - \mathcal{P}_{T_{v^\star}^{J^*}}) \big((v_k - \gamma_J \nabla G^*(v_k) + \gamma_J L \bar{x}_{k+1} - v_{k+1}) - (v^\star - \gamma_J \nabla G^*(v^\star) + \gamma_J L x^\star - v^\star)\big) \\
&= o(\|v_k - v^\star\| + \gamma_J \|L\| \|x_k - x^\star\|).
\end{aligned}$$

Therefore, from (8.8.5), apply Lemma 2.6.1 to $(v_{k+1} - v^\star)$ and $(v_k - v^\star)$, we get

$$\begin{aligned}
&(\mathrm{Id}_m + H_{\overline{J^*}})(v_{k+1} - v^\star) \\
&= (\mathrm{Id}_m - \gamma_J H_{G^*})(v_k - v^\star) + \gamma_J \overline{L}(\bar{x}_{k+1} - x^\star) + o(\|v_k - v^\star\| + \gamma_J \|L\| \|x_k - x^\star\|).
\end{aligned} \tag{8.8.6}$$

Then similar to (8.8.4), we get from (8.8.6)

$$\begin{aligned}
v_{k+1} - v^\star &= W_{\overline{J^*}} \overline{H}_{G^*}(v_k - v^\star) + \gamma_J W_{\overline{J^*}} \overline{L}(\bar{x}_{k+1} - x^\star) + o(\|v_k - v^\star\| + \gamma_J \|L\| \|x_k - x^\star\|) \\
&= W_{\overline{J^*}} \overline{H}_{G^*}(v_k - v^\star) + (1+\theta)\gamma_J W_{\overline{J^*}} \overline{L}(x_{k+1} - x^\star) - \theta \gamma_J W_{\overline{J^*}} \overline{L}(x_k - x^\star) \\
&\quad + o(\|v_k - v^\star\| + \gamma_J \|L\| \|x_k - x^\star\|) \\
&= W_{\overline{J^*}} \overline{H}_{G^*}(v_k - v^\star) - \theta \gamma_J W_{\overline{J^*}} \overline{L}(x_k - x^\star) \\
&\quad + (1+\theta)\gamma_J W_{\overline{J^*}} \overline{L} \big(W_{\overline{R}} \overline{H}_F (x_k - x^\star) - \gamma_R W_{\overline{R}} \overline{L}^*(v_k - v^\star)\big) \\
&\quad + o(\|x_k - x^\star\| + \gamma_R \|L\| \|v_k - v^\star\|) + o(\|v_k - v^\star\| + \gamma_J \|L\| \|x_k - x^\star\|) \\
&= \big(W_{\overline{J^*}} \overline{H}_{G^*} - (1+\theta)\gamma_J \gamma_R W_{\overline{J^*}} \overline{L} W_{\overline{R}} \overline{L}^*\big)(v_k - v^\star) \\
&\quad + \big((1+\theta)\gamma_J W_{\overline{J^*}} \overline{L} W_{\overline{R}} \overline{H}_F - \theta \gamma_J W_{\overline{J^*}} \overline{L}\big)(x_k - x^\star) \\
&\quad + o(\|x_k - x^\star\| + \gamma_R \|L\| \|v_k - v^\star\|) + o(\|v_k - v^\star\| + \gamma_J \|L\| \|x_k - x^\star\|).
\end{aligned} \tag{8.8.7}$$

Now we consider the small o-terms. For the 2 small o-terms in (8.8.3) and (8.8.6). First, let a_1, a_2 be two constants, then we have

$$|a_1| + |a_2| = \sqrt{(|a_1| + |a_2|)^2} \leq \sqrt{2(a_1^2 + a_2^2)} = \sqrt{2} \left\| \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\|.$$

Denote $b = \max\{1, \gamma_J \|L\|, \gamma_R \|L\|\}$, then

$$\begin{aligned}
&(\|v_k - v^\star\| + \gamma_J \|L\| \|x_k - x^\star\|) + (\|x_k - x^\star\| + \gamma_R \|L\| \|v_k - v^\star\|) \\
&\leq 2b(\|x_k - x^\star\| + \|v_k - v^\star\|) \leq 2\sqrt{2} b \left\| \begin{pmatrix} x_k - x^\star \\ v_k - v^\star \end{pmatrix} \right\|.
\end{aligned}$$

Combining this with (8.8.4) and (8.8.7), and ignoring the constants of the small o-term leads to the claimed result. □

Proof of Proposition 8.3.4.

(i) When $\theta = 1$, M_{PD} becomes

$$M_{\text{PD}} = \begin{bmatrix} W_{\overline{R}}\overline{H}_F & -\gamma_R W_{\overline{R}}\overline{L}^* \\ 2\gamma_J W_{\overline{J^*}}\overline{L}W_{\overline{R}}\overline{H}_F - \gamma_J W_{\overline{J^*}}\overline{L} & W_{\overline{J^*}}\overline{H}_{G^*} - 2\gamma_R\gamma_J W_{\overline{J^*}}\overline{L}W_{\overline{R}}\overline{L}^* \end{bmatrix} \tag{8.8.8}$$

Next we show that M_{PD} is *averaged non-expansive*.

First define the following matrices

$$\boldsymbol{A} = \begin{bmatrix} H_J/\gamma_R & \overline{L}^* \\ -\overline{L} & H_{R^*}/\gamma_J \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} H_F & 0 \\ 0 & H_{G^*} \end{bmatrix}, \quad \boldsymbol{\mathcal{V}} = \begin{bmatrix} \mathrm{Id}_n/\gamma_R & -\overline{L}^* \\ -\overline{L} & \mathrm{Id}_m/\gamma_J \end{bmatrix}, \tag{8.8.9}$$

where we have $\boldsymbol{A}$ is maximal monotone [41], $\boldsymbol{B}$ is $\min\{\beta_F, \beta_G\}$-cocoercive, and $\boldsymbol{\mathcal{V}}$ is ν-positive definite with $\nu = (1 - \sqrt{\gamma_J\gamma_R\|L\|^2})\min\{\frac{1}{\gamma_R}, \frac{1}{\gamma_J}\}$.

Now we have

$$\boldsymbol{\mathcal{V}} + \boldsymbol{A} = \begin{bmatrix} \frac{\mathrm{Id}_n + H_J}{\gamma_R} & 0 \\ -2\overline{L} & \frac{\mathrm{Id}_m + H_{R^*}}{\gamma_J} \end{bmatrix} \implies (\boldsymbol{\mathcal{V}} + \boldsymbol{A})^{-1} = \begin{bmatrix} \gamma_R W_{\overline{R}} & 0 \\ 2\gamma_J\gamma_R W_{\overline{J^*}}\overline{L}W_{\overline{R}} & \gamma_J W_{\overline{J^*}} \end{bmatrix},$$

and

$$\boldsymbol{\mathcal{V}} - \boldsymbol{B} = \begin{bmatrix} \frac{1}{\gamma_R}\overline{H}_F & -\overline{L}^* \\ -\overline{L} & \frac{1}{\gamma_J}\overline{H}_{G^*} \end{bmatrix}.$$

As a result, we get

$$\begin{aligned} (\boldsymbol{\mathcal{V}} + \boldsymbol{A})^{-1}(\boldsymbol{\mathcal{V}} - \boldsymbol{B}) &= \begin{bmatrix} \gamma_R W_{\overline{R}} & 0 \\ 2\gamma_J\gamma_R W_{\overline{J^*}}\overline{L}W_{\overline{R}} & \gamma_J W_{\overline{J^*}} \end{bmatrix} \begin{bmatrix} \frac{1}{\gamma_R}\overline{H}_F & -\overline{L}^* \\ -\overline{L} & \frac{1}{\gamma_J}\overline{H}_{G^*} \end{bmatrix} \\ &= \begin{bmatrix} W_{\overline{R}}\overline{H}_F & -\gamma_R W_{\overline{R}}\overline{L}^* \\ 2\gamma_J W_{\overline{J^*}}\overline{L}W_{\overline{R}}\overline{H}_F - \gamma_J W_{\overline{J^*}}\overline{L} & W_{\overline{J^*}}\overline{H}_{G^*} - 2\gamma_J\gamma_R W_{\overline{J^*}}\overline{L}W_{\overline{R}}\overline{L}^* \end{bmatrix}, \end{aligned}$$

which is exactly (8.8.8).

