



**HAL**  
open science

# Étude de méthodes inertielles en optimisation et leur comportement sous conditions de géométrie

Hippolyte Labarrière

► **To cite this version:**

Hippolyte Labarrière. Étude de méthodes inertielles en optimisation et leur comportement sous conditions de géométrie. Mathématiques générales [math.GM]. INSA de Toulouse, 2023. Français. NNT : 2023ISAT0017 . tel-04268678

**HAL Id: tel-04268678**

**<https://theses.hal.science/tel-04268678>**

Submitted on 2 Nov 2023

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

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



# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Institut National des Sciences Appliquées de Toulouse (INSA de Toulouse)*

---

---

Présentée et soutenue le *20/09/2023* par :

**Hippolyte LABARRIÈRE**

Étude de méthodes inertielles en optimisation et leur comportement sous conditions de géométrie

---

---

### JURY

JEAN-FRANÇOIS AUJOL

Université de Bordeaux

Directeur de thèse

ALEXANDRE D'ASPREMONT

CNRS Délégation Paris-Centre

Président du jury

CHARLES DOSSAL

INSA Toulouse

Directeur de thèse

JÉRÔME MALICK

CNRS Grenoble

Rapporteur

JUAN PEYPOUQUET

University of Groningen

Rapporteur

NELLY PUSTELNIK

CNRS Délégation Rhône Auvergne

Examinatrice

AUDE RONDEPIERRE

INSA Toulouse

Directrice de thèse

---

### École doctorale et spécialité :

*MITT : Domaine Mathématiques : Mathématiques appliquées*

### Unité de Recherche :

*IMT- Institut de mathématiques de Toulouse*

### Directeur(s) de Thèse :

*Jean-François AUJOL, Charles DOSSAL et Aude RONDEPIERRE*

### Rapporteurs :

*Jérôme MALICK et Juan PEYPOUQUET*



*L'appétit vient en mangeant, la  
réforme c'est pas sorcier, le cam-  
pement léger en plein air, non !*

Cupidon



---

## Remerciements

Ces trois années de thèse sont passées à une vitesse folle et c'est en grande partie grâce aux personnes que j'ai pu côtoyer. Tant dans le cadre du travail qu'en dehors, j'ai eu la chance d'être extrêmement bien entouré.

Avant de remercier celles et ceux qui ont fait ces trois ans, j'aimerais témoigner de toute ma reconnaissance à Jérôme et Juan qui ont accepté de rapporter cette thèse **pendant les vacances d'été**. D'ordinaire pourtant concis, je vous ai donné de quoi vous occuper longuement sur la plage. Merci beaucoup pour vos retours pertinents et votre bienveillance. J'aimerais également remercier chaleureusement Alexandre et Nelly pour leur participation à mon jury de thèse.

Comment commencer ces remerciements autrement que par Aude, Charles et Jean-François ? Initialement, ma décision de commencer cette thèse avait été guidée, d'une part, par le goût de la recherche mais surtout par le pressentiment que mes deux enseignants de l'époque étaient sympathiques (et accessoirement compétents). Force est de constater que je ne m'étais pas trompé et que Jean-François avait également ces caractéristiques. C'était un régal de travailler, discuter et simplement partager du temps avec vous. Entre sessions au tableau intensives et papotages, le temps est passé très vite et j'ai du mal à réaliser que trois ans et demi se sont écoulés. Je passe maintenant le relais à Julien qui, j'en suis sûr, fera une super thèse avec cette équipe de choc !

Un grand merci à Luca avec qui j'ai eu la chance de travailler. C'est toujours un bonheur d'interagir avec des personnes aussi positives et j'ai hâte que l'on retravaille ensemble, que ce soit sur des maths ou des danses hongroises de Brahms. Merci beaucoup à Vãn Hào que j'ai eu l'honneur d'encadrer en stage. J'ai été bluffé par ton volume de travail et ta gentillesse. Le seul reproche que je peux faire est que je suis désormais habitué à ce qu'un stagiaire soit totalement autonome.

Je voudrais également remercier Lorenzo et Silvia chez qui la suite de l'histoire devrait s'écrire. Je suis terriblement impatient à l'idée de découvrir un nouveau projet, un nouveau laboratoire et une nouvelle vie.

Merci beaucoup à mes collègues du GMM qui ont contribué à rendre cette thèse si agréable. Merci à Hugo avec qui j'ai partagé un bureau cossu à l'isolation relative (que l'on pourrait qualifier de "chambre froide/sauna tout en un"). C'était un plaisir de passer cette année et demi en ta compagnie et je suis sûr que la fin de ta thèse se déroulera bien. Merci à la super clique de doctorants/post-doctorants/ex-doctorants : Alban, Amandine, Clément, Evgeniia, Iain, Julie, Mahmoud, Mustafa, Thibault, Franck, Ousmane, Nathanaël, Romain, Sarah, Mouhamad... Je me suis bien amusé pendant les pauses avec vous et j'espère que la tradition des quizzes de géo perdurera (Worldle, Countrydle, Flagle mes amours). Je vous souhaite le meilleur pour la suite et une bonne rédaction à celles et ceux qui sont en plein dedans. Merci à celles et ceux qui étaient pour la plupart mes anciens profs : Violaine, Benjamin, Fred De Gournay (que, j'avoue, j'ai longtemps trouvé bizarre), Fred Couderc, Robin, Lorick (malgré ton Pikachu toxique), Loïc (malgré ton Ness toxique et nos différends à propos de la Mer Caspienne), Laure, Béatrice, Romain, Sandrine

S., Cathy, Olivier, Juliette, Géraldine, Mélisande, Florent, David (B. et S.), Pascal, Simona, Philippe, Jérôme, Aldéric, Adeline sans oublier Sandrine B. et Marlène. Commencer à vous tutoyer a été complexe mais j'ai passé de très bons moments avec vous. Merci aussi à Pierre qui a certes quitté le navire INSA mais m'a initié à l'ultimate pour le plus grand malheur de mes poumons. Je salue également les élèves à qui j'ai pu donner cours pendant ces trois ans. Les débuts ont été compliqués (je n'oublierai jamais les corrections en visio sans tablette... jamais) mais ça a été un vrai plaisir d'échanger avec vous.

Un petit mot pour celles et ceux que j'ai pu rencontrer en conférence, la communauté de l'optimisation française a l'air fort sympathique et il me tarde de participer aux prochains événements. Mention spéciale à Vassilis, mon prédécesseur de thèse de qui je suis les traces jusqu'en Italie.

Passons à un pan non négligeable de ces trois années (voire de ces vingt-quatre années) : la famille. Un grand merci à mes parents qui m'ont toujours soutenu malgré une incompréhension totale (et logique) de mon sujet de thèse. Je vous dois beaucoup et je suis ravi de voir que votre retraite creusoise (et la vie de grands-parents maintenant) vous sied à merveille. Je tiens d'ailleurs à vous remercier une nouvelle fois pour m'avoir offert ce vélo glané à la disqueuse, c'est une de mes anecdotes qui ont le plus de succès. Je remercie également très fort Apo et Rapha, divines frangines qui m'ont fait tant grandir et sans qui je serais certainement plus bas de plafond. Merci à Léonor, nouvelle venue qui en quelques mois a réussi à me vexer (si, pleurer à ma seule vue, c'est vexant) mais qui n'en demeure pas moins le bébé le plus mignon du monde (objectivement). Merci à Mehdi, père de cette belle chose mais surtout la seule personne au monde qui connaît le script de chaque épisode de Kaamelott mais qui n'aligne pas les références lourdes (rarissime). Merci à Mamyleine, grand-mère aimante qui m'a fait aimer très tôt Saint-Saens et Sarasate et a également inventé la légendaire confiture de prunes **avec noyaux**. Merci également à Véro, Jean-Marc, Raphaël, Michaela, Irma, Brigitte, Iba, Rachel, Mathieu, Caro, Louis, Estelle, Danielle, Marie, Alex, Lou, Adja, Yacine, François, Hauviette, Ruslan, Vadim, Alexandra, Victor et une pensée pour Papilu, Papiné, Mamine et Jean.

Je suis extrêmement reconnaissant envers mes amis avec qui j'ai passé des moments mémorables. J'ai la chance d'être entouré par des personnes bienveillantes, drôles et très différentes. J'espère n'oublier personne mais j'avoue que ça n'est pas facile !

Merci à Hugo/Ujo/Hugo bleu/Copain de gaming que je n'ai pas quitté depuis le week-end d'intégration de première année il y a huit ans (ça nous rajeunit pas). Le temps a passé mais nos conversations n'ont toujours aucun sens, entre fous rires et empilements abusifs de références de niche. On n'est certainement pas sortables mais peu importe, pour un pote pareil je n'hésiterai pas à offrir un diplôme. Merci à Elo/Fit avec qui tout a commencé par l'élaboration d'une montgolfière, une rencontre somme toute banale. Ton addiction au reggaeton ne m'a pas empêché de passer des super moments avec toi. Avec un tel sens de l'écoute, je flaire que le domaine de la psycho ne serait pas une mauvaise idée mais je dis ça comme ça... Merci à Élé/Miche/LNA, binôme et confidente inséparable fût un temps et toujours

---

présente des années après. C'est toujours un plaisir de se poser discuter de choses sérieuses ou de La Flamme (toujours très sérieux) avec toi et je t'attends de pied ferme à Gênes. Merci à Val/Mimouche, coach assermenté et colocataire à son insu. Je mets de côté mes ressentiments liés à un certain épisode faisant intervenir un cellulaire pour me remémorer la quantité astronomique de matchs matés ensemble, les bonnes blagues de beaufs à moitié assumées, les trois mois où mon père t'a maudit pour avoir pris mon vélo en otage (j'avoue j'étais responsable) ou encore le confort de ton canapé. Merci à Laulau/Jess, l'experte du combiné mots fléchés/rage cage. J'ai découvert ces dernières années une personne aux qualités humaines qui n'ont d'égal que son aptitude au tackle. Tu es une des rares personnes<sup>1</sup> qui m'a fait regarder un Saint-Étienne - Rodez (0 - 2, coup dur). Merci à Mouns/la Mouns/Saramouni, qui a réussi avec brio la transition de +1 à titulaire en puissance. Je retiendrai à jamais que tu es LA personne sur qui compter pour les déménagements ou pour les analyses rugbystiques bien senties (j'attends des saranalyses par message en Italie). Petit bémol pour ta toxicité envers Juanchi Mallia mais je passe l'éponge pour cette fois. Merci à Nico, compère musicien doctorant (et maintenant docteur !). Tu es certainement la personne la plus curieuse que je connaisse, j'espère que ça te portera aux sommets de Météo France !

Un énorme merci à une bande de potes immense que j'ai appris à connaître assez récemment et qui est pourtant déjà tellement importante pour moi : la Flav. Merci à Guigz/la bête des terrains/la blatte/la tarentule/la terrasse chauffée/la jonquille qui est celui par qui j'ai été intégré à ce groupe de haut standing. Notre amitié a commencé par des soirées, s'est poursuivie par des fêtes ou encore par des bringues. Quelques années plus tard j'ai un frère avec qui lâcher mon venin sur Wayne Barnes, faire le zouave à Arkose et tout raffer au Trivial Pursuit. Merci à Axel/Jums, mon rugbyman préféré, recordman mondial des thésards avec le plus d'heures par semaine mais surtout mon jumeau (désolé encore pour le tour de magie dans la tarte au citron meringuée) ; merci à Chloé, thésarde itinérante et maquilleuse en formation (j'oublie pas le massacre de la soirée mexicaine) avec qui le rire est garanti ; merci Sophie pour ta folie et les dingeries sorties avec une nonchalance exceptionnelle, j'espère qu'on fera bientôt des duos endiablés ; merci à Kody, *the crazy biker* un peu trop fou au goût de son omelette, toujours un plaisir de parler bagarre et de t'apprendre des jolis mots français ; merci à Cam/Cam Cam, la seule supportrice de Castres que je tolère mais également un être humain trop cool (merci encore pour Royan c'était fou) ; merci à Fanny, fière représentante du 31 aux multiples talents, mention spéciale pour le chant verlan et l'élaboration de jeux ; merci à Paul, influenceur Strava qui me fait déprimer par ses sorties incessantes (ça vaut pour toi aussi Clara), ta gentillesse et la forme de tes mollets sont uniques ; merci à Clara, ma plus grande rivale en limbo et à mon sens la joueuse de rugby la plus redoutable du circuit international ; merci à Gatou, ma blagnacaise préférée et certainement le meilleur public pour faire des blagues moyennes ; merci à Rémy, comparse doctorant parfois un peu trop généreux (être dans les remerciements c'est pareil que deuxième

---

1. avec Alexchuk.



auteur t'inquiète pas) ; merci à David, le feu-follet du Colorado membre du fan-club de Stephen Becker ; merci à Valou, mon coach padel quasi professionnel (les croisés, tu connais) mais surtout chercheur éminent ; merci à Nathan, alpiniste et explorateur fou dont la reconversion comme fossoyeur n'est pas très loin à mes yeux (la plage de Royan peut en témoigner) ; merci groupé à Auke, Max et Hellen, l'équipe des *Derniers inscrits, premiers arrivés* avec qui j'ai eu la chance de partager 70 km (sans préparation adéquate évidemment) et grâce à qui j'ai pu aller au bout, des personnes incroyables sportivement et humainement ; merci (ou peut-être pas) à Noa qui m'a embarqué dans cette galère, profitant de mon inconscience ; merci Nico, le Tadej français qui, à mon grand regret, ne m'a pas accompagné jusqu'au bout de mes aventures capillaires ; merci à Méli, partenaire de beer-pong avec qui j'ai tout gagné, guide touristique de renom et grimpeuse invétérée (Ailefroide le feu) ; merci à Romi à qui je prédis une carrière en ultimate pleine de succès (ou aux joyeux combinards) ; merci à Orane, ex-cogérante du Chicken Ben's Saint Michel et consœur du club select des bruns ténébreux ; merci à Flo, parti à Paris quand je suis arrivé (belle ambiance) mais dont la bienveillance n'est égalée que par le volume sonore nocturne ; merci à Babtou/Bebou, pilote désormais chez Mercedes dont le futur s'annonce brillant ; merci à Coumat pour ta poésie, ta discrétion et ton style vestimentaire si raffiné (je fais des rêves de ton peignoir).

Je voudrais également remercier une clique d'artistes haute en couleur avec laquelle j'ai notamment fait un superbe voyage culturel et touristique en Hongrie (le 7/7 reste gravé à jamais). Merci à Samenka, passé d'humble sétois à trader parisien si vite, et ce, en poursuivant des carrières de rugbyman et footballeur professionnel (avertissement comportement tout de même, Sam tu sais ce que ça veut dire) ; merci à Guillo<sup>2</sup> dont le sommeil quelquefois lourd ne l'a pas empêché de trouver la fête (donde esta la fiesta ?) pour le plus grand bonheur des petits et des grands ; merci à Alexchuk, spécialiste de high-kick sur voiture, d'escalade sur maison et de sprint sur Burger King à qui je souhaite le plus prompt retour en Ligue 1 ; merci à Théozerski, leader parfois corrosif mais qui devient un agneau autour d'un (voire deux) kebab ; merci à Ricard, aficionado du métro barcelonais et de Renault Modus qui va faire les beaux jours de l'Union Bordeaux-Bègles (tant que c'est pas Castres ça va) ; merci à Jo, la légende du padel que c'est toujours un plaisir de voir malgré son emploi du temps de ministre (ou sénateur).

Je remercie aussi ceux qui sont là depuis presque une éternité et avec qui j'ai partagé la passion de la musique et de l'humour fin. Merci Léo/Lapoule/Bibi/Stanley, le frère avec qui j'ai grandi pendant des années avec tant d'anecdotes légendaires à la clé (l'épisode de l'homme nu restant numéro 1). J'ai hâte de la prochaine session ping-pong de minimum 5h et des prochaines tribulations avec Vivi. Merci à Simon, musicien trop fort qui me permet de briller en société ("oui, j'ai un ami à l'orchestre du Capitole") et qui est en plus drôle et sympa. Merci à Cèpe que j'ai pu voir évoluer dans une trajectoire hors du commun durant ces dernières années (from trompette to ingénieur) : je vous annonce que le prochain président sera

---

2. aucun lien avec Guigz susmentionné.

---

ariégeois. Merci à Polo et Fleur qui ont été comme des sœurs depuis plus de dix ans et à Fred et Virginie qui m'ont quasiment adopté. Merci à Andry, le grand frère passé de Rodez à la Suisse (sacré progression) que j'espère plus voir en Italie.

Merci à Vic, Diane et Sophie, la sororie avec qui j'ai partagé une rue mais surtout des super moments. D'abord enfant quand je vous inventais des jeux relativement médiocres, puis des années plus tard avec les légendaires kempferies qui resteront comme un évènement incontournable chaque année. Je retiendrai entre autre les activités variées qui y sont proposées (notamment le ping-pong) et surtout le casting de qualité. J'en profite pour remercier également Lili, Elsa, Raph, Mei, Eng, Yohan, Clem, Benben, Ferrer, Maïa et tous les autres. Merci aussi à Pierre-Bernard et Martine avec qui j'ai hâte de partager la prochaine randonnée.

Bien que ces trois années aient été marquées par une épidémie, elles ont tout de même été l'occasion de faire la fête et rencontrer des personnes géniales. Merci à Robin et Yohan, rencontrés lors d'un cluster Covid et avec qui le courant est passé en environ trois heures. Merci à Ophé/Fef/Bebouuu, ma presque coloc pour qui j'ai du mal à trouver les mots, LSF oblige, mais avec qui j'ai vécu des tranches de vie exceptionnelles (dont la recherche d'un petit sachet de tabac). Merci à Seph, comédienne au sérieux inébranlable qui a su encourager ma vocation de guide spécialisé du Machu Picchu. Merci à Emma, teufeuse assermentée sans qui je n'aurais certainement jamais rencontré Peato Mauvaka. Merci aussi à Lauriane, Trut, Mathilde, Enza, Léa, reparties vers le grand Nord si vite.

Merci à tous les copains de l'INSA grâce à qui j'ai adoré mes études. Merci à Jojo/Chouan, binôme germanique (der neue Profi) aux multiples casquettes (DJ, bénévole 100 tours, coach TPLM, champion SSF2), merci à Pierre, compère supporter du Stade Toulousain présent même quand on finissait 12ème (signe de loyauté évident), merci à Sophia, handballeuse briseuse de rétine et exemple de persévérance qui va faire une thèse légendaire, merci à Bast, camarade doctorant grimpeur à qui je souhaite une très belle fin de thèse, merci à Seb, l'oiseau de nuit insaisissable et ami proche de Kendji Girac, merci à Youss de ne m'en avoir pas (trop) voulu malgré la luxation d'un pouce et quelques roustes sur Mario Kart, merci à Quentin pour ces sessions gaming impromptues et fort sympathiques. Merci également à Othmane, Em, Biche, Oumaima, Julien, Nour, Nada, Cédric, Villegas, Rémi, Tinmar, Dels, Jumouette, Tonio et j'en passe et des meilleurs.

Un grand merci à Clara, Nicole, Magali, Céline et Jérôme, mes professeurs de violon qui m'ont énormément appris tant au niveau musical qu'au niveau humain. Merci à Claudine qui m'a également beaucoup apporté. Merci à Tommaso, ami italien depuis le tout tout début et que j'espère voir à Gênes. Merci aussi à Alessandro et Marina pour leur gentillesse. Merci à toutes les personnes que j'ai pu cotoyer au Touch rugby, à l'ultimate et à l'escalade et avec qui je me suis régalez.

Enfin, cette thèse ne serait peut-être pas si bien passée si certaines expériences ne m'avaient pas endurci. Merci donc à mes cambrioleurs qui m'ont appris à gérer des situations de crise en me rendant visite alors que j'étais là, merci au conservatoire de Toulouse qui m'a fait réaliser que les études supérieures n'étaient qu'un hobby et qui m'a poussé à me débrouiller seul, merci à Katia qui m'a appris le self-contrôle

et à ne pas me vexer et merci au Domino's pizza de Seattle qui a mis mon corps à rude épreuve (il a gagné).

Enfin, j'aimerais avoir une grosse pensée pour Axelle et Myriam.

# Table des matières

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Motivations</b>  | <b>1</b>  |
| 1.1      | L'optimisation, qu'est-ce que c'est ?                                     | 1         |
| 1.2      | Problématiques en optimisation convexe                                    | 2         |
| 1.3      | Contexte  | 3         |
| 1.4      | Contributions et plan du manuscrit  | 7         |
| 1.5      | Publications associées à ce manuscrit                                     | 8         |
| <b>2</b> | <b>Chapitre préliminaire</b>  | <b>11</b> |
| 2.1      | Géométrie des fonctions convexes  | 12        |
| 2.1.1    | Convexité et caractère Lipschitz  | 12        |
| 2.1.2    | Géométrie des fonctions convexes  | 15        |
| 2.2      | Algorithmes d'optimisation  | 19        |
| 2.2.1    | Algorithmes pour l'optimisation de fonctions différentiables              | 19        |
| 2.2.2    | Algorithmes pour l'optimisation non lisse                                 | 23        |
| 2.3      | Optimisation et systèmes dynamiques                                       | 26        |
| 2.3.1    | Historique  | 26        |
| 2.3.2    | Apports du parallèle continu/discret                                      | 29        |
| 2.3.3    | Analyse de Lyapunov   | 30        |
| <b>3</b> | <b>Stratégies de restart</b>  | <b>35</b> |
| 3.1      | Présentation du restart   | 36        |
| 3.1.1    | Contexte  | 36        |
| 3.1.2    | Techniques de restart   | 38        |
| 3.2      | Restart via une estimation automatique du paramètre de croissance         | 43        |
| 3.2.1    | Présentation de la stratégie  | 43        |
| 3.2.2    | Garanties théoriques du schéma  | 48        |
| 3.2.3    | Expériences numériques  | 50        |
| 3.3      | Un schéma de restart avec backtracking : Free-FISTA                       | 54        |
| 3.3.1    | Contexte  | 54        |
| 3.3.2    | Le backtracking appliqué à FISTA  | 55        |
| 3.3.3    | Free-FISTA  | 60        |
| 3.3.4    | Expériences numériques  | 66        |
| 3.A      | Résultats annexes et preuves  | 75        |
| 3.A.1    | Étude de l'Algorithme 3 avec une mise à jour des itérations plus générale | 75        |
| 3.A.2    | Variations de l'Algorithme 3  | 77        |
| 3.A.3    | Preuves techniques  | 80        |

|          |   |            |
|----------|---|------------|
| <b>4</b> | <b>Maîtrise de l’inertie par un terme lié à la hessienne</b>              | <b>93</b>  |
| 4.1      | Contexte . . . . .  | 94         |
| 4.1.1    | Rappels sur les systèmes dynamiques . . . . .                             | 94         |
| 4.1.2    | Un terme lié à la hessienne pour atténuer les oscillations . . . . .      | 94         |
| 4.1.3    | Méthodes numériques associées . . . . .                                   | 96         |
| 4.2      | Résultats de convergence sous conditions de croissance . . . . .          | 97         |
| 4.2.1    | Fonctions à croissance quadratique . . . . .                              | 97         |
| 4.2.2    | Fonctions plates . . . . .  | 101        |
| 4.3      | Observations numériques . . . . .   | 104        |
| 4.3.1    | Atténuation des oscillations sur un problème de moindres carrés . . . . . | 104        |
| 4.3.2    | Étude de la valeur critique de $\beta$ . . . . .                          | 105        |
| 4.A      | Preuves . . . . .   | 110        |
| <b>5</b> | <b>Schémas inertiels et non unicité des minimiseurs</b>                   | <b>133</b> |
| 5.1      | Problématique . . . . .   | 134        |
| 5.2      | Revue de la littérature . . . . .   | 135        |
| 5.2.1    | Systèmes dynamiques . . . . .   | 135        |
| 5.2.2    | Schémas numériques . . . . .  | 138        |
| 5.3      | Non unicité dans le cas continu . . . . .                                 | 142        |
| 5.3.1    | Stratégie . . . . .   | 142        |
| 5.3.2    | Résultats de convergence sans unicité du minimiseur . . . . .             | 146        |
| 5.4      | Impact de la non unicité sur les schémas numériques . . . . .             | 148        |
| 5.4.1    | Convergence de V-FISTA . . . . .  | 149        |
| 5.4.2    | Convergence de FISTA . . . . .  | 152        |
| 5.4.3    | Preuves . . . . .   | 154        |
| 5.A      | Preuves techniques . . . . .  | 170        |
| <b>6</b> | <b>Conclusion et perspectives</b>   | <b>183</b> |
| 6.1      | Récapitulatif . . . . .   | 183        |
| 6.2      | Perspectives . . . . .  | 184        |

# Motivations

---

## 1.1 L'optimisation, qu'est-ce que c'est ?

Mes trois années de thèse ont été consacrées à des problématiques d'optimisation. Avant de rentrer dans le vif du sujet, un bref résumé de ce qu'est ce domaine des mathématiques s'impose.

En mathématiques, résoudre un problème d'optimisation consiste à poser une question relativement simple : comment obtenir le moins possible (ou le plus possible) d'une quantité ? Pour répondre à cette question, il faut avoir la main sur des paramètres qui font varier la dite quantité. L'objectif final est de déterminer les paramètres optimaux pour minimiser (ou maximiser) cette quantité. Pour écrire le problème sous forme mathématique, on peut dire que l'on résout :

$$\min_{x \in X} F(x), \tag{1.1.1}$$

où  $F$  est la quantité à minimiser et  $X$  est l'ensemble des paramètres acceptables.

Ce type de problème est très courant dans la plupart des domaines scientifiques mais également dans la vie de tous les jours. La mise en place d'un trajet pour aller d'un point à un autre est un bon exemple : si l'objectif est de se rendre le plus rapidement possible à un point, on peut considérer que l'on cherche à minimiser le temps de trajet qui dépend de l'itinéraire ; si le but est d'économiser le plus d'argent possible, la quantité à minimiser est alors l'argent dépensé en péage ou en essence. En fait, il est souvent possible de formuler des questions non triviales comme des problèmes d'optimisation : comment minimiser le temps de course en fonction de son entraînement et de sa stratégie de course pour un coureur, comment maximiser la qualité de son manuscrit en fonction du temps disponible pour un doctorant, comment minimiser le carburant embarqué dans une fusée tout en accomplissant le trajet souhaité...

Le rôle d'un optimiseur n'est généralement pas de formuler la question initiale en termes mathématiques. Cette tâche relève de la modélisation et requiert l'expertise liée au domaine auquel la question appartient. Un optimiseur s'intéresse à des problèmes de la forme (1.1.1) et son but peut être d'élaborer des schémas efficaces permettant de donner une approximation précise de la solution ou bien d'analyser le comportement de méthodes déjà existantes pour ce problème.

Le fait de comprendre le comportement de différents algorithmes d'optimisation en fonction des hypothèses du problème auquel on s'intéresse est crucial. Cela permet de sélectionner une méthode efficace pour résoudre ce problème en ayant des garanties théoriques. Concrètement, l'efficacité d'une méthode se mesure à sa

capacité à obtenir une solution à partir d'un nombre de calculs réduit et donc rapidement. Lorsque l'on considère des problèmes en grande dimension, c'est-à-dire qui font intervenir une grande quantité de données, le gain de temps peut être significatif et s'accompagne d'économies énergétiques.

## 1.2 Problématiques en optimisation convexe

En pratique, il est difficile d'étudier directement le problème (1.1.1) sans aucune hypothèse. C'est pourquoi il existe une grande variété de sous-domaines de l'optimisation, chacun basé sur un jeu d'hypothèses.

Mes travaux de thèse portent sur l'optimisation de fonctions convexes composites dans un cadre déterministe. Ainsi, la fonction  $F$  à minimiser est supposée convexe avec une structure particulière (voir Section 2.1.1) et aucune notion d'aléatoire n'intervient. Ces hypothèses excluent donc les fonctions non convexes (c'est le cas des fonctions de coût des réseaux de neurones par exemple) ou encore les méthodes d'optimisation stochastiques.

Dans ce cadre, je me suis concentré avec mes directeurs de recherche sur l'analyse de convergence d'algorithmes d'optimisation sous des hypothèses de croissance. Définissons les différents concepts que cela implique :

- pour résoudre le problème (1.1.1), on peut appliquer un **algorithme d'optimisation** (on utilisera également les appellations **schéma d'optimisation** ou **méthode d'optimisation**). En règle générale, ces méthodes sont itératives : à partir d'un point initial  $x_0 \in X$ , une suite  $(x_k)_{k \in \mathbb{N}}$  est générée de manière à ce que  $x_k$  s'approche d'une solution de (1.1.1) quand  $k$  devient grand. On appelle cette suite de points les **itérés** de la méthode.
- l'**analyse de convergence** d'un algorithme d'optimisation regroupe une grande quantité de questions. La plus basique est : à quelle condition une méthode converge-t-elle vers une solution de mon problème ? Dans mes travaux, je m'intéresse plus particulièrement à la **vitesse de convergence** (ou **taux de convergence**) de schémas d'optimisation, c'est-à-dire comment décroît la valeur de  $F(x_k) - F^*$  où  $F^*$  est la valeur minimale atteinte par  $F$  sur l'ensemble  $X$ . En pratique, les résultats que l'on obtient sont des vitesses de convergence **au pire cas** et peuvent être valables à nombre d'itérations fini ou asymptotiquement (quand  $k$  tend vers l'infini). Ces informations permettent de comparer plusieurs schémas différents dans leur capacité à donner une réponse rapidement. Il existe d'autres axes pour l'analyse de la convergence d'un schéma comme la stabilité de ce schéma en présence de perturbations.
- pour obtenir des résultats de convergence forts pour un schéma, on peut considérer des hypothèses plus restrictives sur la fonction  $F$  comme des **hypothèses de croissance**. Réduire la classe de fonctions à laquelle on s'intéresse donne accès à plus d'outils mathématiques et à une analyse plus fine. Par contre, si les hypothèses sont trop fortes, les résultats démontrés ne sont valables que dans des cas marginaux. Les **hypothèses de croissance** sont

des conditions sur la géométrie de la fonction  $F$ , c'est-à-dire sur la façon dont  $F$  se comporte sur l'ensemble  $X$ . Ces hypothèses sont détaillées en Section 2.1.2.

La question générale à laquelle j'ai voulu répondre est : **quel schéma d'optimisation appliquer en fonction des hypothèses sur la fonction à minimiser ?** Il s'avère que les algorithmes de premier ordre les plus efficaces sont dits **inertiels** car ils font intervenir une notion d'élan. Cette inertie est généralement bénéfique mais peut aussi générer des comportements pathologiques, surtout quand  $F$  vérifie des hypothèses de croissance. La problématique de cette thèse peut donc être reformulée : **comment maîtriser l'inertie lorsque  $F$  vérifie une hypothèse de croissance plus forte que la convexité ?** Mes travaux se sont appuyés sur le parallèle qui peut être fait entre systèmes dynamiques et schémas numériques (voir Section 2.3) mais également sur des stratégies plus algorithmiques.

### 1.3 Contexte

Afin de contextualiser les travaux menés durant cette thèse, deux tableaux récapitulatifs des résultats de convergence démontrés dans la littérature sont fournis. Ces tableaux visent à décrire de façon succincte le comportement de différents schémas (Tableau 1.1) et systèmes dynamiques (Tableau 1.2) en fonction d'hypothèses de géométrie. Pour faciliter la lecture, certaines hypothèses ne sont pas précisées, notamment sur les choix de paramètres. Pour plus de détails, j'encourage les lecteurs à consulter les articles ou chapitres référencés.

**Hypothèses de géométrie** Les hypothèses énoncées dans ces tableaux sont détaillées dans la Section 2.1.2. Celles-ci vont de la forte convexité (notée  $\mathcal{S}_\mu$ ) à la simple convexité. Ces tableaux traitent également des fonctions à croissance quadratique ( $\mathcal{G}_\mu^2$ ) et des fonctions à croissance faible ( $\mathcal{G}^\gamma$  avec  $\gamma > 2$ ), c'est-à-dire les fonctions vérifiant une inégalité du type :

$$\forall x \in \text{dom}(F), \quad \frac{K}{2} d(x, X^*)^\gamma \leq F(x) - F^*, \quad (1.3.1)$$

pour un certain  $K > 0$  et où  $\gamma = 2$  et  $K = \mu$  pour l'hypothèse  $\mathcal{G}_\mu^2$ . Ces hypothèses peuvent être définies localement, c'est-à-dire dans un voisinage de l'ensemble de minimiseurs  $X^*$ . Les résultats énoncés restent valables dans ce cas car ils sont asymptotiques et les systèmes et schémas considérés assurent aux trajectoires solutions et itérés d'atteindre et rester dans un voisinage de  $X^*$ .

**Commentaires sur le Tableau 1.1** Le Tableau 1.1 donne un aperçu non exhaustif des résultats de convergence de méthodes d'optimisation de premier ordre pour des fonctions différentiables et/ou composites. Les différents algorithmes sont comparés à partir des garanties théoriques sur la décroissance asymptotique de l'erreur  $F(x_k) - F^*$  au pire cas sous plusieurs hypothèses différentes. Plus de détails sur ces



méthodes sont fournis dans la Section 2.2. Notons que les schémas aux philosophies proches sont regroupés comme les méthodes de type Heavy-Ball ou les algorithmes de restart de FISTA.

Les hypothèses de géométrie sur la fonction  $F$  à minimiser sont ordonnées de la plus forte (forte convexité) à la plus faible (convexité). Lorsqu'une case du tableau est remplie par une flèche ( $\rightarrow$ ), cela signifie que la décroissance de l'erreur garantie par la méthode considérée pour l'hypothèse faite sur  $F$  est la même que pour une hypothèse plus faible. Prenons l'exemple de la méthode de descente de gradient et de son alternative en optimisation composite, Forward-Backward : le taux de convergence de l'erreur pour ces méthodes si  $F$  est fortement convexe est le même que si  $F$  vérifie  $\mathcal{G}_\mu^2$  et ne nécessite pas de connaître  $\mu$ . Lorsqu'on affaiblit ces hypothèses, la décroissance garantie pour l'erreur devient polynomiale. On peut également lire ce tableau de manière verticale afin de déterminer la ou les méthodes ayant les meilleures garanties pour une hypothèse : pour les hypothèses  $\mathcal{S}_\mu$  et  $\mathcal{G}_\mu^2$  où  $\mu$  est connu, les algorithmes les plus rapides sont de type Heavy-Ball ; pour l'hypothèse  $\mathcal{G}_\mu^2$  sans connaître  $\mu$ , les schémas de restart sont les plus efficaces ; FISTA (comme ses variations) est le meilleur choix pour les fonctions vérifiant  $\mathcal{G}^\gamma$  pour  $\gamma > 2$  et les fonctions convexes.

**Systèmes dynamiques** Le Tableau 1.2 énonce des résultats de convergence pour les trajectoires de systèmes dynamiques. Ces systèmes peuvent être mis en lien avec des schémas d'optimisation et ont des propriétés de convergence proches. L'histoire et les apports de cette approche continue sont développés dans la Section 2.3. Les systèmes que l'on considère dans le Tableau 1.2 sont les suivants :

- le système de **flot de gradient** :

$$\forall t \geq t_0, \quad \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{FG})$$

avec  $t_0 \geq 0$ ,

- le système de **Heavy-Ball avec friction** :

$$\forall t \geq t_0, \quad \ddot{x}(t) + \alpha \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{HBF})$$

avec  $t_0 \geq 0$  et  $\alpha > 0$ .

- le système avec **friction asymptotiquement évanescence** :

$$\forall t \geq t_0, \quad \ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{AVD})$$

avec  $t_0 \geq 0$  et  $\alpha > 0$ .

- le système avec **friction asymptotiquement évanescence et friction liée à la hessienne** (voir Chapitre 4) :

$$\forall t \geq t_0, \quad \ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \beta H_F(x(t)) \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{DIN-AVD})$$

avec  $t_0 \geq 0$ ,  $\alpha > 0$  et  $\beta > 0$ .

| Algorithmes  | $S_\mu$                               | $\mathcal{G}_\mu^2$  | $\mathcal{G}_\mu^2, \mu$ inconnu  | $\mathcal{G}^\gamma, \gamma > 2$   | Convexit           |
|--|---------------------------------------|--|---|--|--------------------|
| Descente de gradient <sup>1</sup><br>Forward-Backward [54]     | $\rightarrow$                         | $\rightarrow$  | $e^{-\kappa k}$ [64]  | $k^{-\frac{\gamma}{\gamma-2}}$ [64]  | $k^{-1}$           |
| Heavy-Ball de Polyak <sup>1</sup> [112]                        | $e^{-2\sqrt{\kappa}k}$ si $C^2$ [112] | $\rightarrow$  | $\rightarrow$   | $\rightarrow$  | $k^{-1}$ [65]      |
| M thode du triple  lan <sup>1</sup> [133]                      | $e^{-2\sqrt{\kappa}k}$ [133]          | ?  | ?   | ?  | ?                  |
| ITEM <sup>1</sup> [128]  | $e^{-2\sqrt{\kappa}k}$ [128]          | $\rightarrow$  | $\rightarrow$   | $\rightarrow$  | $k^{-2}$ [128]     |
| V-FISTA [105, 31]  | $e^{-\sqrt{\kappa}k}$ [105, 31]       | $e^{-\frac{9}{20}\sqrt{\kappa}k}$ si<br>$\kappa < 10^{-2}$ (Chap. 5)   | $e^{-\tau\kappa k}$ pour $\tau > 0$<br>(Chap. 5)  | ?  | ?                  |
| Heavy-Ball d'Aujol et al. [25]                                 | $\rightarrow$                         | $e^{(-(2-\sqrt{2})\sqrt{\kappa}+O(\kappa))k}$<br>si $\mathcal{U}^*$<br>$e^{(-\kappa+\varepsilon+O(\kappa))k}$ [25] | ?   | ?  | ?                  |
| FISTA [103, 32]  | $k^{-\frac{2\gamma}{3}}$ [12]         | $\rightarrow$  | $k^{-\frac{2\gamma}{3}}$ si $\mathcal{U}^*$ [7, 26]<br>$k^{-\frac{2\gamma}{3}}$ (Chap. 5) | $k^{-\frac{2\gamma}{\gamma-2}}$ si $\mathcal{F}_\gamma$ et $\mathcal{U}^*$<br>[7]<br>$k^{-\frac{2\gamma}{\gamma-2}}$ (Chap. 5) | $k^{-2}$ [103, 32] |
| Restart automatique de FISTA [28]<br>(Algorithme 3)            | $\rightarrow$                         | $\rightarrow$  | $e^{-\frac{1}{12}\sqrt{\kappa}k}$ (Chap. 3)   | ?  | ?                  |
| Free-FISTA<br>(Algorithme 6)                                   | $\rightarrow$                         | $\rightarrow$  | $e^{-\frac{\sqrt{\tau}}{12}\sqrt{\kappa}k}$ sans<br>conna tre $L$ (Chap. 3)               | ?  | ?                  |
| Autres sch mas de restart<br>[105, 102, 1, 3, 2, 117, 61, 108] | $\rightarrow$                         | $e^{-\frac{1}{\varepsilon}\sqrt{\kappa}k}$ [105]   | $e^{-\frac{1}{2\sqrt{2}}\sqrt{\kappa}k}$ [117]<br>D tails dans le<br>Tableau 3.1          | ?  | ?                  |

TABLEAU 1.1 – Ordre de d croissance asymptotique de l'erreur  $F(x_k) - F^*$  pour plusieurs sch mas d'optimisation en fonction de la g om trie de la fonction  $F$ . On suppose que  $F$  est convexe et   gradient  $L$ -Lipschitz (ou que  $F$  v rifie  $\mathcal{H}_L$  dans le cas non-diff rentiable, cf. Hypoth se 2.0.1). On note  $\mathcal{U}^*$  l'hypoth se d'unicit  du minimiseur et  $\mathcal{F}_\gamma$  est l'hypoth se de platitude introduite dans la Section 2.1.2.4. La constante  $\kappa$  vaut  $\frac{\mu}{L}$ .