From Lemma 4.6.1 we know that $M_{\text{PD}} : \boldsymbol{\mathcal{K}}_{\boldsymbol{\mathcal{V}}} \to \boldsymbol{\mathcal{K}}_{\boldsymbol{\mathcal{V}}}$ is *averaged non-expansive*, hence it is convergent [16]. Then we have the induced matrix norm

$$\lim_{k\to\infty} \|M_{\text{PD}}^k - M_{\text{PD}}^\infty\|_{\mathcal{V}} = \lim_{k\to\infty} \|M_{\text{PD}} - M_{\text{PD}}^\infty\|_{\mathcal{V}}^k = 0.$$

Since we are in the finite dimensional space and $\boldsymbol{\mathcal{V}}$ is an isomorphism, then the above limit implies that

$$\lim_{k\to\infty} \|M_{\text{PD}} - M_{\text{PD}}^\infty\|^k = 0,$$

which means that $\rho(M_{\text{PD}} - M_{\text{PD}}^\infty) < 1$. The rest of the proof is classical using the spectral radius formula, see *e.g.* [18, Theorem 2.12(i)].

(ii) When J and R^* are locally polyhedral, then $W_{\overline{R}} = \mathrm{Id}_n, W_{\overline{J^*}} = \mathrm{Id}_m$, altogether with $F = 0, G = 0$, for any $\theta \in [0, 1]$, we have

$$M_{\text{PD}} = \begin{bmatrix} \mathrm{Id}_n & -\gamma_R\overline{L}^* \\ \gamma_J\overline{L} & \mathrm{Id}_m - \gamma_R\gamma_J(1+\theta)\overline{L}\overline{L}^* \end{bmatrix}. \tag{8.8.10}$$

With the SVD of $\overline{L}$, for M_{PD}, we have

$$\begin{aligned} M_{\text{PD}} &= \begin{bmatrix} \text{Id}_n & -\gamma_R \overline{L}^* \\ \gamma_J \overline{L} & \text{Id}_m - (1+\theta)\gamma_R\gamma_J \overline{L}\overline{L}^* \end{bmatrix} \\ &= \begin{bmatrix} YY^* & -\gamma_R Y \Sigma_{\overline{L}}^* X^* \\ \gamma_J X \Sigma_{\overline{L}} Y^* & XX^* - (1+\theta)\gamma_R\gamma_J X \Sigma_{\overline{L}}^2 X^* \end{bmatrix} \\ &= \begin{bmatrix} Y & \\ & X \end{bmatrix} \underbrace{\begin{bmatrix} \text{Id}_n & -\gamma_R \Sigma_L^* \\ \gamma_J \Sigma_L & \text{Id}_m - (1+\theta)\gamma_R\gamma_J \Sigma_L^2 \end{bmatrix}}_{\mathcal{F}_\Sigma} \begin{bmatrix} Y^* & \\ & X^* \end{bmatrix}. \end{aligned} \tag{8.8.11}$$

Since we assume that $\text{rank}(\overline{L}) = l \leq p$, then Σ_L can be represented as

$$\Sigma_L = \begin{bmatrix} \Sigma_l & 0_{l,n-l} \\ 0_{m-l,l} & 0_{m-l,n-l} \end{bmatrix},$$

where $\Sigma_l = (\sigma_j)_{j=1,\dots,l}$. Back to $\mathcal{F}_\Sigma$, we have

$$\mathcal{F}_\Sigma = \begin{bmatrix} \text{Id}_l & 0_{l,n-l} & -\gamma_R \Sigma_l & 0_{l,m-l} \\ 0_{n-l,l} & \text{Id}_{n-l} & 0_{n-l,l} & 0_{n-l,m-l} \\ \gamma_J \Sigma_l & 0_{l,n-l} & \text{Id}_l - (1+\theta)\gamma_R\gamma_J \Sigma_l^2 & 0_{l,m-l} \\ 0_{m-l,l} & 0_{m-l,n-l} & 0_{m-l,l} & \text{Id}_{m-l} \end{bmatrix}.$$

Let's study the eigenvalues of $\mathcal{F}_\Sigma$,

$$\begin{aligned} |\mathcal{F}_\Sigma - \rho \text{Id}_{m+n}| &= \left| \begin{bmatrix} (1-\rho)\text{Id}_l & 0_{l,n-l} & -\gamma_R \Sigma_l & 0_{l,m-l} \\ 0_{n-l,l} & (1-\rho)\text{Id}_{n-l} & 0_{n-l,l} & 0_{n-l,m-l} \\ \gamma_J \Sigma_l & 0_{l,n-l} & (1-\rho)\text{Id}_l - (1+\theta)\gamma_R\gamma_J \Sigma_l^2 & 0_{l,m-l} \\ 0_{m-l,l} & 0_{m-l,n-l} & 0_{m-l,l} & (1-\rho)\text{Id}_{m-l} \end{bmatrix} \right| \\ &= (1-\rho)^{m+n-2l} \left| \begin{bmatrix} (1-\rho)\text{Id}_l & -\gamma_R \Sigma_l \\ \gamma_J \Sigma_l & (1-\rho)\text{Id}_l - (1+\theta)\gamma_R\gamma_J \Sigma_l^2 \end{bmatrix} \right|. \end{aligned}$$

Since $(-\gamma_R \Sigma_l)((1-\rho)\text{Id}_l) = ((1-\rho)\text{Id}_l)(-\gamma_R \Sigma_l)$, then by [155, Theorem 3], we have

$$\begin{aligned} |\mathcal{F}_\Sigma - \rho \text{Id}_{a+b}| &= (1-\rho)^{m+n-2l} \left| \begin{bmatrix} (1-\rho)\text{Id}_l & -\gamma_R \Sigma_l \\ \gamma_J \Sigma_l & (1-\rho)\text{Id}_l - (1+\theta)\gamma_R\gamma_J \Sigma_l^2 \end{bmatrix} \right| \\ &= (1-\rho)^{m+n-2l} \left| \left[(1-\rho)\big((1-\rho)\text{Id}_l - (1+\theta)\gamma_R\gamma_J \Sigma_l^2\big) + \gamma_R\gamma_J \Sigma_l \Sigma_l \right] \right| \\ &= (1-\rho)^{m+n-2l} \left| \left[(1-\rho)^2 \text{Id}_l - (1-\rho)(1+\theta)\gamma_R\gamma_J \Sigma_l^2 + \gamma_R\gamma_J \Sigma_l \Sigma_l \right] \right| \\ &= (1-\rho)^{m+n-2l} \prod_{j=1}^{l} \big(\rho^2 - (2-(1+\theta)\gamma_J\gamma_R\sigma_j^2)\rho + (1-\theta\gamma_J\gamma_R\sigma_j^2)\big). \end{aligned}$$