<sup>1</sup> : M thode d finie pour les fonctions diff rentiables.

| EDO                    | Convexité      | $\mathcal{G}^1, \gamma > 2$  | $\mathcal{G}^2_\mu$  | $S_\mu$                            |
|------------------------|----------------|--|--|------------------------------------|
| (FG)                   | $t^{-1}$       | $t^{-\frac{\gamma-2}{\gamma-2}}$   | $e^{-\mu t}$   | $e^{-\mu t}$                       |
| (HBF)                  | $t^{-1}$ [4]   | $t^{-\frac{\gamma-2}{\gamma-2}}$ si $C^2$ [36]                             | $e^{-\mu t}$ si $C^2$ [36]   | $e^{-2\sqrt{\mu}t}$ si $C^2$ [112] |
|                        |                |  | $e^{-(2-\sqrt{2})\sqrt{\mu}t}$ si $\mathcal{U}^*$ [25]                             |                                    |
| (AVD)                  | $t^{-2}$ [127] | $t^{-\frac{2\gamma}{\gamma-2}}$ si $\mathcal{F}_\gamma$ et coercivité [27] | $e^{-(2-\sqrt{2})\sqrt{\mu}t}$ si $\mathcal{R}^*$ (Chap. 5)                        | $e^{-\sqrt{2}\mu t}$ [24]          |
|                        |                |  | $t^{-\frac{2\alpha}{3}}$ si $\mathcal{U}^*$ [27, 26]                               |                                    |
| (DIN-AVD) <sup>1</sup> | $t^{-2}$ [21]  | $t^{-\frac{2\gamma}{\gamma-2}}$ si $\mathcal{R}^*$ (Chap. 5)               | $t^{-\frac{2\alpha}{3}}$ si $\mathcal{R}^*$ (Chap. 5)                              | $t^{-\frac{2\alpha}{3}}$ [16]      |
|                        |                |  | $t^{-\alpha+\varepsilon}$ pour tout $\varepsilon > 0$ si $\mathcal{U}^*$ (Chap. 4) |                                    |

TABLEAU 1.2 – Ordre de décroissance asymptotique de l'erreur  $F(x(t)) - F^*$  le long de la trajectoire de systèmes dynamiques en fonction de la géométrie de la fonction  $F$ . La fonction  $F$  est toujours supposée convexe différentiable. On note  $\mathcal{U}^*$  l'hypothèse d'unicité du minimiseur,  $\mathcal{F}_\gamma$  est une hypothèse de platitude introduite dans la Section 2.1.2.4 et  $\mathcal{R}^*$  est l'hypothèse de régularité de second ordre de l'ensemble  $X^*$ .

<sup>1</sup> : Cette EDO requiert que  $F$  soit  $C^2$ .

## 1.4 Contributions et plan du manuscrit

Ce manuscrit est divisé en quatre chapitres, sans compter cette introduction. Les contributions mentionnées dans cette section sont présentes en rouge dans le Tableau 1.1 et le Tableau 1.2.

Le Chapitre 2 est dédié à l'introduction de notions nécessaires pour les chapitres suivants. Une brève présentation des méthodes classiques en optimisation convexe différentiable et non-différentiable est fournie après avoir introduit plusieurs notions sur la géométrie des fonctions convexes. Une troisième partie est consacrée au lien entre schémas numériques et systèmes dynamiques et ses aspects.

Le Chapitre 3 traite des méthodes de restart. Les travaux qui y sont présentés font suite au constat suivant : lorsque la fonction  $F$  à minimiser vérifie une hypothèse de croissance (forte convexité par exemple), les méthodes les plus rapides doivent être calibrées en fonction du paramètre de croissance  $\mu$  qui peut être difficile à estimer. Les stratégies de restart visent à répondre à la question : **comment profiter des hypothèses de croissance d'une fonction lorsque le paramètre de croissance n'est pas connu ?** Après avoir présenté le concept des stratégies de restart ainsi que les méthodes existant dans la littérature, j'introduis l'Algorithme 3 pour les fonctions à croissance quadratique. Contrairement aux schémas de type Heavy-Ball, ce schéma ne nécessite **aucune estimation du paramètre de croissance** et assure une **décroissance exponentielle rapide de l'erreur** (Corollaire 3.2.9), la notion de décroissance exponentielle rapide ou lente étant introduite en Section 3.1.1. Je propose ensuite le schéma Free-FISTA (Algorithme 6) qui combine des stratégies de restart et de backtracking pour s'affranchir de la connaissance du paramètre de croissance **et de la constante de Lipschitz du gradient de  $F$**  (ou de la partie différentiable de  $F$  dans le cas non lisse). Cette méthode est la seule à ma connaissance permettant de garantir une **décroissance rapide de l'erreur** (Corollaire 3.3.8) sans s'appuyer sur la connaissance de ces paramètres.

Le Chapitre 4 donne lieu à l'étude du système dynamique (DIN-AVD) [21] faisant intervenir un terme de friction lié à la hessienne de  $F$ . L'ajout de ce terme a pour but **l'atténuation des oscillations** liées à l'inertie mise en jeu et des schémas numériques de premier ordre peuvent être dérivés de cette EDO. L'étude proposée dans ce chapitre porte sur les propriétés de convergence des trajectoires de ce système lorsque  $F$  vérifie des hypothèses de croissance. Dans ce contexte, on montre que l'introduction du terme lié à la hessienne permet d'obtenir des **propriétés d'in-tégrabilité sur le gradient de la fonction  $F$**  qui sont de plus en plus fortes à mesure que les hypothèses de croissance sont renforcées. Comme les hypothèses considérées sont liées à l'inégalité de Łojasiewicz, ces propriétés ont un impact sur la décroissance de l'erreur et le Théorème 4.2.2 assure notamment que pour des fonctions à croissance quadratique, **les garanties théoriques de décroissance de l'erreur le long des trajectoires de (DIN-AVD) sont meilleures que celles du système (AVD)**. Ces résultats témoignent de l'apport de ce terme additionnel lorsque  $F$  satisfait des hypothèses de géométrie plus fortes que la convexité.

Le Chapitre 5 porte sur l'hypothèse d'unicité du minimiseur dans l'analyse de

convergence de méthodes inertielles. Dans la littérature, les résultats de convergence les plus forts pour des systèmes dynamiques et des schémas numériques lorsque  $F$  vérifie une condition de croissance s'appuient, sauf exception, sur l'hypothèse que  $F$  a un unique minimiseur. Les questions posées dans ce chapitre sont donc : **à quoi est due cette hypothèse ? Est-il possible d'obtenir des résultats de convergence rapide en s'en affranchissant ?** Après avoir détaillé les résultats existants, je donne une stratégie pour contourner cette hypothèse dans le cas continu : j'introduis une hypothèse de géométrie sur l'ensemble de minimiseurs  $X^*$  (cette hypothèse est notée  $\mathcal{R}^*$  dans le Tableau 1.2) permettant d'**étendre plusieurs résultats de convergence sans unicité du minimiseur**. Dans un deuxième temps, je démontre que cette hypothèse d'unicité du minimiseur n'est pas nécessaire à l'obtention de résultats de convergence rapide dans le cas discret en étudiant V-FISTA [31] et FISTA [32]. Le Théorème 5.4.2 assure que la décroissance de l'erreur pour les itérés de V-FISTA avec un choix d'inertie pertinent est **exponentielle rapide** si on considère une fonction à croissance quadratique. Les résultats obtenus ne s'appuient sur **aucune hypothèse sur l'ensemble des minimiseurs et des garanties sont données pour un choix de paramètre sous-optimal** (Corollaire 5.4.4). Enfin, une analyse de FISTA est proposée sous des hypothèses de croissance plus faibles que la forte convexité. Le résultat principal est donné dans le Théorème 5.4.6 qui assure que **les itérés de FISTA convergent fortement dès que  $F$  vérifie une hypothèse de croissance faible  $\mathcal{G}^\gamma$  avec  $\gamma > 2$**  (hypothèse définie en Section 2.1.2) pour un terme inertiel suffisamment grand.

## 1.5 Publications associées à ce manuscrit

Les travaux menés durant cette thèse ont conduit à l'écriture des articles suivants :

1. **FISTA Restart Using an Automatic Estimation of the Growth Parameter.** Jean-François Aujol, Charles Dossal, Hippolyte Labarrière, Aude Rondepierre. Soumis en 2021 à JOTA (*Journal of Optimization Theory and Applications*), en révision mineure.
2. **Fast Convergence of Inertial Dynamics with Hessian-Driven Damping under Geometry Assumptions.** Jean-François Aujol, Charles Dossal, Văn Hào Hoàng, Hippolyte Labarrière, Aude Rondepierre. Soumis en 2022 à AMOP (*Applied Mathematics & Optimization*), accepté.
3. **Parameter-Free FISTA by Adaptive Restart and Backtracking.** Jean-François Aujol, Luca Calatroni, Charles Dossal, Hippolyte Labarrière, Aude Rondepierre. Soumis en 2023 à SIOPT (*SIAM Journal on Optimization*).
4. **Fast Convergence of Heavy-Ball Dynamics and Derived Scheme Without Uniqueness of the Minimizer.** Jean-François Aujol, Charles Dossal, Hippolyte Labarrière, Aude Rondepierre. Bientôt soumis.
5. **Strong Convergence of FISTA under a Weak Growth Condition.**

Jean-François Aujol, Charles Dossal, Hippolyte Labarrière, Aude Rondépierre. Bientôt soumis.



# Chapitre préliminaire

---

Ce chapitre pose les bases mathématiques nécessaires pour les chapitres suivants. Je définis d'abord la classe de fonctions étudiée dans ce manuscrit ainsi que des conditions de croissance, allant de la forte convexité à des hypothèses de platitude. Par la suite, des méthodes classiques sont présentées pour la minimisation d'une fonction différentiable puis composite. Enfin, je détaille le parallèle existant entre systèmes dynamiques et schémas d'optimisation en donnant un historique et en expliquant l'intérêt de ce cadre continu.

## Contents

---

|  |           |
|--|-----------|
| <b>2.1 Géométrie des fonctions convexes</b> . . . . .                        | <b>12</b> |
| 2.1.1 Convexité et caractère Lipschitz . . . . .                             | 12        |
| 2.1.2 Géométrie des fonctions convexes . . . . .                             | 15        |
| <b>2.2 Algorithmes d'optimisation</b> . . . . .                              | <b>19</b> |
| 2.2.1 Algorithmes pour l'optimisation de fonctions différentiables . . . . . | 19        |
| 2.2.2 Algorithmes pour l'optimisation non lisse . . . . .                    | 23        |
| <b>2.3 Optimisation et systèmes dynamiques</b> . . . . .                     | <b>26</b> |
| 2.3.1 Historique . . . . .   | 26        |
| 2.3.2 Apports du parallèle continu/discret . . . . .                         | 29        |
| 2.3.3 Analyse de Lyapunov . . . . .  | 30        |

---

Ce manuscrit traite principalement le problème d'optimisation suivant :

$$\min_{x \in \mathbb{R}^N} F(x), \quad (2.0.1)$$

où  $N \in \mathbb{N}^*$  et  $F$  est une fonction convexe composite, c'est-à-dire qu'elle vérifie l'hypothèse ci-dessous.

**Hypothèse 2.0.1.** Une fonction  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  vérifie l'hypothèse  $\mathcal{H}_L$  pour un certain  $L > 0$  si elle peut s'écrire  $F = f + h$  où :

- la fonction  $f$  est convexe et différentiable à gradient  $L$ -Lipschitz avec  $L > 0$ . Son gradient sera noté  $\nabla f$ .
- la fonction  $h$  est convexe, propre et semi-continue inférieurement.

De plus,  $F$  admet un ensemble de minimiseur  $X^*$  non vide.

Cette hypothèse s'appuie sur plusieurs notions que l'on définit dans la section ci-dessous.



## 2.1 Géométrie des fonctions convexes

### 2.1.1 Convexité et caractère Lipschitz

Commençons par définir la notion de convexité.

**Définition 2.1.1.** Une fonction  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  est **convexe** si et seulement si

$$\forall (x, y) \in \text{dom}(F) \times \text{dom}(F), \forall \lambda \in [0, 1], F(\lambda x + (1 - \lambda)y) \leq \lambda F(x) + (1 - \lambda)F(y), \quad (2.1.1)$$

où  $\text{dom}(F) = \{x \in \mathbb{R}^N, F(x) \neq +\infty\}$  est le domaine de définition  $F$ .

Pour reformuler, une fonction est convexe si tout segment entre deux points de son graphe se situe au-dessus de ce graphe. On peut voir sur la Figure 2.1 un exemple de fonction convexe et un exemple de fonction non convexe.

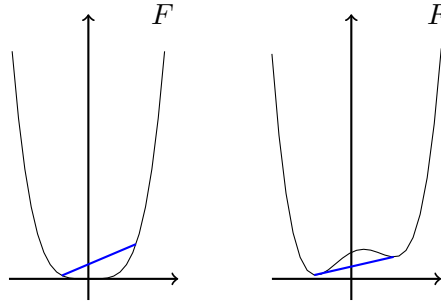


FIGURE 2.1 – Graphes d’une fonction convexe (à gauche) et d’une fonction non convexe (à droite).

On peut également définir la convexité de la manière suivante pour les fonctions différentiables :

**Définition 2.1.2.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction différentiable. La fonction  $F$  est **convexe** si et seulement si

$$\forall (x, y) \in \text{dom}(F) \times \text{dom}(F), F(x) \geq F(y) + \langle \nabla F(y), x - y \rangle, \quad (2.1.2)$$

où  $\nabla F$  est le gradient de  $F$ .

Cette caractérisation de la convexité traduit le fait que tout point du graphe de  $F$  se situe au-dessus du plan tangent à  $F$  en tout autre point.

Notons que la convexité d’une fonction  $F$  n’implique pas l’existence d’une solution au problème (2.0.1). Par exemple, la fonction  $F : x \mapsto e^x$  est convexe et bornée inférieurement par 0 sur  $\mathbb{R}$  mais n’atteint jamais son minimum (voir Figure 2.2).

Une autre notion importante en optimisation est le caractère Lipschitz d’une fonction.

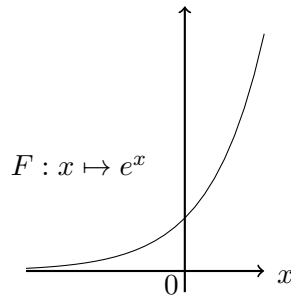


FIGURE 2.2 – Graphe de la fonction exponentielle  $x \mapsto e^x$  qui n’admet pas de minimum sur  $\mathbb{R}$ .

**Définition 2.1.3.** Une fonction  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}^M$  (où  $M \in \mathbb{N}^*$ ) est  **$L$ -Lipschitz** pour un certain  $L > 0$  si

$$\forall (x, y) \in \mathbb{R}^N \times \mathbb{R}^N, \quad \|\phi(x) - \phi(y)\| \leq L\|x - y\|, \quad (2.1.3)$$

*Remarque 2.1.4.* Si une fonction  $\phi$  est  $L$ -Lipschitz pour un certain  $L > 0$ , alors  $\phi$  est également  $\hat{L}$ -Lipschitz pour tout  $\hat{L} > L$ . Usuellement, on dira que la constante de Lipschitz d’une fonction  $\phi$  est la plus petite valeur  $L$  telle que  $\phi$  est  $L$ -Lipschitz.

Quand on étudie des algorithmes d’optimisation de premier ordre, on considère généralement des fonctions à gradient  $L$ -Lipschitz. Cela permet d’obtenir un contrôle sur les variations de ces fonctions et de calibrer le pas de ces méthodes. Cette hypothèse donne également accès à l’inégalité ci-dessous qui intervient dans les preuves de convergence de ces schémas.

**Proposition 2.1.5.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe, différentiable et à gradient  $L$ -Lipschitz. Alors pour tout  $(x, y) \in \text{dom}(F) \times \text{dom}(F)$ , on a :

$$F(x) \leq F(y) + \langle \nabla F(y), x - y \rangle + \frac{L}{2}\|x - y\|^2. \quad (2.1.4)$$

Cette proposition nous permet de majorer toute fonction convexe, différentiable à gradient  $L$ -Lipschitz en tout point de son domaine par une fonction quadratique passant par tout point de son graphe. Si  $F$  vérifie ces hypothèses, pour tout  $y \in \text{dom}(F)$  on connaît donc un minorant affine ( $x \mapsto F(y) + \langle \nabla F(y), x - y \rangle$ ) par la Définition 2.1.2) et un majorant quadratique ( $x \mapsto F(y) + \langle \nabla F(y), x - y \rangle + \frac{L}{2}\|x - y\|^2$ ) par la Proposition 2.1.5) de  $F$  passant par le point  $(y, F(y))$ . Ceci est illustré dans la Figure 2.3.

Enfin, pour s’intéresser à l’optimisation non lisse, nous avons besoin des définitions suivantes.

**Définition 2.1.6.** Une fonction  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  est **propre** si elle n’est pas identiquement égale à  $+\infty$  et qu’elle ne vaut jamais  $-\infty$ .

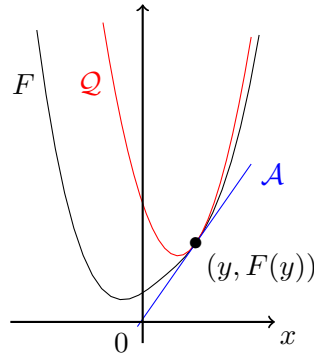


FIGURE 2.3 – Graphe d’une fonction  $F$  convexe, différentiable et à gradient  $L$ -Lipschitz, d’un majorant quadratique passant par  $(y, F(y))$  noté  $\mathcal{Q}$  (en rouge) et d’un minorant affine passant par  $(y, F(y))$  noté  $\mathcal{A}$  (en bleu).

**Définition 2.1.7.** Une fonction  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  est **semi-continue inférieurement** si pour tout  $\alpha \in \mathbb{R}$ , l’ensemble de niveaux  $\{x \in \mathbb{R}^N \mid F(x) \leq \alpha\}$  est fermé.

Un outil d’analyse important en optimisation convexe non lisse est le sous-différentiel pour les fonctions convexes qui a été introduit par Moreau [98] et Rockafellar [119] au début des années 1960.

**Définition 2.1.8.** Soit  $h$  une fonction convexe, propre et semi-continue inférieurement. Le sous-différentiel de  $h$  noté  $\partial h$  est défini tel que :

$$\forall x \in \mathbb{R}^N, \quad \partial h(x) = \{d \in \mathbb{R}^N \mid \forall y \in \mathbb{R}^N, h(y) \geq h(x) + \langle y - x, d \rangle\}.$$

On peut interpréter le sous-différentiel comme l’ensemble des directions des hyperplans passant par le point  $(x, h(x))$  et qui restent sous le graphe. On donne un exemple en dimension 1 sur la Figure 2.4.

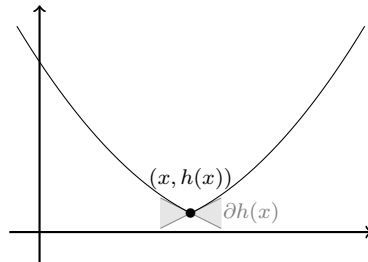


FIGURE 2.4 – Exemple d’une fonction non différentiable en un point  $x$  et son sous-différentiel en ce point, noté  $\partial h(x)$ .

Notons que si  $h$  est une fonction différentiable sur  $\mathbb{R}^N$ , alors on a pour tout  $x \in \mathbb{R}^N$ ,  $\partial h(x) = \{\nabla h(x)\}$ . D’autre part, la semi-continuité inférieure assure l’existence d’un sous-gradient en tout point.

### 2.1.2 Géométrie des fonctions convexes

Mes travaux de thèse ont porté sur des fonctions qui vérifient l'hypothèse  $\mathcal{H}_L$  pour un certain  $L > 0$ . Les résultats obtenus par la suite reposent également sur des hypothèses additionnelles sur la fonction  $F$  à minimiser.

On dit que ces hypothèses, tout comme l'hypothèse de convexité, portent sur la **géométrie** de  $F$ , c'est-à-dire sur la manière dont  $F$  se comporte sur son domaine de définition. En optimisation, faire des hypothèses plus fortes que la convexité permet de construire des schémas d'optimisation adaptés pour obtenir des meilleures vitesses de convergence. La contrepartie est que plus une hypothèse de géométrie est forte, plus l'ensemble de fonctions concernées est réduit.

*Remarque 2.1.9.* Les hypothèses de géométrie présentées dans cette section sont dites globales, c'est-à-dire qu'elles doivent être vérifiées sur le domaine de définition de  $F$  tout entier. On peut cependant définir ces mêmes hypothèses localement, c'est-à-dire au voisinage de l'ensemble de minimiseurs de la fonction étudiée. Les hypothèses locales sont moins restrictives que leur version globale.

Les résultats démontrés durant ma thèse s'appuient principalement sur des hypothèses globales. Toutefois, plusieurs de ces résultats sont **asymptotiques** (valables lorsque le nombre d'itérations ou le temps est suffisamment grand) pour des schémas et systèmes qui garantissent que leurs itérés ou trajectoires rejoignent un voisinage de  $X^*$  au bout d'un certain temps. On pourrait donc les étendre pour des hypothèses locales.

#### 2.1.2.1 Forte convexité

Une condition de géométrie très étudiée dans la littérature est la **forte convexité**. En effet, de nombreux schémas ont été proposés spécifiquement pour des fonctions fortement convexes. On définit la forte convexité de la manière suivante :

**Définition 2.1.10.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$ . On dit que la fonction  $F$  est  $\mu$ -fortement convexe pour  $\mu > 0$  si la fonction  $x \mapsto F(x) - \frac{\mu}{2}\|x\|^2$  est convexe. On écrira que  $F$  satisfait  $\mathcal{S}_\mu$ .

Si  $F$  vérifie quelques hypothèses supplémentaires, on peut utiliser la définition suivante.

**Définition 2.1.11.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe, propre et semi-continue inférieurement. La fonction  $F$  est  $\mu$ -fortement convexe pour un certain  $\mu > 0$  si pour tout  $(x, y) \in \text{dom}(F) \times \text{dom}(F)$  et  $d \in \partial F(y)$  :

$$F(x) \geq F(y) + \langle d, x - y \rangle + \frac{\mu}{2}\|x - y\|^2. \quad (2.1.5)$$

Si  $F$  est différentiable, on a :

$$F(x) \geq F(y) + \langle \nabla F(y), x - y \rangle + \frac{\mu}{2}\|x - y\|^2. \quad (2.1.6)$$

Les fonctions fortement convexes vérifient notamment les deux propriétés ci-dessous.

**Proposition 2.1.12.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction  $\mu$ -fortement convexe pour un certain  $\mu > 0$ . La fonction  $F$  vérifie les propriétés suivantes :*

1.  $F$  admet un unique minimiseur  $x^* \in \text{dom}(F)$ .
2. Pour tout  $x \in \text{dom}(F)$ ,

$$F(x) - F^* \geq \frac{\mu}{2} \|x - x^*\|^2. \quad (2.1.7)$$

Cette hypothèse permet de construire des méthodes rapides en prenant en compte le paramètre de croissance  $\mu$  et la constante de Lipschitz  $L$ , et plus précisément le rapport  $\frac{L}{\mu}$  que l'on définit comme le conditionnement de la fonction  $F$ . Il s'agit d'une hypothèse restrictive en pratique et c'est pourquoi on s'intéressera principalement à des hypothèses plus faibles.

### 2.1.2.2 Quasi-forte convexité

Dans [102], Necoara, Nesterov et Glineur introduisent la notion de **quasi-forte convexité** qui est une relaxation de la forte convexité. La volonté des auteurs est d'affaiblir l'hypothèse tout en conservant des propriétés permettant d'obtenir des vitesses de convergence rapides.

**Définition 2.1.13.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe, propre, semi-continue inférieurement et ayant un ensemble de minimiseurs non vide noté  $X^*$ . La fonction  $F$  est  $\mu$ -quasi-fortement convexe pour un certain  $\mu > 0$  si pour tout  $x \in \text{dom}(F)$  et  $d \in \partial F(x)$  :

$$F^* \geq F(x) + \langle d, P_{X^*}(x) - x \rangle + \frac{\mu}{2} \|P_{X^*}(x) - x\|^2, \quad (2.1.8)$$

où  $P_{X^*}$  est la projection sur  $X^*$ . On écrira que  $F$  satisfait  $\mathcal{Q}_\mu$ .

Une différence majeure avec la forte convexité est que cette hypothèse peut être satisfaite par des fonctions n'ayant pas un unique minimiseur.

### 2.1.2.3 Croissance quadratique

Mes travaux se sont principalement basés sur des fonctions ayant une **croissance quadratique** ce qui est une hypothèse plus faible que les deux précédentes. On la formule de la manière suivante.

**Définition 2.1.14.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction ayant un ensemble de minimiseurs non vide noté  $X^*$ . La fonction  $F$  satisfait une hypothèse de croissance quadratique de paramètre  $\mu > 0$  si pour tout  $x \in \text{dom}(F)$  :

$$\frac{\mu}{2} d(x, X^*)^2 \leq F(x) - F^*, \quad (2.1.9)$$

où  $d(x, X^*)$  est la distance euclidienne entre  $x$  et  $X^*$ . On écrira que  $F$  satisfait  $\mathcal{G}_\mu^2$ .

Une fonction vérifiant  $\mathcal{G}_\mu^2$  peut être bornée inférieurement par une parabole paramétrée par  $\mu$  : plus sa valeur est faible, plus la fonction  $F$  est plate et mal conditionnée (on rappelle que le conditionnement vaut  $\frac{L}{\mu}$ ). On pourra également parler de fonctions 2-conditionnées [64]. On a la relation suivante entre les hypothèses définies jusqu'ici [102] :

$$\mathcal{S}_\mu \subset \mathcal{Q}_\mu \subset \mathcal{G}_\mu^2$$

Comme on considère des fonctions convexes, l'hypothèse  $\mathcal{G}_\mu^2$  peut être reliée à l'inégalité de Łojasiewicz [89, 88] qui est un outil d'analyse important dans l'analyse de convergence de systèmes dynamiques continus ou discrets. Cette propriété est d'abord définie pour les fonctions différentiables avant d'être étendue au cas non lisse par Bolte, Daniilidis, Lewis et Shiota dans [39, 40].

**Définition 2.1.15.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe, propre et semi-continue inférieurement. La fonction  $F$  vérifie l'inégalité de Łojasiewicz d'exposant  $p$  si pour tout minimiseur  $x^* \in X^*$ , il existe  $c > 0$  et  $\varepsilon > 0$  tels que pour tout  $x \in B(x^*, \varepsilon)$  :

$$F(x) - F^* \leq cd(0, \partial F(x))^{\frac{1}{p}}. \quad (2.1.10)$$

On dira que  $F$  vérifie globalement l'inégalité de Łojasiewicz d'exposant  $p$  si l'inégalité (2.1.10) est satisfaite pour tout  $x \in \text{dom}(F)$ .

Bolte, Nguyen, Peypouquet et Suter dans [41, Théorème 5] puis Garrigos, Villa et Rozasco montrent dans [64, Proposition 3.3] que la condition de croissance quadratique  $\mathcal{G}_\mu^2$  implique l'inégalité de Łojasiewicz globale d'exposant  $\frac{1}{2}$  :

**Proposition 2.1.16.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe, propre et semi-continue inférieurement. Si  $F$  vérifie  $\mathcal{G}_\mu^2$  pour un certain  $\mu > 0$ , alors l'inégalité de Łojasiewicz (2.1.10) est satisfaite globalement avec  $p = \frac{1}{2}$  et  $c = \frac{2}{\mu}$ .

Les auteurs montrent également l'équivalence entre ces deux hypothèses mais seulement sur les ensembles  $\partial F$ -invariants.

Notons que l'inégalité de Łojasiewicz globale d'exposant  $\frac{1}{2}$  est appelée condition de Polyak-Łojasiewicz [111, 77].

Comme la quasi-forte convexité, cette hypothèse ne nécessite pas d'avoir un unique minimiseur contrairement à la forte convexité. On peut citer plusieurs fonctions usuelles non fortement convexes qui vérifient cette hypothèse comme la fonction associée au problème LASSO [41, Corollaire 9] ou la divergence de Kullback-Leibler [50, Lemme 1.2].

#### 2.1.2.4 Hypothèses de platitude

Les hypothèses de géométrie présentées jusqu'ici visent à décrire des fonctions qui ont des propriétés favorables à l'obtention de vitesses de convergence rapides. On pourra parler d'hypothèses de pente raide ou de *sharpness* en anglais, étant donné qu'elles décrivent des fonctions qui ont un comportement proche de  $x \mapsto \frac{\mu}{2}\|x\|^2$ . On

peut cependant noter que cet aspect *sharp* est très atténué quand  $\mu$  est faible. Dans cette section, on introduit des hypothèses de platitude ou *flatness* en anglais. L'hypothèse  $\mathcal{G}^\gamma$  introduite ci-dessous peut être trouvée dans la littérature sous le nom de *Hölderian error bound* [41, 87, 120] ou  $\gamma$ -conditionnement [64].

**Définition 2.1.17.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$ . La fonction  $F$  satisfait  $\mathcal{G}^\gamma$  pour un certain  $\gamma > 2$  s'il existe  $K > 0$  tel que pour tout  $x \in \text{dom}(F)$  :

$$\frac{K}{2}d(x, X^*)^\gamma \leq F(x) - F^*. \quad (2.1.11)$$

Si  $\gamma = 2$ , cette hypothèse est en fait la condition de croissance quadratique  $\mathcal{G}_\mu^2$  introduite précédemment. En prenant  $\gamma > 2$ , la fonction  $x \mapsto \frac{K}{2}d(x, X^*)^\gamma$  devient plus plate au voisinage de ses minimiseurs. Ainsi, quand l'hypothèse est définie localement, on peut voir  $\mathcal{G}^\gamma$  comme une hypothèse de croissance de plus en plus faible à mesure que  $\gamma$  est grand. Comme l'hypothèse de croissance quadratique  $\mathcal{G}_\mu^2$ , l'hypothèse  $\mathcal{G}^\gamma$  peut être reliée à l'inégalité de Łojasiewicz pour un exposant dépendant de  $\gamma$ . Le lien avec l'inégalité de Łojasiewicz est détaillé dans la proposition ci-dessous qui provient de [41, Théorème 5].

**Proposition 2.1.18.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe, propre et semi-continue inférieurement. Si  $F$  vérifie  $\mathcal{G}^\gamma$  pour un certain  $\gamma > 2$ , alors l'inégalité de Łojasiewicz (2.1.10) est vérifiée globalement pour  $p = 1 - \frac{1}{\gamma}$  et  $c = \left(\frac{2}{K}\right)^{\frac{1}{\gamma-1}}$  où  $K > 0$  permet de vérifier (2.1.11).

On introduit ci-dessous l'hypothèse  $\mathcal{F}_\gamma$  qui est une généralisation de la notion de convexité. On définit l'hypothèse dans le cas  $F$  différentiable.

**Définition 2.1.19** (Hypothèse de platitude (*flatness*)). Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe différentiable. La fonction  $F$  satisfait  $\mathcal{F}_\gamma$  pour  $\gamma \geq 1$  si pour tout  $x^* \in X^*$ ,

$$\forall x \in \text{dom}(F), \quad F(x) - F^* \leq \frac{1}{\gamma} \langle \nabla F(x), x - x^* \rangle. \quad (2.1.12)$$

On peut noter que  $\mathcal{F}_1$  est une simple hypothèse de convexité. Cette condition de géométrie a d'abord été proposée par Cabot, Engler et Gadat dans [46] avant d'être notamment étudiée par Su, Boyd et Candès [127] et Aujol, Dossal et Rondepierre [27]. Une propriété notable est que toute fonction  $F$  telle que  $x \mapsto (F(x) - F^*)^{\frac{1}{\gamma}}$  est convexe vérifie  $\mathcal{F}_\gamma$ . Cela peut être exploité pour déterminer des résultats de convergence améliorés [127]. Pour avoir une meilleure intuition de ce que signifie cette hypothèse, notons que si une fonction  $F$  vérifie  $\mathcal{F}_\gamma$ , alors pour tout minimiseur  $x^*$ , il existe  $M > 0$  et  $\nu > 0$  tels que :

$$\forall x \in B(x^*, \nu), \quad F(x) - F^* \leq M \|x - x^*\|^\gamma, \quad (2.1.13)$$

où  $B(x^*, \nu)$  désigne la boule de centre  $x^*$  et de rayon  $\nu$  (voir [27, Lemme 2.2]). Cette hypothèse permet donc de s'assurer que  $F$  ne croît pas trop vite.

### 2.1.2.5 Autres géométries

Les hypothèses présentées jusqu'ici sont celles que j'ai manipulées pendant ma thèse. Cependant, elles ne représentent qu'une partie de ce qui existe dans la littérature. En effet, il existe une très grande variété d'hypothèses reliées entre elles et qui peuvent différer par des nuances très subtiles.

Une des conditions les plus anciennes est la condition d'*error bound* introduite par Luo et Tseng dans [92]. Dans [139, 138], Zhang, Cheng et Yin utilisent les notions de forte convexité restreinte (*restricted strong convexity*) et l'inégalité de la sécante restreinte (*restricted secant inequality*). On peut relier ces différentes hypothèses à la condition de croissance quadratique sous certaines conditions comme montré dans [137]. Les liens entre ces variations de la forte convexité sont détaillés dans les travaux de Karimi, Nutini et Schmidt [77], Necoara, Nesterov et Glineur [102] ou Garrigos, Rozasco et Villa [64].

L'inégalité de Kurdyka-Łojasiewicz, faisant suite aux travaux de Kurdyka [82] et Łojasiewicz [89], est également une hypothèse très présente en optimisation, puisqu'elle permet d'analyser la convergence de schémas dans le cas non convexe. Dans le cas convexe, cette condition est équivalente à une condition *error bound* [41].

## 2.2 Algorithmes d'optimisation

### 2.2.1 Algorithmes pour l'optimisation de fonctions différentiables

L'ensemble des fonctions vérifiant l'hypothèse  $\mathcal{H}_L$  (Hypothèse 2.0.1) contient notamment les fonctions convexes, bornées inférieurement et différentiables à gradient  $L$ -Lipschitz (cas  $h = 0$ ). Dans ce contexte, on a la relation d'équivalence suivante :

$$x^* \text{ est un minimiseur de } \iff \nabla F(x^*) = 0.$$

la fonction  $F$ .

#### 2.2.1.1 Descente de gradient

Lorsqu'on considère une fonction vérifiant  $\mathcal{H}_L$  et différentiable, il existe un grand nombre de méthodes pour déterminer la solution de (2.0.1) parmi lesquelles les méthodes de premier ordre. Il s'agit de schémas itératifs qui nécessitent la connaissance du gradient de  $F$  et un point initial  $x_0$  pour construire une suite  $(x_k)_{k \in \mathbb{N}}$  convergent vers une solution de (2.0.1). La méthode de premier ordre la plus classique est la **descente de gradient** à pas constant qui est définie de la façon suivante. Soit  $x_0 \in \mathbb{R}^N$ ,

$$\forall k \in \mathbb{N}, \quad x_{k+1} = x_k - s \nabla F(x_k), \quad (\text{DG})$$

où  $s > 0$  est le pas. On peut noter que tout minimiseur  $x^*$  de  $F$  est un point fixe de l'opérateur  $I - s \nabla F$ . À condition que l'on choisisse  $s \in (0, \frac{2}{L})$ , la méthode de descente de gradient à pas constant converge globalement, c'est-à-dire  $\lim_{k \rightarrow +\infty} \|\nabla F(x_k)\| = 0$  quel que soit le point initial  $x_0 \in \mathbb{R}^N$ . Cette même condition garantit que la



descente de gradient est un algorithme de descente c'est-à-dire que pour tout  $k \in \mathbb{N}$ ,  $F(x_{k+1}) \leq F(x_k)$ . Si on renforce cette condition et que l'on choisit  $s \leq \frac{1}{L}$ , alors on a

$$F(x_k) - F^* = \mathcal{O}(k^{-1}), \quad (2.2.1)$$

où  $F^* = \min_{x \in \mathbb{R}^N} F(x)$ . Une démonstration de ces propriétés est donnée dans l'Exemple 2.3.3.

### 2.2.1.2 Méthode de Newton

Si en plus de connaître le gradient de  $F$ , on a accès à la matrice hessienne de  $F$  notée  $H_F$ , on pourra utiliser la **méthode de Newton** qui est une méthode de second ordre. À partir d'un point initial  $x_0 \in \mathbb{R}^N$ , on définit :

$$\forall k \in \mathbb{N}, \quad x_{k+1} = x_k - H_F(x_k)^{-1} \nabla F(x_k). \quad (\text{Newton})$$

Ce schéma converge très rapidement à condition que le point initial  $x_0$  soit suffisamment proche d'une solution  $x^*$ . Lorsqu'elle peut être appliquée, la méthode de Newton est souvent une des méthodes les plus efficaces, en exploitant la géométrie de  $F$ . On peut noter cependant que chaque itération demande d'inverser la matrice  $H_F$  qui peut être mal conditionnée. Le coût de chaque itération est donc significativement plus élevé que pour la méthode de descente de gradient. Il existe des alternatives à la méthode de Newton comme les méthodes Quasi-Newton où la hessienne (ou son inverse) est remplacée par une approximation. Ce type de méthodes est pertinent lorsque la hessienne est inconnue, trop coûteuse à calculer ou mal conditionnée.

Dans ce manuscrit, on ne fera généralement pas l'hypothèse que  $F$  est différentiable deux fois donc ces méthodes ne seront pas applicables.

### 2.2.1.3 Méthodes inertielles

Si la méthode de descente de gradient est considérée comme lente, il existe des méthodes de premier ordre plus efficaces.

En 1964, Polyak propose dans [112] un schéma visant à minimiser les fonctions fortement convexes (voir Section 2.1.2) et dérivables deux fois. Cet algorithme appelé **méthode Heavy-Ball de Polyak** s'écrit comme suit. Soit  $x_0 \in \mathbb{R}^N$ ,

$$\forall k \in \mathbb{N}, \quad x_{k+1} = x_k + \alpha(x_k - x_{k-1}) - s \nabla F(x_k), \quad (\text{HBP})$$

où  $\alpha > 0$  et  $x_{-1} = x_0$ . Ce schéma est le premier à introduire de l'**inertie** dans un algorithme d'optimisation de premier ordre. En effet, en prenant en compte la direction de descente précédente via le terme  $\alpha(x_k - x_{k-1})$ , ce schéma génère un élan qui permet d'atteindre des vitesses rapides. À condition que  $F$  soit fortement convexe et deux fois dérivable, il existe un jeu de paramètres  $(\alpha, s)$  dépendant de la géométrie de  $F$  qui garantit une décroissance exponentielle de  $F(x_k) - F^*$ . En revanche, si  $F$  est convexe sans hypothèse de géométrie supplémentaire, cette méthode ne permet

pas de garantir mieux que  $F(x_k) - F^* = \mathcal{O}(k^{-1})$ , c'est-à-dire le même taux de convergence que la descente de gradient.