For the eigenvalues ρ, clearly, except the 1's, we have for $j = 1, ..., l$

$$\rho_j = \frac{\big(2-(1+\theta)\gamma_J\gamma_R\sigma_j^2\big) \pm \sqrt{(1+\theta)^2\gamma_J^2\gamma_R^2\sigma_j^4 - 4\gamma_J\gamma_R\sigma_j^2}}{2}.$$

Since $\gamma_J\gamma_R\sigma_j^2 \leq \gamma_J\gamma_R\|L\|^2 < 1$, then ρ_j are complex and

$$|\rho_j| = \frac{1}{2}\sqrt{\big(2-(1+\theta)\gamma_J\gamma_R\sigma_j^2\big)^2 - \big((1+\theta)^2\gamma_J^2\gamma_R^2\sigma_j^4 - 4\gamma_J\gamma_R\sigma_j^2\big)} = \sqrt{1-\theta\gamma_J\gamma_R\sigma_j^2} < 1.$$

As a result, we also obtain the M_{PD}^{∞}, which reads

$$M_{\text{PD}}^{\infty} = \begin{bmatrix} Y & \\ & X \end{bmatrix} \begin{bmatrix} 0_l & & & \\ & \text{Id}_{n-l} & & \\ & & 0_l & \\ & & & \text{Id}_{m-l} \end{bmatrix} \begin{bmatrix} Y^* & \\ & X^* \end{bmatrix}.$$

Define the matrix $\mathcal{P}_S \overset{\text{def}}{=} \begin{bmatrix} \mathcal{P}_{S_{x^\star}^J} & \\ & \mathcal{P}_{S_{v^\star}^{J^*}} \end{bmatrix}$. For (8.3.10), suppose it is not true. Then there exists a matrix M'_{PD} such that

$$M_{\text{PD}}^{\infty} = \mathcal{P}_S + M'_{\text{PD}} \quad \text{and} \quad \mathcal{P}_S M'_{\text{PD}} = 0. \tag{8.8.12}$$

Since $\dim(T_{x^\star}^J) = \dim(T_{v^\star}^{R^*}) = p$, this means that

$$\text{rank}(\mathcal{P}_S) = (n-p) + (m-p) = m + n - 2p.$$

From (8.3.8) we have also $\text{rank}(M_{\text{PD}}^{\infty}) = m + n - 2p$. Since $\mathcal{P}_S M'_{\text{PD}} = 0$, then (8.8.12) implies that $\text{rank}(M_{\text{PD}}^{\infty}) = \text{rank}(\mathcal{P}_S) + \text{rank}(M'_{\text{PD}})$. As a result, we have $\text{rank}(M'_{\text{PD}}) = 0$, hence $M'_{\text{PD}} = 0$ which is a contradiction.

If $n = m$ and moreover $L = \text{Id}_n$, then $(\sigma_j)_{j=1,\dots,p}$ corresponds to *cosine* the *principal angles* between the tangent spaces $T_{x^\star}^J$ and $T_{v^\star}^{R^*}$ [73]. □

Proof of Corollary 8.6.1. Owing to the result of [62], condition (8.6.7) guarantees the convergence of the algorithm.

(i) the identification result follows naturally from Theorem 8.2.1.

(ii) the result follows Proposition 8.3.2, Corollary 7.4.9 and Theorem 8.3.6. First, for the update of x_k of (8.6.1), we have

$$\begin{aligned} &x_{k+1} - x^\star \\ &= W_{\overline{R}} \overline{H}_F (x_k - x^\star) - \gamma_R W_{\overline{R}} \textstyle\sum_i \overline{L}_i^* (v_{i,k} - v_i^\star) + o(\|x_k - x^\star\| + \gamma_R \textstyle\sum_i \|L_i\| \|v_{i,k} - v_i^\star\|). \end{aligned} \tag{8.8.13}$$

Then the update of $v_{i,k+1}$, for each $i = 1, \dots, m$, similar to (8.8.7), we get

$$\begin{aligned} &v_{i,k+1} - v_i^\star \\ &= \big(W_{\overline{J_i^*}} \overline{H}_{G_i^*} - (1+\theta)\gamma_{J_i} \gamma_R W_{\overline{J_i^*}} \overline{L}_i W_{\overline{R}} \overline{L}_i^*\big)(v_{i,k} - v_i^\star) \\ &\quad + \big(2\gamma_{J_i} W_{\overline{J_i^*}} \overline{L}_i W_{\overline{R}} \overline{H}_F - \gamma_{J_i} W_{\overline{J_i^*}} \overline{L}_i\big)(x_k - x^\star) \\ &\quad + o(\|x_k - x^\star\| + \gamma_R \textstyle\sum_i \|L_i\| \|v_{i,k} - v_i^\star\|) + o(\|v_{i,k} - v_i^\star\| + \gamma_{J_i} \|L_i\| \|x_k - x^\star\|). \end{aligned} \tag{8.8.14}$$

Now consider the small o-terms. For the 2 small o-terms in (8.8.3) and (8.8.6). First, let $a_0, a_1, \dots, a_m$ be $m+1$ constants, then we have

$$\textstyle\sum_{i=0}^m |a_i| = \sqrt{(\sum_{i=0}^m |a_i|)^2} \le \sqrt{(m+1) \sum_{i=0}^m |a_i|^2} = \sqrt{m+1} \|(a_0, \dots, a_m)^T\|.$$

Denote $b = \max\{1, \sum_i \sigma_i \|L_i\|, \gamma_R \|L_1\|, \dots, \gamma_R \|L_m\|\}$, then

$$\begin{aligned} &\textstyle\sum_i \left(\|v_{i,k} - v_i^\star\| + \sigma_i \|L_i\| \|x_k - x^\star\|\right) + \left(\|x_k - x^\star\| + \gamma_R \sum_i \|L_i\| \|v_{i,k} - v_i^\star\|\right) \\ &\le 2b(\|x_k - x^\star\| + \textstyle\sum_i \|v_{i,k} - v_i^\star\|) \le 2b\sqrt{m+1} \|\boldsymbol{z}_k - \boldsymbol{z}^\star\|. \end{aligned}$$

Combining this with (8.8.13) and (8.8.14), and ignoring the constants of the small o-term, we have that the fixed-point iteration (8.6.3) is equivalent to

$$\boldsymbol{z}_{k+1} - \boldsymbol{z}^\star = M_{\text{PD}} (\boldsymbol{z}_k - \boldsymbol{z}^\star) + o(\|\boldsymbol{z}_k - \boldsymbol{z}^\star\|).$$

The rest of the proof follows the proof of Theorem 8.3.6. □

Chapter 9

Conclusion and Perspectives

This manuscript provides new results on operator splitting methods along three main standpoints: global convergence rates (Chapter 3); new provably convergent inertial operator splitting schemes (Chapter 4); and local convergence behaviour of proximal splitting algorithms when solving structure non-smooth optimization problems (Chapters 6-8).

Our results are not just unifying with an unprecedented level of generality, but also allow to tackle many problems and algorithms that are not covered by the current literature. Consequently, they naturally cover existing results as very special cases (*e.g.* ℓ_1-norm and FB algorithm), and provide a theoretical and insightful justification to long-standing phenomena that have been observed in practice for other problems and algorithms, but were until now open problems. This is in particular true for the local convergence behaviour that we investigated through the key and insightful notion of partial smoothness.

All our theoretical results were supported by comprehensive numerical experiments.

Take-away messages: several conclusions and take-away messages can be drawn from this work:

(i) our results reveal that under an appropriate non-degeneracy condition, partial smoothness allows to transform any globally non-linear proximal splitting method operating on an Euclidean space into a locally (almost) linearized iteration on a Riemannian manifold. This necessitated key tools from both non-smooth analysis, convex analysis and Riemannian geometry.