*Remarque 2.2.1.* La méthode Heavy-Ball de Polyak n'est pas un algorithme de descente : rien ne garantit que  $F(x_{k+1}) \leq F(x_k)$  pour tout  $k \in \mathbb{N}$ .

*Remarque 2.2.2.* Il existe d'autres schémas type Heavy-Ball comme le schéma de Nesterov fortement convexe [105, 106], le schéma de Siegel [126] et les schémas de Aujol, Dossal et Rondepierre [25, 24]. Leur appellation commune provient du fait qu'il s'agit de différentes discrétisations du système de Heavy-ball avec friction :

$$\ddot{x}(t) + \alpha\dot{x}(t) + \nabla F(x(t)) = 0,$$

avec  $\alpha > 0$ . Plus de détails seront donnés en Section 2.3.1.

En 1983, Nesterov introduit dans [103] le schéma que l'on appellera **méthode de gradient accéléré de Nesterov**. Soit  $x_0 \in \mathbb{R}^N$ , le schéma s'écrit de la façon suivante :

$$\forall k \in \mathbb{N}, \begin{cases} x_{k+1} = y_k - s\nabla F(y_k) \\ y_{k+1} = x_{k+1} + \alpha_{k+1}(x_{k+1} - x_k), \end{cases} \quad (2.2.2)$$

où  $s > 0$ ,  $y_0 = x_0$  et  $(\alpha_k)_{k \in \mathbb{N}^*}$  vérifie :

$$\forall k \in \mathbb{N}, \begin{cases} t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2} \\ \alpha_{k+1} = \frac{t_k - 1}{t_{k+1}}, \end{cases} \quad (2.2.3)$$

avec  $t_0 = 1$ . La suite  $(\alpha_k)_{k \in \mathbb{N}}$  ainsi définie peut être approchée pour tout  $k \in \mathbb{N}^*$  par  $\frac{k-1}{k+2}$  comme illustré sur la Figure 2.5. Bien que ces deux choix pour la suite  $(\alpha_k)_{k \in \mathbb{N}}$  assurent un comportement similaire de la méthode, chacun offre des outils d'analyse différents : la définition originale (2.2.3) est la plus naturelle pour construire des schémas de backtracking (voir Section 3.3.2) et le choix  $\frac{k-1}{k+2}$  favorise la mise en parallèle de cette méthode avec un systèmes dynamique (voir Section 2.3).

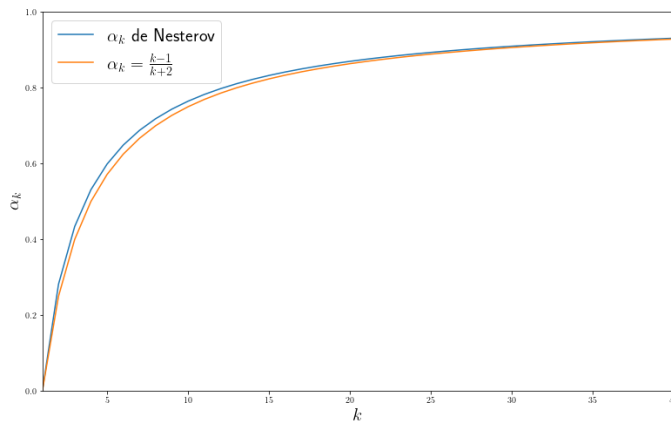


FIGURE 2.5 – Comparaison entre la suite  $(\alpha_k)_{k \in \mathbb{N}^*}$  définie dans (2.2.3) et la suite égale à  $\frac{k-1}{k+2}$  pour tout  $k \in \mathbb{N}^*$ .

Nesterov [103] montre que si  $F$  est convexe, différentiable à gradient  $L$ -Lipschitz, bornée inférieurement et  $s \leq \frac{1}{L}$ , alors la suite  $(x_k)_{k \in \mathbb{N}}$  vérifie  $F(x_k) - F^* = \mathcal{O}(k^{-2})$ . Dans [105, 106], Nesterov montre que cette vitesse est optimale à une constante près pour cette classe de fonction. Dans [55], D'Aspremont propose un schéma qui atteint le même taux sous l'hypothèse que le calcul du gradient est inexact.

La différence majeure entre ce schéma et les méthodes type Heavy-Ball réside dans la gestion de l'inertie. En effet, les méthodes Heavy-Ball font intervenir un coefficient inertiel constant  $\alpha$  là où la méthode de gradient accéléré de Nesterov dispose d'un coefficient croissant  $\alpha_k$  qui tend vers 1 quand  $k$  tend vers  $+\infty$ . Cela signifie que la mémoire des itérations précédentes a de plus en plus d'importance au fil des itérations.

L'analyse de schémas de premier ordre dans le cas différentiable a également été le sujet de travaux plus récents. On peut par exemple citer Kim et Fessler qui proposent dans [80] deux méthodes de premier ordre optimales dans un certain sens. Les auteurs reprennent l'approche de Drori et Teboulle [57] qui consiste à voir l'analyse de performance de schémas de premier ordre comme la résolution d'un problème d'optimisation. Ainsi, ils déterminent le schéma de premier ordre qui leur permet de minimiser la borne au pire cas de  $F(x_k) - F^*$ . Les deux méthodes proposées garantissent le même taux de convergence amélioré :

$$F(x_k) - F^* \leq \frac{L \|x_0 - x^*\|^2}{(k+1)(k+1+\sqrt{2})},$$

pour tout minimiseur  $x^* \in X^*$ . On peut noter que ce résultat est asymptotiquement équivalent à celui du gradient accéléré de Nesterov. Ces méthodes sont optimales dans le cas convexe mais pas nécessairement si  $F$  satisfait des hypothèses plus fortes. De nombreux travaux se portent sur le cas fortement convexe : Nesterov propose dans [105, 106] un schéma de gradient accéléré spécifiquement pour les fonctions fortement convexes ; dans [45], Bubeck, Lee et Singh proposent un algorithme rapide avec une interprétation géométrique ; dans [58], Drusvyatskiy, Fazel et Roy donnent un schéma proche de celui de [45] faisant intervenir la notion de moyennage quadratique ; dans [133], Van Scoy, Freeman et Lynch introduisent la méthode du triple momentum assurant la meilleure vitesse de convergence théorique pour les fonctions fortement convexes lors de la publication du papier ; dans [129], Taylor, Hendrickx et Glineur s'appuient sur l'approche de Drori et Teboulle [57] pour déterminer un schéma qui atteint la vitesse optimale dans le cadre fortement convexe démontrée par Nesterov dans [105, 106] ; dans [128], Taylor et Drori proposent le schéma ITEM (pour *Information-Theoretic Exact Method*) qui assure des vitesses asymptotiques optimales dans le cas convexe et fortement convexe.

L'analyse de la convergence de ces schémas est en général assez différente de ce qui a été fait durant ma thèse. Par exemple, le parallèle entre schéma numérique et système dynamique qui sera détaillé dans la section 2.3.1 n'a pas toujours été mis en évidence pour ces méthodes.

### 2.2.2 Algorithmes pour l'optimisation non lisse

Les méthodes présentées dans la section précédente sont consacrées aux fonctions différentiables. En conséquence, elles ne sont pas directement applicables pour toute fonction satisfaisant  $\mathcal{H}_L$ . Pour s'attaquer à la minimisation de fonctions composites, nous avons besoin d'outils supplémentaires.

Par la suite, on fera l'hypothèse que  $F$  satisfait l'hypothèse  $\mathcal{H}_L$  et que l'on peut donc écrire  $F = f + h$  avec  $f$  une fonction différentiable et  $h$  une fonction en général non différentiable. Comme le gradient de  $F$  n'est pas bien défini dans ce contexte, on a recours à son sous-différentiel. Si on considère une fonction  $F$  convexe, propre et semi-continue inférieurement, on peut écrire une condition d'optimalité de la manière suivante :

$$x^* \text{ est un minimiseur de } \iff 0 \in \partial F(x^*).$$

la fonction  $F$

Comme on fait l'hypothèse que  $F$  satisfait  $\mathcal{H}_L$ , cette condition d'optimalité peut s'écrire  $0 \in \partial h(x^*) + \{\nabla f(x^*)\}$ .

*Remarque 2.2.3.* Si on considère une fonction  $F$  telle que l'on peut calculer  $F(x)$  et au moins une direction  $d \in \partial F(x)$  pour tout  $x \in \mathbb{R}^N$ , alors on peut appliquer des **méthodes de sous-gradients**. Je n'aborderai pas ces schémas dans ce manuscrit mais on pourra consulter [113] et [44] pour plus de détails.

Par la suite, on s'appuiera sur l'opérateur proximal, introduit par Moreau dans [99].

**Définition 2.2.4.** Soit  $h : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$ . L'opérateur proximal de la fonction  $h$  noté  $\text{prox}_h$  est défini tel que :

$$\forall x \in \mathbb{R}^N, \quad \text{prox}_h(x) = \arg \min_{y \in \mathbb{R}^N} h(y) + \frac{1}{2} \|y - x\|^2. \quad (2.2.4)$$

Si  $h : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  est une fonction convexe semi-continue inférieurement, alors  $\text{prox}_h$  existe et est unique pour tout  $x \in \mathbb{R}^N$ . De plus, on a l'équivalence suivante :

$$p = \text{prox}_h(x) \iff x - p \in \partial h(p).$$

On peut noter que pour tout minimiseur  $x^*$  de  $h$ , on a  $\text{prox}_h(x^*) = x^*$ .

L'opérateur proximal peut être utilisé pour minimiser des fonctions non différentiables. En 1970, Martinet [93] introduit la **méthode du point proximal** qui est ensuite popularisée par Rockafellar [118]. Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe semi-continue inférieurement et  $x_0 \in \mathbb{R}^N$ , la méthode du point proximal s'écrit :

$$\forall k \in \mathbb{N}, \quad x_{k+1} = \text{prox}_{tF}(x_k),$$

où  $t > 0$ . On peut montrer que cette méthode converge vers un minimiseur de  $F$  avec un pas constant. Güler analyse la convergence du schéma pour des pas variables

dans [70] et propose des variantes dans [69]. L'utilisation de la méthode du point proximal est assez rare en pratique. En fait, savoir calculer  $\text{prox}_F$  implique généralement que l'on peut directement obtenir une solution. En effet, en faisant tendre  $t$  vers  $+\infty$ , on peut faire converger l'algorithme en une simple itération.

Cette méthode a ouvert la porte à de nombreuses méthodes dites de *splitting* (comme Douglas-Rachford ou l'algorithme des directions alternées) qui s'appuient sur la réécriture de la fonction d'intérêt comme une somme de fonctions traitées de manière différentes. On peut citer les travaux de Passty [110], Lions et Mercier [86], Eckstein [59], Eckstein et Bersektaş [60] ou Combettes et Wajs [54]. Pour une présentation détaillée de ces méthodes, on pourra également consulter les papiers de Tseng [132] ou Combettes et Pesquet [53].

Dans mes travaux, je me suis surtout intéressé à la **méthode du gradient proximal** aussi appelée **Forward-Backward** [54] et à ses variantes. Étant donnée une fonction  $F$  qui satisfait  $\mathcal{H}_L$  pour un certain  $L > 0$  et  $x_0 \in \mathbb{R}^N$ , la méthode du gradient proximal s'écrit de la manière suivante :

$$\forall k \in \mathbb{N}, \quad x_{k+1} = \text{prox}_{sh}(x_k - s\nabla f(x_k)), \quad (\text{FB})$$

où  $s > 0$ . À condition que  $s \in (0, \frac{2}{L})$ , la suite  $(x_k)_{k \in \mathbb{N}}$  converge vers un minimum de  $F$  et  $F(x_k) - F^* = \mathcal{O}(k^{-1})$ . L'appellation *Forward-Backward* provient du fait que chaque itération fait intervenir une étape de descente de gradient explicite (*forward*) et une étape de point proximal qui est implicite (*backward*). On peut voir cette méthode comme une adaptation de la descente de gradient au cas composite ou encore comme une méthode de point fixe à partir de l'opérateur  $T_s$  défini comme :

$$\forall x \in \mathbb{R}^N, \quad T_s(x) = \text{prox}_{sh}(x - s\nabla f(x)). \quad (2.2.5)$$

Au-delà de son rôle majeur dans la méthode de gradient proximal, cet opérateur nous permet de définir l'opérateur de **gradient composite** de  $F$  introduit par Nesterov dans [104] qui est défini par :

$$\forall x \in \mathbb{R}^N, \quad g_s(x) = \frac{1}{s}(x - T_s(x)), \quad (2.2.6)$$

pour  $s > 0$ . Le choix original et usuel pour le gradient composite correspond à  $g_{\frac{1}{L}}$ . Cet opérateur peut servir de substitut au gradient qui n'est en principe pas défini dans le cas composite et lorsque  $F$  est différentiable, on a  $g_s = \nabla F$  pour tout  $s > 0$ .

On peut également utiliser le gradient composite pour définir le critère d'arrêt d'une méthode, c'est-à-dire arrêter cette méthode à condition que  $\|g(x_k)\| \leq \varepsilon$  où  $\varepsilon > 0$  correspond à la précision attendue. Quand  $F$  vérifie une hypothèse de croissance, on peut relier la valeur de  $\|g(x_k)\|$  à celle de  $F(T_s(x_k)) - F^*$  (voir Section 3.2.1.2), impliquant que ce critère d'arrêt nous donne une garantie sur la précision atteinte à la dernière itération.

*Remarque 2.2.5.* On peut construire un schéma type *Forward-Backward* sans avoir besoin que  $\nabla f$  soit à gradient Lipschitz. C'est ce qui est fait par Bauschke, Bolte et Teboulle dans [30] avec l'algorithme NoLips.

Il est possible d'introduire de l'inertie dans la méthode de gradient proximal pour l'accélérer. En 2009, Beck et Teboulle proposent **FISTA** [32] (pour *Fast Iterative Shrinkage-Thresholding Algorithm*) qui est une adaptation de la méthode de gradient accéléré de Nesterov au cas composite. Pour une fonction  $F$  qui vérifie  $\mathcal{H}_L$  pour un certain  $L > 0$  et  $x_0 \in \mathbb{R}^N$ , les itérés de FISTA sont définis par

$$\forall k \in \mathbb{N}, \begin{cases} x_{k+1} = \text{prox}_{sh}(y_k - s\nabla f(y_k)) \\ y_{k+1} = x_{k+1} + \alpha_{k+1}(x_{k+1} - x_k), \end{cases} \quad (2.2.7)$$

où  $y_0 = x_0$ ,  $s > 0$  et  $(\alpha_k)_{k \in \mathbb{N}}$  vérifie (2.2.3). Les auteurs montrent que pour  $s = \frac{1}{L}$ , on a une vitesse de convergence semblable à celle garantie dans le cadre différentiable à savoir  $\mathcal{O}(k^{-2})$ . La question de la convergence de la suite  $(x_k)_{k \in \mathbb{N}}$  est par contre restée sans réponse avant que Chambolle et Dossal [51] ne démontrent que cette suite converge faiblement dans un espace de Hilbert avec le choix  $\alpha_k = \frac{k-1}{k+\alpha-1}$  pour  $\alpha > 3$ . On peut d'ailleurs définir une version plus générale de FISTA : pour  $x_0 \in \mathbb{R}^N$ ,

$$\forall k \in \mathbb{N}, \begin{cases} y_k = x_k + \frac{k}{k+\alpha}(x_k - x_{k-1}), \\ x_{k+1} = \text{prox}_{sh}(y_k - s\nabla f(y_k)) \end{cases} \quad (\text{FISTA})$$

où  $y_0 = x_0$ ,  $s > 0$  et  $\alpha > 0$ . Dans la suite de ce manuscrit, on mentionnera l'algorithme ci-dessus comme FISTA pour un certain choix de  $\alpha$  bien que ce ne soit pas la définition originale.

Notons qu'Attouch et Peypouquet prouvent dans [20] que la décroissance asymptotique de l'erreur quand  $\alpha > 3$  pour FISTA est en fait de l'ordre de  $o(k^{-2})$ . Apidopoulos, Aujol et Dossal [6] et Attouch, Chbani et Riahi [17] démontrent également des vitesses de convergence dans le cas sous-critique  $\alpha < 3$  de l'ordre de  $\mathcal{O}\left(k^{-\frac{2\alpha}{3}}\right)$ . Dans [84], Liang, Luo et Schönlieb proposent des schémas dérivés de FISTA offrant plus de flexibilité avec notamment une méthode adaptée à la forte convexité. On peut noter que FISTA n'est pas le seul schéma accéléré pour les fonctions composites. En 2011, Becker, Bobin et Candès proposent NESTA [34] (pour *Nesterov Acceleration*) qui est un schéma proche de FISTA et s'appuie sur une technique de *smoothing*. La méthode est construite pour la restauration de données parcimonieuses. Dans [121] et [134], Salzo, Villa, Baldassarre et Verri étudient des schémas proximaux inertiels sous l'hypothèse que le calcul de l'opérateur proximal est inexact. Les auteurs proposent un schéma qui assure une décroissance de l'erreur de l'ordre de  $\mathcal{O}(k^{-2})$  sous une condition sur les perturbations. Iutzeler et Malick proposent dans [75] une méthode proximale faisant intervenir un terme d'inertie actif une itération sur deux. Ce schéma permet de contrôler la monotonie de la suite d'itérés et des garanties théoriques sont données.

Des schémas proximaux de type Heavy-Ball ont également été proposés dans le cas fortement convexe parmi lesquels iPiasco de Ochs, Brox et Pock [109] ou V-FISTA de Beck [31]. Le comportement de ces méthodes est proche de celui des schémas analogues en optimisation différentiable.

Il existe également une grande variété d'algorithmes proximaux inertiels pour des

problèmes proches de (2.0.1) comme la recherche de zéros d'opérateurs monotones (Alvarez et Attouch [4], Moudafi et Oliny [100] ou Lorenz et Pock [90]).

## 2.3 Optimisation et systèmes dynamiques

Dans cette partie, nous allons voir un des outils essentiels dans l'analyse de convergence en optimisation à savoir le parallèle qui peut être fait entre algorithmes d'optimisation et systèmes dynamiques.

### 2.3.1 Historique

Le lien entre ces objets mathématiques est relativement naturel et a été notamment utilisé par Polyak dans [112] pour définir son schéma dit Heavy-Ball (HBP). En effet, Polyak a défini cette méthode en discrétisant le système dynamique de **Heavy-Ball avec friction** qui s'écrit :

$$\forall t \geq t_0, \ddot{x}(t) + \alpha \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{HBF})$$

avec  $\alpha > 0$ ,  $t_0 \geq 0$  et  $(x(t_0), \dot{x}(t_0)) = (x_0, \dot{x}_0)$ . Notons que les paramètres  $\alpha$  dans (HBF) et (HBP) n'ont pas exactement le même rôle : si on les note respectivement  $\alpha_c$  et  $\alpha_d$ , alors on a :  $\alpha_c \approx \frac{1-\alpha_d}{\sqrt{s}}$ .

Cette EDO (Équation Différentielle Ordinaire) modélise le mouvement d'un corps (une boule pesante par exemple, ou *heavy ball*) dans un champ de potentiel (la fonction  $F$ ) soumis à la gravité et des forces de friction. Si on définit l'énergie totale du système par

$$\mathcal{E}(t) = \underbrace{F(x(t)) - F^*}_{\text{énergie potentielle}} + \underbrace{\frac{1}{2} \|\dot{x}(t)\|^2}_{\text{énergie cinétique}}, \quad (2.3.1)$$

alors on peut facilement montrer que  $\dot{\mathcal{E}}(t) = -\alpha \|\dot{x}(t)\|^2 \leq 0$ , ce qui signifie qu'il s'agit d'un système dissipatif. Dans [4], Alvarez montre que si  $F$  est bornée inférieurement et que son ensemble de minimiseurs est non vide, alors les trajectoires solutions de (HBF) convergent faiblement vers un minimiseur de  $F$ . Ce résultat est dans la continuité de [71] où Haraux et Jendoubi analysent un système de second ordre plus général.

La méthode de Heavy-Ball de Polyak peut être vue comme une version discrète de (HBF) mais il est important de noter que le pas proposé  $s = \frac{4}{(\sqrt{L} + \sqrt{\mu})^2}$  est grand, ce qui peut faire diverger l'algorithme dès que les hypothèses demandées ne sont pas vérifiées. Par la suite, Nesterov [105, 106], Siegel [126] et Aujol, Dossal et Ronderio [25, 24] proposent des schémas issus de (HBF) pour diverses géométries de  $F$  avec des conditions sur le pas  $s$  plus restrictives assurant une meilleure stabilité. La propriété de minimisation de la fonction  $F$  offerte par ces schémas est directement liée à celle des trajectoires de (HBF). Le système dynamique est analysé en profondeur dans le papier de Attouch, Goudou et Redont [19]. Par la suite, le système de Heavy-Ball avec friction a été étudié sous différentes hypothèses sur  $F$  par Bégout,

Bolte et Jendoubi [36], Siegel [126], Aujol, Dossal et Rondepierre [25, 24] ou Apidopoulos, Ginatta et Villa [8].

De la même manière, on peut associer la méthode de descente de gradient (DG) à l'équation de **flot de gradient** (ou *gradient flow* en anglais) qui est définie par

$$\forall t \geq t_0, \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{FG})$$

avec  $t_0 \geq 0$  et  $x(t_0) = x_0$ . En effet, pour  $s > 0$  suffisamment petit, on peut écrire que la suite  $(x_k)_{k \in \mathbb{N}}$  qui vérifie

$$\forall k \in \mathbb{N}^*, \frac{x_k - x_{k-1}}{s} + \nabla F(x_{k-1}) = 0 \quad (2.3.2)$$

est une discrétisation de la trajectoire solution de (FG) avec  $x(t_0) = x_0$ . Or la suite  $(x_k)_{k \in \mathbb{N}}$  ainsi définie n'est autre que la suite des itérés de la descente de gradient pour un pas  $s > 0$ . Contrairement au système (HBF), (FG) est un système de premier ordre dont les solutions ne dépendent que du point initial  $x_0$ . La notion de vitesse initiale n'entre pas en jeu et il n'y a pas d'effet d'inertie.

En 2014, Su, Boyd et Candès proposent dans [127] une EDO associée à la méthode de gradient accéléré de Nesterov. En particulier, les auteurs montrent que le schéma défini par :

$$\forall k \in \mathbb{N}, \begin{cases} x_{k+1} = y_k - s \nabla F(y_k) \\ y_{k+1} = x_{k+1} + \frac{k}{k + \alpha} (x_{k+1} - x_k), \end{cases} \quad (2.3.3)$$

avec  $\alpha \geq 3$ , est une discrétisation du système dynamique avec **friction asymptotiquement évanescence** (ou *asymptotic vanishing damping* en anglais) :

$$\forall t \geq t_0, \ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{AVD})$$

avec  $t_0 \geq 0$  et  $(x(t_0), \dot{x}(t_0)) = (x_0, \dot{x}_0)$ . Ici, le paramètre  $\alpha$  a le même rôle dans (AVD) et (2.3.3).

Ce système est proche de (HBF) et ne diffère que par le coefficient auquel est multiplié  $\dot{x}(t)$ . Ce coefficient détermine le type de force de friction (ou de viscosité) qui s'applique au système. Dans le système de Heavy-Ball avec friction (HBF), ce coefficient vaut  $\alpha > 0$  et en conséquence, les forces de friction appliquées sont constantes en fonction du temps. Au contraire, le coefficient correspondant dans (AVD) vaut  $\frac{\alpha}{t}$ , impliquant que les forces de friction sont réduites jusqu'à disparaître au cours du temps. Ces discussions sont transférables aux schémas correspondants. L'EDO (AVD) a notamment été étudiée par Attouch, Chbani, Peypouquet et Redont dans [16] (convergence faible des trajectoires, propriété de minimisation, forte convexité...), May dans [96] (convergence faible des trajectoires), Attouch, Chbani et Riahi dans [17] (cas sous-critique  $\alpha \leq 3$ ), Aujol, Dossal et Rondepierre dans [27, 26] (étude de convergence sous hypothèses de géométrie) et Luo et Xiao [91] (hypothèse



de géométrie). Dans [6], Apidopoulos, Aujol et Dossal obtiennent des résultats de convergence pour les trajectoires de l'inclusion différentielle :

$$\forall t \geq t_0, \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \partial F(x(t)) \ni 0, \quad (2.3.4)$$

pour  $t_0 \geq 0$  et  $\alpha \leq 3$ , ce qui est un cadre plus général que le système (AVD). On peut noter que Muelebach et Jordan proposent dans [101] un système dynamique alternatif pour étudier l'accélération de Nesterov.

Ces systèmes dynamiques ne sont pas les seuls dignes d'intérêt. Cabot, Engler et Gadat dans [46, 47] et Attouch, Cabot, Chbani et Riahi dans [11, 13] étudient des systèmes avec un terme inertiel plus général :

$$\forall t \geq t_0, \ddot{x}(t) + \gamma(t)\dot{x}(t) + \nabla F(x(t)) = 0, \quad (2.3.5)$$

pour  $t_0 > 0$ . Cabot et Frankel étudient également un système proche dans [48]. L'étude d'un coefficient de friction de la forme  $\frac{\alpha}{t^\theta}$  est proposée par Balti et May [29] et Jendoubi et May [76].

Plusieurs travaux sont consacrés à l'étude des effets de termes perturbatifs sur le comportement des trajectoires de différents systèmes. Cela consiste en général à étudier le système original en remplaçant le terme de droite (égal à 0 initialement) par une perturbation  $g(t)$ . Le but de tels travaux est de déterminer des conditions sur ce terme  $g$  pour retrouver les propriétés de convergence du cas non perturbé. Des études dans ce cadre ont été proposées dans [16, 17, 123, 11, 29, 18].

De nombreux travaux ont été menés sur l'analyse de systèmes faisant intervenir des forces de friction liées à la hessienne de  $F$ . Un état de l'art de cette partie de la littérature est fourni en Section 4.1.

Dans [125], Shi, Du, Jordan et Su proposent des EDO dites à haute-résolution qui ont pour but de correspondre de manière plus fidèle aux schémas numériques que les EDO originales. En effet, plusieurs schémas différents peuvent être associés au même système dynamique, par exemple la méthode Heavy-Ball de Polyak et le schéma de gradient accéléré de Nesterov en fortement convexe sont deux discrétisations de (HBF). Cela n'est plus le cas lorsque l'on considère des EDO haute-résolution. Pour définir des systèmes dynamiques qui correspondent plus finement à chaque schéma, les auteurs proposent d'introduire des termes de correction pondérés par un paramètre  $\sqrt{s}$  qui est pris suffisamment petit. Le choix  $s = 0$  correspond aux systèmes classiques. Par exemple, l'EDO haute-résolution associée au schéma de gradient accéléré de Nesterov pour les fonctions convexes est :

$$\ddot{x}(t) + \frac{3}{t}\dot{x}(t) + \sqrt{s}H_F(x(t))\dot{x}(t) + \left(1 + \frac{3\sqrt{s}}{2t}\right)\nabla F(x(t)) = 0, \quad (2.3.6)$$

où  $H_F$  désigne la hessienne de  $F$ . On peut considérer que cette approche est un pont entre le monde continu et le monde discret.

### 2.3.2 Apports du parallèle continu/discret

Bien que le parallèle entre schéma numérique et EDO soit naturel, on peut légitimement se poser la question : qu'est-ce que cela apporte ?

Une première réponse est que ce lien entre algorithme et système dynamique permet d'avoir une bonne **intuition physique** du comportement d'un schéma.

Prenons l'exemple du système de Heavy-Ball avec friction (HBF) avec une fonction  $F$  non convexe que l'on trace sur la Figure 2.6. La fonction  $F$  possède deux minima locaux  $x_0^*$  et  $x_1^*$  qui sont proches de la donnée initiale  $x_0$ . Si on prend en considération ce que modélise physiquement (HBF) (une boule lourde se déplaçant dans le potentiel  $F$ ), on peut déduire que la trajectoire  $x(t)$  et sa limite quand  $t$  tend vers  $+\infty$  vont dépendre de la vitesse initiale  $\dot{x}_0$  et du paramètre de friction  $\alpha$ . Supposons que  $\dot{x}_0 = 0$ . Si le paramètre  $\alpha$  est très faible, la boule sera peu freinée par des forces de friction et pourra éventuellement rejoindre le voisinage de  $x_0^*$  et converger vers ce point en oscillant. Si  $\alpha$  est grand, alors la boule sera fortement ralentie par la friction et risque de rester bloquée dans un voisinage de  $x_1^*$  avant de converger vers ce point. Si l'on applique un schéma dérivé de ce système avec cette même donnée initiale, on aura le même type de comportement. En somme, ce parallèle avec le système dynamique (HBF) permet notamment de comprendre pourquoi les schémas dérivés de cette EDO (comme la méthode de Heavy-Ball de Polyak) ne sont pas des schémas de descente, pourquoi des oscillations peuvent apparaître ou encore quel est le rôle du paramètre de friction.

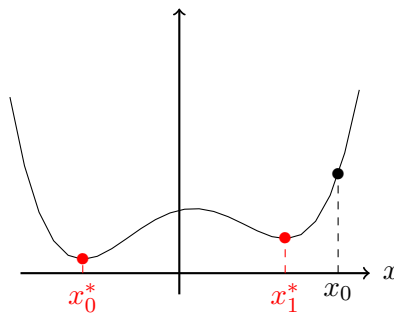


FIGURE 2.6 – Exemple d'une fonction  $F$  non convexe ayant 2 minima locaux.

Une deuxième raison pour laquelle on étudie les systèmes dynamiques associés à des schémas numériques est que **les calculs de convergence de trajectoire sont beaucoup plus abordables dans le cas continu**. En effet, l'analyse en continu nous permet de différentier par rapport à  $t$  et d'éviter les termes de discrétisation. Comme les propriétés de convergence de trajectoire sont en général extensibles aux schémas numériques dérivés, l'analyse en continu constitue ensuite une feuille de route pour l'analyse en discret.

Enfin, ce parallèle peut permettre de **définir de nouveaux schémas**. Si les propriétés de convergence des trajectoires d'un système dynamique semblent inté-

ressantes à exploiter, alors il est pertinent de dériver des schémas numériques à partir de ce système dans le but d'exploiter ces mêmes propriétés. Récemment, l'étude de systèmes dynamiques impliquant la hessienne de  $F$  a été suivie par l'introduction de nouveaux schémas [15, 14] (voir Chapitre 4). De même, des systèmes faisant intervenir un terme de régularisation de Tikhonov [10, 43, 97] ont été analysés, ce terme additionnel assurant de nouvelles propriétés comme la convergence forte des trajectoires.

### 2.3.3 Analyse de Lyapunov

#### 2.3.3.1 Cadre continu

Pour étudier la convergence des trajectoires d'un système dynamique, on utilisera souvent une analyse de Lyapunov. En général, l'idée est de définir une fonctionnelle  $\mathcal{E}$  positive qui décroît le long des trajectoires  $x(t)$  quand  $t$  augmente. On appellera  $\mathcal{E}$  une énergie de Lyapunov. On cherchera ensuite à extraire la propriété de convergence qui nous intéresse à partir de la décroissance de cette énergie : décroissance de  $F(x(t)) - F^*$ , intégrabilité du gradient le long des trajectoires, longueur de la trajectoire... La complexité d'une telle démarche réside dans la définition de la bonne énergie de Lyapunov pour arriver à la conclusion voulue.

Un exemple simple de ce type d'approche est donné ci-dessous.

*Exemple 2.3.1.* Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe différentiable ayant un ensemble de minimiseurs  $X^*$  non vide. On considère les trajectoires du système de flot de gradient, c'est-à-dire définies par :

$$\forall t \geq t_0, \dot{x}(t) + \nabla F(x(t)) = 0, \quad (2.3.7)$$

avec  $t_0 \geq 0$  et  $x(t_0) = x_0$ .

L'énergie de Lyapunov la plus simple dans ce contexte est  $\mathcal{E}(t) = F(x(t))$ . On peut montrer que cette énergie est décroissante en fonction du temps :

$$\forall t \geq t_0, \dot{\mathcal{E}}(t) = \langle \nabla F(x(t)), \dot{x}(t) \rangle = -\|\nabla F(x(t))\|^2 \leq 0. \quad (2.3.8)$$

Ceci permet de conclure que pour tout  $t_2 \geq t_1 \geq t_0$ ,  $\mathcal{E}(t_2) \leq \mathcal{E}(t_1) \leq \mathcal{E}(t_0)$  et donc que  $F(x(t_2)) \leq F(x(t_1)) \leq F(x_0)$ . Cela nous permet de prouver que la valeur de la fonctionnelle décroît le long de la trajectoire  $x$ .

Cependant, cette analyse ne nous donne pas d'information sur la vitesse de décroissance de la fonctionnelle. Pour obtenir un tel résultat, on peut définir une nouvelle énergie  $\mathcal{E}(t) = t(F(x(t)) - F^*) + \frac{1}{2}\|x(t) - x^*\|^2$  où  $F^* = \min_{x \in \mathbb{R}^N} F(x)$  et  $x^*$  est un minimiseur de  $F$ . Si on dérive  $\mathcal{E}$  par rapport au temps, on a :

$$\begin{aligned} \forall t \geq t_0, \dot{\mathcal{E}}(t) &= t\langle \nabla F(x(t)), \dot{x}(t) \rangle + F(x(t)) - F^* + \langle \dot{x}(t), x(t) - x^* \rangle \\ &= -t\|\nabla F(x(t))\|^2 + F(x(t)) - F^* - \langle \nabla F(x(t)), x(t) - x^* \rangle. \end{aligned} \quad (2.3.9)$$

Comme  $F$  est convexe, on a  $F(x(t)) \leq F^* + \langle \nabla F(x(t)), x(t) - x^* \rangle$  et donc  $\dot{\mathcal{E}}(t) \leq -t\|\nabla F(x(t))\|^2 \leq 0$ . On obtient donc que  $\mathcal{E}(t) \leq \mathcal{E}(t_0)$  pour tout  $t \geq t_0$ , ce qui

implique que :

$$\forall t \geq t_0, F(x(t)) - F^* \leq \frac{\mathcal{E}(t_0)}{t}. \quad (2.3.10)$$

Cet exemple simple illustre l'importance de l'énergie étudiée dans l'obtention de résultats de convergence : dans le premier cas, on démontre une propriété de descente des trajectoires quand dans le deuxième cas, on obtient une vitesse de convergence de l'erreur.

*Remarque 2.3.2.* La définition d'une énergie de Lyapunov pertinente varie en fonction des systèmes dynamiques considérés. Dans la littérature [126, 123, 25, 24], pour étudier le système (HBF), l'énergie de Lyapunov considérée pour déterminer des vitesses de convergence sera généralement de la forme suivante :

$$\mathcal{E}(t) = F(x(t)) - F^* + \frac{1}{2} \|\lambda(x(t) - x^*) + \dot{x}(t)\|^2 + \frac{\xi}{2} \|x(t) - x^*\|^2, \quad (2.3.11)$$

où  $\lambda > 0$  et  $\xi \in \mathbb{R}$  sont des paramètres à ajuster en fonction des hypothèses faites sur la géométrie de  $F$ . Pour analyser (AVD), on utilisera plutôt des énergies de la forme :

$$\mathcal{E}(t) = t^2(F(x(t)) - F^*) + \frac{1}{2} \|\lambda(x(t) - x^*) + t\dot{x}(t)\|^2 + \frac{\xi}{2} \|x(t) - x^*\|^2, \quad (2.3.12)$$

avec  $\lambda > 0$  et  $\xi \in \mathbb{R}$  (voir [127, 16, 11, 6, 27, 26]).

### 2.3.3.2 Analyse de schémas numériques

On a vu ci-dessus que l'analyse de Lyapunov permet d'étudier les propriétés de convergence des trajectoires de systèmes dynamiques. Cependant, l'objectif est souvent de démontrer des propriétés de convergence sur des schémas numériques. En ce sens, obtenir des résultats en continu permet d'avoir une idée du comportement des schémas numériques associés mais on cherchera en général à retrouver ces résultats dans le cadre discret.

Quand elle fait suite à une analyse dans le cadre continu, l'étude de la convergence d'un schéma numérique suit les mêmes principes : on construit une énergie de Lyapunov discrète qui décroît (ou que l'on peut borner) au fil des itérations, puis on extrait de cette énergie la propriété de convergence voulue. L'analyse préalable du système dynamique correspondant est une grande aide puisqu'elle donne une direction : si une énergie de Lyapunov donne accès à des propriétés de convergence sur les trajectoires d'un système dynamique, alors il existe certainement une discrétisation de cette énergie qui permet d'obtenir des résultats analogues en discret (sous certaines conditions notamment sur le pas).

La difficulté est qu'il existe de nombreuses discrétisations possibles d'une même énergie et aucune méthode "automatique" n'existe pour trouver la bonne énergie

discrète :

$$F(x(t)) - F^* \rightarrow \begin{cases} F(x_k) - F^*? \\ F(x_{k+1}) - F^*? \\ \theta F(x_k) + (1 - \theta)F(x_{k+1}) - F^*? \end{cases}$$

$$\dot{x}(t) \rightarrow \begin{cases} x_k - x_{k-1}? \\ x_{k+1} - x_k? \\ \frac{x_{k+1} - x_{k-1}}{2}? \end{cases}$$

De plus, les calculs sont plus fastidieux à mener dans le cadre discret puisque la dérivée n'est plus accessible et certains termes de discrétisation doivent être gérés. Ceci est souligné dans l'exemple suivant.

*Exemple 2.3.3.* Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction convexe différentiable, à gradient  $L$ -Lipschitz et ayant un ensemble de minimiseurs  $X^*$  non vide. On considère  $(x_k)_{k \in \mathbb{N}}$  la suite définie par la méthode de descente de gradient :

$$\forall k \in \mathbb{N}^*, x_k = x_{k-1} - s \nabla F(x_{k-1}), \quad (2.3.13)$$

avec un pas  $s > 0$ . Comme on l'a vu précédemment, la suite  $(x_k)_{k \in \mathbb{N}}$  est une discrétisation de la trajectoire du système de flot de gradient. On peut donc étendre l'analyse proposée dans l'exemple 2.3.1.

Pour montrer que la descente de gradient est un schéma de descente, on peut étudier l'énergie discrète définie par  $\mathcal{E}_k = F(x_k)$  pour tout  $k \in \mathbb{N}$ . Ainsi, on a :

$$\mathcal{E}_{k+1} - \mathcal{E}_k = F(x_{k+1}) - F(x_k) \leq \langle \nabla F(x_k), x_{k+1} - x_k \rangle + \frac{L}{2} \|x_{k+1} - x_k\|^2, \quad (2.3.14)$$

en utilisant le caractère  $L$ -Lipschitz du gradient de  $F$ . Sachant que la suite  $(x_k)_{k \in \mathbb{N}}$  vérifie (2.3.13), on a  $x_{k+1} - x_k = -s \nabla F(x_k)$  et donc :

$$F(x_{k+1}) - F(x_k) \leq s^2 \left( \frac{L}{2} - \frac{1}{s} \right) \|\nabla F(x_k)\|^2. \quad (2.3.15)$$

On peut donc conclure que si  $s \leq \frac{2}{L}$ , alors  $F(x_{k+1}) \leq F(x_k)$  pour tout  $k \in \mathbb{N}$ . On trouve un résultat analogue à ce qui a été fait en continu mais cela repose sur le caractère  $L$ -Lipschitz de  $\nabla F$  (qui n'était pas nécessaire en continu) et une condition sur le pas doit être vérifiée.

Pour obtenir une vitesse de décroissance de la fonctionnelle, on peut définir une énergie discrète proche de celle donnée en continu :

$$\forall k \in \mathbb{N}, \mathcal{E}_k = sk (F(x_k) - F^*) + \frac{1}{2} \|x_k - x^*\|^2. \quad (2.3.16)$$

On a ainsi pour tout  $k \in \mathbb{N}$  :

$$\mathcal{E}_{k+1} - \mathcal{E}_k = s(k+1) (F(x_{k+1}) - F(x_k)) + s (F(x_k) - F^*) + \frac{1}{2} \|x_{k+1} - x^*\|^2 - \frac{1}{2} \|x_k - x^*\|^2, \quad (2.3.17)$$

et comme  $\|x_{k+1} - x^*\|^2 = \|x_k - x^*\|^2 + 2\langle x_k - x^*, x_{k+1} - x_k \rangle + \|x_{k+1} - x_k\|^2$ , alors :

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &= s(k+1)(F(x_{k+1}) - F(x_k)) + s(F(x_k) - F^*) \\ &\quad + \langle x_k - x^*, x_{k+1} - x_k \rangle + \frac{1}{2}\|x_{k+1} - x_k\|^2, \end{aligned} \quad (2.3.18)$$

et donc

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &= s(k+1)(F(x_{k+1}) - F(x_k)) + s(F(x_k) - F^*) \\ &\quad - s\langle x_k - x^*, \nabla F(x_k) \rangle + \frac{s^2}{2}\|\nabla F(x_k)\|^2. \end{aligned} \quad (2.3.19)$$

En utilisant la convexité de  $F$  et le résultat démontré précédemment, on a que :

$$\begin{aligned} \forall k \in \mathbb{N}, \mathcal{E}_{k+1} - \mathcal{E}_k &\leq s^2(k+1) \left( \frac{Ls}{2} - 1 \right) \|\nabla F(x_k)\|^2 + \frac{s^2}{2} \|\nabla F(x_k)\|^2 \\ &\leq s^3 \left( k \left( \frac{L}{2} - \frac{1}{s} \right) + \frac{1}{2} \left( L - \frac{1}{s} \right) \right) \|\nabla F(x_k)\|^2. \end{aligned} \quad (2.3.20)$$

On trouve donc que pour  $s \leq \frac{1}{L}$ , la suite  $(\mathcal{E}_k)_{k \in \mathbb{N}}$  est décroissante. Ainsi, comme pour tout  $k \in \mathbb{N}^*$  on a  $\mathcal{E}_k \leq \mathcal{E}_0$ , alors :

$$\forall k \in \mathbb{N}^*, F(x_k) - F^* \leq \frac{\|x_0 - x^*\|^2}{2sk}. \quad (2.3.21)$$

Les résultats démontrés dans cet exemple sont bien connus et peuvent être facilement retrouvés sans passer par le système dynamique correspondant. On peut cependant remarquer que malgré une démarche identique à ce qui est fait en continu, la démonstration de ces résultats sur le schéma numérique est plus technique. La complexité des calculs augmentant avec la complexité des schémas et hypothèses, le recours à l'analyse en continu est inévitable dans certains cas.



# Stratégies de restart

Ce chapitre est consacré aux méthodes de restart de schémas inertiels et plus particulièrement aux stratégies de restart de FISTA. Ces schémas sont particulièrement pertinents lors de la minimisation de fonctions à croissance quadratique (voire fortement convexes) et visent à tirer parti des propriétés d'accélération de FISTA permises par l'inertie tout en évitant certains inconvénients liés aux oscillations. Pour cette classe de fonctions, les méthodes inertielles les plus efficaces nécessitent d'estimer précisément le paramètre de croissance  $\mu$  de la fonctionnelle à minimiser pour offrir des vitesses de convergence rapide. Cependant, l'estimation du paramètre de croissance est généralement difficile ce qui complique la paramétrisation de ce type de méthodes.

Face à ce problème, j'introduis un schéma de restart (Algorithme 3) qui ne requiert **aucune information sur le paramètre de croissance** et qui assure une **décroissance exponentielle rapide de l'erreur asymptotiquement** (Corollaire 3.2.9). Le schéma s'appuie sur l'**estimation automatique du paramètre de croissance** de la fonction d'intérêt pour adapter son comportement. Il a fait l'objet de l'écriture d'un article actuellement en révision mineure [28].

L'Algorithme 3 s'appuie sur la connaissance de la constante de Lipschitz  $L$  du gradient de la fonction  $F$  (ou de la partie différentiable de  $F$ ) comme la plupart des méthodes classiques. Il arrive toutefois que cette constante soit **mal estimée voire inaccessible**. Cela a motivé l'élaboration de **Free-FISTA** (Algorithme 6), une méthode combinant une stratégie de restart avec du **backtracking** sur la constante de Lipschitz  $L$ . Cette méthode permet donc de s'affranchir de la connaissance de  $L$  et  $\mu$  et garantit une **vitesse de convergence rapide** (Corollaire 3.3.8). Ce travail a été mené avec mes directeurs de thèse ainsi que Luca Calatroni et a fait l'objet de l'écriture d'un article [22].

L'annexe contient des explorations sur le thème du restart et des preuves techniques.

## Contents

|            |  |           |
|------------|--|-----------|
| <b>3.1</b> | <b>Présentation du restart</b> . . . . .   | <b>36</b> |
| 3.1.1      | Contexte . . . . .   | 36        |
| 3.1.2      | Techniques de restart . . . . .  | 38        |
| <b>3.2</b> | <b>Restart via une estimation automatique du paramètre de croissance</b> . . . . . | <b>43</b> |
| 3.2.1      | Présentation de la stratégie . . . . .   | 43        |
| 3.2.2      | Garanties théoriques du schéma . . . . .   | 48        |
| 3.2.3      | Expériences numériques . . . . .   | 50        |



---

|            |  |           |
|------------|--|-----------|
| <b>3.3</b> | <b>Un schéma de restart avec backtracking : Free-FISTA . . .</b>                       | <b>54</b> |
| 3.3.1      | Contexte . . . . .   | 54        |
| 3.3.2      | Le backtracking appliqué à FISTA . . . . .   | 55        |
| 3.3.3      | Free-FISTA . . . . .   | 60        |
| 3.3.4      | Expériences numériques . . . . .   | 66        |
| <b>3.A</b> | <b>Résultats annexes et preuves . . . . .</b>  | <b>75</b> |
| 3.A.1      | Étude de l’Algorithme 3 avec une mise à jour des itérations<br>plus générale . . . . . | 75        |
| 3.A.2      | Variations de l’Algorithme 3 . . . . .   | 77        |
| 3.A.3      | Preuves techniques . . . . .   | 80        |

---

## 3.1 Présentation du restart

### 3.1.1 Contexte

Lorsque l’on considère un problème d’optimisation composite convexe, les méthodes de premier ordre offrant les meilleures vitesses de convergence sont très souvent des méthodes inertielles. En prenant en compte la direction de descente précédente, ce type de schéma génère un élan qui, en principe, accélère la convergence de l’algorithme. Si cette inertie est globalement bénéfique, elle peut aussi déclencher un comportement oscillatoire de la suite d’itérés. Comme cela est précisé dans la Section 2.3.1, les schémas inertiels usuels peuvent être vus comme la discrétisation de la trajectoire d’un corps lourd dans un champ de potentiel plus ou moins visqueux (où les forces de friction n’ont pas le même comportement en fonction du schéma). En ce sens, il semble naturel de voir apparaître des oscillations autour du point le plus bas.

Pour la minimisation d’une fonction  $F$  convexe, différentiable à gradient  $L$ -Lipschitz pour un certain  $L > 0$ , on sait que la meilleure vitesse asymptotique de décroissance de la fonctionnelle est de l’ordre de  $k^{-2}$  si aucune hypothèse additionnelle n’est faite [105]. Cette vitesse est atteinte par FISTA [32] introduit par Beck et Teboulle sur le modèle du schéma de gradient accéléré de Nesterov [103] pour les fonctions composites vérifiant  $\mathcal{H}_L$ . On rappelle que ce schéma peut s’écrire comme suit :

$$\forall k \in \mathbb{N}, \begin{cases} y_k = x_k + \frac{k}{k+\alpha}(x_k - x_{k-1}), \\ x_{k+1} = \text{prox}_{sh}(y_k - s\nabla f(y_k)), \end{cases} \quad (3.1.1)$$

avec  $x_0 \in \mathbb{R}^N$  et où  $\alpha = 3$  approxime le choix original de Nesterov. Dans ce schéma, l’inertie prend de plus en plus d’importance au cours des itérations puisque  $\frac{k}{k+\alpha}$  est croissant. Ce schéma est très efficace dans le cas convexe puisqu’il permet notamment de s’adapter aux fonctions plates en exploitant de plus en plus l’inertie.

En revanche, quand on s’intéresse à une fonction  $F$  qui vérifie des hypothèses plus fortes comme  $\mathcal{G}_\mu^2$  (croissance quadratique) ou  $\mathcal{S}_\mu$  (forte convexité) pour  $\mu > 0$ ,

FISTA n'est plus la meilleure solution. Pour ces classes de fonctions qui ont un comportement proche de  $x \mapsto \frac{\mu}{2}\|x\|^2$ , les schémas les plus efficaces sont ceux de type Heavy-Ball (c'est-à-dire associés au système de Heavy-Ball avec friction (HBF)). Ces méthodes s'appuient sur des forces de friction constantes dans le temps, calibrées en fonction des paramètres  $L$  et  $\mu$  de la fonction  $F$  à minimiser. La quantité qui joue un grand rôle pour extraire la meilleure performance de telles méthodes est  $\kappa := \frac{\mu}{L}$ , qui est en fait l'inverse du conditionnement de la fonction  $F$  (dans certains travaux,  $\kappa$  désigne le conditionnement mais on gardera la même notation tout le long de ce manuscrit). Les méthodes Heavy-Ball sont donc théoriquement les plus rapides **à condition d'avoir accès à la valeur de  $\kappa$**  ce qui est rarement le cas en pratique. Les vitesses que ces schémas permettent d'obtenir sont de la forme  $\mathcal{O}\left(e^{-K\sqrt{\frac{\mu}{L}}k}\right)$  où  $K > 0$ . C'est ce que l'on appellera par la suite une **décroissance exponentielle rapide** en opposition à une **décroissance exponentielle lente** qui s'écrit sous la forme  $\mathcal{O}\left(e^{-K\frac{\mu}{L}k}\right)$  avec  $K > 0$ . Comme on s'intéresse généralement à des problèmes mal conditionnés, on a que  $\kappa \ll 1$  et par conséquent, la présence de la racine carrée au-dessus de  $\kappa$  assure une décroissance nettement plus rapide. Notons que sous les hypothèses  $\mathcal{G}_\mu^2$  et  $\mathcal{S}_\mu$ , la méthode du gradient proximal assure une décroissance exponentielle lente [64].

Dans le même cadre, FISTA avec le choix original  $\alpha = 3$  ne permet pas d'atteindre théoriquement ces vitesses et conserve son taux polynomial de  $\mathcal{O}(k^{-2})$ . En pratique, appliquer cette méthode à des fonctions vérifiant  $\mathcal{G}_\mu^2$  ou  $\mathcal{S}_\mu$  donne lieu à des **oscillations** liées à l'inertie qui devient trop importante au fil des itérations. Cela est illustré sur la Figure 3.1. Notons tout de même que FISTA a l'avantage de ne pas dépendre de la connaissance de  $\mu$  contrairement aux méthodes type Heavy-Ball. Dans [26], Aujol, Dossal et Rondepierre montrent qu'à nombre d'itérations fini, il est possible d'obtenir des vitesses comparables à ce qui est garanti par les méthodes Heavy-Ball en calibrant le paramètre  $\alpha$  par rapport à la précision voulue (à condition d'avoir accès à une estimation correcte de  $F(x_0) - F^*$ ).

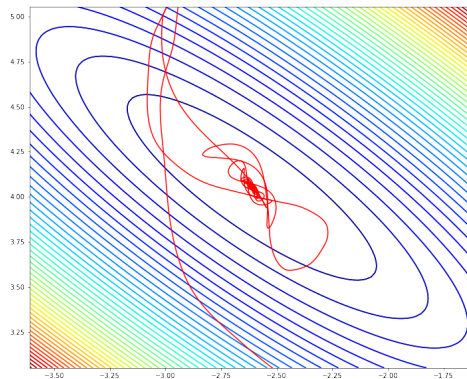


FIGURE 3.1 – Projection de la trajectoire des itérés de FISTA pour la résolution d'un problème de moindres carrés (fortement convexe,  $N = 20$ ). Le point initial se situe en dehors de la fenêtre d'affichage (en bas à droite).

### 3.1.2 Techniques de restart

On se place dans le cadre où  $F$  vérifie  $\mathcal{H}_L$  et une condition de croissance quadratique globale  $\mathcal{G}_\mu^2$ , c'est-à-dire qu'il existe  $\mu > 0$  tel que pour tout  $x \in \mathbb{R}^N$  :

$$\frac{\mu}{2}d(x, X^*)^2 \leq F(x) - F^*. \quad (3.1.2)$$

Dans cette section, l'appellation FISTA désigne le schéma (3.1.1) pour le choix  $\alpha = 3$ .

Un moyen pour éviter les oscillations et accélérer la convergence de FISTA est de fixer le coefficient d'inertie à zéro au moment opportun : les méthodes suivant ce principe sont appelées des méthodes de **redémarrage de FISTA** ou de **restart de FISTA**. Cette stratégie est un moyen simple d'atténuer les effets néfastes de l'inertie mais soulève la question : comment décider quand redémarrer ? Plusieurs solutions existent et on pourra généralement réécrire les schémas de restart sous la forme de l'Algorithme 1 où la particularité de chaque schéma est la condition de restart.

---

**Algorithme 1** : Schéma de restart de FISTA générique

---

**Entrée:**  $x_0 \in \mathbb{R}^N$ ,  $y_0 = x_0$ ,  $k = 0$ ,  $i = 0$ .

**répéter**

$k = k + 1, i = i + 1$

$x_k = \text{prox}_{sh}(y_{k-1} - s\nabla f(y_{k-1}))$

**si** La condition de restart est vérifiée **alors**

$i = 1$

**fin si**

$y_k = x_k + \frac{i-1}{i+2}(x_k - x_{k-1})$

**jusqu'à** La condition de sortie est vérifiée

**retourne**  $x_k$

---

À noter que l'on doit une des premières apparitions de la notion de restart à Powell en 1977 dans [115] pour accélérer l'algorithme de gradient conjugué.

#### 3.1.2.1 Restart heuristique

Une façon intuitive de décider du moment auquel redémarrer FISTA est de **détecter l'apparition d'oscillations**. C'est ce qui est proposé par O'Donoghue et Candès [108] via deux conditions de restart heuristiques. La première est basée sur la valeur de la fonctionnelle à chaque itération :

$$F(x_k) > F(x_{k-1}). \quad (3.1.3)$$

En utilisant cette règle pour redémarrer FISTA, on coupe l'inertie dès que la valeur de la fonctionnelle augmente entre deux itérés successifs. Cette idée peut se rapprocher de ce que proposent Beck et Teboulle [33] avec Monotone FISTA ou la variation

introduite par Zibetti, Helou, Regatte et Herman [140]. La deuxième condition est basée sur la valeur du gradient de  $F$  quand  $F$  est différentiable mais elle peut être adaptée au cas composite avec l'utilisation du gradient composite. Elle est introduite sous la forme suivante pour  $F$  différentiable :

$$\langle \nabla F(y_{k-1}), x_k - x_{k-1} \rangle > 0. \quad (3.1.4)$$

mais on peut la réécrire de manière plus générale :

$$\langle x_k - y_{k-1}, x_k - x_{k-1} \rangle < 0. \quad (3.1.5)$$

Ici, on vérifie si la direction de descente  $x_k - x_{k-1}$  va dans la même direction que le gradient (composite dans le cas non-différentiable) de  $F$  évalué en  $y_{k-1}$ . Si c'est le cas, cela peut signifier que l'on a rebroussé chemin par rapport au point décalé  $y_{k-1}$  lors de l'étape de descente de gradient (ou gradient proximal). On peut en déduire que l'inertie devient trop importante et donc qu'il vaut mieux la remettre à zéro.

Ces schémas de restart sont très performants en pratique et faciles à mettre en place. Ils ont été aussi étudiés par Giselsson et Boyd [66] et sont une solution assez répandue pour accélérer des méthodes inertielles [79, 84].

Cependant, il n'existe pas de garantie théorique de l'amélioration de la vitesse de convergence par rapport à FISTA.

### 3.1.2.2 Restart périodique

On peut également décider de **redémarrer FISTA périodiquement** à partir d'un nombre d'itérations fixé en avance. Il est alors relativement simple d'obtenir des vitesses de convergence améliorées. L'idée est de s'appuyer sur l'inégalité suivante vérifiée par les itérés de FISTA :

$$\forall k \in \mathbb{N}^*, F(x_k) - F^* \leq \frac{2}{s(k+1)^2} \|x_0 - x^*\|^2, \quad (3.1.6)$$

pour  $s \leq \frac{1}{L}$  et pour tout minimiseur  $x^*$  [51, Theorem 3.1]. Alors, si  $F$  vérifie  $\mathcal{G}_\mu^2$  pour  $\mu > 0$  et  $s = \frac{1}{L}$ , on a pour tout  $k \in \mathbb{N}^*$ ,

$$F(x_k) - F^* \leq \frac{2L}{(k+1)^2} d(x_0, X^*)^2 \leq \frac{4L}{\mu(k+1)^2} (F(x_0) - F^*). \quad (3.1.7)$$

On peut ainsi montrer que si  $k \geq k^* := \lfloor \frac{2}{\sqrt{\gamma}} \sqrt{\frac{L}{\mu}} \rfloor$  pour  $\gamma > 0$ , alors

$$F(x_k) - F^* \leq \gamma (F(x_0) - F^*). \quad (3.1.8)$$

En conséquence, si un redémarrage est effectué toutes les  $k^*$  itérations, alors si  $k$  est un multiple de  $k^*$ , on a :

$$F(x_k) - F^* \leq \gamma^{\frac{k}{k^*}} (F(x_0) - F^*) \leq e^{\frac{\ln(\gamma)\sqrt{\gamma}}{2}\sqrt{\kappa}k} (F(x_0) - F^*). \quad (3.1.9)$$

Quelques calculs supplémentaires permettent d'arriver à la conclusion qu'un restart périodique toutes les  $\lfloor \frac{2}{\sqrt{\gamma}} \sqrt{\frac{L}{\mu}} \rfloor$  itérations assure une vitesse de  $\mathcal{O}\left(e^{\frac{\ln(\gamma)\sqrt{\gamma}}{2}\sqrt{\kappa k}}\right)$ .

En optimisant cette vitesse de convergence par rapport à  $\gamma$ , on obtient finalement que le restart périodique optimal garantit la décroissance exponentielle rapide de l'erreur  $\mathcal{O}\left(e^{-\frac{1}{e}\sqrt{\kappa k}}\right)$  pour un redémarrage toutes les  $\lfloor 2e\sqrt{\frac{L}{\mu}} \rfloor$  itérations. Ce type de schéma de restart a été notamment étudié dans [104, 108, 137, 87, 102] pour FISTA sous différentes hypothèses de croissance mais également pour d'autres schémas comme PARNES [68] ou la méthode de descente de gradient optimisée [79].

En redémarrant périodiquement FISTA, on peut donc atteindre des vitesses comparables à celles des méthodes type Heavy-Ball. Néanmoins, ces vitesses ne peuvent être atteintes qu'à condition de connaître une bonne estimation de  $\kappa$ . Le restart périodique fait donc face aux difficultés suivantes :

- si  $\kappa$  n'est pas accessible, il est alors difficile de fixer une période de restart assurant les vitesses théoriques améliorées à moins de tâtonner. Notons qu'effectuer des restarts périodiquement avec une période fixée à la main peut permettre de significativement accélérer FISTA en pratique (ou des algorithmes proches [35]) mais cela relève du restart heuristique.
- si  $\kappa$  est connu, alors il est plus pertinent d'appliquer des méthodes type Heavy-Ball puisqu'elles peuvent atteindre des vitesses plus rapides que le restart périodique optimal.

Dans [120], Roulet et D'Aspremont introduisent des schémas de restart planifié (*scheduled restart*) qui généralisent le restart périodique. Les auteurs considèrent des hypothèses de géométrie plus générales que  $\mathcal{G}_\mu^2$  ou  $\mathcal{H}_L$ , englobant ainsi des fonctions plus plates. Il est montré que pour des fonctions vérifiant des hypothèses du type  $\mathcal{G}^\gamma$  avec  $\gamma > 2$  (hypothèse de croissance faible), des taux améliorés peuvent être atteints en redémarrant FISTA pour un nombre d'itérations qui croît exponentiellement au cours des restarts. Les meilleures vitesses étant obtenues à condition de connaître les paramètres de géométrie de  $F$ , une stratégie adaptative est proposée, basée sur de la recherche en grille (*grid-search*) de ces paramètres avec des garanties théoriques. Cette approche est également détaillée dans [56] et étendue à l'algorithme de Frank-Wolfe dans [78].

### 3.1.2.3 Restart adaptatif

Cette section regroupe plusieurs schémas de restart qui peuvent avoir des philosophies très différentes. On les différencie ici des schémas heuristiques car ils s'accompagnent de preuves de convergence théoriques.

Su, Boyd et Candès proposent dans [127] une méthode de restart adaptatif basée sur des comparaisons de vitesse au cours des itérations. Les auteurs montrent une vitesse de convergence exponentielle pour les trajectoires du système dynamique correspondant mais ne donnent pas de résultat sur le schéma en tant que tel. Les résultats présentés en continu ne permettent pas de savoir si la décroissance exponentielle de l'erreur est rapide ou lente.

Dans [62, 61], Fercoq et Qu introduisent des schémas de restart adaptatifs pour FISTA (et d'autres méthodes) sous l'hypothèse que  $F$  satisfait  $\mathcal{G}_\mu^2$ . En particulier, le schéma AdaRES proposé dans [61] permet d'assurer une décroissance exponentielle rapide de l'erreur sans avoir besoin d'une estimation exacte de  $\mu$ . À partir d'une estimation  $\mu_0$  de la constante de croissance  $\mu$ , les auteurs donnent des garanties théoriques de la convergence rapide du schéma. En revanche, les taux de convergence dépendent de la qualité de l'estimation de  $\mu_0$  :

- si  $\mu_0 \leq \mu$ , la vitesse de convergence au pire cas est  $\mathcal{O}\left(e^{-\frac{1}{2\sqrt{e}}\sqrt{\frac{\mu_0}{\mu}}\sqrt{\frac{\mu}{L}}k}\right)$ ,
- si  $\mu_0 > \mu$ , la vitesse de convergence au pire cas est  $\mathcal{O}\left(e^{-\frac{\sqrt{2}-1}{2\sqrt{e}(2-\sqrt{\frac{\mu}{\mu_0}})}\sqrt{\frac{\mu}{L}}k}\right)$ .

Un tel schéma est donc intéressant à condition d'avoir moyen d'estimer de manière suffisamment fine le paramètre de croissance.

Alamo, Krupa et Limon proposent plusieurs méthodes pour la minimisation de fonctions qui vérifient  $\mathcal{G}_\mu^2$  et  $\mathcal{H}_L$ . Dans [3], un schéma de restart de FISTA est introduit, se basant sur des évaluations de la valeur de  $F$  à chaque itéré pour définir à quelles itérations redémarrer. L'algorithme ne s'appuie pas sur une estimation préalable de  $\mu$  et garantit une décroissance exponentielle rapide de l'erreur  $\mathcal{O}\left(e^{-\frac{1}{16}\sqrt{\frac{\mu}{L}}k}\right)$ . Notons que le fait d'évaluer  $F$  à chaque itération peut alourdir le temps de calcul. Il existe en général des techniques pour diminuer ce coût mais cela dépend de la fonction à minimiser et une implémentation de la méthode adaptée spécifiquement au problème est nécessaire. Le schéma proposé dans [1] est basé sur la valeur du gradient composite, permettant d'éviter les problèmes liés au coût de calcul de  $F$ . Par contre, le taux de convergence prouvé par les auteurs  $\mathcal{O}\left(e^{-\frac{1}{4e(1+\sqrt{\mu+1})}\frac{\mu}{L}k}\right)$  peut être vu comme exponentiel lent puisque  $e^{-\frac{1}{4e(1+\sqrt{\mu+1})}\frac{\mu}{L}k} > e^{-\frac{1}{8e}\frac{\mu}{L}k}$  pour tout  $k > 0$ . Enfin, Alamo, Krupa et Limon introduisent un schéma de restart adapté à FISTA mais qui peut se généraliser à d'autres algorithmes accélérés [2]. Dans la lignée de la méthode introduite dans [3], la stratégie est basée sur des évaluations de la fonctionnelle à chaque itération et assure une vitesse de  $\mathcal{O}\left(e^{-\frac{\ln(15)}{4e}\sqrt{\frac{\mu}{L}}k}\right)$ .

Dans [117], un schéma de restart pour une variété de méthodes accélérées est proposé avec des extensions incluant du calcul parallèle. Renegar et Grimmer prouvent des vitesses exponentielles rapides pour ce schéma appliqué à FISTA avec une décroissance de l'erreur de  $\mathcal{O}\left(e^{-\frac{1}{2\sqrt{2}}\sqrt{\frac{\mu}{L}}k}\right)$  à condition de connaître  $F^*$ . En pratique, cette hypothèse est forte et cette valeur est très rarement accessible en amont. Des alternatives basées sur du calculs parallèle sont proposées permettant de se passer de l'estimation de  $F^*$ . L'analyse des auteurs s'étend également à l'hypothèse  $\mathcal{G}^\gamma$  pour  $\gamma > 2$ .

### 3.1.2.4 Récapitulatif

Les méthodes de restart présentées dans les sections précédentes sont regroupées dans le Tableau 3.1 pour les fonctions composites qui vérifient une hypothèse de

croissance quadratique. On peut voir en bas du tableau les résultats théoriques garantis par les schémas proposés avec mes collaborateurs qui sont détaillés par la suite.

TABLEAU 3.1 – Résumé des différents schémas de restart de FISTA et de leur taux de convergence pour une fonction  $F$  qui vérifie  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  avec  $L > 0$  connu en amont et  $\mu > 0$ .

| Schéma                            | Référence  | Taux de convergence  | Remarques   |
|-----------------------------------|--|--|---|
| Restart heuristique               | O'Donoghue et Candès [108]<br>Giselsson et Boyd [66] | $\mathcal{O}(k^{-2})$  | Rapide en pratique<br>Pas de garantie théorique d'une convergence améliorée |
| Restart périodique optimal        | [104, 108, 137, 87, 102]                             | $\mathcal{O}\left(e^{-\frac{1}{e}\sqrt{\frac{\mu}{L}}k}\right)$  | Nécessité de connaître $\mu$ en amont.                                      |
| <i>Speed restart scheme</i>       | Su, Boyd et Candès [127]                             | $\mathcal{O}(k^{-2})$  | Preuve de convergence améliorée en continu mais pas en discret.             |
| AdaRES                            | Fercoq et Qu [61]                                    | $\mathcal{O}\left(e^{-\frac{1}{2\sqrt{e}}\sqrt{\frac{\mu_0}{\mu}}\sqrt{\frac{\mu}{L}}k}\right)$<br>si $\mu_0 \leq \mu$<br>$\mathcal{O}\left(e^{-\frac{\sqrt{2}-1}{2\sqrt{e}(2-\sqrt{\frac{\mu_0}{\mu}})}\sqrt{\frac{\mu}{L}}k}\right)$<br>si $\mu_0 > \mu$ | Nécessité d'estimer $\mu$<br>Les taux dépendent de cette estimation.        |
| LCR-FISTA                         | Alamo, Krupa et Limon [3]                            | $\mathcal{O}\left(e^{-\frac{1}{16}\sqrt{\frac{\mu}{L}}k}\right)$   | Nécessité d'évaluer $F$ régulièrement.                                      |
| Restart basé sur le gradient      | Alamo, Krupa et Limon [1]                            | $\mathcal{O}\left(e^{-\frac{1}{4e(1+\sqrt{\mu+1})}\sqrt{\frac{\mu}{L}}k}\right)$   | -   |
| Restart basé sur la fonctionnelle | Alamo, Krupa et Limon [2]                            | $\mathcal{O}\left(e^{-\frac{\ln(15)}{4e}\sqrt{\frac{\mu}{L}}k}\right)$   | Nécessité d'évaluer $F$ régulièrement.                                      |
| Restart à la Polyak               | Renegar et Grimmer [117]                             | $\mathcal{O}\left(e^{-\frac{1}{2\sqrt{2}}\sqrt{\frac{\mu}{L}}k}\right)$  | Nécessité de connaître $F^*$ .<br>Alternatives avec calcul parallèle.       |
| Algorithme 3                      | -  | $\mathcal{O}\left(e^{-\frac{1}{12}\sqrt{\frac{\mu}{L}}k}\right)$   | Nécessité d'évaluer $F$ occasionnellement.                                  |
| Free-FISTA (Algorithme 6)         | -  | $\mathcal{O}\left(e^{-\frac{\sqrt{\rho}}{12}\sqrt{\frac{\mu}{L}}k}\right)$<br>avec $\rho \in (0, 1)$   | Inclut une stratégie de backtracking sur $L$ .                              |

## 3.2 Restart via une estimation automatique du paramètre de croissance

Cette section est consacrée au schéma de restart que j'ai proposé en collaboration avec mes directeurs de thèse Jean-François Aujol, Charles Dossal et Aude Rondepierre [28].

Dans cette partie, on s'intéresse à des fonctions qui vérifient  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  que l'on suppose connu et  $\mu > 0$  inconnu. Rappelons qu'une fonction  $F$  vérifie  $\mathcal{H}_L$  si elle possède au moins un minimiseur sur  $\mathbb{R}^N$  et peut s'écrire  $F = f + h$  où  $f$  est convexe différentiable à gradient  $L$ -Lipschitz et  $h$  est convexe propre et semi-continue inférieurement. De plus,  $F$  satisfait  $\mathcal{G}_\mu^2$  pour  $\mu > 0$  si et seulement si pour tout  $x \in \text{dom}(F)$ ,

$$\frac{\mu}{2}d(x, X^*)^2 \leq F(x) - F^*, \quad (3.2.1)$$

où  $X^*$  est l'ensemble des minimiseurs de  $F$  et  $F^* = \min_{x \in \mathbb{R}^N} F(x)$ .

Dans ce contexte, on sait que le paramètre de croissance  $\mu$  est crucial pour obtenir des méthodes de premier ordre qui assurent des vitesses de convergence rapides. Les schémas de type Heavy-Ball offrent les meilleurs résultats théoriques mais font intervenir un paramètre de friction qui doit être calibré par rapport à  $\kappa := \frac{\mu}{L}$ . De même, le schéma de restart périodique optimal s'appuie sur cette valeur pour atteindre le meilleur taux de convergence (voir Section 3.1.2.2).

Il existe cependant des méthodes de restart alternatives assurant des vitesses rapides à partir d'une estimation potentiellement mauvaise de  $\mu$  [61], voire sans aucune estimation [1, 3, 2, 117] (voir Section 3.1.2.3). Le schéma présenté dans cette partie s'inscrit dans la lignée de ces schémas et ne requiert aucune information préalable sur le paramètre de croissance.

### 3.2.1 Présentation de la stratégie

L'idée principale de cet algorithme est d'utiliser les garanties théoriques de FISTA pour estimer automatiquement le paramètre de croissance  $\mu$ . Les redémarrages de FISTA sont ensuite calibrés par rapport à la valeur de ces estimations.

On décrit la stratégie mise en place pour estimer  $\mu$  dans la section suivante.

#### 3.2.1.1 FISTA et ses propriétés

On utilisera la formulation présentée dans l'Algorithme 2 pour les itérés de FISTA où la notation  $y^+$  désigne le vecteur issu d'une étape de gradient proximal appliquée à  $y$  avec un pas égal à  $\frac{1}{L}$  :

$$\forall y \in \mathbb{R}^N, \quad y^+ = \text{prox}_{\frac{1}{L}h} \left( y - \frac{1}{L} \nabla f(y) \right). \quad (3.2.2)$$



**Algorithme 2** : FISTA( $z, n$ )**Entrée:**  $z \in \mathbb{R}^N, n \in \mathbb{N}$  $y_0 = x_0 = z, k = 0;$ **répéter** $k = k + 1;$  $x_k = y_{k-1}^+;$  $y_k = x_k + \frac{k-1}{k+2}(x_k - x_{k-1});$ **jusqu'à**  $k \geq n.$ **retourne**  $r = x_k$ 

Pour tout  $k \in \mathbb{N}^*$ , on écrit que  $x_k$  le vecteur obtenu après  $k$  itérations de FISTA à partir de  $x_0 \in \mathbb{R}^N$  vérifie :

$$x_k = \text{FISTA}(x_0, k). \quad (3.2.3)$$

Les itérés de FISTA ont des propriétés de convergence que l'on exploite dans notre schéma de restart. On s'appuie sur la proposition suivante.

**Proposition 3.2.1.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . La suite  $(x_k)_{k \in \mathbb{N}}$  générée par FISTA (Algorithme 2) vérifie les deux points suivants :*

1. Pour tout  $k \in \mathbb{N}^*$ ,

$$F(x_k) - F^* \leq \frac{4L}{\mu(k+1)^2} (F(x_0) - F^*), \quad (3.2.4)$$

2. Pour tout  $k \in \mathbb{N}^*$ ,

$$F(x_k) \leq F(x_0). \quad (3.2.5)$$

**Preuve.**

Le premier point a déjà été justifié dans la Section 3.1.2.2 en utilisant [51, Theorem 3.1]. Le deuxième point est direct en appliquant [51, Lemma 3.1].  $\square$

En prenant en considération le premier point de la Proposition 3.2.1, une idée naturelle pour estimer le paramètre de croissance est d'exploiter la borne suivante :

$$\forall k \in \mathbb{N}^*, \quad \mu \leq \frac{4L}{(k+1)^2} \frac{F(x_0) - F^*}{F(x_k) - F^*}. \quad (3.2.6)$$

Cependant, comme la valeur de  $F^*$  est rarement connue en amont, cette borne est difficilement calculable. En revanche, le deuxième point de la Proposition 3.2.1 assure que la fonction  $u \mapsto \frac{F(x_0) - u}{F(x_k) - u}$  est croissante sur  $[F^*, F(x_k)]$ . Ceci implique que pour tout  $y \in \mathbb{R}^N$  tel que  $F(y) < F(x_k)$ , on a :

$$\mu \leq \frac{4L}{(k+1)^2} \frac{F(x_0) - F(y)}{F(x_k) - F(y)}. \quad (3.2.7)$$

C'est à partir de cette inégalité que l'on construit une suite d'estimation  $(\tilde{\mu}_j)_{j \geq 2}$ .

3.2.1.2 Schéma de restart

Pour un vecteur initial  $r_0 \in \mathbb{R}^N$ , notre schéma de restart s'écrit comme suit.

---

**Algorithme 3** : Schéma de restart [28]

---

**Entrée:**  $r_0 \in \mathbb{R}^N$ ,  $j = 1$ ,  $C > 4$ .

$n_0 = \lfloor 2C \rfloor$ ;

$r_1 = \text{FISTA}(r_0, n_0)$ ;

$n_1 = \lfloor 2C \rfloor$ ;

**répéter**

$j = j + 1$ ;

$r_j = \text{FISTA}(r_{j-1}, n_{j-1})$ ;

$\tilde{\mu}_j = \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}$ ;

**si**  $n_{j-1} \leq C \sqrt{\frac{L}{\tilde{\mu}_j}}$  **alors**

$n_j = 2n_{j-1}$ ;

**sinon**

$n_j = n_{j-1}$ ;

**fin si**

**jusqu'à**  $\|g(r_j)\| \leq \varepsilon$ .

**retourne**  $r = r_j^+$ .

---

L'algorithme fait intervenir trois suites importantes :

- $(r_j)_{j \geq 0}$  est la suite contenant les vecteurs résultant de chaque exécution complète de FISTA. Autrement dit,  $r_j$  est le vecteur obtenu après  $j$  redémarrages de FISTA.
- $(n_j)_{j \geq 0}$  est la suite qui contient le nombre d'itérations avant chaque restart de FISTA. On fait  $n_j$  itérations de FISTA avant le  $j + 1$ -ème restart de FISTA :

$$r_{j+1} = \text{FISTA}(r_j, n_j). \quad (3.2.8)$$

- $(\tilde{\mu}_j)_{j \geq 2}$  est la suite des estimations du paramètre de croissance.

Une estimation du paramètre de croissance est calculée à chaque redémarrage à partir de  $j \geq 2$  en utilisant la valeur de la fonction  $F$  évaluée aux points  $(r_i)_{0 \leq i \leq j}$  :

$$\tilde{\mu}_j = \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}. \quad (3.2.9)$$

La suite  $(\tilde{\mu}_j)_{j \geq 2}$  ainsi définie vérifie le lemme suivant.

**Lemme 3.2.2.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . La suite  $(\tilde{\mu}_j)_{j \geq 2}$  définie dans l'Algorithme 3 avec  $C > 4$  et  $\varepsilon > 0$  vérifie l'inégalité suivante :*

$$\forall j \geq 2, \quad \tilde{\mu}_j \geq \tilde{\mu}_{j+1} \geq \mu. \quad (3.2.10)$$

**Preuve.**

En appliquant l'inégalité (3.2.7) à la suite  $(r_j)_{j \geq 0}$ , on a que pour  $i \geq 1$  et  $y \in \mathbb{R}^N$  tel que  $F(y) \leq F(r_i)$ ,

$$\mu \leq \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(y)}{F(r_i) - F(y)}. \quad (3.2.11)$$

Or, le deuxième point de la Proposition 3.2.1 garantit que pour tout  $j > i$ ,  $F(r_j) \leq F(r_i)$  et donc :

$$\mu \leq \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}. \quad (3.2.12)$$

Comme cette inégalité est vraie pour tout  $i < j$ , on peut passer au minimum pour obtenir que  $\mu \leq \tilde{\mu}_j$ .

Le caractère décroissant de la suite peut aussi être démontré. Pour tout  $j \geq 2$ ,

$$\tilde{\mu}_{j+1} = \min_{\substack{i \in \mathbb{N}^* \\ i < j+1}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_{j+1})}{F(r_i) - F(r_{j+1})} \leq \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_{j+1})}{F(r_i) - F(r_{j+1})}. \quad (3.2.13)$$

La fonction  $u \mapsto \frac{F(r_{i-1}) - u}{F(r_i) - u}$  étant croissante sur  $[F^*, F(r_i)]$  pour tout  $i < j$ , l'inégalité  $F(r_{j+1}) \leq F(r_j)$  garantie par la Proposition 3.2.1 implique que :

$$\tilde{\mu}_{j+1} \leq \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} = \tilde{\mu}_j. \quad (3.2.14)$$

□

Une fois qu'une estimation de  $\mu$  est calculée, le nombre d'itérations avant le prochain redémarrage est mis à jour. Pour cela, on suit une procédure de doublement des itérations si la condition suivante est vérifiée :

$$n_{j-1} \leq C \sqrt{\frac{L}{\tilde{\mu}_j}}. \quad (3.2.15)$$

La constante  $C$  est un paramètre propre à la méthode qui agit sur la fréquence de restart : si  $C$  est petit, la condition ci-dessus sera rarement vérifiée et le nombre d'itérations de FISTA par exécution sera rarement doublé. En conséquence, une grande valeur de  $C$  favorise les redémarrages fréquents. Dans le cas contraire, les restarts seront très occasionnels. La valeur de ce paramètre est donc essentielle pour définir un schéma efficace. Il est montré par la suite que l'on peut optimiser  $C$  pour avoir les meilleures garanties théoriques possibles, la valeur optimale étant  $\hat{C} \approx 6.38$ .