(ii) Our global and local analysis explained the commonly observed phase transition phenomenon between sub-linear to linear convergence regimes, which was one of the main motivations of this work. Moreover, the local analysis (based on partial smoothness) revealed the role of the structure of the manifold in the rate, and allowed to optimize the parameters of the algorithm to locally accelerate its local linear rate. Even more, identification of the active manifold opened the door to higher-order acceleration (*e.g.* Newton or conjugate gradient) in a Riemannian setting.

(iii) In certain circumstances, partial smoothness allows to have finite termination as we showed for DR in the polyhedral case.

(iv) Our analysis revealed that the supposedly globally faster (in the objective function) FISTA-type method, is eventually slower than FB. This precisely justifies the use of restarting in practice by many authors even beyond the strongly convex case.

(v) Under partial smoothness, and once global convergence is ensured for a proximal splitting algorithms, convexity is irrelevant for the local behaviour of the algorithm. The reason underlying this key conclusion is message (i) above.

Our research program will not stop here, and many open questions are yet to be answered.

Upper bound of finite identification For the finite activity identification, knowing when identification happens is of paramount practical interest. Though in Proposition 6.2.3 and 7.3.3, we provided

two insightful bounds, they still depend on $x^\star$ which is unknown in practice. We also provide some practical guidelines in Section 6.2.2, though they are not fully theoretically justified. Therefore, designing theoretically grounded and implementable identification bounds which are independent of $x^\star$ is a direction of further research.

Beyond non-degeneracy In our analysis, the non-degeneracy condition (*e.g.* ($\mathrm{ND}_{\mathrm{FB}}$), ($\mathrm{ND}_{\mathrm{DR}}$) and ($\mathrm{ND}_{\mathrm{PD}}$)) is a sufficient condition for finite activity identification of proximal splitting methods. However, in practice local linear convergence still can be observed when this condition fails, which means that the non-degeneracy can be relaxed.

Beyond restricted injectivity In Chapter 6, a restricted injectivity assumption (RI) is imposed for the local linear convergence rate analysis of FB-type methods. This allowed to get precise and explicit rate estimates. In Theorem 6.3.8 we have show that this condition can be removed if the non-smooth component R is locally polyhedral around the minimizer. In Corollary 8.5.1, we also demonstrated that this condition can be removed for general partly smooth functions but the linear convergence is not anymore on $\|x_k - x^\star\|$ (see (8.5.2)). Moreover, things will become more complicated when inertia is involved, since we need to prove that the matrix M_{FB} is convergent.

Beyond optimization In Chapter 5-8 we considered only the case of functions and optimization. In [77], the authors generalized the notion of partial smoothness to set-valued operators. Therefore, it would be interesting to generalize our results to monotone inclusion problems. The key again is finite identification.

Finite dimension Another restriction of partial smoothness is that it is defined on finite dimensional Euclidean space. Extensions to the infinite dimensional setting would be important.

Beyond proximal splitting algorithms It would be also very interesting to extend our results to analyse the local behaviour of splitting methods based on Bregman distances for solving composite convex optimization and monotone inclusions problems; see *e.g.* [135] and references therein. This would for instance allow to solve problems ($\mathcal{P}_{\mathrm{opt}}$) without a Lipschitzian assumption on the smooth part. Extension to other algorithms that are not in the proximal splitting family is another direction of research.

Partial smoothness versus other assumptions In this manuscript, we related the rates obtained through partial smoothness (Chapter 6-7) to those with metric sub-regularity (Chapter 3). Recent results by [31] have established local linear convergence results for the LASSO problem based on error bounds and the Kurdyka-Łojasiewicz inequality (with exponent 1/2). It would be interesting to investigate the connections between the rates obtained with these two seemingly different concepts.

List of Publications

In preparation

(1) J. Liang, J. Fadili and G. Peyré, *Local Linear Convergence of the Primal-Dual splitting methods*.

(2) J. Liang, J. Fadili and G. Peyré, *Partly smooth monotone operators and a modified Forward–Backward–Forward splitting method*.

Preprints

(3) J. Liang, J. Fadili and G. Peyré, *Local Convergence Properties of Douglas–Rachford and ADMM*, submitted to Journal of Optimization Theory and Applications, *arXiv:1606.02118*.

(4) J. Liang, J. Fadili and G. Peyré, *Activity Identification and Local Linear Convergence of Forward–Backward-type methods*, submitted to SIAM Journal on Optimization, *arXiv:1503.03703*.

Journal Papers

(5) J. Liang, J. Fadili and G. Peyré, *Convergence Rates with Inexact Non-expansive Operators*, Mathematical Programming, 159(1):403–434, September 2016.

(6) J. Liang, X. Zhang, *Retinex by Higher Order Total Variation L^1 Decomposition*, Journal of Mathematical Imaging and Vision, 52(3):345–355, 2015.

(7) J. Liang, J. Ma and X. Zhang, *Seismic Data Restoration via Data-Driven Framelet*, Geophysics, 79(3):65–74, March 2014.

(8) J. Liang, J. Li, Z. Shen and X. Zhang, *Wavelet Frame based Color Image Demosaicing*, Inverse problems and Imaging, 7(3):777–794, August 2013.

Conference Proceedings

(9) J. Liang, J. Fadili and G. Peyré, *A Multi-step Inertial Forward–Backward Splitting Method for Non-convex Optimization*, Advances in Neural Information Processing Systems (**NIPS**), 2016. *arXiv:1606.02116*.

(10) J. Liang, J. Fadili, G. Peyré, *Identification en Temps Fini et Convergence Linéaire Locale de l'Algorithme Proximal Implicite-Explicite Inertiel*, Colloque sur le Traitement du Signal et des Images (**GRETSI**), 2015.

(11) J. Liang, J. Fadili, G. Peyré and R. Luke, *Activity Identification and Local Linear Convergence of Douglas–Rachford/ADMM under Partial Smoothness*, 5th Int. Conf. on Scale Space and Variational Methods in Computer Vision (**SSVM**), 2015.

(12) J. Liang, J. Fadili and G. Peyré, *Locally Linear Convergence of Forward–Backward under Partial Smoothness*, Advances in Neural Information Processing Systems (**NIPS**), 2014.

(13) J. Liang, J. Fadili and G. Peyré, *On the Convergence Rates of Proximal Splitting Algorithms*, IEEE Int. Conf. on Image Processing (**ICIP**), 2014. (**Top 10% Papers**).

(14) J. Liang, J. Fadili and G. Peyré, *Iteration-Complexity of a Generalized Forward–Backward Splitting Algorithm*, IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (**ICASSP**), 2014.