Cette façon de définir la suite  $(n_j)_{j \geq 0}$  permet de bien contrôler son comportement, notamment grâce au lemme ci-dessous.

**Lemme 3.2.3.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . La suite  $(n_j)_{j \geq 0}$  définie dans l'Algorithme 3 avec  $C > 4$  et  $\varepsilon > 0$  vérifie l'inégalité suivante :*

$$\forall j \geq 0, \quad n_j \leq 2C \sqrt{\frac{L}{\mu}}. \quad (3.2.16)$$

**Preuve.**

Ce lemme est une conséquence directe du Lemme 3.2.2. En effet, pour  $j \geq 2$ , on définira  $n_j$  comme le double de  $n_{j-1}$  si la condition (3.2.15) est vérifiée. Or, le Lemme 3.2.2 garantit que si cette condition est vérifiée, alors :

$$n_{j-1} \leq C \sqrt{\frac{L}{\mu}}, \quad (3.2.17)$$

et donc  $n_j \leq 2C \sqrt{\frac{L}{\mu}}$ . Si la condition n'est pas vérifiée, on conserve le même nombre d'itérations  $n_j = n_{j-1}$ . Sachant que  $n_0 = n_1 = \lfloor 2C \rfloor \leq 2C \sqrt{\frac{L}{\mu}}$  (car  $L \geq \mu$  pour toute fonction  $F$ ), une simple récurrence nous amène à la conclusion voulue.  $\square$

*Remarque 3.2.4.* Plutôt que de doubler le nombre d'itérations à chaque fois que la condition (3.2.15) est vérifiée, on pourrait choisir de multiplier le nombre d'itérations par un paramètre  $\gamma > 1$ . En utilisant cette règle de mise à jour des itérations et en choisissant  $n_0 = n_1 = \lfloor \gamma C \rfloor$ , le contrôle obtenu sur la suite  $(n_j)_{j \geq 0}$  devient

$$\forall j \geq 0, \quad n_j \leq \gamma C \sqrt{\frac{L}{\mu}}. \quad (3.2.18)$$

On peut montrer que le choix  $\gamma = 2$  permet d'obtenir les meilleures vitesses théoriques (voir Section 3.A.1).

Après avoir mis à jour le nombre d'itérations, FISTA est relancé depuis le dernier itéré  $r_j$  pour  $n_j$  itérations et ainsi de suite. L'algorithme s'interrompt si un critère d'arrêt est satisfait :

$$\|g(r_j)\| \leq \varepsilon, \quad (3.2.19)$$

où  $g$  représente le gradient composite, c'est-à-dire  $g : x \mapsto L(x - x^+)$  et  $\varepsilon > 0$  correspond au niveau de précision attendu par l'utilisateur. Notons qu'évaluer cette condition ne demande pas véritablement de calcul additionnel puisque  $r_j^+$  est calculé quoi qu'il en soit par la suite. L'utilisation de cette condition de sortie est motivée par le lemme ci-dessous qui relie la valeur du gradient composite à celle de  $F$  via l'hypothèse  $\mathcal{G}_\mu^2$ . La preuve de ce lemme est donnée en Section 3.A.3.6.

**Lemme 3.2.5.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Pour tout  $x \in \mathbb{R}^N$ , on a :*

$$F(x^+) - F^* \leq \frac{8}{\mu} \|g(x)\|^2. \quad (3.2.20)$$

Grâce à ce résultat, on sait que le vecteur de sortie  $r$  donnée par l'Algorithme 3 vérifie :

$$F(r) - F^* \leq \frac{8\varepsilon^2}{\mu}. \quad (3.2.21)$$

*Remarque 3.2.6.* Les calculs additionnels liés à l'estimation du paramètre de croissance n'ont lieu qu'à chaque restart, et ont donc lieu dans le pire des cas toute les  $\lfloor 2C \rfloor$  itérations. Les itérations de FISTA ne demandent aucun calcul supplémentaire lié au schéma de restart.

*Remarque 3.2.7.* Le nombre total d'itérations pour obtenir le vecteur  $r = r_j^+$  est de  $1 + \sum_{i=0}^j n_i$ , où la dernière itération est l'étape de gradient proximal appliquée à  $r_j$ .

### 3.2.2 Garanties théoriques du schéma

Maintenant que le fonctionnement du schéma a été détaillé, on peut s'intéresser à ses propriétés de convergence.

**Théorème 3.2.8.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  les suites définies dans l'Algorithme 3 avec  $C > 4$  et  $\varepsilon > 0$ . Le nombre d'itérations  $1 + \sum_{i=0}^j n_i$  nécessaire pour assurer  $\|g(r_j)\| \leq \varepsilon$  est borné et vérifie l'inégalité suivante :*

$$\sum_{i=0}^j n_i \leq \frac{4C}{\ln\left(\frac{C^2}{4} - 1\right)} \sqrt{\frac{L}{\mu}} \left( 2 \ln\left(\frac{C^2}{4} - 1\right) + \ln\left(1 + \frac{16}{C^2 - 16} \frac{2L(F(r_0) - F^*)}{\varepsilon^2}\right) \right). \quad (3.2.22)$$

#### Éléments de preuve.

La preuve complète de ce théorème est donnée en Annexe 3.A.3.1 de ce chapitre. Comme la démonstration est technique, on donne les arguments utilisés et la démarche ci-dessous. L'idée de la preuve a été inspirée par l'article d'Alamo, Krupa et Limon [3].

La première étape consiste à montrer qu'un doublement du nombre d'itérations a lieu au moins tous les  $T$  redémarrages pour un  $T$  bien choisi. Pour cela, on utilise un raisonnement par l'absurde :

- On suppose qu'aucun doublement des itérations n'a été fait entre les  $s+1$ -ème et  $s+T$ -ème restarts pour un certain  $s > 0$ .
- On montre que dans une telle situation,  $(F(r_{j-1}) - F(r_j))_{j \in \llbracket s+1, s+T \rrbracket}$  décroît géométriquement.
- On applique le Lemme 3.A.4 qui nous permet de majorer  $\|g(r_{j-1})\|$  en fonction de  $F(r_{j-1}) - F(r_j)$  pour tout  $j \in \llbracket s+1, s+T \rrbracket$ .
- On utilise la décroissance géométrique de  $(F(r_{j-1}) - F(r_j))_{j \in \llbracket s+1, s+T \rrbracket}$  pour montrer que la condition de sortie est vérifiée pour  $j = s+T$ .

La deuxième étape consiste à manipuler ce résultat pour montrer que le nombre d'itérations nécessaire pour activer le critère d'arrêt est borné supérieurement par  $2Tn_j$ . La conclusion du Théorème 3.2.8 découle du Lemme 3.2.3 qui nous permet de majorer la suite  $(n_j)_{j \geq 0}$ .  $\square$

Ce théorème permet d'avoir une borne supérieure sur le nombre d'itérations nécessaires pour atteindre une précision  $\varepsilon > 0$ . On peut déduire de ce résultat un

taux de convergence pour ce schéma de restart avec le corollaire ci-dessous. Une preuve est fournie en Section 3.A.3.2.

**Corollaire 3.2.9.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Pour  $C > 4$  et  $\varepsilon > 0$ , les suites  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  de l'Algorithme 3 vérifient*

$$F(r_j^+) - F^* = \mathcal{O} \left( e^{-\frac{\ln\left(\frac{C^2-1}{4C}\right)}{4C} \sqrt{\frac{\mu}{L}} \sum_{i=0}^j n_i} \right). \quad (3.2.23)$$

En particulier, si  $C = \hat{C}$  où  $\hat{C}$  maximise  $\frac{\ln\left(\frac{C^2-1}{4C}\right)}{4C}$  (soit  $\hat{C} \approx 6.38$ ), alors il existe  $K > \frac{1}{12}$  tel que  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  vérifient

$$F(r_j^+) - F^* = \mathcal{O} \left( e^{-K \sqrt{\frac{\mu}{L}} \sum_{i=0}^j n_i} \right). \quad (3.2.24)$$

En pratique, comme le paramètre  $C$  n'est pas intuitif à définir (il influe sur la fréquence de redémarrage mais ne représente rien de concret), on utilise la valeur offrant la meilleure vitesse théorique  $\hat{C} \approx 6.38$ . Numériquement, cette valeur permet généralement d'obtenir les meilleures vitesses de convergence pour le schéma.

On peut comparer le résultat de convergence offert par l'Algorithme 3 à celui d'autres schémas sous les mêmes hypothèses sur  $F$ .

Le taux de convergence (3.2.24) montre que notre schéma assure une décroissance exponentielle **rapide** de l'erreur si on se réfère à la discussion de la Section 3.1.1. En effet, le nombre d'itérations nécessaires pour atteindre une précision donnée avec l'Algorithme 3 est proportionnel à  $\sqrt{\frac{L}{\mu}}$  contrairement à la méthode de gradient proximal ou le schéma de restart basé sur le gradient de Alamo, Krupa et Limon [1] qui demandent un nombre d'itérations proportionnel à  $\frac{L}{\mu}$ . Sachant que  $\frac{L}{\mu}$  est en fait le conditionnement de la fonction  $F$  et qu'il est donc généralement très grand, une telle différence est significative.

Ce résultat de convergence est légèrement meilleur que celui assuré par le schéma de restart de Alamo, Krupa et Limon [3] ( $\mathcal{O}\left(e^{-\frac{1}{16}\sqrt{\frac{\mu}{L}}k}\right)$ ) mais moins bon que celui donné par leur schéma le plus récent [2] qui assure une décroissance de l'erreur en  $\mathcal{O}\left(e^{-\frac{\ln(15)}{4e}\sqrt{\frac{\mu}{L}}k}\right)$  où  $\frac{\ln(15)}{4e} > 0.249$ . Rappelons que ces méthodes demandent des évaluations régulières de la fonction  $F$  notamment au sein des itérations de FISTA. Ces calculs additionnels peuvent être implémentés de manière intelligente pour réduire significativement leur coût mais cela demande un travail supplémentaire de l'utilisateur. Notre schéma est moins impacté par cette question puisque les calculs additionnels sont occasionnels. Le taux de convergence (3.2.24) est également

meilleur asymptotiquement que celui garanti par AdaRES de Fercoq et Qu [61] à partir du moment où l'utilisateur fournit une estimation  $\mu_0$  qui est supérieure à  $4\mu$  ou inférieure à  $\frac{\mu}{15}$ . Sachant que le paramètre de croissance  $\mu$  est souvent difficile à estimer, le fait de ne pas dépendre d'une première estimation de  $\mu$  est un avantage non négligeable. On peut noter que le schéma proposé par Renegar et Grimmer dans [117] assure un meilleur taux  $\mathcal{O}\left(e^{-\frac{1}{2\sqrt{2}}}\right)$  mais s'appuie sur la valeur de  $F^*$  qui est rarement disponible ou sur des calculs en parallèle pour estimer cette valeur.

La vitesse au pire cas de l'Algorithme 3 est naturellement surpassée par celle de schémas s'appuyant sur la valeur théorique de  $\frac{\mu}{L}$  comme le schéma type Heavy-Ball introduit par Aujol, Dossal et Rondepierre dans [25] ( $\mathcal{O}\left(e^{-(2-\sqrt{2})(1-\sqrt{\frac{\mu}{L}})\sqrt{\frac{\mu}{L}k}}\right)$  si  $F$  a un unique minimiseur), le restart périodique optimal ( $\mathcal{O}\left(e^{-\frac{1}{e}\sqrt{\frac{\mu}{L}k}}\right)$ ) ou V-FISTA [31] qui peut assurer une décroissance de  $\mathcal{O}\left(e^{-K\sqrt{\frac{\mu}{L}k}}\right)$  avec  $K = 0.45$  si  $\frac{L}{\mu} \geq 100$  (voir Corollaire 5.4.3).

Il est important de préciser que la discussion ci-dessus concerne des vitesses théoriques asymptotiques au pire cas. Ainsi, les taux mentionnés ne reflètent pas nécessairement la performance des schémas en pratique.

### 3.2.3 Expériences numériques

Dans cette section, deux expériences numériques sont présentées visant à illustrer le comportement de l'Algorithme 3. On se placera dans le cas où  $F$  vérifie  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  où  $L$  est connu mais aucune estimation de  $\mu$  n'est disponible. De plus, dans les deux exemples donnés, la valeur de  $F^*$  n'est pas accessible au préalable donc la méthode de Renegar et Grimmer [117] basée sur cette valeur n'est pas directement applicable. On pourrait cependant appliquer leurs schémas alternatifs faisant intervenir de la parallélisation.

Les méthodes que l'on compare sont les suivantes : Forward-Backward, FISTA [32], le restart empirique avec la règle sur la fonctionnelle [108], le restart empirique avec la règle sur le gradient composite [108], les schémas de Alamo, Krupa et Limon [3, 1, 2] et l'Algorithme 3 avec  $C = 6.38$  c'est-à-dire le choix théorique optimal.

#### 3.2.3.1 Moindres carrés avec régularisation $\ell^1$ (LASSO)

Dans ce premier exemple, on cherche à résoudre un problème de moindres carrés avec un terme de régularisation  $\ell^1$  qui est aussi appelé LASSO (pour *Least Absolute Shrinkage and Selection Operator* [130]). Il s'agit d'un problème très répandu dans différents domaines comme en statistiques ou en traitement de l'image. À partir d'une matrice  $A \in \mathcal{M}_{M \times N}(\mathbb{R})$  et d'un vecteur  $b \in \mathbb{R}^M$  avec  $M > 0$  et  $N > 0$ , on cherche à déterminer  $\tilde{x} \in \mathbb{R}^N$  qui résout le problème ci-dessous :

$$\min_{x \in \mathbb{R}^N} F(x) := \underbrace{\frac{1}{2} \|Ax - b\|_2^2}_{:=f(x)} + \underbrace{\lambda \|x\|_1}_{:=h(x)}, \quad (3.2.25)$$

où  $\lambda > 0$  est le paramètre de régularisation. La matrice  $A$  correspond à la matrice d'entrée et  $b$  est le vecteur de sortie. L'idée lorsque l'on résout ce type de problème est de déterminer un vecteur  $\tilde{x}$  qui permet de rattacher ces deux objets de façon fiable ( $A\tilde{x} \approx b$ ) tout en faisant en sorte que la valeur de ses coefficients soit suffisamment faible en valeur absolue ( $\|\tilde{x}\|_1 \leq \varepsilon$ ). Le paramètre  $\lambda$  permet de donner plus ou moins de poids au terme d'attache aux données  $f$  par rapport au terme de régularisation  $h$ . Quand  $\lambda$  tend vers zéro, on retrouve un problème de moindres carrés standard.

La fonction  $F$  définie pour ce problème vérifie bien les hypothèses nécessaires pour appliquer notre algorithme puisque  $\nabla f$  est  $L$ -Lipschitz où  $L$  est la plus grande valeur propre de  $A^T A$  et il est montré dans [41, Corollaire 9] que  $F$  vérifie  $\mathcal{G}_\mu^2$  avec  $\mu > 0$ . De plus, on peut facilement calculer le gradient de  $f$  :

$$\forall x \in \mathbb{R}^N, \quad \nabla f(x) = A^T(Ax - b). \quad (3.2.26)$$

D'autre part, l'opérateur proximal de  $h$  est le seuillage doux, c'est-à-dire :

$$\forall x \in \mathbb{R}^N, \quad \forall i \in \llbracket 1, N \rrbracket, \quad (\text{prox}_h(x))_i = \begin{cases} x_i + \lambda & \text{si } x_i < -\lambda, \\ 0 & \text{si } x_i \in [-\lambda, \lambda], \\ x_i - \lambda & \text{si } x_i > \lambda. \end{cases} \quad (3.2.27)$$

On peut donc appliquer les méthodes de premier ordre présentées précédemment. Il est important de préciser que les méthodes comparées ici ne sont pas les plus efficaces pour résoudre un problème LASSO. Les méthodes de descente par coordonnées sont plus performantes dans ce contexte et il existe des schémas spécifiquement conçus pour le LASSO comme CELER [94] qui donnent des résultats très efficaces.

On compare les différents schémas de restart sur un problème LASSO où  $A$ ,  $b$  et  $x_0$  sont générés aléatoirement où  $M = N = 1500$ . Le paramètre de régularisation est choisi arbitrairement égal à 20. L'évolution de l'erreur en fonction du nombre d'itérations pour chaque méthode est donnée par la Figure 3.2.

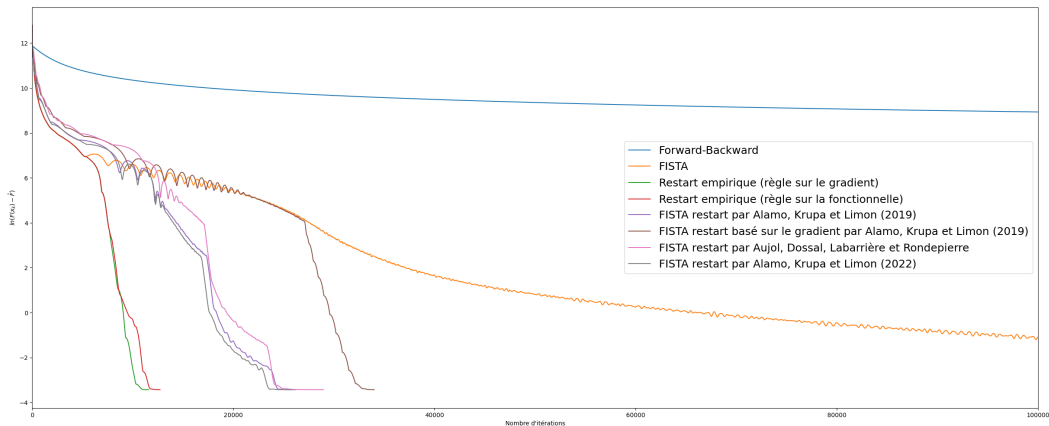


FIGURE 3.2 – Évolution de l'erreur en fonction du nombre d'itérations pour plusieurs méthodes de restart sur un problème LASSO.



Cet exemple met en lumière la différence de performance entre Forward-Backward et FISTA mais également entre FISTA et les méthodes de restart. Ici, les méthodes de redémarrage empiriques sont les plus rapides en terme de nombre d'itérations devant les schémas basés sur la valeur de la fonctionnelle proposés par Alamo, Krupa et Limon et l'Algorithme 3. Le schéma d'Alamo, Krupa et Limon qui est construit autour du gradient composite est moins rapide, ce qui peut souligner la différence entre décroissance exponentielle rapide et lente. On notera qu'ici, calculer la fonctionnelle  $F$  n'est pas très coûteux à condition d'implémenter efficacement chaque méthode. En effet, comme la fonctionnelle à minimiser est de la forme  $\phi(Ax) + h(x)$  on peut appliquer FASTA qui est une légère adaptation de FISTA introduite par Goldstein, Studer et Baraniuk dans [67] permettant d'économiser du temps de calcul. C'est pour cette raison que la comparaison entre les méthodes par rapport à l'erreur en fonction du nombre d'itérations est pertinent ici.

### 3.2.3.2 Problème d'inpainting

Ce deuxième exemple est un problème de traitement d'image appelé problème d'inpainting. L'objectif est de reconstruire le plus fidèlement possible une image qui a été préalablement détériorée. Si on définit  $x^0 \in \mathcal{M}_{N \times N}(\mathbb{R})$  l'image originale, alors on définit l'image détériorée  $y = Mx^0$  où  $M : \mathcal{M}_{N \times N}(\mathbb{R}) \rightarrow \mathcal{M}_{N \times N}(\mathbb{R})$  est un opérateur de masquage. Cet opérateur détruit une partie des pixels de l'image d'entrée et laisse intacts les autres. On considère que  $M$  est connu ici.



FIGURE 3.3 – Exemple d'un problème d'inpainting : image détériorée  $y$  (à gauche) et solution du problème (à droite).

On considère le problème d'optimisation ci-dessous pour construire une approximation de l'image originale  $x^0$  :

$$\min_{x \in \mathcal{M}_{N \times N}(\mathbb{R})} F(x) := \underbrace{\frac{1}{2} \|Mx - y\|_2^2}_{:=f(x)} + \underbrace{\lambda \|Tx\|_1}_{:=h(x)}, \quad (3.2.28)$$

où  $\lambda > 0$  est le paramètre de régularisation et  $T$  est une transformée en ondelettes.

L'utilisation de la transformée en ondelettes a pour but de construire une image qui a une structure cohérente. En effet, une image cohérente pour l'oeil humain a en général une représentation parcimonieuse dans une base d'ondelettes bien choisie (peu de coefficients ont un module significatif). Il paraît donc pertinent d'introduire un terme de régularisation sur les coefficients en ondelettes.

La fonction  $F$  vérifie bien l'hypothèse  $\mathcal{H}_L$  car  $\nabla f : x \mapsto Mx - y$  est une fonction 1-Lipschitz. On peut également montrer que  $F$  satisfait  $\mathcal{G}_\mu^2$  pour un certain  $\mu > 0$  qui n'est pas facilement estimable. Cet exemple est en fait un problème LASSO où le véritable paramètre que l'on optimise est  $w := Tx$  la transformée en ondelettes de  $x$ .

On compare la performance des différents schémas de restart dans la Figure 3.4. Les tests ci-dessous sont effectués pour une image de  $225 \times 225$  pixels avec la base d'ondelettes Daubechies 4 et  $\lambda = 2$ .

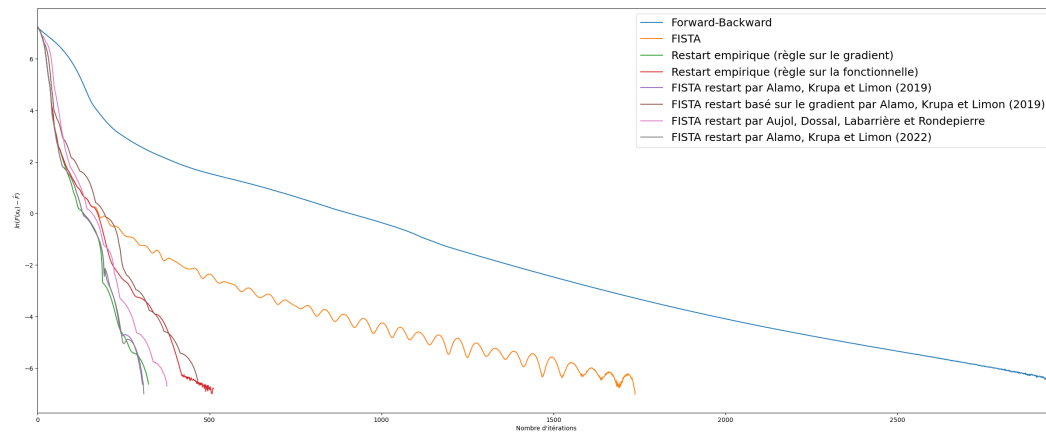


FIGURE 3.4 – Évolution de l'erreur en fonction du nombre d'itérations pour plusieurs méthodes de restart sur un problème d'inpainting.

Comme dans l'exemple précédent, une hiérarchie claire se dégage entre Forward-Backward, FISTA et les méthodes de restart. Ici, les schémas de restart ont des performances très comparables si on se réfère à la décroissance de l'erreur en fonction du nombre d'itérations. Si on compare la performance de chaque méthode en fonction du temps de calcul, on peut toutefois avoir une interprétation différente. La Figure 3.5 affiche la décroissance de l'erreur en fonction du temps de calcul pour une implémentation brute de chacune des méthodes : pour ce problème, il est possible de réduire la complexité de chaque itération nécessitant de calculer  $F$  en conservant et réutilisant certaines quantités (notamment la transformée en ondelettes de  $x$ ). Le graphe ci-dessous est obtenu lorsque ce travail d'optimisation n'est pas fait et donc quand le coût de  $F$  est élevé.

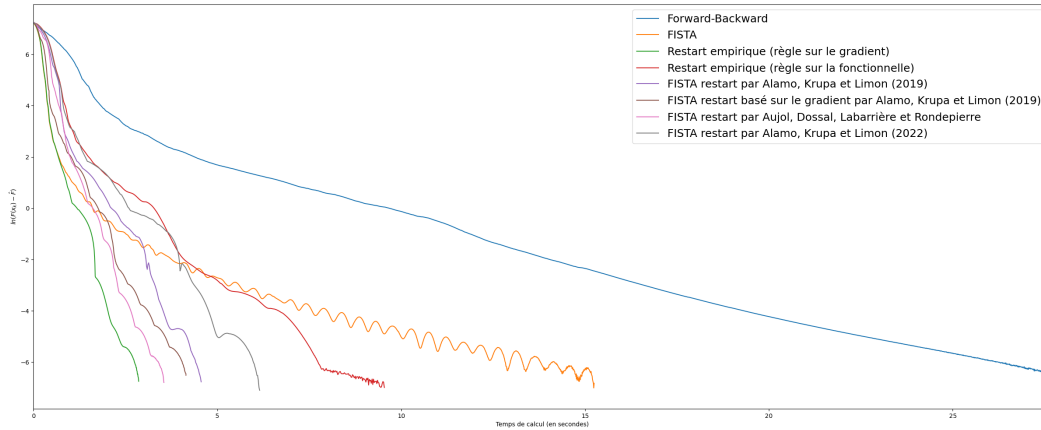


FIGURE 3.5 – Évolution de l'erreur en fonction du temps de calcul pour plusieurs méthodes de restart sur un problème d'inpainting pour une implémentation brute des méthodes.

On peut voir que lorsque  $F$  coûte cher à évaluer, certaines méthodes parmi lesquelles celles introduites par Alamo, Krupa et Limon dans [3, 2], l'efficacité de certain schémas adaptatifs est réduite. On peut ainsi remarquer qu'en comparaison, l'Algorithme 3 n'est quasiment pas touché puisqu'il ne nécessite que très peu de calculs additionnels.

### 3.3 Un schéma de restart avec backtracking : Free-FISTA

Dans cette partie, j'introduis une méthode mise au point avec Jean-François Aujol, Luca Calatroni, Charles Dossal et Aude Rondepierre. Ce travail est un prolongement de celui présenté dans la partie précédente.

On s'appuie sur les mêmes hypothèses de géométrie que précédemment c'est-à-dire  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . On considère cependant que la constante de Lipschitz  $L$  n'est pas nécessairement connue. Pour faire face à cette difficulté, on choisit d'utiliser une stratégie de **backtracking** permettant d'estimer la constante de Lipschitz  $L$  combinée à la philosophie du schéma précédent (Algorithme 3). La philosophie du backtracking est expliquée dans la section suivante. On se concentrera ensuite sur les propriétés de cette stratégie appliquée à FISTA avant d'introduire le schéma Free-FISTA (Algorithme 6). Enfin, on donnera des garanties théoriques et des exemples d'applications numériques.

#### 3.3.1 Contexte

Jusqu'ici, toutes les méthodes présentées s'appuient sur un pas de descente  $\tau$  qui doit être calibré par rapport à la fonction  $F$  à minimiser. En particulier, c'est la constante de Lipschitz  $L$  du gradient de  $f$  (ou de  $F$  dans le cas différentiable) qui joue un rôle majeur. Comme elle permet un contrôle sur les variations de  $F$ , cette constante nous donne un moyen de restreindre le pas pour qu'il ne soit pas

trop grand et que la suite d'itérés ne diverge pas. Pour la méthode de descente de gradient, un pas inférieur ou égal à  $\frac{2}{L}$  permet d'avoir un algorithme de descente. Pour FISTA, on choisit en général un pas de  $\frac{1}{L}$ , c'est notamment le choix qui est fait dans la section précédente.

Dans certains cas, la constante de Lipschitz  $L$  est difficile à estimer précisément. Ainsi, fixer le pas de descente est délicat puisqu'un pas trop grand peut empêcher l'algorithme de converger, et un pas trop petit le ralentit. Armijo introduit en 1966 [9] une méthode pour faire face à ce problème lors de l'application de la descente de gradient. Cela consiste à tester des valeurs de pas en évaluant la fonctionnelle à chaque itération. À partir d'une première valeur de pas potentiellement trop grande, la valeur de  $\frac{1}{\tau}(F(x_k) - F(x_k - \tau \nabla F(x_k)))$  est calculée et tant qu'elle n'est pas suffisamment grande, le pas  $\tau$  est réduit. Cette stratégie permet d'obtenir des résultats de convergence sans connaître  $L$  mais le coût d'une itération est plus élevé que pour une descente de gradient classique.

Plus récemment, on compte de nombreux schémas incluant une stratégie de backtracking : dans [32], Beck et Teboulle proposent une version de FISTA avec backtracking d'Armijo ; Nesterov introduit dans [104] un schéma accéléré avec backtracking qui peut prendre en compte la forte convexité avec des résultats de convergence ; dans [122], Scheinberg, Goldfarb et Bai donnent une stratégie de backtracking pour FISTA avec une garantie théorique de la décroissance de l'erreur ; dans [85], Lin et Xiao introduisent un schéma pour l'optimisation composite fortement convexe avec du backtracking sur  $L$  et une estimation automatique du paramètre de croissance avec des garanties théoriques ; Calatroni et Chambolle [49] proposent une version backtrackée de FISTA dans le cadre fortement convexe qui, contrairement à ce que propose Armijo, autorise à augmenter le pas au cours des itérations (cette philosophie est introduite dans [122]) ; dans le même esprit, Rebegoldi et Calatroni proposent dans [116] une version de FISTA faisant intervenir du *scaling*, du backtracking et qui est robuste à des calculs inexacts de l'opérateur proximal ; Liang, Luo et Schoenlieb [84] proposent des variations de FISTA notamment adaptées à la forte convexité et qui s'appuient sur du backtracking sur  $L$ . La liste ci-dessus n'est pas exhaustive et l'utilisation de stratégies de backtracking est très répandue.

La méthode que l'on introduit est la seule à ma connaissance qui permet d'obtenir **une décroissance exponentielle rapide** de l'erreur sur un problème convexe composite avec l'hypothèse de croissance quadratique  $\mathcal{G}_\mu^2$  **sans avoir besoin de connaître  $L$  ou  $\mu$** . Le résultat qui s'approche le plus du notre est celui de Lin et Xiao [85] qui proposent un schéma qui estime automatiquement  $L$  et  $\mu$  avec une vitesse de convergence rapide mais dans le cadre fortement convexe.

### 3.3.2 Le backtracking appliqué à FISTA

Le schéma que l'on propose est une méthode de restart d'une version de FISTA avec backtracking. Cette adaptation de FISTA est construite autour de la formulation originale proposée par Beck et Teboulle [32] (Algorithme 4) qui est la plus na-

turelle pour dériver des schémas avec backtracking. Les auteurs proposent d'ailleurs une version de FISTA avec backtracking dans le même papier. Rappelons que l'opérateur  $T_\tau$  qui apparaît dans l'Algorithme 4 est défini comme suit pour tout  $\tau > 0$  :

$$\forall \tau > 0, T_\tau(x) = \text{prox}_{\tau h}(x - \tau \nabla f(x)).$$

---

**Algorithme 4** : Version originale de FISTA [32]

---

**Entrée:**  $x_0 \in \mathbb{R}^N$ ,  $n > 0$ ,  $L > 0$ .

$x_{-1} = x_0$ ,  $t_0 = 1$ ,  $k = 0$ ,  $\tau = 1/L$ .

**répéter**

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2};$$

$$\beta_{k+1} = \frac{t_k - 1}{t_{k+1}};$$

$$y_{k+1} = x_k + \beta_{k+1}(x_k - x_{k-1});$$

$$x_{k+1} = T_\tau(y_{k+1});$$

$$k = k + 1;$$

**jusqu'à**  $k \geq n$ .

**retourne**  $x_k$

---

La stratégie sur laquelle on s'appuie est analogue à celle adoptée dans [122, 49, 50] : à chaque itération, on effectue un test pour une valeur de pas  $\tau$  et on réduit  $\tau$  jusqu'à que le test soit validé ; si l'utilisateur l'autorise, on augmente ensuite la valeur du dernier pas accepté à l'itération suivante. Cette procédure est décrite dans l'Algorithme 5 qui fait intervenir la divergence de Bregman  $D_f$  définie comme :

$$\forall (x, y) \in \text{dom}(F) \times \text{dom}(F), \quad D_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle. \quad (3.3.1)$$

---

**Algorithme 5** : FISTA\_adaBT( $x_0, n, L_0, L_{min}; \rho, \delta$ )

---

**Entrée:**  $x_0 \in \mathbb{R}^N$ ,  $n > 0$ ,  $L_0 > 0$ ,  $L_{min} \ll L$ ,  $\rho \in (0, 1)$ ,  $\delta \in (0, 1]$ .

 $\tau_0 = 1/L_0$ ,  $x_{-1} = x_0$ ,  $t_0 = 1$ ,  $k = 0$ .

**répéter**

$$\tau_{k+1}^0 = \min \left\{ \frac{\tau_k}{\delta}, \frac{1}{L_{min}} \right\};$$

 $i = 0$ ;

**répéter**

$$\tau_{k+1} = \rho^i \tau_{k+1}^0;$$

$$t_{k+1} = \frac{1 + \sqrt{1 + 4 \frac{\tau_k}{\tau_{k+1}} t_k^2}}{2};$$

$$\beta_{k+1} = \frac{t_k - 1}{t_{k+1}};$$

$$y_{k+1} = x_k + \beta_{k+1}(x_k - x_{k-1});$$

$$x_{k+1} = T_{\tau_{k+1}}(y_{k+1});$$

 $i = i + 1$ ;

**jusqu'à**  $D_f(x_{k+1}, y_{k+1}) \leq \|x_{k+1} - y_{k+1}\|^2 / 2\tau_{k+1}$ .

 $k = k + 1$ ;

**jusqu'à**  $k \geq n$ .

**retourne**  $x_k, 1/\tau_k$ 


---

Cette version de FISTA est une légère adaptation du schéma de Scheinberg, Goldfarb et Bai [122] ou du schéma de Calatroni et Chambolle [49] dans le cas non fortement convexe. On peut remarquer que l'Algorithme 5 s'appuie sur la formulation originale de FISTA de Beck et Teboulle [32] et non pas sur celle avec le paramètre d'inertie  $\frac{k}{k+\alpha}$ . Décrivons **le rôle des paramètres** :

- $\rho \in (0, 1)$  est le paramètre d'Armijo. C'est cette constante qui régit la manière dont on réduit la valeur du pas quand la condition d'acceptation ci-dessous n'est pas acceptée :

$$D_f(x_{k+1}, y_{k+1}) \leq \frac{\|x_{k+1} - y_{k+1}\|^2}{2\tau_{k+1}}. \quad (3.3.2)$$

En effet, à chaque fois que le test n'est pas validé, le pas est multiplié par  $\rho$ . Ainsi, si  $\rho$  est proche de 1, la valeur du pas diminue assez peu à chaque test non validé, ce qui peut entraîner un nombre d'itérations de backtracking important pour trouver un pas acceptable. Au contraire, si  $\rho$  est proche de 0, on réduit grandement la valeur du pas en cas d'échec de la condition (3.3.2) et on fixe potentiellement un pas trop petit. En pratique, on utilisera des valeurs comprises entre 0.75 et 0.9.

- $\delta \in (0, 1]$  est le paramètre qui gère la façon dont le pas peut augmenter à chaque itération. Si  $\delta = 1$ , on applique une stratégie de backtracking monotone semblable à la règle d'Armijo puisque la valeur du pas au cours des itérations est nécessairement décroissante. Pour un choix de  $\delta < 1$ , on augmente la valeur du pas accepté à l'itération précédente au début de chaque

itération en choisissant  $\frac{\tau_k}{\delta}$  comme pas de départ (si cette valeur est inférieure à  $\frac{1}{L_{min}}$ ). Autoriser le pas à augmenter à chaque itération permet au schéma de **s'adapter à la géométrie locale de la fonction  $F$** . En effet, la constante de Lipschitz  $L$  telle qu'elle est définie dans ce manuscrit est une constante globale. Sa valeur permet d'avoir une idée de comment la fonction  $F$  varie sur l'ensemble  $\mathbb{R}^N$  tout entier. Par contre, elle ne donne pas d'information sur la géométrie locale de  $F$  : la fonction  $F$  peut avoir une constante de Lipschitz très élevée mais être très plate au voisinage de ses minimiseurs, ce qui signifie que définir le pas égal à  $\frac{1}{L}$  est un choix non optimal lorsque les itérés s'approchent de  $X^*$ . Le meilleur choix pour le pas peut également dépendre de la direction de descente. En ce sens, le choix  $\delta < 1$  permet de s'adapter à la géométrie locale.

- $L_0$  et  $L_{min}$  sont deux estimations préalables de  $L$  qui ne jouent pas le même rôle. La constante  $L_0$  est celle qui permettra au schéma d'initialiser le pas pour FISTA  $\tau_0 = \frac{1}{L_0}$ . Idéalement,  $L_0$  doit être choisi proche de la vraie valeur de  $L$  pour éviter des itérations de backtracking mais l'algorithme converge quelle que soit sa valeur. La constante  $L_{min}$  est plus artificielle et a un intérêt théorique puisqu'elle permet de s'assurer que la suite  $(\tau_k)_{k \geq 0}$  est bornée. Cette propriété sera nécessaire pour obtenir des résultats de convergence sur notre schéma de restart. Elle doit vérifier la condition  $L_{min} < L$  qui n'est pas restrictive puisque l'on peut choisir  $L_{min}$  aussi faible qu'on le souhaite.

L'Algorithme 5 est construit autour de l'inégalité suivante qui est vérifiée pour tout  $\bar{x} \in \mathbb{R}^N$  avec  $\hat{x} := T_\tau \bar{x}$  et  $\tau > 0$  :

$$\forall x \in \mathbb{R}^N, F(\hat{x}) + \frac{\|x - \hat{x}\|^2}{2\tau} + \left( \frac{\|\hat{x} - \bar{x}\|^2}{2\tau} - D_f(\hat{x}, \bar{x}) \right) \leq F(x) + \frac{\|x - \bar{x}\|^2}{2\tau}. \quad (3.3.3)$$

En choisissant  $x = \bar{x}$  dans l'inégalité ci-dessus, on peut voir que la valeur de la fonctionnelle  $F$  diminue entre  $\bar{x}$  et  $\hat{x}$  à partir du moment où

$$D_f(\hat{x}, \bar{x}) \leq \frac{\|\hat{x} - \bar{x}\|^2}{2\tau}, \quad (3.3.4)$$

Cette condition est toujours vérifiée si  $\tau \in (0, \frac{1}{L}]$ , ce qui donne le lemme de descente [51, Lemme 3.1] très utile pour prouver la convergence de FISTA. L'idée de l'Algorithme 5 est de forcer l'inégalité (3.3.4) à être vraie via la condition (3.3.2) et la règle de mise à jour du pas  $\tau_k$ . On peut montrer qu'avec une telle règle de backtracking, on a pour tout  $k \geq 0$  :

$$\tau_k \geq \min \left\{ \tau_0, \frac{\rho}{L} \right\}, \quad (3.3.5)$$

et ainsi, si on note  $L_k := \frac{1}{\tau_k}$  pour tout  $k \geq 0$ ,

$$L_k \leq \max \left\{ L_0, \frac{L}{\rho} \right\}. \quad (3.3.6)$$

En notant  $\bar{\rho} := \min \left\{ \rho, \frac{L}{L_0} \right\}$ , on a donc  $\tau_k \geq \frac{\bar{\rho}}{L}$  et  $L_k \leq \frac{L}{\bar{\rho}}$ .

Je recommande la lecture de [49, Section 4] pour plus de détails sur l'analyse de ce schéma. On utilise le résultat suivant qui est une conséquence de [49, Théorème 4.1].

**Théorème 3.3.1.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  pour  $L > 0$ . Soient  $(x_k)_{k \geq 0}$  et  $(\tau_k)_{k \geq 0}$  les suites définies dans l'Algorithme 5 avec  $L_{\min} < L$ ,  $\rho \in (0, 1)$  et  $\delta \in (0, 1]$ . On a pour tout  $k > 0$  et pour tout  $x^* \in X^*$  :*

$$F(x_k) - F^* \leq \frac{2\bar{L}_k}{k^2} \|x_0 - x^*\|^2, \quad (3.3.7)$$

où  $L_k := \frac{1}{\tau_k}$  et

$$\bar{L}_k := \left( \frac{1}{k} \sum_{i=1}^k \frac{1}{\sqrt{L_i}} \right)^{-2}. \quad (3.3.8)$$

Plus précisément, nous exploitons la proposition ci-dessous.

**Proposition 3.3.2.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(x_k)_{k \geq 0}$  et  $(\tau_k)_{k \geq 0}$  les suites définies dans l'Algorithme 5 avec  $L_{\min} < L$ ,  $\rho \in (0, 1)$  et  $\delta \in (0, 1]$ . On a alors les résultats suivants :*

1. Pour tout  $k \in \mathbb{N}^*$ ,

$$F(x_k) - F^* \leq \frac{4L}{\bar{\rho}\mu k^2} (F(x_0) - F^*), \quad (3.3.9)$$

avec  $\bar{\rho} = \min \left\{ \rho, \frac{L}{L_0} \right\}$ .

2. Pour tout  $k \in \mathbb{N}^*$ ,

$$F(x_k) \leq F(x_0). \quad (3.3.10)$$

**Preuve.**

Compte tenu du contrôle que l'on connaît sur la suite  $L_k$ , on peut déduire que  $\bar{L}_k \leq \frac{L}{\bar{\rho}}$ . D'après le Théorème 3.3.1, la suite  $(x_k)_{k \geq 0}$  donnée par l'Algorithme 5 vérifie donc

$$\forall k > 0, \quad F(x_k) - F^* \leq \frac{2L}{\bar{\rho}k^2} \|x_0 - x^*\|^2, \quad (3.3.11)$$

pour tout  $x^* \in X^*$ .

Comme  $F$  satisfait  $\mathcal{G}_\mu^2$  pour un certain  $\mu > 0$ , on obtient que

$$\forall k > 0, \quad F(x_k) - F^* \leq \frac{4L}{\bar{\rho}\mu k^2} (F(x_0) - F^*). \quad (3.3.12)$$

Le deuxième point provient du fait que la condition suivante est vraie pour tout  $k \geq 0$

$$D_f(x_{k+1}, y_{k+1}) \leq \frac{\|x_{k+1} - y_{k+1}\|^2}{2\tau_{k+1}}. \quad (3.3.13)$$



En effet, l'inégalité (3.3.3) devient donc :

$$\begin{aligned}
F(x_{k+1}) + \frac{\|x_{k+1} - x_k\|^2}{2\tau_{k+1}} &\leq F(x_k) + \frac{\|y_{k+1} - x_k\|^2}{2\tau_{k+1}} \\
&\leq F(x_k) + \frac{(t_k - 1)^2 \|x_k - x_{k-1}\|^2}{t_{k+1}^2 2\tau_{k+1}} \\
&\leq F(x_k) + \frac{(t_k - 1)^2 \tau_k}{t_{k+1}^2 \tau_{k+1}} \frac{\|x_k - x_{k-1}\|^2}{2\tau_k}. \quad (3.3.14)
\end{aligned}$$

Par définition de la suite  $(t_k)_{k \geq 0}$ , on a que  $\tau_{k+1} t_{k+1} (t_{k+1} - 1) = \tau_k t_k^2$ . Par conséquent :

$$\frac{(t_k - 1)^2 \tau_k}{t_{k+1}^2 \tau_{k+1}} = \frac{(t_k - 1)^2 t_{k+1} (t_{k+1} - 1)}{t_k^2 t_{k+1}^2} \leq 1,$$

ainsi, par (3.3.14) on a :

$$F(x_{k+1}) + \frac{\|x_{k+1} - x_k\|^2}{2\tau_{k+1}} \leq F(x_k) + \frac{\|x_k - x_{k-1}\|^2}{2\tau_k}, \quad (3.3.15)$$

pour tout  $k \geq 0$ , ce qui garantit le deuxième résultat.  $\square$

Ces résultats ont la même forme que ceux de la Proposition 3.2.1 et on peut donc appliquer un raisonnement similaire. En effet, le premier point assure que pour tout  $k > 0$

$$\mu \leq \frac{4L}{\bar{\rho} k^2} \frac{F(x_0) - F^*}{F(x_k) - F^*}, \quad (3.3.16)$$

et comme le deuxième point implique que  $u \mapsto \frac{F(x_0) - u}{F(x_k) - u}$  est une fonction croissante sur  $[F^*, F(x_k))$ , on a pour tout  $y \in \mathbb{R}^N$  tel que  $F(y) < F(x_k)$  :

$$\mu \leq \frac{4L}{\bar{\rho} k^2} \frac{F(x_0) - F(y)}{F(x_k) - F(y)}. \quad (3.3.17)$$

Cependant, on se positionne dans le contexte où la constante  $L$  n'est pas connue exactement et cette borne ne sera donc pas estimable. C'est pourquoi on choisira plutôt d'estimer le conditionnement  $\kappa := \frac{\mu}{L}$ . Dans cette même logique, on fera l'hypothèse que l'on a une estimation basse  $L_0 < \frac{L}{\rho}$  (donc  $\bar{\rho} = \rho$ ) puisque dans le cas contraire,  $\bar{\rho}$  n'est pas accessible ( $\bar{\rho} = \frac{L}{L_0}$ ). On s'appuiera donc sur l'inégalité

$$\kappa \leq \frac{4}{\rho k^2} \frac{F(x_0) - F(y)}{F(x_k) - F(y)} \quad (3.3.18)$$

pour construire notre schéma.

### 3.3.3 Free-FISTA

Free-FISTA est un schéma de restart de l'Algorithme 5 qui suit les mêmes principes que l'Algorithme 3. Il est construit autour de l'estimation du paramètre  $\kappa = \frac{\mu}{L}$

qui est l'inverse du conditionnement. La valeur des différentes estimations influe sur la fréquence des redémarrages. On donne son expression dans l'Algorithme 6.

---

**Algorithme 6** : Free-FISTA
 

---

**Entrée:**  $x_0 \in \mathbb{R}^N$ ,  $n > 0$ ,  $L_0 > 0$ ,  $L_{min} \ll L$ ,  $\rho \in (0, 1)$ ,  $\delta \in (0, 1]$ ,  $C > 4/\sqrt{\rho}$ .

$n_0 = \lfloor 2C \rfloor$  ;

$(r_1, \tilde{L}_1) = \text{FISTA\_adaBT}(r_0, n_0, L_0, L_{min}; \rho, \delta)$  ;

$n_1 = \lfloor 2C \rfloor$  ;

$(\hat{r}_1, \hat{L}_1) = \text{FB\_BT}(r_1, \tilde{L}_1; \rho)$  ;

**répéter**

$j = j + 1$  ;

$(r_j, \tilde{L}_j) = \text{FISTA\_adaBT}(\hat{r}_{j-1}, n_{j-1}, \hat{L}_{j-1}, L_{min}; \rho, \delta)$  ;

$\tilde{\kappa}_j = \min_{i < j} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}$  ;

**si**  $n_{j-1} \leq \frac{C}{\sqrt{\tilde{\kappa}_j}}$  **alors**

$n_j = 2n_{j-1}$  ;

**sinon**

$n_j = n_{j-1}$  ;

**fin si**

$(\hat{r}_j, \hat{L}_j) = \text{FB\_BT}(r_j, \tilde{L}_j; \rho)$  ;

**jusqu'à**  $\|g_{1/\hat{L}_j}(r_j)\| \leq \varepsilon$ .

**retourne**  $r = \hat{r}_j$  .

---

### 3.3.3.1 Le critère d'arrêt

Parmi les différences avec le schéma de restart proposé précédemment (Algorithme 3), on doit relever que le critère d'arrêt est adapté pour faire face à l'hypothèse que  $L$  n'est pas directement accessible. En effet, comme  $L$  n'est pas connu, on ne peut pas calculer directement le gradient composite  $g$  tel qu'il est défini dans la section précédente. On définit donc le gradient composite généralisé comme

$$g_\tau : y \mapsto \frac{y - T_\tau(y)}{\tau}.$$

Notons que le gradient composite classique utilisé dans la section précédente correspond à  $g_{1/L}$ .

On définit le critère d'arrêt de Free-FISTA comme la condition  $\|g_{1/\hat{L}_j}(r_j)\| \leq \varepsilon$  où  $\hat{L}_j$  est obtenu via une itération de Forward-Backward avec backtracking d'Armijo (Algorithme 7) :

$$(\hat{r}_j, \hat{L}_j) = \text{FB\_BT}(r_j, \tilde{L}_j; \rho). \quad (3.3.19)$$

---

**Algorithme 7** : Étape de Forward-Backward avec backtracking d'Armijo,  $\text{FB\_BT}(r, L_0; \rho)$

---

**Entrée:**  $i = 0, r \in \mathbb{R}^N, L_0 > 0, \rho \in (0, 1)$ .

**répéter**

$$\tau = \frac{\rho^i}{L_0};$$

$$\hat{r} = T_\tau r;$$

$$i = i + 1;$$

**jusqu'à**  $D_f(\hat{r}, r) \leq \|\hat{r} - r\|^2 / 2\tau$ .

**retourne**  $\hat{r}, \frac{L_0}{\rho^{i-1}}$ .

---

Au vu de la définition de chaque variable, on obtient que  $g_{1/\hat{L}_j}(r_j) = \hat{L}_j(r_j - \hat{r}_j)$  ce qui signifie que l'évaluation du critère d'arrêt a un coût de calcul très faible.

Ce choix pour la condition de sortie est motivé par le lemme ci-dessous qui est une version plus générale du Lemme 3.2.5. La preuve se trouve en Section 3.A.3.7.

**Lemme 3.3.3.** *Soit  $F$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  avec  $L > 0$  et  $\mu > 0$ . Alors pour tout  $x \in \mathbb{R}^N$  et  $\tau > 0$  on a :*

$$F(T_\tau x) - F^* \leq \frac{2(1 + L\tau)^2}{\mu} \|g_\tau(x)\|^2. \quad (3.3.20)$$

En appliquant ce lemme, on obtient que si le critère d'arrêt est satisfait pour  $j > 0$ , alors :

$$F(\hat{r}_j) - F^* \leq \frac{(1 + L/\hat{L}_j)^2 \varepsilon^2}{2\mu}. \quad (3.3.21)$$

Comme pour toute itération  $k$  de FISTA avec backtracking on a  $\tau_k \leq \frac{1}{L_{\min}}$ , on déduit que  $\tilde{L}_j \geq L_{\min}$ . De plus, la procédure utilisée dans l'Algorithme 7 assure que  $\hat{L}_j \geq \tilde{L}_j$  (le pas ne peut pas être augmenté dans cette étape). On obtient finalement que

$$F(\hat{r}_j) - F^* \leq \frac{(1 + L/L_{\min})^2 \varepsilon^2}{2\mu}. \quad (3.3.22)$$

### 3.3.3.2 Structure du schéma

La structure de Free-FISTA est globalement la même que celle de l'Algorithme 3. On utilise notamment des variables très similaires :

- $(r_j)_{j \geq 0}$  est la suite contenant les vecteurs résultant de chaque exécution complète de FISTA avec backtracking. Free-FISTA s'appuie sur une suite additionnelle  $(\hat{r}_j)_{j \geq 1}$  qui joue un rôle pour obtenir des garanties sur la décroissance de l'erreur à partir du critère d'arrêt (voir section précédente). Pour  $j > 0$ ,  $\hat{r}_j$  est en fait le vecteur obtenu à partir d'une étape de Forward-Backward (ou gradient proximal) avec backtracking d'Armijo sur  $r_j$  et c'est à partir de ce point que l'on relance FISTA backtracké à l'itération suivante.

- $(n_j)_{j \geq 0}$  est la suite qui contient le nombre d'itérations avant chaque restart de FISTA avec backtracking. On détaille la règle de mise à jour par la suite.
- $(\tilde{\kappa}_j)_{j \geq 2}$  est la suite des estimations du conditionnement.
- $(\tilde{L}_j)_{j \geq 1}$  contient les estimations de la constante de Lipschitz données par l'Algorithme 5 à chaque restart. Le choix qui est fait ici est de prendre l'inverse du pas de la dernière itération de FISTA avec backtracking. On aurait pu choisir  $\tilde{L}_j$  de la manière suivante :

$$\tilde{L}_j = \left( \frac{1}{\frac{1}{n_{j-1}} \sum_{k=1}^{n_{j-1}} \frac{1}{\sqrt{L_k}}} \right)^2. \quad (3.3.23)$$

On préfère cependant choisir la dernière estimation qui semble plus adaptée à la géométrie locale. Dans les deux cas, on peut facilement montrer que  $\tilde{L}_j \leq \frac{L}{\bar{\rho}}$  avec  $\bar{\rho} = \min \left\{ \rho, \frac{L}{L_0} \right\}$ . De manière similaire,  $(\hat{L}_j)_{j \geq 1}$  est la suite qui contient les estimations de  $L$  obtenues après l'étape de Forward-Backward additionnelle.

À chaque redémarrage, on calcule une estimation du conditionnement  $\kappa$  :

$$\forall j \geq 2, \tilde{\kappa}_j = \min_{i < j} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}, \quad (3.3.24)$$

qui vérifie le lemme ci-dessous.

**Lemme 3.3.4.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soit  $\rho \in (0, 1)$ . Si  $L_0 < \frac{L}{\rho}$ , alors la suite  $(\tilde{\kappa}_j)_{j \geq 2}$  définie dans l'Algorithme 6 avec les paramètres  $C > 4/\sqrt{\rho}$ ,  $L_0$ ,  $L_{\min} < L$ ,  $\rho$ ,  $\delta \in (0, 1]$  et  $\varepsilon > 0$  vérifie l'inégalité suivante :*

$$\forall j \geq 2, \tilde{\kappa}_j \geq \tilde{\kappa}_{j+1} \geq \kappa. \quad (3.3.25)$$

**Preuve.**

Appliquons (3.3.18) à l'Algorithme 6. Étant donné la définition de la suite  $(r_j)_{j \geq 0}$ , on a que pour  $i \geq 1$  et  $y \in \mathbb{R}^N$  tel que  $F(y) \leq F(r_i)$ ,

$$\kappa \leq \frac{4}{\rho n_{i-1}^2} \frac{F(\hat{r}_{i-1}) - F(y)}{F(r_i) - F(y)}. \quad (3.3.26)$$

De plus, on a que pour tout  $i \geq 1$ ,  $F(\hat{r}_i) \leq F(r_i)$  et donc

$$\kappa \leq \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(y)}{F(r_i) - F(y)}. \quad (3.3.27)$$

Le deuxième point de la Proposition 3.3.2 garantit que pour tout  $j > i$ ,  $F(r_j) \leq F(\hat{r}_i) \leq F(r_i)$  et donc :

$$\kappa \leq \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}. \quad (3.3.28)$$

Cette inégalité étant vraie pour tout  $i < j$ , on peut passer au minimum pour obtenir que  $\kappa \leq \tilde{\kappa}_j$ .

De plus, la suite  $(\tilde{\kappa}_j)_{j \geq 2}$  est décroissante. On a que pour tout  $j \geq 2$ ,

$$\tilde{\kappa}_{j+1} = \min_{\substack{i \in \mathbb{N}^* \\ i < j+1}} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_{j+1})}{F(r_i) - F(r_{j+1})} \leq \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_{j+1})}{F(r_i) - F(r_{j+1})}. \quad (3.3.29)$$

La fonction  $u \mapsto \frac{F(r_{i-1})-u}{F(r_i)-u}$  étant croissante sur  $[F^*, F(r_i))$  pour tout  $i < j$ , l'inégalité  $F(r_{j+1}) \leq F(r_j)$  garantie par la Proposition 3.3.2 implique que :

$$\tilde{\kappa}_{j+1} \leq \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} = \tilde{\kappa}_j. \quad (3.3.30)$$

□

Le nombre d'itérations pour chaque exécution de FISTA avec backtracking est mis à jour en fonction de cette estimation de la même manière que dans l'Algorithme 3. La condition  $n_{j-1} \leq \frac{C}{\sqrt{\tilde{\kappa}_j}}$  est évaluée et si elle est vérifiée, le nombre d'itérations est doublé. Sinon, il reste inchangé. Le paramètre  $C$  joue le même rôle que dans l'Algorithme 3, c'est-à-dire qu'il impacte la fréquence des redémarrages (plus il est grand, moins les restarts sont fréquents). Cela permet à la suite  $(n_j)_{j \geq 0}$  de satisfaire le lemme suivant. La preuve est exactement la même que celle du Lemme 3.2.3.

**Lemme 3.3.5.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soit  $\rho \in (0, 1)$ . Si  $L_0 < \frac{L}{\rho}$ , alors la suite  $(n_j)_{j \geq 0}$  définie dans l'Algorithme 6 avec les paramètres  $C > 4/\sqrt{\rho}$ ,  $L_0$ ,  $L_{\min} < L$ ,  $\rho$ ,  $\delta \in (0, 1]$  et  $\varepsilon > 0$  vérifie l'inégalité suivante :*

$$\forall j \geq 0, \quad n_j \leq 2C \sqrt{\frac{L}{\mu}}. \quad (3.3.31)$$

### 3.3.3.3 Garanties théoriques

Les propriétés des différentes variables de Free-FISTA nous permettent d'obtenir le résultat de convergence suivant.

**Théorème 3.3.6.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  les suites définies dans l'Algorithme 3 avec  $\rho \in (0, 1)$ ,  $\delta \in (0, 1]$ ,  $C > 4/\sqrt{\rho}$ ,  $L_0 > 0$ ,  $L_{\min} < L$  et  $\varepsilon > 0$ . Si  $L_0 < \frac{L}{\rho}$ , alors le nombre d'itérations  $j + \sum_{i=0}^j n_i$  nécessaire pour assurer  $\|g_{1/\hat{L}_j}(r_j)\| \leq \varepsilon$  est borné et vérifie l'inégalité suivante :*

$$\sum_{i=0}^j n_i \leq \frac{4C}{\ln\left(\frac{C^2\rho}{4} - 1\right)} \sqrt{\frac{L}{\mu}} \left( 2 \ln\left(\frac{C^2\rho}{4} - 1\right) + \ln\left(1 + \frac{16}{C^2\rho - 16} \frac{2L(F(r_0) - F^*)}{\rho\varepsilon^2}\right) \right). \quad (3.3.32)$$

La preuve de ce théorème suit le même raisonnement que celle du Théorème 3.2.8. Elle est donnée en Section 3.A.3.3.

*Remarque 3.3.7.* Le nombre total d'itérations de l'Algorithme 6 si le dernier itéré est  $\hat{r}_j$  est  $j + \sum_{i=0}^j n_i$  car à chaque étape  $i$ , l'algorithme exécute  $n_i$  itérations de FISTA et une itération de Forward-Backward avec backtracking. Notons que l'on aurait pu définir l'itération additionnelle de Forward-Backward comme la première itération de l'exécution de FISTA auquel cas le nombre d'itérations total serait  $1 + \sum_{i=0}^j n_i$ . On privilégie la formulation donnée dans l'Algorithme 6 par souci de clarté.

Ce résultat garantit que le nombre d'itérations nécessaire pour atteindre la condition de sortie est borné et proportionnel à  $\sqrt{\frac{L}{\mu}}$ . Cette borne est valable à condition que l'estimation initiale  $L_0$  soit inférieure à  $\frac{L}{\rho}$ , c'est-à-dire à condition que celle-ci soit suffisamment basse. On privilégie donc le fait de choisir un pas initial trop grand pour ensuite le réduire si cela est nécessaire. Cependant, on peut montrer que Free-FISTA converge même si cette condition n'est pas vérifiée. De plus, en choisissant  $\delta < 1$ , les estimations de  $L$  sont autorisées à diminuer et pour  $j$  suffisamment grand, on aura  $\hat{L}_j < \frac{L}{\rho}$  ce qui nous permettra d'obtenir un résultat de convergence semblable à celui du Théorème 3.3.6.

Notons que le nombre d'itérations mentionné dans cette section ne prend pas en compte les itérations de backtracking.

Ce théorème implique le corollaire suivant qui donne une vitesse de décroissance asymptotique de l'erreur dans le pire des cas. Une preuve est fournie en Section 3.A.3.4.

**Corollaire 3.3.8.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  les suites définies dans l'Algorithme 3 avec  $\rho \in (0, 1)$ ,  $\delta \in (0, 1]$ ,  $C > 4/\sqrt{\rho}$ ,  $L_0 > 0$ ,  $L_{\min} < L$  et  $\varepsilon > 0$ . Si  $L_0 < \frac{L}{\rho}$ , alors on a :*

$$F(\hat{r}_j) - F^* = \mathcal{O} \left( e^{-\frac{\ln\left(\frac{C^2\rho-1}{4C}\right)}{4C} \sqrt{\kappa} \sum_{i=0}^j n_i} \right). \quad (3.3.33)$$

*En particulier, si on choisit  $C$  pour maximiser  $\frac{\ln\left(\frac{C^2\rho-1}{4C}\right)}{4C}$ , c'est-à-dire  $\hat{C} \approx 6.38/\sqrt{\rho}$ , alors il existe  $K > \frac{1}{12}$  tel que  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  vérifient*

$$F(\hat{r}_j) - F^* = \mathcal{O} \left( e^{-\sqrt{\rho}K\sqrt{\kappa} \sum_{i=0}^j n_i} \right). \quad (3.3.34)$$

Le taux de convergence donné dans l'équation (3.2.24) pourrait suggérer que le paramètre d'Armijo  $\rho$  doit être choisi très proche de 1. Cependant, on rappelle que le nombre d'itérations  $1 + \sum_{i=0}^j n_i$  ne prend pas en compte les itérations de backtracking. Ainsi, si  $\rho$  est proche de 1, la borne théorique assure que la vitesse de décroissance de l'erreur au pire cas est meilleure en terme de nombre d'itérations de

FISTA backtracké en comparaison avec une valeur plus faible de  $\rho$ . En revanche, le coût de chaque itération de FISTA est susceptible d'être élevé puisqu'avec  $\rho$  proche de 1, le pas n'est que très peu réduit à chaque itération de backtracking.

Cette vitesse de décroissance de l'erreur au pire cas est moins bonne que celle de l'Algorithme 3 puisque le nombre d'itérations nécessaire pour atteindre une certaine précision est divisé par  $\sqrt{\rho} < 1$  et les itérations de backtracking ne sont pas prises en considération. Cependant, la procédure de backtracking permet de s'affranchir de la connaissance de  $L$  ou de s'adapter à la géométrie locale de  $F$ . Le seul autre schéma proposant une estimation automatique de  $L$  et  $\mu$  [85] assure une décroissance de l'erreur en  $\mathcal{O}\left(e^{-\sqrt{\rho}\sqrt{\frac{\mu}{L}k}}\right)$  mais pour une fonction  $F$  fortement convexe. Le schéma est construit autour d'une adaptation de la méthode de gradient accéléré de Nesterov dans le cas fortement convexe [105] (ou de V-FISTA [31]).

### 3.3.4 Expériences numériques

Pour illustrer l'apport du backtracking combiné à une stratégie de restart, plusieurs exemples numériques sont présentés dans cette section. On compare Free-FISTA avec la version originale de FISTA, le schéma de restart introduit précédemment (Algorithme 3) et la version de FISTA avec backtracking (Algorithme 5).

Les deux premiers exemples sont des situations favorables à l'application de Free-FISTA et le troisième souligne des limitations liées aux schémas de restart.

#### 3.3.4.1 Régression logistique avec régularisation $\ell^2$ - $\ell^1$

On s'intéresse dans un premier temps à un problème de classification où on considère une matrice d'entrées  $A \in \mathbb{R}^{m \times n}$  et le vecteur de sortie  $b \in \{-1, 1\}^m$  pour  $n > 0$  et  $m > 0$ . Le problème d'optimisation que l'on résout est le suivant :

$$\min_{x \in \mathbb{R}^n} F(x) := \underbrace{\frac{\lambda_1}{2\|A^T b\|_\infty} \sum_{j=1}^m \ln\left(1 + e^{-b_j a_j^T x}\right)}_{:=f(x)} + \frac{\lambda_2}{2}\|x\|^2 + \underbrace{\|x\|_1}_{:=h(x)}, \quad (3.3.35)$$

où  $a_j = (A_{i,j})_i \in \mathbb{R}^m$  est la  $j$ -ème ligne de  $A$ ,  $\lambda_1 > 0$  et  $\lambda_2 > 0$ . Par définition, le vecteur  $x^*$  solution de ce problème vise à satisfaire  $\mathbb{P}(b_i = 1|a_i) = \frac{1}{1+e^{-a_i^T x^*}}$  pour tout  $i \in \llbracket 1, m \rrbracket$ . Notons que cette régularisation également appelée *elastic net* permet de rendre  $F$  plus fortement convexe via la régularisation  $\ell^2$  et de favoriser les solutions parcimonieuses via le terme de régularisation  $\ell^1$ . Ainsi, la fonction  $F$  a une croissance quadratique au voisinage de ses minimiseurs et vérifie donc  $\mathcal{G}_\mu^2$  pour un paramètre de croissance  $\mu$  qui n'est pas facilement estimable. En revanche, une estimation haute de la constante de Lipschitz de  $\nabla f$  peut être calculée :

$$\hat{L} = \frac{\lambda_1 \|A^T b\|^2}{8\|A^T b\|_\infty} + \lambda_2, \quad (3.3.36)$$

qui peut être très grande lorsque  $\|A^T b\| \gg 1$ .

On résout ce problème pour un jeu de données généré aléatoirement avec  $n = 30000$  et  $m = 100$ . On compare donc les méthodes suivantes :

- FISTA [32] avec un pas de  $\tau = \frac{1}{L}$  ;
- FISTA restart (Algorithme 3 avec un pas de  $\tau = \frac{1}{L}$ ,
- FISTA\_adaBT (Algorithme 5) avec  $\rho = 0.8$ ,  $\delta = 0.95$ ,  $L_0 = 1$  (estimation basse choisie arbitrairement),  $L_{min} = 10^{-16}$  (précision machine),
- Free-FISTA (Algorithme 6) avec  $\rho = 0.8$ ,  $\delta = 0.95$ ,  $L_0 = 1$  (estimation basse choisie arbitrairement),  $L_{min} = 10^{-16}$  (précision machine), et  $C = 6.38/\sqrt{\rho}$ .

On choisit  $\lambda_1 = 10$ ,  $\lambda_2 = 3$  et  $x_0 \leftarrow \mathcal{U}([-1, 1])$ . On obtient  $\hat{L} \geq 900000$  comme borne supérieure de la constante de Lipschitz de  $\nabla F$ . Une estimation de la solution de (3.3.35) est calculée préalablement en appliquant Free-FISTA pour un grand nombre d'itérations. Cela nous permet de représenter l'évolution de  $\ln(F(x_k) - \hat{F})$  où  $\hat{F}$  correspond à cette estimation.

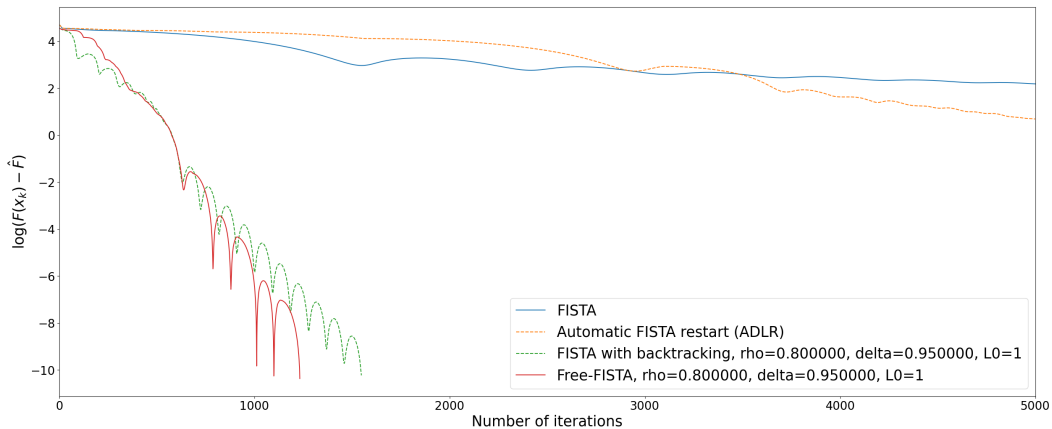


FIGURE 3.6 – Évolution de l'erreur en fonction du nombre d'itérations de FISTA (on ne compte pas les itérations de backtracking) pour le problème de régression logistique.

La Figure 3.6 nous montre les vitesses de convergence de chaque méthode présentée ci-dessus en fonction du nombre d'itérations de FISTA sans compter les itérations de backtracking. On peut noter que l'utilisation du backtracking permet d'améliorer les performances de FISTA et FISTA restart en termes de nombre d'itérations. De plus, la combinaison du backtracking et du restart semble pertinente ici puisque Free-FISTA présente le meilleur de taux de convergence. Cependant, comme les étapes de backtracking ne sont pas prises en compte, on ne peut pas directement conclure sur la performance du schéma en terme de temps de calcul. C'est pourquoi on affiche les courbes de convergence en fonction du temps de calcul sur la Figure 3.7. On peut observer que les calculs additionnels liés au backtracking n'empêchent pas Free-FISTA d'être la méthode plus rapide.



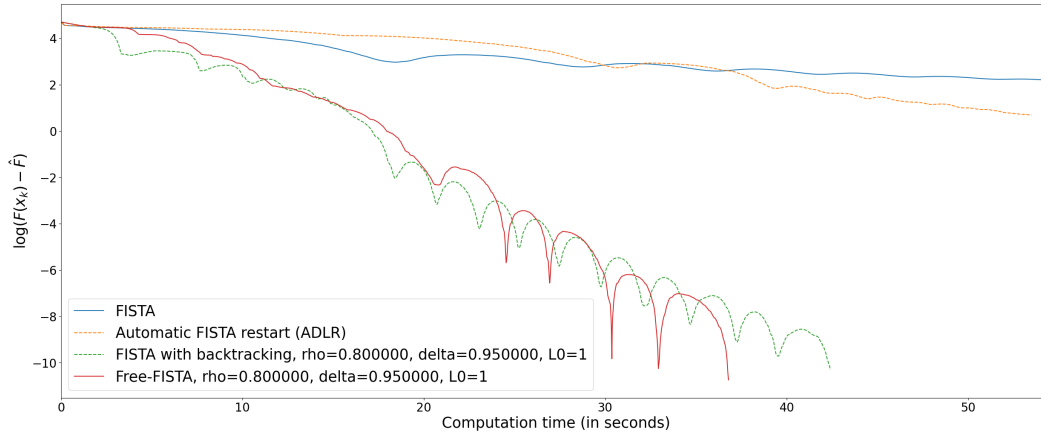


FIGURE 3.7 – Évolution de l’erreur en fonction du temps de calcul pour le problème de régression logistique.

On affiche sur la Figure 3.8 l’évolution de l’erreur en fonction du temps de calcul pour Free-FISTA avec différents choix de paramètres. On choisit  $\rho = 0.8$ ,  $\delta \in \{0.95, 0.995\}$  and  $L_0 \in \{1, \hat{L}\}$  où  $\hat{L}$  est l’estimation haute donnée dans l’équation (3.3.36) et 1 est une estimation basse arbitraire. Ce graphe montre que la performance de Free-FISTA n’est pas très impactée par la variation de ces paramètres dans ce contexte. Le choix  $\delta = 0.95$  paraît plus pertinent que  $\delta = 0.995$  dans cette situation et cela peut s’expliquer : la constante de Lipschitz de  $\nabla f$  est estimée très grossièrement par  $\hat{L}$  ou 1 et il est donc intéressant de choisir un  $\delta$  plus faible pour explorer des choix de pas différents plus efficacement. Il est intéressant de noter que l’algorithme est efficace lorsque l’on choisit  $L_0 = \hat{L}$  alors que l’on peut supposer que dans ce cas  $L_0 > \frac{L}{\rho}$  et l’hypothèse nécessaire à l’application du Théorème 3.3.6 n’est donc pas satisfaite. En fait, le choix de  $\delta < 1$  nous permet de diminuer notre estimation de  $L$  à chaque itération comme on va le voir ci-dessous.

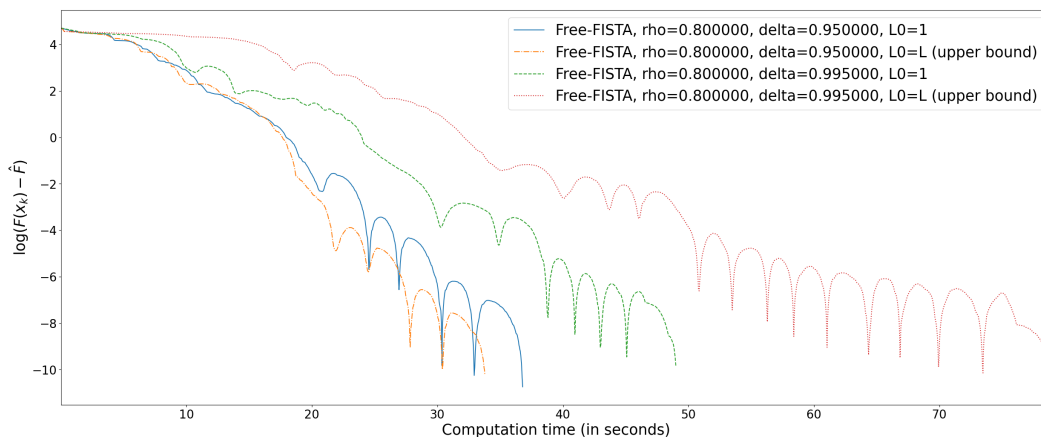


FIGURE 3.8 – Évolution de l’erreur en fonction du temps de calcul pour les itérés de Free-FISTA avec différents jeux de paramètres.

La Figure 3.9 donne un aperçu des estimations de la constante de Lipschitz  $L$  en fonction du nombre d'itérations pour chaque jeu de paramètres. Comme on l'a mentionné ci-dessus, la borne théorique  $\hat{L}$  s'avère très grande par rapport aux estimations calculées par Free-FISTA au cours des itérations :  $\hat{L} \geq 900000$  alors que les dernières estimations de Free-FISTA valent approximativement 3000. Cet écart important explique la différence de performance entre les schémas avec et sans backtracking sur la 3.6 puisqu'une estimation très grande de  $L$  a pour conséquence un pas très petit. On peut remarquer également sur la Figure 3.9 qu'une valeur plus faible de  $\delta$  autorise plus de variations au niveau des estimations de  $L$  ce qui permet une plus grande flexibilité.

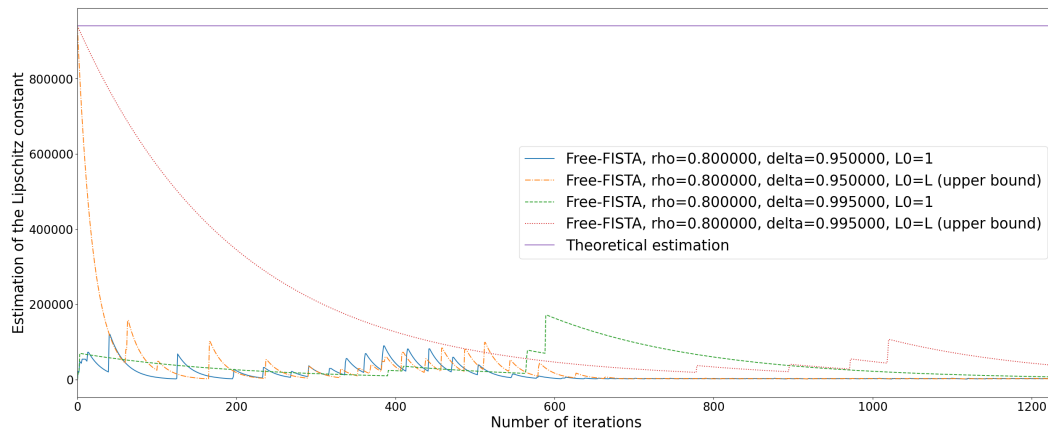


FIGURE 3.9 – Estimation de la constante de Lipschitz de  $\nabla f$  en fonction du nombre d'itérations de FISTA pour la régression logistique.

On compare l'impact qu'a le choix de  $L_0$  sur le nombre d'itérations de backtracking par itération de FISTA dans la Figure 3.10. On peut remarquer qu'en prenant une estimation basse de  $L$  pour  $L_0$ , cela force la première itération de FISTA à faire un grand nombre d'étapes de backtracking pour trouver un  $\tilde{L}$  tel que le pas  $\frac{1}{\tilde{L}}$  est admissible. Après cette première itération très coûteuse, l'estimation  $\tilde{L}$  ainsi obtenue permet de choisir un pas intéressant et le nombre d'itérations de backtracking diminue considérablement par la suite. En choisissant  $L_0$  comme une estimation haute de  $L$ , le nombre d'itérations requis est bien moins important mais l'estimation de  $L$  reste très haute pendant assez longtemps, ce qui implique que l'on choisit des pas assez petits. Dans cet exemple, les deux approches donnent des performances très proches : le grand nombre d'itérations de backtracking dans un cas est compensé par des pas sous-optimaux dans l'autre.

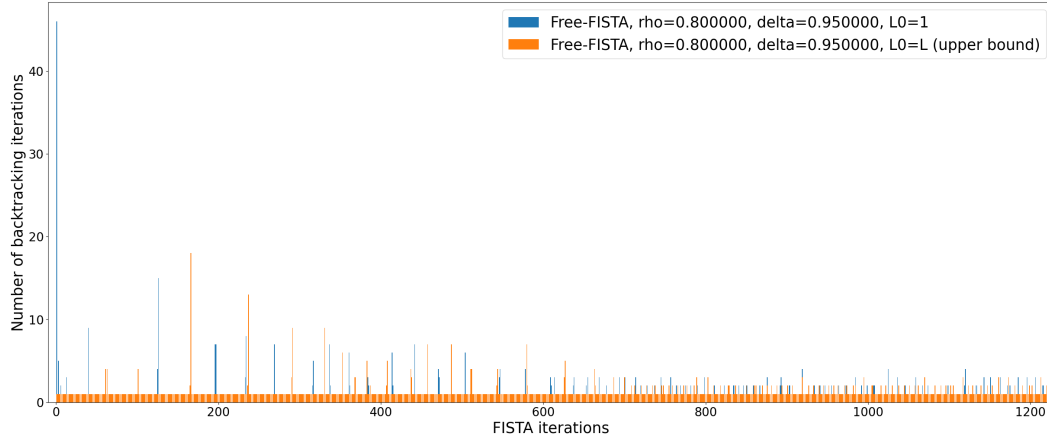


FIGURE 3.10 – Nombre d’itérations de backtracking par itération de FISTA pour le problème de régression logistique.