List of Notations

General definitions

$\mathbb{R}$: the set of real numbers
$\mathbb{R}_+$: positive real numbers
$\overline{\mathbb{R}}$: $]-\infty, +\infty[\cup\{+\infty\}$, the extended real value
ℓ^1_+: non-negative summable sequence
$\mathbb{N}$: set of non-negative integers
$\mathbb{N}_+$: set of positive integers
$\mathbb{R}^n, \mathbb{R}^m$: finite dimensional real Euclidean spaces
$\mathcal{H}, \mathcal{G}$: real Hilbert spaces
$\Gamma_0(\mathcal{H}), \Gamma_0(\mathbb{R}^n)$: the set of proper convex and lower semi-continuous functions on $\mathcal{H}$ and $\mathbb{R}^n$ respectively
Id: identity operator on $\mathcal{H}$ or $\mathbb{R}^n$
$\mathrm{Id}_n, \mathrm{Id}_m$: identity operators on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively
$\mathcal{K}, L, L_i$: linear operators
$\mathbb{B}_r(x^\star)$: a ball centered at $x^\star$ with radius $r > 0$

Sets related

$\mathcal{S}$: a non-empty (closed) subset of $\mathcal{H}$ or $\mathbb{R}^n$
$\iota_{\mathcal{S}}(\cdot)$: indicator function
$\mathrm{dist}(\cdot, \mathcal{S})$: distance function to a convex set $\mathcal{S}$
$\sigma_{\mathcal{S}}(\cdot)$: support function
$\mathcal{N}_{\mathcal{S}}(\cdot)$: normal cone operator
$\mathcal{P}_{\mathcal{S}}(\cdot)$: projection operator onto $\mathcal{S}$
$\mathrm{ri}(\mathcal{S})$: relative interior of $\mathcal{S}$
$\mathrm{rbd}(\mathcal{S})$: relative boundary of $\mathcal{S}$
$\mathrm{span}(\mathcal{S})$: smallest linear subspace of $\mathbb{R}^n$ that contains $\mathcal{S}$
$\mathrm{aff}(\mathcal{S})$: smallest affine subspace that contains $\mathcal{S}$, a.k.a. affine hull of $\mathcal{S}$
$\mathrm{par}(\mathcal{S}) = \mathbb{R}_+(\mathcal{S} - \mathcal{S})$: the subspace parallel to $\mathrm{aff}(\mathcal{S})$

Functions related

F, G, J, R: functions of $\Gamma_0(\mathcal{H})$ or $\Gamma_0(\mathbb{R}^n)$
$\mathrm{dom}(J)$: domain of J
$\mathrm{epi}(J)$: epigraph of J
J^*: Fenchel conjugate of J
∇F: gradient of F
$\mathrm{prox}_{\gamma J}$: proximity operator of J with $\gamma > 0$
$(J \mathbin{\overset{\vee}{\wedge}} G)(x)$: infimal convolution of function J and G
${}^{\gamma}J(x)$: Moreau envelope of J parameterised by $\gamma > 0$
∂J: sub-differential of function J

T_x: $T_x \stackrel{\text{def}}{=} \operatorname{par}\big(\partial J(x)\big)^{\perp}$

Monotone operators

A, B, C, D: maximal monotone operators on $\mathcal{H}$
A^{-1}: inverse of A
$\operatorname{gra}(A)$: graph of A
$\operatorname{dom}(A)$: domain of A
$\operatorname{ran}(A)$: range of A
$\operatorname{zer}(A)$: zeros of A
$\mathcal{J}_{\gamma A}$: resolvent of A
$\mathcal{R}_{\gamma A}$: reflection of A
${}^{\gamma}A$: Yosida approximation of A
$C \square D$: parallel sum of C and D

Non-expansive operators

$\mathcal{F}$: non-expansive operator
$\mathcal{A}(\alpha)$: the class of α-averaged non-expansive operators
$\mathcal{A}(\frac{1}{2})$: the class of firmly non-expansive operators
$\operatorname{fix}(\mathcal{F})$: the set of fixed points of $\mathcal{F}$

Matrices

M: real square matrix in $\mathbb{R}^{n \times n}$
$\rho(M)$: spectral radius of M
M^{∞}: limit of the power $\lim_{k \to +\infty} M^k$
M_{λ}: relaxed matrix

Manifolds

$\mathcal{M}$: C^2-smooth manifold
$\mathcal{T}_{\mathcal{M}}(x)$: tangent space to $\mathcal{M}$ at a point x in $\mathcal{M}$
$\operatorname{PSF}_x(\mathcal{M})$: The class of partly smooth functions at x relative to $\mathcal{M}$

List of Figures

List of Tables

Bibliography

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[2] P.-A. Absil, R. Mahony, and J. Trumpf. An extrinsic look at the Riemannian Hessian. In *Geometric Science of Information*, pages 361–368. Springer, 2013.

[3] A. Agarwal, S. Negahban, and M. J. Wainwright. Fast global convergence of gradient methods for high-dimensional statistical recovery. *The Annals of Statistics*, 40(5):2452–2482, 2012.

[4] F. Alvarez. On the minimizing property of a second order dissipative system in Hilbert spaces. *SIAM Journal on Control and Optimization*, 38(4):1102–1119, 2000.

[5] F. Alvarez and H. Attouch. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Analysis*, 9(1-2):3–11, 2001.

[6] F. J. Aragòn and M. H. Geoffroy. Characterization of metric regularity of subdifferentials. *Journal of Convex Analysis*, 15(2):365–380, 2008.

[7] K.-J. Arrow, L. Hurwicz, H. Uzawa, H.-B. Chenery, S.-M. Johnson, S. Karlin, and T. Marschak. Studies in linear and non-linear programming. 1959.

[8] T. Aspelmeier, C. Charitha, and D. R. Luke. Local linear convergence of the ADMM/Douglas–Rachford algorithms without strong convexity and application to statistical imaging. *SIAM J. Imaging Science, to appear*, 2016.

[9] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality. *Mathematics of Operations Research*, 35(2):438–457, 2010.

[10] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, Forward–Backward splitting, and regularized Gauss–Seidel methods. *Mathematical Programming*, 137(1-2):91–129, 2013.

[11] H. Attouch, Z. Chbani, J. Peypouquet, and P. Redont. Fast convergence of inertial dynamics and algorithms with asymptotic vanishing damping. *Mathematical Programming*, pages 1–53, 2016.

[12] H. Attouch and J. Peypouquet. The rate of convergence of Nesterov's accelerated Forward–Backward method is actually $o(k^{-2})$. Technical Report arXiv:1510.08740, 2015.

[13] H. Attouch, J. Peypouquet, and P. Redont. On the fast convergence of an inertial gradient-like dynamics with vanishing viscosity. Technical Report arXiv:1507.04782, 2015.

[14] J. B. Baillon. Un théoreme de type ergodique pour les contractions non linéaires dans un espace de Hilbert. *CR Acad. Sci. Paris Sér. AB*, 280:1511–1514, 1975.

[15] J. B. Baillon and G. Haddad. Quelques propriétés des opérateurs angle-bornés etn-cycliquement monotones. *Israel Journal of Mathematics*, 26(2):137–150, 1977.

[16] H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2011.

[17] H. Bauschke, J.Y.B. Cruz, T.A. Nghia, H.M. Phan, and X. Wang. The rate of linear convergence of the Douglas–Rachford algorithm for subspaces is the cosine of the Friedrichs angle. *J. of Approx. Theo.*, 185(63–79), 2014.

[18] H. H. Bauschke, J.Y. Bello Cruz, T.A. Nghia, H. M. Phan, and X. Wang. Optimal rates of convergence of matrices with applications. arxiv:1407.0671, 2014.

[19] H. H. Bauschke, M. N. Dao, D. Noll, and H. M. Phan. On Slater's condition and finite convergence of the Douglas–Rachford algorithm for solving convex feasibility problems in euclidean spaces. *Journal of Global Optimization*, 65(2):329–349, 2016.

[20] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang. Restricted normal cones and the method of alternating projections: theory. *Set-Valued and Variational Analysis*, 21(3):431–473, 2013.

[21] H. H. Bauschke and W. M. Moursi. On the order of the operators in the Douglas–Rachford algorithm. *Optimization Letters*, 10(3):447–455, 2016.

[22] H. H. Bauschke, D. Noll, and H. M. Phan. Linear and strong convergence of algorithms involving averaged nonexpansive operators. *Journal of Mathematical Analysis and Applications*, 421(1):1–20, 2015.

[23] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *Image Processing, IEEE Transactions on*, 18(11):2419–2434, 2009.

[24] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[25] S. Becker and J. Fadili. A quasi-Newton proximal splitting method. In *Advances in Neural Information Processing Systems*, pages 2618–2626, 2012.

[26] D. Boley. Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs. *SIAM Journal on Optimization*, 23(4):2183–2207, 2013.

[27] J. Bolte, A. Daniilidis, and A. Lewis. A nonsmooth morse–sard theorem for subanalytic functions. *Journal of mathematical analysis and applications*, 321(2):729–740, 2006.

[28] J. Bolte, A. Daniilidis, and A. Lewis. The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. *SIAM Journal on Optimization*, 17(4):1205–1223, 2007.

[29] J. Bolte, A. Daniilidis, A. Lewis, and M. Shiota. Clarke subgradients of stratifiable functions. *SIAM Journal on Optimization*, 18(2):556–572, 2007.

[30] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet. Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. *Transactions of the American Mathematical Society*, 362(6):3319–3363, 2010.

[31] J. Bolte, T. P. Nguyen, J. Peypouquet, and B. Suter. From error bounds to the complexity of first-order descent methods for convex functions. *arXiv preprint arXiv:1510.08234*, 2015.

[32] R. I. Boţ, E. R. Csetnek, and C. Hendrich. Inertial Douglas–Rachford splitting for monotone inclusion problems. *Applied Mathematics and Computation*, 256:472–487, 2015.

[33] R. I. Boţ, E. R. Csetnek, and S. C. László. An inertial Forward–Backward algorithm for the minimization of the sum of two non-convex functions. *EURO Journal on Computational Optimization*, pages 1–23, 2014.

[34] N. Boumal and P.-A. Absil. Rtrmc: A riemannian trust-region method for low-rank matrix completion. In *Advances in neural information processing systems*, pages 406–414, 2011.

[35] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds. *The Journal of Machine Learning Research*, 15(1):1455–1459, 2014.

[36] K. Bredies and D. A. Lorenz. Linear convergence of iterative soft-thresholding. *Journal of Fourier Analysis and Applications*, 14(5-6):813–837, 2008.

[37] K. Bredies and H. Sun. Preconditioned Douglas–Rachford splitting methods for convex-concave saddle-point problems. *SIAM Journal on Numerical Analysis*, 53(1):421–444, 2015.

[38] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Math. Stud. Elsevier, New York, 1973.

[39] H. Brézis and P. L. Lions. Produits infinis de résolvantes. *Israel Journal of Mathematics*, 29(4):329–345, 1978.

[40] L. M. Briceño-Arias. Forward–Douglas–Rachford splitting and forward-partial inverse method for solving monotone inclusions. *Optimization*, 64(5):1239–1261, 2015.

[41] L. M. Briceño-Arias and P. L. Combettes. A monotone+ skew splitting model for composite monotone inclusions in duality. *SIAM Journal on Optimization*, 21(4):1230–1250, 2011.

[42] R. S. Burachik, S. Scheimberg, and B. F. Svaiter. Robustness of the hybrid extragradient proximal-point algorithm. *Journal of Optimization Theory and Applications*, 111(1):117–136, 2001.

[43] Regina S. Burachik and Alfredo N. Iusem. *Set-valued Mappings and Enlargements of Monotone Operators*. Optimization and Its Applications. Springer, 2008.

[44] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11, 2011.

[45] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2009.

[46] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.

[47] E. J. Candès and T. Tao. The power of convex relaxation: near-optimal matrix completion. *IEEE Trans. Inform. Theory*, 56(5):2053–2080, 2010.

[48] A. Chambolle and J. Darbon. On total variation minimization and surface evolution using parametric maximum flows. *International journal of computer vision*, 84(3):288–307, 2009.

[49] A. Chambolle and J. Darbon. A parametric maximum flow approach for discrete total variation regularization. In *Image Processing and Analysis with Graphs*. CRC Press, 2012.

[50] A. Chambolle and C. Dossal. On the convergence of the iterates of the "fast iterative shrinkage/thresholding algorithm". *Journal of Optimization Theory and Applications*, 166(3):968–982, 2015.

[51] A. Chambolle and T. Pock. A first-order Primal–Dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[52] I. Chavel. *Riemannian geometry: a modern introduction*, volume 98. Cambridge University Press, 2006.

[53] G. Chen and M. Teboulle. A proximal-based decomposition method for convex minimization problems. *Mathematical Programming*, 64(1-3):81–101, 1994.

[54] P. Chen, J. Huang, and X. Zhang. A Primal–Dual fixed-point algorithm for convex separable minimization with applications to image restoration. *Inverse Problems*, 29(2):025011, 2013.

[55] P. L. Combettes. Fejér monotonicity in convex optimization. In A. Christodoulos Floudas and M. Panos Pardalos, editors, *Encyclopedia of Optimization*, pages 1016–1024. Springer, Boston, MA, 2001.

[56] P. L. Combettes. Quasi-Fejérian analysis of some optimization algorithms. *Studies in Computational Mathematics*, 8:115–152, 2001.

[57] P. L. Combettes. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, 53(5-6):475–504, 2004.

[58] P. L. Combettes, L. Condat, J.-C. Pesquet, and B. C. Vũ. A Forward–Backward view of some Primal–Dual optimization methods in image recovery. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 4141–4145. IEEE, 2014.

[59] P. L. Combettes and J.-C. Pesquet. A proximal decomposition method for solving convex variational inverse problems. *Inverse Problems*, 24(6):065014, 2008.

[60] P. L. Combettes and J. C. Pesquet. Primal–Dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators. *Set-Valued and variational analysis*, 20(2):307–330, 2012.

[61] P. L. Combettes and B. C. Vũ. Variable metric quasi-Fejér monotonicity. *Nonlinear Analysis: Theory, Methods & Applications*, 78:17–31, 2013.

[62] P. L. Combettes and B. C. Vũ. Variable metric Forward–Backward splitting with applications to monotone inclusions in duality. *Optimization*, 63(9):1289–1318, 2014.

[63] P. L. Combettes and V. R. Wajs. Signal recovery by proximal Forward–Backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.

[64] P. L. Combettes and I. Yamada. Compositions and convex combinations of averaged nonexpansive operators. *Journal of Mathematical Analysis and Applications*, 425(1):55–70, 2015.

[65] R. Cominetti, J. A. Soto, and J. Vaisman. On the rate of convergence of Krasnosel'skiĭ–Mann iterations and their connection with sums of Bernoullis. *Israel Journal of Mathematics*, 199(2):757–772, 2014.

[66] L. Condat. A Primal–Dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *Journal of Optimization Theory and Applications*, pages 1–20, 2012.

[67] L. Condat. A direct algorithm for 1d total variation denoising. *IEEE Signal Processing Letters*, 20(11):1054–1057, 2013.

[68] M. Coste. *An introduction to o-minimal geometry*. Istituti editoriali e poligrafici internazionali Pisa, 2000.

[69] A. Daniilidis, D. Drusvyatskiy, and A. S. Lewis. Orthogonal invariance and identifiability. *to appear in SIAM J. Matrix Anal. Appl.*, 2014.

[70] A. Daniilidis, W. Hare, and J. Malick. Geometrical interpretation of the predictor-corrector type algorithms in structured optimization problems. *Optimization: A Journal of Mathematical Programming & Operations Research*, 55(5-6):482–503, 2009.

[71] P. L. Davies and A. Kovac. Local extremes, runs, strings and multiresolution. *Ann. Statist.*, 29:1–65, 2001.

[72] D. Davis and W. Yin. Convergence rate analysis of several splitting schemes. Technical report, arXiv:1406.4834, August 2014.