On donne ci-dessous une deuxième application de Free-FISTA au problème (3.3.35) inspirée par Fercoq et Qu [61]. On considère le jeu de données *dorothea* ( $n = 100000$  et  $m = 800$ ) avec  $\lambda_1 = 10$  et  $\lambda_2 = \frac{\lambda_1 \|A^T b\|^2}{80 \|A^T b\|_\infty n} = 0.9097$ . Le Tableau 3.2 donne le temps de calcul nécessaire pour atteindre une précision de  $\varepsilon = 10^{-5}$  en appliquant les méthodes présentées précédemment à ce problème. Cette expérience confirme l’intérêt du backtracking pour ce problème et de sa combinaison avec du restart puisque Free-FISTA donne les meilleurs résultats. On a choisi ici une estimation basse arbitraire  $L_0 = 1$  pour rentrer dans le cadre du Théorème 3.3.6.

| Algorithme   | $\rho$ | $\delta$ | Temps de calcul (sec.) |
|--------------|--------|----------|------------------------|
| FISTA        | -      | -        | 28594                  |
| Algorithme 3 | -      | -        | 12825                  |
| FISTA_adaBT  | 0.85   | 0.95     | 3292                   |
|              | 0.8    | 0.95     | 2348                   |
| Free-FISTA   | 0.85   | 0.95     | 1173                   |
|              | 0.8    | 0.95     | <b>989</b>             |

TABLEAU 3.2 – Temps de calcul nécessaire (moyenne sur 5 occurrences) pour la résolution de (3.3.35) à partir du jeu de données *dorothea* ( $n = 100000$  et  $m = 800$ ) avec  $\lambda_1 = 10$ ,  $\lambda_2 = \frac{\lambda_1 \|A^T b\|^2}{80 \|A^T b\|_\infty n} = 0,9097$  et  $\varepsilon = 10^{-5}$ .

## 3.3.4.2 Problème d'inpainting

Dans ce deuxième exemple, on étudie le problème d'inpainting déjà décrit en Section 3.2.3.2. Rappelons que la fonction à minimiser s'écrit comme suit

$$\arg \min_x F(x) := \underbrace{\frac{1}{2} \|Mx - y\|^2}_{:=f(x)} + \underbrace{\lambda \|Tx\|_1}_{:=h(x)}, \quad (3.3.37)$$

où  $M$  est un opérateur de masquage,  $x^0$  est l'image originale,  $y = Mx^0$  est l'image détériorée et  $T$  est une transformée en ondelettes.

Comme cela a été précisé précédemment, la fonction  $F$  satisfait  $\mathcal{G}_\mu^2$  pour un paramètre  $\mu$  difficile à estimer et  $\nabla f$  est une fonction 1-Lipschitz. Dans un tel contexte, comme la constante de Lipschitz est connue, il pourrait sembler inutile d'utiliser une stratégie de backtracking. On montre cependant que Free-FISTA est efficace dans cette situation puisque le backtracking permet de s'adapter à la géométrie locale de  $F$  pour un coût assez faible. Les tests ci-dessous sont effectués pour une image de  $225 \times 225$  pixels avec la base d'ondelettes Daubechies 4 et  $\lambda = 2$ . On compare les mêmes méthodes que dans la section précédente.

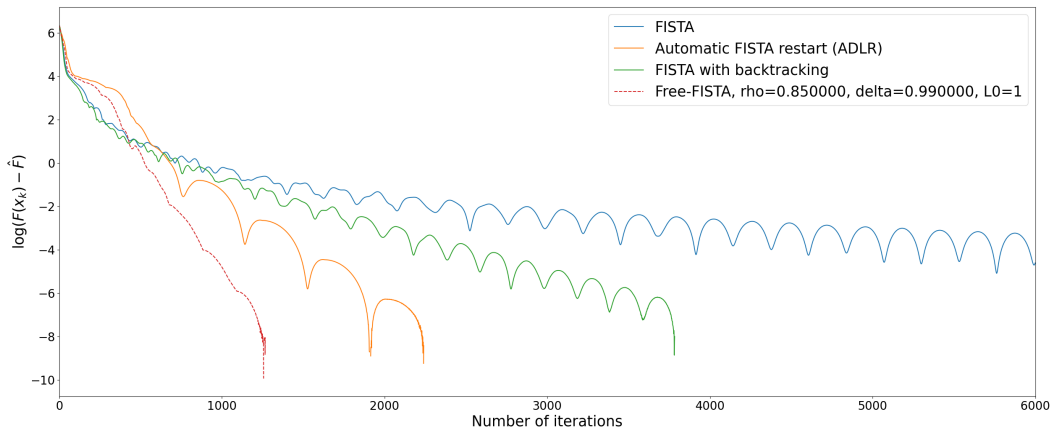


FIGURE 3.11 – Évolution de l'erreur en fonction du nombre d'itérations de FISTA sans compter les itérations de backtracking pour le problème d'inpainting.

On peut voir sur la Figure 3.11 que la stratégie de backtracking améliore légèrement la convergence de FISTA avec ou sans restart. L'apport du backtracking est moins significatif que pour le problème de régression logistique car notre estimation théorique de la constante de Lipschitz  $L = 1$  est précise. La Figure 3.12 illustre que le coût additionnel lié aux itérations de backtracking n'affecte pas l'efficacité des schémas correspondant en terme de temps de calcul. Cela est dû au faible coût que représente l'évaluation de  $f$  dans cet exemple.

Dans la Figure 3.13, on compare la performance de Free-FISTA pour différentes valeurs de  $\delta$  avec l'Algorithme 3. On peut observer que les meilleurs résultats sont obtenus pour une valeur de  $\delta$  très proche de 1 contrairement à l'exemple précédent.

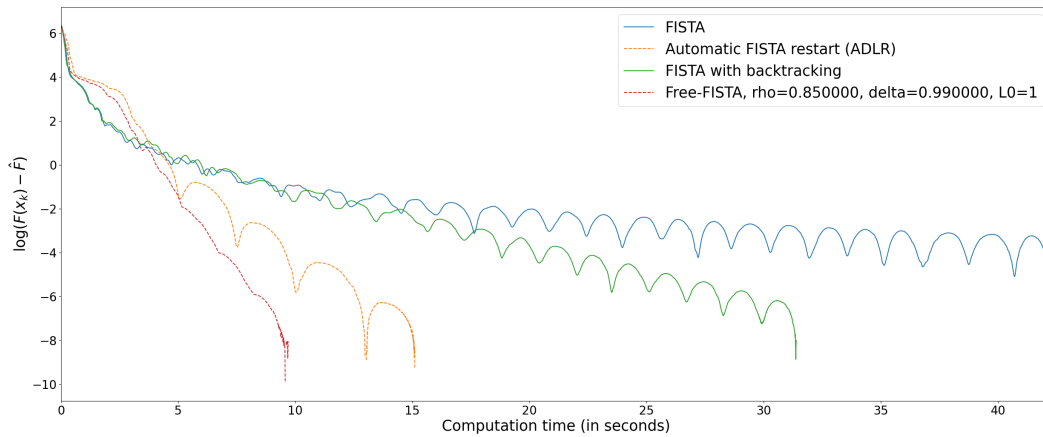


FIGURE 3.12 – Évolution de l’erreur en fonction du temps de calcul pour le problème d’inpainting.

En effet, si  $\delta$  est petit ( $\delta = 0.95$ ), beaucoup d’itérations de backtracking évitables ont lieu pour compenser le fait que l’on augmente le pas trop brutalement à chaque itération. Cela ralentit donc l’algorithme. Ce phénomène est souligné dans la Figure 3.14 et la Figure 3.15. En choisissant  $\delta = 0.99$ , les estimations de  $L$  varient moins au cours du temps et le nombre d’itération de backtracking par itération de FISTA diminue grandement.

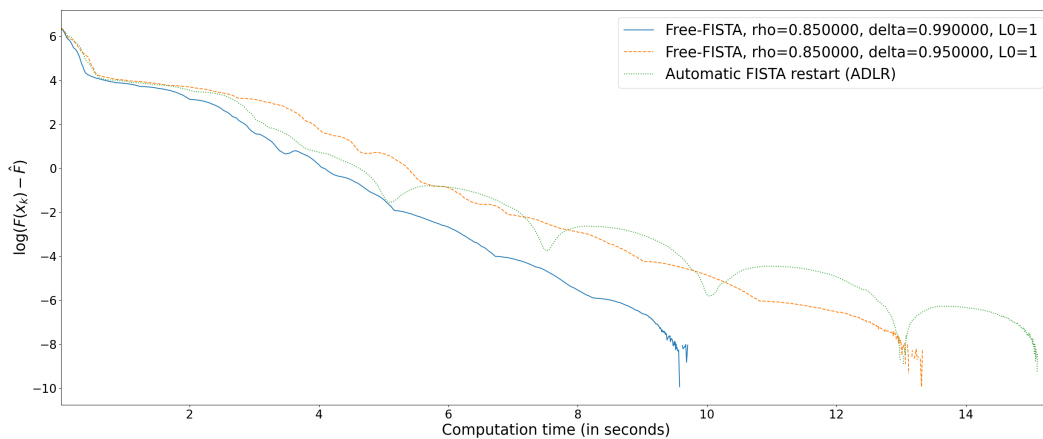


FIGURE 3.13 – Évolution de l’erreur en fonction du temps de calcul pour Free-FISTA et l’Algorithme 3 avec plusieurs jeux de paramètres.

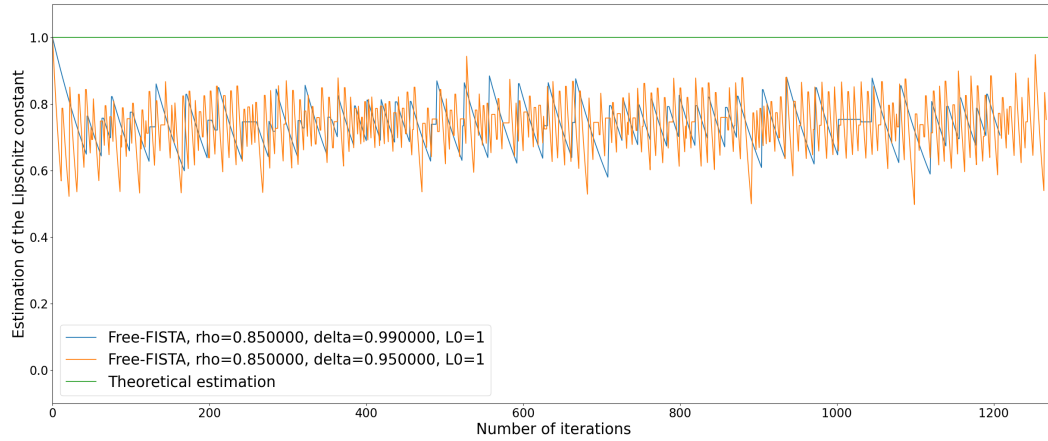


FIGURE 3.14 – Estimation de la constante de Lipschitz de  $\nabla f$  en fonction du nombre d'itérations de FISTA pour le problème d'inpainting.

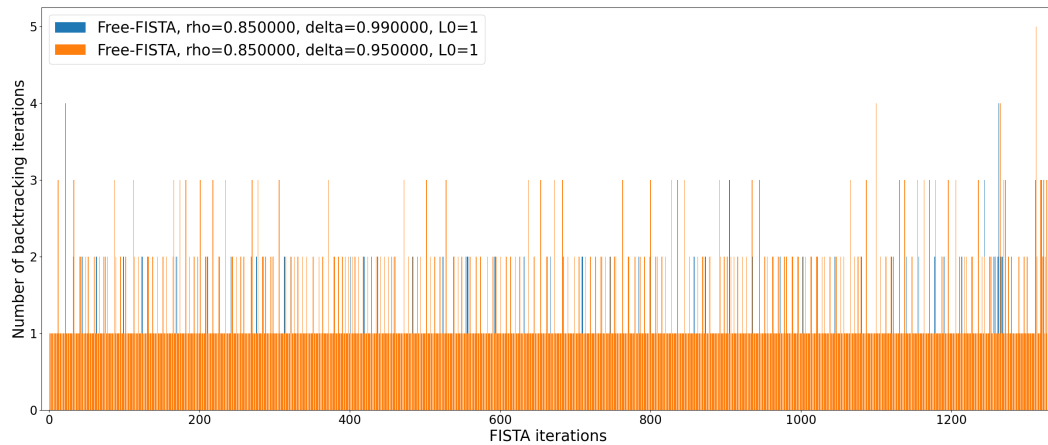


FIGURE 3.15 – Nombre d'itérations de backtracking par itération de FISTA pour le problème d'inpainting.

### 3.3.4.3 Super-résolution en microscopie

Ce dernier exemple est un problème de traitement d'image que l'on peut rencontrer dans des applications de microscopie par fluorescence [83, 116]. Le but est de reconstituer une image abîmée par un opérateur de flou et du bruit poissonien. Soit  $z \in \mathbb{R}_+^m$  l'image floutée et bruitée, on cherche à retrouver une image parcimonieuse et à coefficients positifs  $x \in \mathbb{R}_+^n$  telle que  $z = \mathcal{P}(MHx + b) \in \mathbb{R}^m$  avec  $m = q^2n$ ,  $q > 1$ , où  $M \in \mathbb{R}^{m \times n}$  est un opérateur de sous-échantillonnage de facteur  $q$ ,  $H \in \mathbb{R}^{n \times n}$  est un opérateur de convolution calculé pour une *point spread function* (PSF) donnée,  $b = \bar{b}e_m \in \mathbb{R}_{>0}^m$  est un terme de background constant et positif<sup>1</sup> et

1. On note  $e_d$  le vecteur qui ne contient que des 1 dans  $\mathbb{R}^d$ .

$\mathcal{P}(w)$  qui correspond à une réalisation d'un bruit poissonnien en dimension  $m$  avec pour paramètre  $w \in \mathbb{R}_+^m$ .

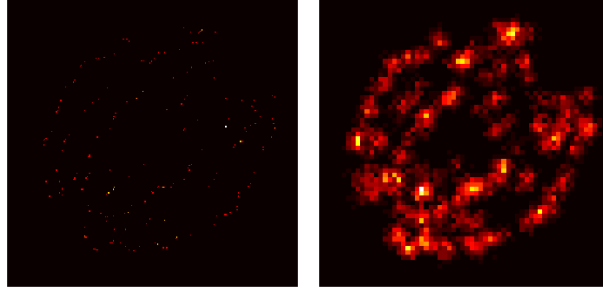


FIGURE 3.16 – Image originale tirée du jeu de données SMLM ISBI 2013 (à gauche) et image bruitée, floutée et sous-échantillonnée (à droite).

Pour prendre en compte le bruit poissonnien qui affecte nos données, la divergence de Kullback-Leibler [38] est un terme d'attache aux données consistant et pertinent. On définit cette fonctionnelle de la manière suivante :

$$KL(MHx + b; z) := \sum_{i=1}^m \left( z_i \ln \frac{z_i}{(MHx)_i + \bar{b}} + (MHx)_i + \bar{b} - z_i \right), \quad (3.3.38)$$

avec la convention que  $0 \ln 0 = 0$ . On ajoute un terme de régularisation  $\ell^1$  pour forcer la solution à être parcimonieuse, et on impose la non-négativité de la solution via une fonction indicatrice  $\iota_{\geq 0}(\cdot)$  sur l'orthant positif. Ainsi, le problème considéré est le suivant :

$$\min_{x \in \mathbb{R}^n} F(x) := \underbrace{KL(MHx + b, z)}_{:=f(x)} + \lambda \underbrace{\|x\|_1 + \iota_{\geq 0}(x)}_{:=h(x)}. \quad (3.3.39)$$

On peut calculer le gradient de  $f$  :

$$\nabla f(x) = (MH)^T e_m - (MH)^T \left( \frac{z}{MHx + b} \right). \quad (3.3.40)$$

D'après [72, 116], on a que  $\nabla f$  est  $L$ -Lipchitz sur  $\{x : x \geq 0\}$  où on peut calculer une estimation haute de cette constante de Lipschitz :

$$\hat{L} = \frac{\max z_i}{\bar{b}^2} \max((MH)^T e_m) \max(MH e_n). \quad (3.3.41)$$

Cette estimation théorique de  $L$  peut devenir significativement grande, en particulier lorsque  $\bar{b} \ll 1$ . De plus, comme cela est montré dans [50], la divergence de Kullback-Leibler (3.3.38) est localement 2-conditionnée et il existe donc  $\mu > 0$  tel que  $F$  satisfait  $\mathcal{G}_\mu^2$ . Comme ce paramètre  $\mu$  est inconnu et l'estimation  $\hat{L}$  peut être significativement plus grande que  $L$ , l'application de Free-FISTA semble approprié. On applique FISTA\_adaBT (Algorithme 5) et Free-FISTA pour obtenir les résultats

présentés dans la Figure 3.17. Ici, la stratégie de restart incluse dans Free-FISTA n'apporte pas d'avantage par rapport à la version de FISTA avec simple backtracking. En effet, l'algorithme converge plus lentement en terme d'itérations et donc en terme de temps de calcul (Free-FISTA s'appuie sur quelques calculs supplémentaires pour effectuer ses redémarrages). Ce phénomène s'explique certainement par la géométrie de  $F$  et le paramètre de croissance  $\mu$  qui est extrêmement faible. Le fait que les itérés de FISTA avec backtracking ne présentent aucune forme de comportement oscillatoire montre que  $F$  est une fonction relativement plate et dans un tel contexte, couper l'inertie n'est pas bénéfique.

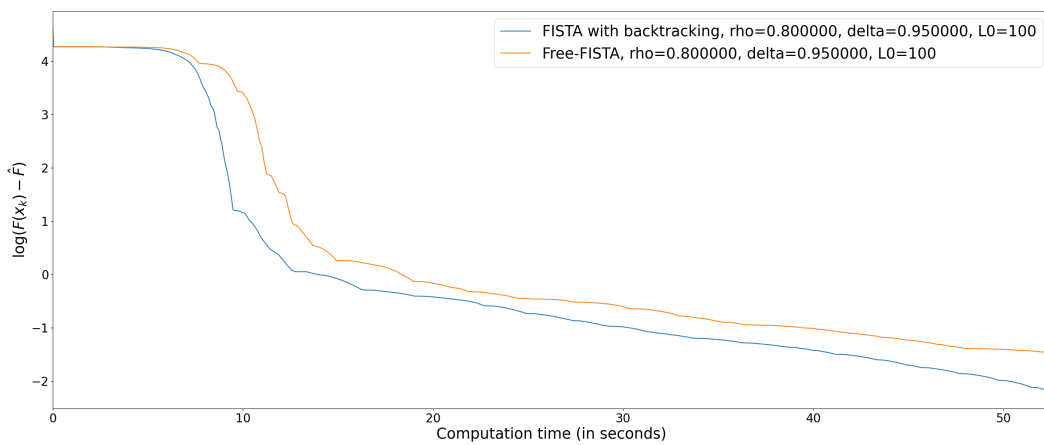


FIGURE 3.17 – Évolution de l'erreur en fonction du temps de calcul pour le problème (3.3.39).

## 3.A Résultats annexes et preuves

### 3.A.1 Étude de l'Algorithme 3 avec une mise à jour des itérations plus générale

Dans cette partie, on analyse une version légèrement plus générale de l'Algorithme 3 pour montrer que le choix de mise à jour des itérations est optimal dans un certain sens.



---

**Algorithme 8** : Version alternative de l'Algorithme 3
 

---

**Entrée:**  $r_0 \in \mathbb{R}^N$ ,  $j = 1$ ,  $C > 4$ ,  $\gamma > 1$ .

 $n_0 = \lfloor \gamma C \rfloor$ ;

 $r_1 = \text{FISTA}(r_0, n_0)$ ;

 $n_1 = \lfloor \gamma C \rfloor$ ;

**répéter**
 $j = j + 1$ ;

 $r_j = \text{FISTA}(r_{j-1}, n_{j-1})$ ;

$$\tilde{\mu}_j = \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)};$$
**si**  $n_{j-1} \leq C \sqrt{\frac{L}{\tilde{\mu}_j}}$  **alors**
 $n_j = \gamma n_{j-1}$ ;

**sinon**
 $n_j = n_{j-1}$ ;

**fin si**
**jusqu'à**  $\|g(r_j)\| \leq \varepsilon$ .

**retourne**  $r = r_j^+$ .
 

---

Pour retrouver le schéma de restart original, il suffit de choisir  $\gamma = 2$ . Le fait que l'Algorithme 8 soit légèrement plus général ne nous empêche pas d'utiliser la même approche que dans le cas  $\gamma = 2$ . Les contrôles sur la suite  $(\tilde{\mu}_j)_{j \geq 2}$  énoncés dans le Lemme 3.2.2 sont toujours valables et le Lemme 3.2.3 doit être légèrement modifié.

**Lemme 3.A.1.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . La suite  $(n_j)_{j \geq 0}$  définie dans l'Algorithme 8 avec  $C > 4$ ,  $\gamma > 1$  et  $\varepsilon > 0$  vérifie l'inégalité suivante :

$$\forall j \geq 0, \quad n_j \leq \gamma C \sqrt{\frac{L}{\mu}}. \quad (3.A.1)$$

En suivant la même feuille de route que pour la preuve du Théorème 3.2.8, on peut montrer que l'on a le contrôle ci-dessous. Les différences avec la preuve initiale sont présentées dans la Section 3.A.3.5.

**Théorème 3.A.2.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  les suites définies dans l'Algorithme 8 avec  $C > 4$ ,  $\gamma > 1$  et  $\varepsilon > 0$ . Le nombre d'itérations  $1 + \sum_{i=0}^j n_i$  nécessaire pour assurer  $\|g(r_j)\| \leq \varepsilon$  est borné et vérifie l'inégalité suivante :

$$\sum_{i=0}^j n_i \leq \frac{\gamma^2}{\gamma - 1} \frac{C}{\ln\left(\frac{C^2}{4} - 1\right)} \sqrt{\frac{L}{\mu}} \left( 2 \ln\left(\frac{C^2}{4} - 1\right) + \ln\left(1 + \frac{4\gamma^2}{C^2 - 4\gamma^2} \frac{2L(F(r_0) - F^*)}{\varepsilon^2}\right) \right). \quad (3.A.2)$$

Ainsi, avec les mêmes arguments que ceux utilisés pour démontrer le Corollaire 3.2.9, on peut montrer que si  $C > 4$  et  $\gamma > 1$ , les suites  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  de

l'Algorithme 8 vérifient :

$$F(r_j^+) - F^* = \mathcal{O} \left( e^{-\frac{\gamma-1}{\gamma^2} \frac{\ln\left(\frac{C^2}{4}-1\right)}{C}} \sqrt{\frac{\mu}{L}} \sum_{i=0}^j n_i \right). \quad (3.A.3)$$

Ceci montre que  $\gamma$  et  $C$  peuvent être optimisés indépendamment pour obtenir la meilleure constante possible. En particulier, comme la fonction  $\gamma \mapsto \frac{\gamma-1}{\gamma^2}$  atteint son maximum pour  $\gamma = 2$ , on peut en déduire que c'est théoriquement le meilleur choix.

### 3.A.2 Variations de l'Algorithme 3

Le schéma de restart présenté en Section 3.2 a l'avantage d'offrir des résultats théoriques de convergence rapide pour des fonctions variant  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  sans avoir besoin de connaître le paramètre de croissance  $\mu$ . Cependant, le schéma peut être ralenti sur les premières itérations à cause de redémarrages trop rapides. En effet, des restarts ont lieu toutes les  $\lfloor 2C \rfloor$  itérations de FISTA sur les premières boucles de restart où le meilleur choix théorique pour  $C$  est d'environ 6.38. Cela permet au schéma d'estimer le paramètre de croissance  $\mu$  et d'ensuite mettre à jour le nombre d'itérations avant les redémarrages suivants tout en gardant un contrôle sur ce nombre d'itérations.

Dans cette section, je propose deux alternatives à l'Algorithme 3 qui visent à résoudre ce problème.

#### 3.A.2.1 Une légère modification de l'Algorithme 3

La première variation proposée est une légère modification de l'Algorithme 3 qui est écrite ci-dessous.

**Algorithme 9** : Variation de l'Algorithme 3

---

**Entrée:**  $r_0 \in \mathbb{R}^N, j = 2, n_{init} > 0$ ;

$n_0 = n_{init}$  ;

$r_1 = \text{FISTA}(r_0, n_0)$  ;

$n_1 = n_{init}$  ;

$r_2 = \text{FISTA}(r_1, n_1)$  ;

$\tilde{\mu}_2 = \frac{4L}{(n_{init}+1)^2} \frac{F(r_0)-F(r_2)}{F(r_1)-F(r_2)}$  ;

$n_2 = 2C \sqrt{\frac{L}{\tilde{\mu}_2}}$  ;

**répéter**

$j = j + 1$  ;

$r_j = \text{FISTA}(r_{j-1}, n_{j-1})$  ;

$\tilde{\mu}_j = \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4L}{(n_{i-1}+1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}$  ;

**si**  $n_{j-1} \leq C \sqrt{\frac{L}{\tilde{\mu}_j}}$  **alors**

$n_j = 2n_{j-1}$  ;

**sinon**

$n_j = n_{j-1}$  ;

**fin si**

**jusqu'à**  $\|g(r_j)\| \leq \varepsilon$ .

**retourne**  $r = r_j$

---

L'idée derrière ce schéma est de rajouter un paramètre  $n_{init} > 0$  à l'algorithme que l'utilisateur doit fixer. Ce paramètre détermine le nombre d'itérations de FISTA qui doivent être exécutés avant le premier puis le second redémarrage. En prenant  $n_{init}$  suffisamment grand, on peut espérer à la fois éviter un redémarrage prématuré et obtenir une première itération  $\tilde{\mu}_2$  qui soit relativement précise. En effet, mes expériences numériques montrent que l'estimation du paramètre de croissance  $\mu$  via la technique proposée ici :

$$\tilde{\mu}_2 = \frac{4L}{(n_{init} + 1)^2} \frac{F(r_0) - F(r_2)}{F(r_1) - F(r_2)}, \quad (3.A.4)$$

est meilleure à mesure que  $n_{init}$  augmente. Le risque est par contre de choisir  $n_{init}$  trop grand, exécuter FISTA trop longtemps et réduire les bénéfices du restart. Ce schéma permet d'assurer le théorème suivant :

**Théorème 3.A.3.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(r_j)_{j \geq 0}$  et  $(n_j)_{j \geq 0}$  les suites définies dans l'Algorithme 9 avec  $C > 4$ ,  $n_{init} > 0$  et  $\varepsilon > 0$ . Le nombre d'itérations  $1 + \sum_{i=0}^j n_i$  nécessaire pour assurer  $\|g(r_j)\| \leq \varepsilon$  est borné et vérifie l'inégalité suivante :*

$$\sum_{i=0}^j n_i \leq 2n_{init} + \frac{4C}{\ln\left(\frac{C^2}{4} - 1\right)} \sqrt{\frac{L}{\mu}} \left( 2 \ln\left(\frac{C^2}{4} - 1\right) + \ln\left(1 + \frac{16}{C^2 - 16} \frac{2L(F(r_0) - F^*)}{\varepsilon^2}\right) \right). \quad (3.A.5)$$

**Éléments de preuve.**

Après le deuxième redémarrage, l'Algorithme 9 est identique au schéma original (Algorithme 3). On sait donc que  $2n_{init}$  itérations sont effectuées avant que l'Algorithme 9 se comporte comme son modèle. En utilisant les mêmes arguments que pour démontrer le Lemme 3.2.2, on a que  $\tilde{\mu}_2 \geq \mu$  et donc  $n_2 \leq 2C$ . On est donc assuré que  $n_j \leq 2C\sqrt{\frac{L}{\mu}}$  et en appliquant le même raisonnement que pour démontrer le Théorème 3.2.8, on obtient que le nombre d'itérations nécessaires après les deux premiers redémarrages pour satisfaire  $\|g(r_j)\| \leq \varepsilon$  a la même borne supérieure que le nombre d'itérations requis pour satisfaire le même critère d'arrêt avec l'Algorithme 3. On obtient donc la borne énoncée.  $\square$

Ce résultat garantit que le comportement asymptotique de l'Algorithme 9 est le même que celui de l'Algorithme 3.

Cependant, cette méthode n'améliore pas numériquement la convergence de l'Algorithme 3 de manière significative et consistante. Sa performance est assez sensible aux variations du paramètre  $n_{init}$  dont la valeur optimale dépend vraisemblablement de  $\mu$ . C'est pourquoi j'ai préféré privilégier la formulation de l'Algorithme 3.

**3.A.2.2 Une version parallèle de l'Algorithme 3**

Cette deuxième alternative est plus expérimentale que la première. L'idée du schéma est d'utiliser la parallélisation pour déléguer les tâches présentes dans l'Algorithme 3 à différents processus. En particulier, j'ai voulu séparer un processus principal ayant pour objectif la minimisation efficace de  $F$  sans calcul additionnel et des processus esclaves responsables de l'estimation du paramètre de croissance. La combinaison de calcul parallèle et de stratégie de restart a déjà été proposée par Renegar et Grimmer dans [117] mais avec une philosophie différente.

On écrit les tâches assignées aux deux types de processus ci-dessous.

---

**Algorithme 10** : Processus esclave

---

**Entrée:**  $x_0 \in \mathbb{R}^N, j = 1, n > 0$ .

$$\mu_1 = L$$

$$x_1 = \text{FISTA}(x_0, n)$$

**répéter**

$$k = k + 1$$

$$x_k = \text{FISTA}(x_{k-1}, n)$$

$$\tilde{\mu}_k = \min_{\substack{i \in \mathbb{N}^* \\ i < k}} \frac{4L}{(n+1)^2} \frac{F(x_{i-1}) - F(x_k)}{F(x_i) - F(x_k)}$$

Mise à jour de l'estimation  $\tilde{\mu}$  si  $\tilde{\mu}_k \leq \tilde{\mu}$

**jusqu'à** Arrêt du processus principal

---

---

**Algorithme 11** : Processus principal
 

---

**Entrée:**  $r_0 \in \mathbb{R}^N, j = 1$ 
 $r_1 = \text{FISTA}(r_0)$  jusqu'à ce qu'une estimation  $\tilde{\mu}$  de  $\mu$  soit accessible et  $i$  vérifie  $i > C\sqrt{\frac{L}{\tilde{\mu}}}$ .

 $n_1 = \lfloor 2C\sqrt{\frac{L}{\tilde{\mu}}} \rfloor$ 
**répéter**
 $j = j + 1$ 
 $r_j = \text{FISTA}(r_{j-1}, n_{j-1})$ 
**si**  $n_{j-1} \leq C\sqrt{\frac{L}{\tilde{\mu}}}$  **alors**
 $n_j = 2n_{j-1}$ 
**sinon**
 $n_j = n_{j-1}$ 
**fin si**
**jusqu'à**  $\|g(r_j)\| \leq \varepsilon$ 
**retourne**  $r = r_j$ 


---

De par sa philosophie, il s'agit d'un schéma parallèle asynchrone qui peut s'appuyer sur un nombre  $n_e \geq 1$  de processus esclaves. Une extension possible à cette méthode serait de remplacer l'itéré courant du processus principal par l'itéré d'un processus esclave dès lors que celui ci assure une valeur plus basse de la fonctionnelle.

Des tests numériques ont été mis en place mais n'étant pas spécialiste de programmation en calcul parallèle, les résultats n'ont pas été concluant : l'ajout de processus supplémentaires semble avoir trop d'impact sur le temps de calcul, ce qui peut être lié à l'implémentation de la méthode ou à la philosophie de la méthode. D'un point de vue théorique, il semble très complexe d'obtenir des résultats de convergence en suivant les mêmes principes que pour la preuve du Théorème 3.2.8. Enfin, une autre difficulté liée à ce schéma est qu'il fait intervenir des paramètres additionnels comme le nombre de processus esclave ou le nombre d'itération avant estimation de  $\mu$  pour chaque processus esclave. Faire des tests numériques dans ces conditions et sans bagage théorique a donc été un frein au développement d'une telle méthode.

### 3.A.3 Preuves techniques

#### 3.A.3.1 Preuve du Théorème 3.2.8

Soient  $C > 4$  et  $\varepsilon > 0$ . On définit  $T = 1 + \left\lceil \frac{\ln\left(1 + \frac{16}{C^2 - 16} \frac{2(F(r_0) - F^*)}{L\varepsilon^2}\right)}{\ln\left(\frac{C^2}{4} - 1\right)} \right\rceil$ . Dans un premier temps, on montre qu'un doublement des itérations a lieu au moins tous les  $T$  restarts.

On raisonne par l'absurde. Supposons que le nombre d'itérations n'a pas été doublé de  $j = s + 1$  jusqu'à  $j = s + T$  pour un certain  $s > 0$ . Cela signifie que l'on

a :

$$\forall j \in \llbracket s+1, s+T \rrbracket, \quad n_{j-1} > C \sqrt{\frac{L}{\tilde{\mu}_j}}, \quad (3.A.6)$$

et donc :

$$\forall j \in \llbracket s+1, s+T \rrbracket, \quad n_j = n_s. \quad (3.A.7)$$

Par conséquent :

$$\begin{aligned} \forall j \in \llbracket s+2, s+T \rrbracket, \quad \tilde{\mu}_j &= \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} \\ &\leq \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} \\ &\leq \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{4L}{n_{i-1}^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} \\ &\leq \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{4L}{n_s^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} \\ &\leq \frac{4L}{n_s^2} \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}, \end{aligned}$$

puisque le nombre d'itérations n'est pas doublé pour  $j \geq s$ . On peut en déduire que

$$\forall j \in \llbracket s+2, s+T \rrbracket, \quad \tilde{\mu}_j \leq \frac{4L}{n_s^2} \frac{F(r_{j-2}) - F(r_j)}{F(r_{j-1}) - F(r_j)}. \quad (3.A.8)$$

En combinant les inégalités (3.A.6), (3.A.7) et (3.A.8), on montre que :

$$n_s > C \sqrt{\frac{L}{\frac{4L}{n_s^2} \frac{F(r_{j-2}) - F(r_j)}{F(r_{j-1}) - F(r_j)}}}} = n_s \frac{C}{2} \sqrt{\frac{F(r_{j-1}) - F(r_j)}{F(r_{j-2}) - F(r_j)}}}. \quad (3.A.9)$$

Ceci implique donc que

$$F(r_{j-2}) - F(r_j) > \frac{C^2}{4} (F(r_{j-1}) - F(r_j)), \quad (3.A.10)$$

et ainsi,

$$F(r_{j-2}) - F(r_{j-1}) > \left( \frac{C^2}{4} - 1 \right) (F(r_{j-1}) - F(r_j)). \quad (3.A.11)$$

Comme  $C > 2$  on peut écrire que

$$F(r_{j-1}) - F(r_j) < \frac{4}{C^2 - 4} (F(r_{j-2}) - F(r_{j-1})). \quad (3.A.12)$$

Considérons le cas  $j = s + 1$  :

$$\begin{aligned}\tilde{\mu}_{s+1} &= \min_{\substack{i \in \mathbb{N}^* \\ i < s+1}} \frac{4L}{(n_{i-1} + 1)^2} \frac{F(r_{i-1}) - F(r_{s+1})}{F(r_i) - F(r_{s+1})} \\ &\leq \frac{4L}{(n_{s-1} + 1)^2} \frac{F(r_{s-1}) - F(r_{s+1})}{F(r_s) - F(r_{s+1})} \\ &\leq \frac{4L}{\left(\frac{n_s}{2} + 1\right)^2} \frac{F(r_{s-1}) - F(r_{s+1})}{F(r_s) - F(r_{s+1})} \\ &\leq \frac{16L}{n_s^2} \frac{F(r_{s-1}) - F(r_{s+1})}{F(r_s) - F(r_{s+1})},\end{aligned}$$

car  $n_s \leq 2n_{s-1}$ .

De la même manière que précédemment, on peut montrer que

$$F(r_s) - F(r_{s+1}) < \frac{16}{C^2 - 16} (F(r_{s-1}) - F(r_s)). \quad (3.A.13)$$

On introduit ci-dessous le Lemme 3.A.4 qui relie la valeur du gradient composite  $g$  à celle de  $F$ . Une preuve est donnée en Section 3.A.3.8 :

**Lemme 3.A.4.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  pour  $L > 0$ . La suite  $(r_j)_{j \geq 0}$  générée par l'Algorithme 3 avec  $C > 4$  et  $\varepsilon > 0$  vérifie*

$$\forall j \geq 1, \quad \frac{1}{2L} \|g(r_{j-1})\|^2 \leq F(r_{j-1}) - F(r_j). \quad (3.A.14)$$

Par le Lemme 3.A.4 et les inégalités (3.A.13) et (3.A.12), on peut écrire les inégalités ci-dessous

$$\begin{aligned}\frac{1}{2L} \|g(r_{s+T-1})\|^2 &\leq F(r_{s+T-1}) - F(r_{s+T}) \\ &\leq \frac{4}{C^2 - 4} (F(r_{s+T-2}) - F(r_{s+T-1})) \\ &\leq \left(\frac{4}{C^2 - 4}\right)^{T-1} \left(\frac{16}{C^2 - 16}\right) (F(r_{s-1}) - F(r_s)) \\ &\leq \left(\frac{4}{C^2 - 4}\right)^{T-1} \left(\frac{16}{C^2 - 16}\right) (F(r_0) - F^*) \\ &\leq \left(\frac{4}{C^2 - 4}\right)^{\left\lceil \frac{\ln\left(1 + \frac{16}{C^2 - 16} \frac{2L(F(r_0) - F^*)}{\varepsilon^2}\right)}{\ln\left(\frac{C^2}{4} - 1\right)} \right\rceil} \left(\frac{16}{C^2 - 16}\right) (F(r_0) - F^*) \\ &\leq \left(\frac{4}{C^2 - 4}\right)^{\frac{\ln\left(1 + \frac{16}{C^2 - 16} \frac{2L(F(r_0) - F^*)}{\varepsilon^2}\right)}{\ln\left(\frac{C^2}{4} - 1\right)}} \left(\frac{16}{C^2 - 16}\right) (F(r_0) - F^*) \\ &\leq \frac{1}{1 + \frac{16}{C^2 - 16} \frac{2L(F(r_0) - F^*)}{\varepsilon^2}} \left(\frac{16}{C^2 - 16}\right) (F(r_0) - F^*) \\ &\leq \frac{\varepsilon^2}{2L}.\end{aligned}$$

Ceci signifie donc que si  $T$  restarts ont eu lieu sans doublement du nombre d'itération, alors le critère d'arrêt  $\|g(r_j)\| \leq \varepsilon$  a été satisfait. On peut en déduire que le nombre d'itérations est doublé au moins tous les  $T$  redémarrages et que pour tout  $s \geq 1$ , il existe  $j \in \llbracket s+1, s+T \rrbracket$  tel que

$$n_{j-1} < C \sqrt{\frac{L}{\bar{\mu}_j}} \quad (3.A.15)$$

ou tel que le critère d'arrêt est vérifié. Ainsi,  $n_j = 2n_{j-1}$  et comme  $(n_j)_{j \geq 0}$  est une suite croissante, on a  $n_{s+T} \geq n_j = 2n_{j-1} \geq 2n_s$ . Il en résulte que

$$\forall s \geq 1, n_s \leq \frac{n_{s+T}}{2}. \quad (3.A.16)$$

Réécrivons  $j$  comme  $j = m + nT$  où  $0 \leq m < T$  et  $n \geq 0$ . Comme la suite  $(n_j)_{j \geq 0}$  est croissante, on a

$$\sum_{i=0}^j n_i = \sum_{i=0}^{m+nT} n_i = \sum_{i=0}^m n_i + \sum_{l=0}^{n-1} \sum_{i=1}^T n_{m+i+lT} \quad (3.A.17)$$

$$\leq Tn_m + T \sum_{l=1}^n n_{m+lT} = T \sum_{l=0}^n n_{m+lT} = T \sum_{l=0}^n n_{j-lT}. \quad (3.A.18)$$

L'équation (3.A.16) assure que  $n_{j-T} \leq \frac{n_j}{2}$  et par conséquent

$$n_{j-lT} \leq \left(\frac{1}{2}\right)^l n_j, \quad \forall l \in \llbracket 0, n \rrbracket. \quad (3.A.19)$$

On obtient donc les inégalités suivantes :

$$\sum_{i=0}^j n_i \leq T \sum_{l=0}^n n_{j-lT} \leq T \sum_{l=0}^n \left(\frac{1}{2}\right)^l n_j \leq T \sum_{l=0}^{\infty} \left(\frac{1}{2}\right)^l n_j = 2Tn_j. \quad (3.A.20)$$

Le résultat ci-dessus combiné au Lemme 3.2.3 permet de conclure que pour  $j > 0$

$$\sum_{i=0}^j n_i \leq 2Tn_j \leq 4C \sqrt{\frac{L}{\mu}} T \leq 4C \sqrt{\frac{L}{\mu}} \left( 1 + \left| \frac{\ln \left( 1 + \frac{16}{C^2-16} \frac{2L(F(r_0)-F^*)}{\varepsilon^2} \right)}{\ln \left( \frac{C^2}{4} - 1 \right)} \right| \right) \quad (3.A.21)$$

$$\leq \frac{4C}{\ln \left( \frac{C^2}{4} - 1 \right)} \sqrt{\frac{L}{\mu}} \left( 2 \ln \left( \frac{C^2}{4} - 1 \right) + \ln \left( 1 + \frac{16}{C^2-16} \frac{2L(F(r_0)-F^*)}{\varepsilon^2} \right) \right). \quad (3.A.22)$$

□



### 3.A.3.2 Preuve du Corollaire 3.2.9

Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(r_j)_{j \in \mathbb{N}}$  et  $(n_j)_{j \in \mathbb{N}}$  les suites définies dans l'Algorithme 3 avec  $C > 4$  et  $\varepsilon > 0$ .