[73] L. Demanet and X. Zhang. Eventual linear convergence of the Douglas–Rachford iteration for basis pursuit. *Mathematics of Computation*, 85(297):209–238, 2016.

[74] A. L. Dontchev, M. Quincampoix, and N. Zlateva. Aubin criterion for metric regularity. *Journal of Convex Analysis*, 13:281–297, 2006.

[75] A. L. Dontchev and R. T. Rockafellar. *Implicit functions and solution mappings: A view from variational analysis*. Springer, 2009.

[76] J. Douglas and Jr. H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American mathematical Society*, 82(2):421–439, 1956.

[77] D. Drusvyatskiy and A. S. Lewis. Optimality, identifiability, and sensitivity. *Mathematical Programming*, 147(1-2):467–498, 2014.

[78] J. Eckstein and D. P. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1-3):293–318, 1992.

[79] E. Esser, X. Zhang, and T. F. Chan. A general framework for a class of first-order Primal–Dual algorithms for convex optimization in imaging science. *SIAM Journal on Imaging Sciences*, 3(4):1015–1046, 2010.

[80] M. Fortin and R. Glowinski. *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems*. Access Online via Elsevier, 2000.

[81] P. Frankel, G. Garrigos, and J. Peypouquet. Splitting methods with variable metric for Kurdyka–Łojasiewicz functions and general convergence rates. *Journal of Optimization Theory and Applications*, 165(3):874–900, 2015.

[82] D. Gabay. Applications of the method of multipliers to variational inequalities. In M. Fortin and R. Glowinski, editors, *Augmented Lagrangian Methods: Applications to the Solution of Boundary-Value Problems*, pages 299–331. Elsevier, North-Holland, Amsterdam, 1983.

[83] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1):17–40, 1976.

[84] P. Giselsson and S. Boyd. Metric selection in Douglas–Rachford splitting and ADMM. *arXiv preprint arXiv:1410.8479*, 2014.

[85] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, volume 9. SIAM, 1989.

[86] M. Gu, L.-H. Lim, and C. J. Wu. ParNes: a rapidly convergent algorithm for accurate recovery of sparse and approximately sparse signals. *Numerical Algorithms*, 64(2):321–347, 2012.

[87] E. Hale, W. Yin, and Y. Zhang. Fixed-point continuation for ℓ_1-minimization: methodology and convergence. *SIAM Journal on Optimization*, 19(3):1107–1130, 2008.

[88] W. L. Hare. Identifying active manifolds in regularization problems. In H. H. Bauschke, R. S., Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, volume 49 of *Springer Optimization and Its Applications*, chapter 13. Springer, 2011.

[89] W. L. Hare and A. S. Lewis. Identifying active constraints via partial smoothness and prox-regularity. *Journal of Convex Analysis*, 11(2):251–266, 2004.

[90] W. L. Hare and A. S. Lewis. Identifying active manifolds. *Algorithmic Operations Research*, 2(2):75–82, 2007.

[91] B. He and X. Yuan. On convergence rate of the Douglas–Rachford operator splitting method. *Mathematical Programming, under revision*, 2011.

[92] B. He and X. Yuan. Convergence analysis of Primal–Dual algorithms for a saddle-point problem: from contraction perspective. *SIAM Journal on Imaging Sciences*, 5(1):119–149, 2012.

[93] B. He and X. Yuan. On non-ergodic convergence rate of Douglas–Rachford alternating direction method of multipliers. Technical report, Technical report, Tech. rep, 2012.

[94] R. Hesse and D. R. Luke. Nonmonotone fixed point mappings and convergence of first-order algorithms. *Preprint*, 2012.

[95] R. Hesse and D. R. Luke. Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. *SIAM Journal on Optimization*, 23(4):2397–2419, 2013.

[96] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis And Minimization Algorithms*, volume I and II. Springer, 2001.

[97] D. S. Hochbaum. An efficient algorithm for image segmentation, Markov random fields and related problems. *Journal of the ACM (JACM)*, 48(4):686–701, 2001.

[98] K. Hou, Z. Zhou, A. M.-C. So, and Z.Q. Luo. On the linear convergence of the proximal gradient method for trace norm regularization. In *Advances in Neural Information Processing Systems*, pages 710–718, 2013.

[99] P. R. Johnstone and P. Moulin. A Lyapunov analysis of FISTA with local linear convergence for sparse optimization. *arXiv preprint arXiv:1502.02281*, 2015.

[100] N.T.B. Kim and D.T. Luc. Normal cones to a polyhedral convex set and generating efficient faces in multiobjective linear programming. *Acta Math. Vietnam.*, 25:101–124, 2000.

[101] M. A. Krasnosel'skiĭ. Two remarks on the method of successive approximations. *Uspekhi Matematicheskikh Nauk*, 10(1):123–127, 1955.

[102] H. Y. Le. Generalized subdifferentials of the rank function. *Optimization Letters*, 7(4):731–743, 2013.

[103] J. M. Lee. *Smooth manifolds*. Springer, 2003.

[104] B. Lemaire. Stability of the iteration method for non expansive mappings. *Serdica Mathematical Journal*, 22(3):331p–340p, 1996.

[105] B. Lemaire. Which fixed point does the iteration method select? In *Recent Advances in Optimization*, pages 154–167. Springer, 1997.

[106] C. Lemaréchal, F. Oustry, and C. Sagastizábal. The U-Lagrangian of a convex function. *Trans. Amer. Math. Soc.*, 352(2):711–729, 2000.

[107] D. Leventhal. Metric subregularity and the proximal point method. *Journal of Mathematical Analysis and Applications*, 360(2):681–688, 2009.

[108] A. S. Lewis. Active sets, non-smoothness, and sensitivity. *SIAM Journal on Optimization*, 13(3):702–725, 2003.

[109] A. S. Lewis, D. R. Luke, and J. Malick. Local linear convergence for alternating and averaged non-convex projections. *Found. Comput. Math.*, 9(4):485–513, 2009.

[110] A. S. Lewis and J. Malick. Alternating projections on manifolds. *Mathematics of Operations Research*, 33(1):216–234, 2008.

[111] A. S. Lewis and H. S. Sendov. Twice differentiable spectral functions. *SIAM Journal on Matrix Analysis and Applications*, 23(2):368–386, 2001.

[112] A. S. Lewis and S. Zhang. Partial smoothness, tilt stability, and generalized Hessians. *SIAM Journal on Optimization*, 23(1):74–94, 2013.

[113] G. Li and B. S. Mordukhovich. Hölder metric subregularity with applications to proximal point method. *SIAM Journal on Optimization*, 22(4):1655–1684, 2012.

[114] J. Liang, J. Fadili, and G. Peyré. Local linear convergence of Forward–Backward under partial smoothness. In *Advances in Neural Information Processing Systems*, pages 1970–1978, 2014.

[115] J. Liang, J. Fadili, and G. Peyré. Activity identification and local linear convergence of Forward–Backward-type methods. arXiv:1503.03703, 2015.

[116] J. Liang, J. Fadili, and G. Peyré. Convergence rates with inexact non-expansive operators. *Mathematical Programming*, 159(1):403–434, September 2016.

[117] J. Liang, J. Fadili, and G. Peyré. A multi-step inertial Forward–Backward splitting method for non-convex optimization. In *Advances in Neural Information Processing Systems*, 2016.

[118] J. Liang, J. Fadili, G. Peyré, and R. Luke. Activity identification and local linear convergence of Douglas–Rachford/ADMM under partial smoothness. In *Scale Space and Variational Methods in Computer Vision*, pages 642–653. Springer, 2015.

[119] P. L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.

[120] D. A. Lorenz and T. Pock. An accelerated Forward–Backward algorithm for monotone inclusions. *arXiv preprint arXiv:1403.3522*, 2014.

[121] F. J. Luque. Asymptotic convergence analysis of the proximal point algorithm. *SIAM Journal on Control and Optimization*, 22:277–293, 1984.

[122] S. Mallat. *A wavelet tour of signal processing*. Academic press, 1999.

[123] W. R. Mann. Mean value methods in iteration. *Proceedings of the American Mathematical Society*, 4(3):506–510, 1953.

[124] B. Martinet. Brève communication. régularisation d'inéquations variationnelles par approximations successives. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 4(R3):154–158, 1970.

[125] C. D. Meyer. *Matrix analysis and applied linear algebra*, volume 2. SIAM, 2000.

[126] S. A. Miller and J. Malick. Newton methods for non-smooth convex minimization: connections among-lagrangian, Riemannian Newton and SQP methods. *Mathematical programming*, 104(2-3):609–633, 2005.

[127] B. Mishra, K. A. Apuroop, and R. Sepulchre. A Riemannian geometry for low-rank matrix completion. *arXiv preprint arXiv:1211.1550*, 2012.

[128] R. DC Monteiro and B. F. Svaiter. On the complexity of the hybrid proximal extragradient method for the iterates and the ergodic mean. *SIAM Journal on Optimization*, 20(6):2755–2787, 2010.

[129] B.S. Mordukhovich. Sensitivity analysis in non-smooth optimization. *Theoretical Aspects of Industrial Design (D. A. Field and V. Komkov, eds.), SIAM Volumes in Applied Mathematics*, 58:32–46, 1992.

[130] J. J. Moreau. Proximité et Dualité dans un espace Hilbertien. *Bulletin de la Société mathématique de France*, 93:273–299, 1965.

[131] A. Moudafi and M. Oliny. Convergence of a splitting inertial proximal method for monotone operators. *Journal of Computational and Applied Mathematics*, 155(2):447–454, 2003.

[132] Y. Nesterov. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Dokl. Akad. Nauk SSSR*, 269(3):543–547, 1983.

[133] Y. Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer, 2004.

[134] Y. Nesterov. Gradient methods for minimizing composite objective function. 2007.

[135] V. Q. Nguyen. *Méthodes d'éclatement basées sur les distances de Bregman pour les inclusions monotones composites et l'optimisation*. PhD thesis, Paris 6, 2015.

[136] P. Ochs, Y. Chen, T. Brox, and T. Pock. iPiano: Inertial proximal algorithm for nonconvex optimization. *SIAM Journal on Imaging Sciences*, 7(2):1388–1419, 2014.

[137] B. O'Donoghue and E. J. Candés. Adaptive restart for accelerated gradient schemes. *Foundations of Computational Mathematics*, pages 1–18, 2012.

[138] N. Ogura and I. Yamada. Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping. *Numerical Functional Analysis and Optimization*, 22, 2002.

[139] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bulletin of the American Mathematical Society*, 73(4):591–597, 1967.

[140] G. B. Passty. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *Journal of Mathematical Analysis and Applications*, 72(2):383–390, 1979.

[141] D. W. Peaceman and H. H. Rachford, Jr. The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, 3(1):28–41, 1955.

[142] S. S. Petrova and A. D. Solov'ev. The origin of the method of steepest descent. *Historia Mathematica*, 24(4):361–375, 1997.

[143] R. Poliquin and R. Rockafellar. Prox-regular functions in variational analysis. *Transactions of the American Mathematical Society*, 348(5):1805–1838, 1996.

[144] R. A. Poliquin and R. T. Rockafellar. Tilt stability of a local minimum. *SIAM Journal on Optimization*, 8(2):287–299, 1998.

[145] R. A. Poliquin, R. T. Rockafellar, and L. Thibault. Local differentiability of distance functions. *Trans. Amer. Math. Soc.*, 352:5231–5249, 2000.

[146] B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964.

[147] B. T. Polyak. *Introduction to optimization*. Optimization Software, 1987.

[148] H. Raguet, J. Fadili, and G. Peyré. A generalized Forward–Backward splitting. *SIAM Journal on Imaging Sciences*, 6(3):1199–1226, 2013.

[149] H. Raguet and L. Landrieu. Preconditioning of a generalized forward-backward splitting and application to optimization on graphs. *SIAM Journal on Imaging Sciences*, 8(4):2706–2739, 2015.

[150] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

[151] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.

[152] R. T. Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1997.

[153] R. T. Rockafellar and R. Wets. *Variational analysis*, volume 317. Springer Verlag, 1998.

[154] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[155] J. R. Silvester. Determinants of block matrices. *The Mathematical Gazette*, pages 460–467, 2000.

[156] S. T. Smith. Optimization techniques on Riemannian manifolds. *Fields institute communications*, 3(3):113–135, 1994.

[157] M. V. Solodov and B. F. Svaiter. A hybrid approximate extragradient–proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Analysis*, 7(4):323–345, 1999.

[158] J. E. Spingarn. Partial inverse of a monotone operator. *Applied Mathematics & Optimization*, 10(1):247–265, 1983.

[159] J.-L. Starck, F. Murtagh, and J. Fadili. *Sparse image and signal processing: wavelets, curvelets, morphological diversity*. Cambridge university press, 2010.

[160] W. Su, S. Boyd, and R. Candes. A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights. In *Advances in Neural Information Processing Systems*, pages 2510–2518, 2014.

[161] T. Sun, R. Barrio, H. Jiang, and L. Cheng. Local linear convergence of a primal-dual algorithm for the augmented convex models. *Journal of Scientific Computing*, pages 1–15, 2016.

[162] B. F. Svaiter. A class of Fejer convergent algorithms, approximate resolvents and the hybrid proximal-extragradient method. *arXiv preprint arXiv:1204.1353*, 2012.

[163] B. F. Svaiter. A class of fejér convergent algorithms, approximate resolvents and the hybrid proximal-extragradient method. *Journal of Optimization Theory and Applications*, 162(1):133–153, 2014.

[164] B. F. Svaiter and R. S. Burachik. ε-enlargements of maximal monotone operators in Banach spaces. *Set-Valued Anal.*, 7:117–132, 1999.

[165] S. Tao, D. Boley, and S. Zhang. Local linear convergence of ISTA and FISTA on the LASSO problem. *SIAM Journal on Optimization*, 26(1):313–336, 2016.

[166] P. Tseng. Alternating projection-proximal methods for convex programming and variational inequalities. *SIAM Journal on Optimization*, 7(4):951–965, 1997.

[167] P. Tseng and S. Yun. A coordinate gradient descent method for non-smooth separable minimization. *Math. Prog. (Ser. B)*, 117, 2009.

[168] S. Vaiter, C. Deledalle, J. Fadili, G. Peyré, and C. Dossal. The degrees of freedom of partly smooth regularizers. *Annals of the Institute of Statistical Mathematics*, pages 1–42, 2015.

[169] S. Vaiter, M. Golbabaee, J. Fadili, and G. Peyré. Model selection with low complexity priors. *Information and Inference*, page iav005, 2015.

[170] S. Vaiter, G. Peyré, and J. Fadili. Model consistency of partly smooth regularizers. *Preprint*, 2014.

[171] L. Van den Dries. *Tame topology and o-minimal structures*, volume 248. Cambridge university press, 1998.

[172] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013.

[173] B. C. Vũ. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Advances in Computational Mathematics*, pages 1–15, 2011.

[174] S. J. Wright. Identifiable surfaces in constrained optimization. *SIAM Journal on Control and Optimization*, 31(4):1063–1079, 1993.