Le résultat que l'on démontre est asymptotique, c'est-à-dire qu'il est valable pour un nombre d'itérations assez grand. On peut donc se placer dans le cas où la condition  $\|g(r_j)\| \leq \varepsilon$  est validée après au moins  $8C\sqrt{\frac{L}{\mu}}$  itérations. On définit  $\psi_\mu : \mathbb{R}_+^* \rightarrow (8C\sqrt{\frac{L}{\mu}}, +\infty)$  telle que :

$$\psi_\mu : \gamma \mapsto \frac{4C}{\ln\left(\frac{C^2}{4} - 1\right)} \sqrt{\frac{L}{\mu}} \left( 2 \ln\left(\frac{C^2}{4} - 1\right) + \ln\left(1 + \frac{16}{C^2 - 16} \frac{2L(F(r_0) - F^*)}{\gamma}\right) \right). \quad (3.A.23)$$

D'après le Théorème 3.2.8, le nombre d'itérations pour vérifier  $\|g(r_j)\| \leq \varepsilon$  satisfait l'inégalité suivante :

$$\sum_{i=0}^j n_i \leq \psi_\mu(\varepsilon^2). \quad (3.A.24)$$

Comme  $\psi_\mu$  est une fonction strictement décroissante et  $\sum_{i=0}^j n_i > 8C\sqrt{\frac{L}{\mu}}$ , on obtient que :

$$\psi_\mu^{-1}\left(\sum_{i=0}^j n_i\right) \geq \varepsilon^2, \quad (3.A.25)$$

où  $\psi_\mu^{-1}$  est la fonction inverse de  $\psi_\mu$ . En appliquant le Lemme 3.2.5 on a :

$$F(r_j^+) - F^* \leq \frac{8}{\mu} \|g(r_j)\|^2 \quad (3.A.26)$$

$$\leq \frac{8\varepsilon^2}{\mu} \quad (3.A.27)$$

$$\leq \frac{8}{\mu} \psi_\mu^{-1}\left(\sum_{i=0}^j n_i\right). \quad (3.A.28)$$

Quelques calculs supplémentaires montrent que :

$$\psi_\mu^{-1} : n \mapsto 2L \frac{16}{C^2 - 16} \frac{1}{e^{-2 \ln\left(\frac{C^2}{4} - 1\right)} e^{\frac{\ln\left(\frac{C^2}{4} - 1\right)}{4C} \sqrt{\frac{\mu}{L}} n} - 1} (F(r_0) - F^*), \quad (3.A.29)$$

et ainsi :

$$F(r_j^+) - F^* \leq \frac{16L}{\mu} \frac{16}{C^2 - 16} \frac{1}{e^{-2 \ln\left(\frac{C^2}{4} - 1\right)} e^{\frac{\ln\left(\frac{C^2}{4} - 1\right)}{4C} \sqrt{\frac{\mu}{L}} \sum_{i=0}^j n_i} - 1} (F(r_0) - F^*). \quad (3.A.30)$$

Par conséquent, on obtient le résultat ci-dessous

$$F(r_j^+) - F^* = \mathcal{O}\left(e^{-\frac{\ln\left(\frac{C^2}{4} - 1\right)}{4C} \sqrt{\frac{\mu}{L}} \sum_{i=0}^j n_i}\right). \quad (3.A.31)$$

Si on veut maximiser la fonction  $C \mapsto \frac{\ln(\frac{C^2}{4}-1)}{4C}$ , on obtient que la valeur optimale est  $\hat{C} \approx 6.38$ . Pour ce choix, on a finalement :

$$F(r_j^+) - F^* = \mathcal{O}\left(e^{-\frac{1}{12}\sqrt{\frac{\rho}{L}}\sum_{i=0}^j n_i}\right). \quad (3.A.32)$$

□

### 3.A.3.3 Preuve du Théorème 3.3.6

Soient  $\rho \in (0, 1)$ ,  $\delta \in (0, 1]$ ,  $C > 4/\sqrt{\rho}$ ,  $L_{min} < L_0 < \frac{L}{\rho}$  et  $\varepsilon > 0$ . La première partie de la preuve consiste à montrer qu'un doublement des itérations a lieu au moins tous les  $T$  restarts où on définit

$$T := 1 + \left\lceil \frac{\ln\left(1 + \frac{16}{C^2\rho-16} \frac{2L(F(r_0)-F^*)}{\rho\varepsilon^2}\right)}{\ln\left(\frac{C^2\rho}{4} - 1\right)} \right\rceil. \quad (3.A.33)$$

Raisonnons par l'absurde. Supposons que le nombre d'itérations n'a pas été doublé de  $j = s + 1$  jusqu'à  $j = s + T$  pour un certain  $s > 0$ . Cela signifie que l'on a :

$$\forall j \in \llbracket s + 1, s + T \rrbracket, \quad n_{j-1} > \frac{C}{\sqrt{\tilde{\kappa}_j}}, \quad (3.A.34)$$

et donc :

$$\forall j \in \llbracket s + 1, s + T \rrbracket, \quad n_j = n_s. \quad (3.A.35)$$

On en déduit ainsi que :

$$\begin{aligned} \forall j \in \llbracket s + 2, s + T \rrbracket, \quad \tilde{\kappa}_j &= \min_{\substack{i \in \mathbb{N}^* \\ i < j}} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} \\ &\leq \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} \\ &\leq \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{4}{\rho n_s^2} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)} \\ &\leq \frac{4}{\rho n_s^2} \min_{\substack{i \in \mathbb{N}^* \\ s < i < j}} \frac{F(r_{i-1}) - F(r_j)}{F(r_i) - F(r_j)}, \end{aligned}$$

puisque le nombre d'itérations n'est pas doublé pour  $j \geq s$ . Ainsi,

$$\forall j \in \llbracket s + 2, s + T \rrbracket, \quad \tilde{\kappa}_j \leq \frac{4}{\rho n_s^2} \frac{F(r_{j-2}) - F(r_j)}{F(r_{j-1}) - F(r_j)}. \quad (3.A.36)$$

En combinant les inégalités (3.A.34), (3.A.35) et (3.A.36), on obtient que :

$$n_s > \frac{C}{\sqrt{\frac{4}{\rho n_s^2} \frac{F(r_{j-2}) - F(r_j)}{F(r_{j-1}) - F(r_j)}}} = n_s \frac{C\sqrt{\rho}}{2} \sqrt{\frac{F(r_{j-1}) - F(r_j)}{F(r_{j-2}) - F(r_j)}}. \quad (3.A.37)$$

Par conséquent,

$$F(r_{j-2}) - F(r_j) > \frac{C^2\rho}{4}(F(r_{j-1}) - F(r_j)), \quad (3.A.38)$$

et ainsi,

$$F(r_{j-2}) - F(r_{j-1}) > \left(\frac{C^2\rho}{4} - 1\right)(F(r_{j-1}) - F(r_j)). \quad (3.A.39)$$

Comme  $C > 2/\sqrt{\rho}$  on peut écrire que

$$F(r_{j-1}) - F(r_j) < \frac{4}{C^2\rho - 4}(F(r_{j-2}) - F(r_{j-1})). \quad (3.A.40)$$

Si  $j = s + 1$ , on a :

$$\begin{aligned} \tilde{\kappa}_{s+1} &= \min_{\substack{i \in \mathbb{N}^* \\ i < s+1}} \frac{4}{\rho n_{i-1}^2} \frac{F(r_{i-1}) - F(r_{s+1})}{F(r_i) - F(r_{s+1})} \\ &\leq \frac{4}{\rho n_{s-1}^2} \frac{F(r_{s-1}) - F(r_{s+1})}{F(r_s) - F(r_{s+1})} \\ &\leq \frac{16}{\rho n_s^2} \frac{F(r_{s-1}) - F(r_{s+1})}{F(r_s) - F(r_{s+1})}, \end{aligned}$$

car  $n_s \leq 2n_{s-1}$ .

En suivant les mêmes calculs que précédemment, on montre que

$$F(r_s) - F(r_{s+1}) < \frac{16}{C^2\rho - 16}(F(r_{s-1}) - F(r_s)). \quad (3.A.41)$$

On introduit ci-dessous le lemme qui relie la valeur du gradient composite  $g$  à celle de  $F$  pour les itérés de Free-FISTA. Une preuve est donnée en Section 3.A.3.9.

**Lemme 3.A.5.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  pour  $L > 0$ . La suite  $(r_j)_{j \geq 0}$  générée par l'Algorithme 6 avec  $\rho \in (0, 1)$ ,  $\delta \in (0, 1]$ ,  $C > 4/\sqrt{\rho}$ ,  $L_{\min} < L_0 < \frac{L}{\rho}$  et  $\varepsilon > 0$  vérifie*

$$\forall j \geq 1, \quad \frac{\rho}{2L} \|g_{1/\hat{L}_j}(r_j)\|^2 \leq F(r_j) - F(r_{j+1}). \quad (3.A.42)$$

Par le Lemme 3.A.5 et les inégalités (3.A.40) et (3.A.41), on obtient le contrôle

suivant sur le gradient composite généralisé :

$$\begin{aligned}
\frac{\rho}{2L} \|g_{1/\hat{L}_{s+T-1}}(r_{s+T-1})\|^2 &\leq F(r_{s+T-1}) - F(r_{s+T}) \\
&\leq \frac{4}{C^2\rho - 4} (F(r_{s+T-2}) - F(r_{s+T-1})) \\
&\leq \left(\frac{4}{C^2\rho - 4}\right)^{T-1} \left(\frac{16}{C^2\rho - 16}\right) (F(r_{s-1}) - F(r_s)) \\
&\leq \left(\frac{4}{C^2\rho - 4}\right)^{T-1} \left(\frac{16}{C^2\rho - 16}\right) (F(r_0) - F^*) \\
&\leq \left(\frac{4}{C^2\rho - 4}\right) \left[ \frac{\ln\left(1 + \frac{16}{C^2\rho - 16} \frac{2L(F(r_0) - F^*)}{\rho\varepsilon^2}\right)}{\ln\left(\frac{C^2\rho - 1}{4}\right)} \right] \left(\frac{16}{C^2\rho - 16}\right) (F(r_0) - F^*) \\
&\leq \left(\frac{4}{C^2\rho - 4}\right) \frac{\ln\left(1 + \frac{16}{C^2\rho - 16} \frac{2L(F(r_0) - F^*)}{\rho\varepsilon^2}\right)}{\ln\left(\frac{C^2\rho - 1}{4}\right)} \left(\frac{16}{C^2\rho - 16}\right) (F(r_0) - F^*) \\
&\leq \frac{1}{1 + \frac{16}{C^2\rho - 16} \frac{2L(F(r_0) - F^*)}{\rho\varepsilon^2}} \left(\frac{16}{C^2\rho - 16}\right) (F(r_0) - F^*) \\
&\leq \frac{\rho\varepsilon^2}{2L}.
\end{aligned}$$

Ainsi, après  $T$  restarts sans doublement du nombre d'itération, le critère d'arrêt  $\|g_{1/\hat{L}_j}(r_j)\| \leq \varepsilon$  est validé. On peut en déduire que le nombre d'itérations est doublé au moins tous les  $T$  redémarrages. En utilisant les mêmes arguments que dans la Section 3.A.3.1, on obtient que pour tout  $j > 0$  tel que le critère d'arrêt n'a pas été vérifié ou a été vérifié pour la première fois :

$$\sum_{i=0}^j n_i \leq 2Tn_j. \quad (3.A.43)$$

Le résultat ci-dessus combiné au Lemme 3.3.5 permet de conclure que

$$\begin{aligned}
\sum_{i=0}^j n_i &\leq 2Tn_j \leq 4C\sqrt{\frac{L}{\mu}}T \leq 4C\sqrt{\frac{L}{\mu}} \left(1 + \left[ \frac{\ln\left(1 + \frac{16}{C^2\rho - 16} \frac{2L(F(r_0) - F^*)}{\rho\varepsilon^2}\right)}{\ln\left(\frac{C^2\rho - 1}{4}\right)} \right]\right) \\
&\leq \frac{4C}{\ln\left(\frac{C^2\rho - 1}{4}\right)} \sqrt{\frac{L}{\mu}} \left(2\ln\left(\frac{C^2\rho - 1}{4}\right) + \ln\left(1 + \frac{16}{C^2\rho - 16} \frac{2L(F(r_0) - F^*)}{\rho\varepsilon^2}\right)\right).
\end{aligned}$$

□

### 3.A.3.4 Preuve du Corollaire 3.3.8

Cette preuve est proche de celle donnée en Section 3.A.3.2.

Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soient  $(r_j)_{j \in \mathbb{N}}$  et  $(n_j)_{j \in \mathbb{N}}$  les suites définies dans l'Algorithme 6 avec  $\rho \in (0, 1)$ ,  $\delta \in (0, 1]$ ,  $C > 4/\sqrt{\rho}$ ,  $L_{\min} < L_0 < \frac{L}{\rho}$  et  $\varepsilon > 0$ .

On suppose que la condition de sortie  $\|g_{1/\hat{L}_j}(r_j)\| \leq \varepsilon$  est validée après au moins  $8C\sqrt{\frac{L}{\mu}}$  itérations. On définit  $\psi_\mu : \mathbb{R}_+^* \rightarrow \left(8C\sqrt{\frac{L}{\mu}}, +\infty\right)$  telle que :

$$\psi_\mu : \gamma \mapsto \frac{4C}{\ln\left(\frac{C^2\rho}{4} - 1\right)} \sqrt{\frac{L}{\mu}} \left( 2 \ln\left(\frac{C^2\rho}{4} - 1\right) + \ln\left(1 + \frac{16}{C^2\rho - 16} \frac{2L(F(r_0) - F^*)}{\rho\gamma}\right) \right). \quad (3.A.44)$$

D'après le Théorème 3.3.6, le nombre d'itérations pour vérifier  $\|g_{1/\hat{L}_j}(r_j)\| \leq \varepsilon$  satisfait l'inégalité suivante :

$$\sum_{i=0}^j n_i \leq \psi_\mu(\varepsilon^2). \quad (3.A.45)$$

Comme  $\psi_\mu$  est une fonction strictement décroissante et  $\sum_{i=0}^j n_i > 8C\sqrt{\frac{L}{\mu}}$ , on obtient que :

$$\psi_\mu^{-1}\left(\sum_{i=0}^j n_i\right) \geq \varepsilon^2, \quad (3.A.46)$$

où  $\psi_\mu^{-1}$  est la fonction inverse de  $\psi_\mu$ . En appliquant le Lemme 3.3.3 on a :

$$F(\hat{r}_j) - F^* \leq \frac{\left(1 + \frac{L}{\hat{L}_j}\right)^2}{2\mu} \|g_{1/\hat{L}_j}(r_j)\|^2, \quad (3.A.47)$$

et comme par construction,  $\hat{L}_j \geq L_{min}$  :

$$F(\hat{r}_j) - F^* \leq \frac{2\left(1 + \frac{L}{L_{min}}\right)^2}{\mu} \varepsilon^2 \quad (3.A.48)$$

$$\leq \frac{2\left(1 + \frac{L}{L_{min}}\right)^2}{\mu} \psi_\mu^{-1}\left(\sum_{i=0}^j n_i\right). \quad (3.A.49)$$

Quelques calculs supplémentaires montrent que :

$$\psi_\mu^{-1} : n \mapsto \frac{2L}{\rho} \frac{16}{C^2\rho - 16} \frac{1}{e^{-2\ln\left(\frac{C^2\rho}{4} - 1\right)} e^{\frac{\ln\left(\frac{C^2\rho}{4} - 1\right)}{\frac{4C}{\mu}} \sqrt{\frac{L}{\mu}} n} - 1} (F(r_0) - F^*), \quad (3.A.50)$$

et ainsi :

$$F(\hat{r}_j) - F^* \leq \frac{4L\left(1 + \frac{L}{L_{min}}\right)^2}{\rho\mu} \frac{16}{C^2\rho - 16} \frac{1}{e^{-2\ln\left(\frac{C^2\rho}{4} - 1\right)} e^{\frac{\ln\left(\frac{C^2\rho}{4} - 1\right)}{\frac{4C}{\mu}} \sqrt{\frac{L}{\mu}} \sum_{i=0}^j n_i} - 1} (F(r_0) - F^*). \quad (3.A.51)$$

Par conséquent, on obtient le résultat ci-dessous

$$F(\hat{r}_j) - F^* = \mathcal{O}\left(e^{-\frac{\ln\left(\frac{C^2\rho}{4} - 1\right)}{\frac{4C}{\mu}} \sqrt{\kappa} \sum_{i=0}^j n_i}\right). \quad (3.A.52)$$

En optimisant la fonction  $C \mapsto \frac{\ln(\frac{C^2\rho-1}{4C})}{4C}$ , on obtient que la valeur optimale est  $\hat{C} \approx 6.38/\sqrt{\rho}$ . Pour ce choix, on a finalement :

$$F(\hat{r}_j) - F^* = \mathcal{O}\left(e^{-\frac{\sqrt{\rho}}{12}\sqrt{\kappa}\sum_{i=0}^j n_i}\right). \quad (3.A.53)$$

□

### 3.A.3.5 Éléments de preuve du Théorème 3.A.2

La preuve du Théorème 3.A.2 étant calquée sur celle du Théorème 3.2.8, tous les calculs ne sont pas détaillés.

Soient  $C > 4$ ,  $\gamma > 1$  et  $\varepsilon > 0$ . On montre ici que le nombre d'itérations est augmenté au moins tous les  $T' = 1 + \left\lceil \frac{\ln\left(1 + \frac{4\gamma^2}{C^2 - 4\gamma^2} \frac{2L(F(r_0) - F^*)}{\varepsilon^2}\right)}{\ln\left(\frac{C^2}{4} - 1\right)} \right\rceil$  redémarrages.

En supposant qu'il n'y a pas d'augmentation du nombre d'itérations entre les restarts  $j = s + 1$  et  $j = s + T$  pour un certain  $s > 0$ , on peut montrer qu'on a les inégalités suivantes :

$$\forall j \in [s + 2, s + T], \quad F(r_{j-1}) - F(r_j) < \frac{4}{C^2 - 4}(F(r_{j-2}) - F(r_{j-1})), \quad (3.A.54)$$

et

$$F(r_s) - F(r_{s+1}) < \frac{4\gamma^2}{C^2 - 4\gamma^2}(F(r_{s-1}) - F(r_s)). \quad (3.A.55)$$

Grâce à ces inégalités et le Lemme 3.A.4 qui est toujours valable ici, on peut finalement montrer que  $\|g(r_{s+T-1})\| \leq \varepsilon$ . Ainsi, pour tout  $s \geq 1$ , on obtient que  $n_s \leq \frac{n_{s+T}}{\gamma}$ . Avec des calculs identiques à ceux pour  $\gamma = 2$ , on a que pour  $j \geq 0$ ,  $\sum_{i=0}^j n_i \leq T \sum_{i=0}^n n_{j-lT}$  où  $0 \leq m < T$  et  $n \geq 0$  tels que  $j = m + nT$ . Or, via les résultats précédents, on peut déduire que pour tout  $l \in \llbracket 0, n \rrbracket$ ,

$$n_{j-lT} \leq \left(\frac{1}{\gamma}\right)^l n_j. \quad (3.A.56)$$

Ceci implique le résultat ci-dessous

$$\sum_{i=0}^j n_i \leq T \sum_{l=0}^{\infty} \left(\frac{1}{\gamma}\right)^l n_j \leq \frac{\gamma}{\gamma - 1} T n_j, \quad (3.A.57)$$

et en appliquant le Lemme 3.A.1, on arrive à la conclusion voulue. □

### 3.A.3.6 Preuve du Lemme 3.3.3

Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  qui vérifie  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . On utilise le lien entre l'hypothèse de croissance quadratique et la propriété de Łojasiewicz

dans le cadre convexe. En particulier, on applique la Proposition 2.1.16 qui indique que la fonction  $F$  vérifie

$$\forall x \in \mathbb{R}^N, \quad F(x) - F^* \leq \frac{2d(0, \partial F(x))^2}{\mu}. \quad (3.A.58)$$

Soit  $x \in \mathbb{R}^N$ . En revenant à la définition de l'opérateur proximal (2.2.4),  $x^+$  est l'unique minimiseur de la fonction définie par

$$z \mapsto h(z) + \frac{L}{2} \|z - x + \frac{1}{L} \nabla f(x)\|^2 \quad (3.A.59)$$

et donc  $x^+$  vérifie

$$0 \in \partial h(x^+) + \{L(x^+ - x) + \nabla f(x)\}. \quad (3.A.60)$$

Par conséquent,

$$g(x) - \nabla f(x) + \nabla f(x^+) \in \partial F(x^+). \quad (3.A.61)$$

Le caractère  $L$ -Lipschitz du gradient de  $f$  nous garantissant que

$$\|g(x) - \nabla f(x) + \nabla f(x^+)\| \leq \|g(x)\| + \|\nabla f(x^+) - \nabla f(x)\| \quad (3.A.62)$$

$$\leq 2\|g(x)\|, \quad (3.A.63)$$

on obtient donc que

$$F(x^+) - F^* \leq \frac{2}{\mu} d(0, \partial F(x^+))^2 \quad (3.A.64)$$

$$\leq \frac{2}{\mu} \|g(x) - \nabla f(x) + \nabla f(x^+)\|^2 \quad (3.A.65)$$

$$\leq \frac{8}{\mu} \|g(x)\|^2. \quad (3.A.66)$$

□

### 3.A.3.7 Preuve du Lemme 3.3.3

Cette preuve est une version plus générale de celle en Section 3.A.3.6. En remplaçant  $\tau$  par  $\frac{1}{L}$ , on retrouve les mêmes calculs. On s'appuie sur l'hypothèse  $\mathcal{G}_\mu^2$  vérifiée par  $F$  qui, d'après la Proposition 2.1.16, implique que :

$$\forall x \in \mathbb{R}^N, \quad F(x) - F^* \leq \frac{2d(0, \partial F(x))^2}{\mu}. \quad (3.A.67)$$

Soient  $x \in \mathbb{R}^N$  et  $\tau > 0$ . Par définition,  $T_\tau x$  est l'unique minimiseur de la fonction ci-dessous

$$z \mapsto h(z) + \frac{1}{2\tau} \|z - x + \tau \nabla f(x)\|^2. \quad (3.A.68)$$

Ainsi,  $T_\tau x$  satisfait

$$0 \in \partial h(T_\tau x) + \left\{ \frac{1}{\tau} (T_\tau x - x) + \nabla f(x) \right\}, \quad (3.A.69)$$

ce qui implique que

$$g_\tau(x) - \nabla f(x) + \nabla f(T_\tau x) \in \partial F(T_\tau x). \quad (3.A.70)$$

Comme  $\nabla f$  est  $L$ -Lipschitz, on déduit que

$$\|g_\tau(x) - \nabla f(x) + \nabla f(T_\tau x)\| \leq \|g_\tau(x)\| + \|\nabla f(T_\tau x) - \nabla f(x)\| \quad (3.A.71)$$

$$\leq \|g_\tau(x)\| + L\|T_\tau x - x\| \quad (3.A.72)$$

$$\leq (1 + L\tau)\|g_\tau(x)\|. \quad (3.A.73)$$

En combinant ces inégalités, on trouve le résultat désiré :

$$F(T_\tau x) - F^* \leq \frac{2}{\mu} d(0, \partial F(T_\tau x))^2 \quad (3.A.74)$$

$$\leq \frac{2}{\mu} \|g_\tau(x) - \nabla f(x) + \nabla f(T_\tau x)\|^2 \quad (3.A.75)$$

$$\leq \frac{2(1 + L\tau)^2}{\mu} \|g_\tau(x)\|^2. \quad (3.A.76)$$

□

### 3.A.3.8 Preuve du Lemme 3.A.4

Pour démontrer ce lemme, on s'appuie sur [51, Lemme 3.1]. En effet, en appliquant ce lemme à  $y \in \mathbb{R}^N$  et  $y^+$ , on a

$$\frac{L}{2} \|y^+ - y\|^2 \leq F(y) - F(y^+). \quad (3.A.77)$$

Soit  $j \geq 1$ , on peut choisir  $y = r_{j-1}$ . Alors, comme  $F(r_{j-1}^+) \leq F(r_j)$  on peut conclure que

$$\frac{1}{2L} \|g(r_{j-1})\|^2 \leq F(r_{j-1}) - F(r_j). \quad (3.A.78)$$

□

### 3.A.3.9 Preuve du Lemme 3.A.5

Comme par définition  $(\hat{r}_j, \hat{L}_j) = \text{FB\_BT}(r_j, \tilde{L}_j; \rho)$  pour tout  $j \geq 1$ , on a donc :

$$D_f(\hat{r}_j, r_j) \leq \frac{\hat{L}_j}{2} \|\hat{r}_j - r_j\|^2, \quad (3.A.79)$$

où  $\hat{r}_j = T_{1/\hat{L}_j} r_j$ . Ainsi, en appliquant (3.3.3) on obtient :

$$F(\hat{r}_j) + \frac{\hat{L}_j}{2} \|\hat{r}_j - x\|^2 \leq F(x) + \frac{\hat{L}_j}{2} \|r_j - x\|^2,$$

pour tout  $x \in \mathbb{R}^N$ . Le choix  $x = r_j$  combiné permet de montrer que

$$\frac{1}{2\hat{L}_j} \|g_{1/\hat{L}_j} r_j\|^2 \leq F(r_j) - F(\hat{r}_j).$$



Comme  $L_0 < \frac{L}{\rho}$ , on a que  $\hat{L}_j \leq \frac{L}{\rho}$  et donc

$$\frac{\rho}{2L} \|g_{1/\hat{L}_j}(r_j)\|^2 \leq F(r_j) - F(\hat{r}_j).$$

Le deuxième point de la Proposition 3.3.2 garantit que  $F(r_{j+1}) \leq F(\hat{r}_j)$ , et par conséquent :

$$\frac{\rho}{2L} \|g_{1/\hat{L}_j}(r_j)\|^2 \leq F(r_j) - F(r_{j+1}).$$

# Maîtrise de l'inertie par l'introduction d'un terme de friction lié à la hessienne

---

Ce chapitre porte sur l'étude du système dynamique (DIN-AVD) qui correspond au système associé au schéma de gradient accéléré de Nesterov (AVD) avec un terme supplémentaire lié à la hessienne de  $F$ . L'introduction de ce terme de friction permet de **réduire le comportement oscillatoire des trajectoires solutions**. De plus, l'EDO (DIN-AVD) peut être discrétisée de manière à obtenir des **schémas de premier ordre**, c'est-à-dire qu'ils ne nécessitent que le gradient de  $F$ . Les propriétés d'atténuation des oscillations de ce système peuvent donc être exploitées via des schémas dérivés.

On s'intéresse ici aux propriétés de convergence des trajectoires de ce système lorsque  $F$  vérifie des hypothèses de croissance. Dans ce contexte, on montre que l'introduction du terme lié à la hessienne permet d'obtenir des **propriétés d'intégrabilité sur le gradient de la fonction  $F$**  qui sont de plus en plus fortes à mesure que l'hypothèse de croissance considérée est restrictive. En s'appuyant sur le lien entre ces hypothèses de croissance et l'inégalité de Łojasiewicz, on obtient que ces propriétés ont un impact sur la décroissance de l'erreur. Le Théorème 4.2.2 assure notamment que pour des fonctions à croissance quadratique, **les garanties théoriques de décroissance de l'erreur le long des trajectoires de (DIN-AVD) sont meilleures que celles du système (AVD)**. Ces résultats témoignent de l'apport de ce terme additionnel lorsque  $F$  satisfait des hypothèses de géométrie plus fortes que la convexité.

Des explorations numériques autour du schéma (IGAHD), dérivé d'un système proche de (DIN-AVD), sont proposées pour un problème de moindres carrés.

L'essentiel de ce travail a été mené avec Jean-François Aujol, Charles Dossal, Văn Hào Hoàng et Aude Rondepierre et a fait l'objet de l'écriture d'un article [23].

## Contents

---

|            |   |           |
|------------|---|-----------|
| <b>4.1</b> | <b>Contexte</b> . . . . .   | <b>94</b> |
| 4.1.1      | Rappels sur les systèmes dynamiques . . . . .                     | 94        |
| 4.1.2      | Un terme lié à la hessienne pour atténuer les oscillations . . .  | 94        |
| 4.1.3      | Méthodes numériques associées . . . . .                           | 96        |
| <b>4.2</b> | <b>Résultats de convergence sous conditions de croissance</b> . . | <b>97</b> |
| 4.2.1      | Fonctions à croissance quadratique . . . . .                      | 97        |

|            |   |            |
|------------|---|------------|
| 4.2.2      | Fonctions plates . . . . .                                      | 101        |
| <b>4.3</b> | <b>Observations numériques . . . . .</b>                        | <b>104</b> |
| 4.3.1      | Atténuation des oscillations sur un problème de moindres carrés | 104        |
| 4.3.2      | Étude de la valeur critique de $\beta$ . . . . .                | 105        |
| <b>4.A</b> | <b>Preuves . . . . .</b>  | <b>110</b> |
| 4.A.1      | Preuve du Théorème 4.2.1 . . . . .                              | 110        |
| 4.A.2      | Preuve du Théorème 4.2.2 . . . . .                              | 115        |
| 4.A.3      | Preuve du Corollaire 4.2.3 . . . . .                            | 116        |
| 4.A.4      | Preuve du Théorème 4.2.4 . . . . .                              | 117        |
| 4.A.5      | Preuve du Corollaire 4.2.5 . . . . .                            | 121        |
| 4.A.6      | Autres preuves . . . . .  | 122        |

## 4.1 Contexte

### 4.1.1 Rappels sur les systèmes dynamiques

Comme cela a été détaillé dans la Section 2.3.1, schémas d'optimisation et systèmes dynamiques sont étroitement liés. Ce parallèle ne date pas d'hier et dès 1964, Polyak propose la méthode Heavy-Ball [112] qui est dérivée du système de Heavy-Ball avec friction :

$$\forall t \geq t_0, \ddot{x}(t) + \alpha \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{HBF})$$

avec  $\alpha > 0$  et  $t_0 \geq 0$ .

En 2014, Su, Boyd et Candès [127] introduisent l'EDO (AVD) :

$$\forall t \geq t_0, \ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{AVD})$$

avec  $\alpha > 0$  et  $t_0 \geq 0$ , dont la discrétisation permet de retrouver le schéma de gradient accéléré de Nesterov [103] (pour le choix  $\alpha = 3$ ).

Ces systèmes dynamiques sont très utiles pour l'étude de la convergence des schémas qui leur sont associés (voir la Section 2.3.2). En effet, les propriétés de convergence des trajectoires d'un système dynamique peuvent généralement être étendues aux itérés des schémas numériques associés sous certaines conditions. Le papier de Su, Boyd et Candès [127] a été marquant dans le paysage de l'optimisation puisque de nombreux travaux s'appuyant sur une analyse de tels systèmes dynamiques lui ont fait suite [36, 126, 25, 24, 16, 96, 17, 27, 26, 91, 6, 11, 13, 29, 76].

### 4.1.2 Un terme lié à la hessienne pour atténuer les oscillations

Le système dynamique auquel on s'intéresse dans ce chapitre est le suivant :

$$\forall t \geq t_0, \ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \beta H_F(x(t)) \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{DIN-AVD})$$

où  $t_0 > 0$ ,  $\alpha > 0$ ,  $\beta > 0$ ,  $x(t_0) \in \mathbb{R}^N$ ,  $\dot{x}(t_0) = 0$  et  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  est convexe et  $C^2$ . On note  $H_F$  la hessienne de  $F$ .

Une des particularités de ce système est qu'il peut être discrétisé de manière à obtenir des schémas de premier ordre, c'est-à-dire qui ne s'appuient pas sur la hessienne. Ainsi, les propriétés de convergence des trajectoires de (DIN-AVD) peuvent être étendues à des méthodes numériques sans nécessiter que la fonction à minimiser soit deux fois dérivable (voir Section 4.1.3).

Cette EDO introduite en 2016 par Attouch, Peypouquet et Redont [20] fait notamment suite à [5]. En effet, Alvarez, Attouch, Bolte et Redont introduisent en 2002 le système inertiel dynamique à la Newton (DIN) :

$$\ddot{x}(t) + \alpha \dot{x}(t) + \beta H_F(x(t)) \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{DIN})$$

qui peut être vu comme une combinaison entre le système de Heavy-Ball avec friction (HBF) et le système dynamique de Newton. Les auteurs étudient les propriétés des trajectoires de ce système et montrent notamment que celles-ci convergent faiblement vers un minimiseur de  $F$  lorsque  $F$  est convexe et a un ensemble de minimiseurs non vide. La présence du terme de friction supplémentaire permet en pratique d'atténuer les oscillations des trajectoires lorsque le paramètre  $\alpha$  n'est pas bien calibré.

Le système (DIN-AVD) proposé par Attouch, Peypouquet et Redont dans [21] combine un terme de friction asymptotiquement évanescant qui est présent dans (AVD) et un terme de friction géométrique rattaché à  $H_F$  déjà présent dans (DIN). Cette combinaison semble particulièrement pertinente puisque le terme de friction évanescant  $\frac{\alpha}{t}$  peut être à l'origine d'oscillations et le terme lié à la hessienne permet d'atténuer celles-ci. Les auteurs démontrent que pour une fonction  $F$  convexe et  $C^2$ , la décroissance de l'erreur se comporte comme  $o(k^{-2})$  pour  $\alpha > 3$  et  $\beta > 0$ . Ce résultat n'est pas le plus significatif puisqu'il est déjà valable pour les trajectoires de (AVD) c'est-à-dire pour  $\beta = 0$ . En revanche, il est montré que pour  $\alpha \geq 3$  et  $\beta > 0$ , on a :

$$\int_{t_0}^{+\infty} t^2 \|\nabla F(x(t))\|^2 dt < +\infty. \quad (4.1.1)$$

Cette propriété de convergence rapide du gradient de  $F$  le long de la trajectoire reflète l'**atténuation des oscillations** que permet l'ajout du terme  $\beta H_F(x(t)) \dot{x}(t)$ . Par la suite, Attouch, Chbani, Fadili et Riahi [14, 15] proposent une version plus générale de (DIN-AVD) :

$$\ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \beta(t) H_F(x(t)) \dot{x}(t) + b(t) \nabla F(x(t)) = 0, \quad (4.1.2)$$

et démontrent des résultats de convergence équivalents à ceux de (DIN-AVD) sous certaines conditions sur  $b$ .

Récemment, Shi, Du, Jordan et Su ont introduit le concept d'EDO haute-résolution dans [125]. Ces systèmes permettent de modéliser plus précisément le comportement de schémas numériques et notamment de différencier plusieurs schémas issus d'un même système. Il est notable que le schéma de gradient accéléré de

Nesterov [103] peut être modélisé par l'EDO haute-résolution suivante [125] :

$$\ddot{x}(t) + \frac{3}{t}\dot{x}(t) + \sqrt{s}H_F(x(t))\dot{x}(t) + \left(1 + \frac{3\sqrt{s}}{2t}\right)\nabla F(x(t)) = 0, \quad (4.1.3)$$

pour  $s > 0$  petit, qui correspond à (4.1.2) pour un certain choix de paramètres. Rappelons que cette EDO est dite haute-résolution car elle fait intervenir un pas  $s$  choisi suffisamment petit qui permet d'approcher fidèlement le comportement du schéma de gradient accéléré de Nesterov.

La stabilité d'un système proche de (DIN-AVD) est étudiée par Attouch, Fadili et Kungurtsev [18] en ajoutant un terme perturbatif. Maulen et Peypouquet proposent également dans [95] un schéma de restart pour (DIN-AVD) inspiré du *speed restart scheme* de Su, Boyd et Candès [127] avec des résultats de convergence.

### 4.1.3 Méthodes numériques associées

Les propriétés de convergence des trajectoires (et en particulier l'atténuation des oscillations liée au terme additionnel) du système (DIN-AVD) et sa version générale (4.1.2) ont conduit à l'élaboration de schémas numériques dérivés dans le but d'obtenir des propriétés similaires. Dans [14, 15], Attouch, Chbani, Fadili et Riahi introduisent ainsi le schéma (IGAHD) (pour *Inertial Gradient Algorithm with Hessian-driven Damping*)

$$\begin{cases} x_k = y_{k-1} - s\nabla F(y_{k-1}), \\ y_k = x_k + \alpha_k(x_k - x_{k-1}) - \beta\sqrt{s}(\nabla F(x_k) - \nabla F(x_{k-1})) - \frac{\beta\sqrt{s}}{k}\nabla F(x_{k-1}), \end{cases} \quad (\text{IGAHD})$$

avec  $\alpha_k = \frac{k-1}{k+\alpha-1}$ ,  $\alpha > 0$ ,  $\beta \geq 0$  et  $s > 0$ , qui est une discrétisation de (4.1.2) pour  $b : t \mapsto 1 + \frac{\beta}{t}$ . On peut remarquer que ce schéma ne requiert pas de calculer la hessienne puisque celle-ci est discrétisée comme une différence de gradient. Ainsi, l'hypothèse  $C^2$  n'est pas nécessaire pour appliquer la méthode à  $F$ . Les auteurs montrent que pour  $\alpha \geq 3$ ,  $s \leq \frac{1}{L}$  et  $\beta \in (0, 2\sqrt{s})$ , on a :

$$F(x_k) - F^* = \mathcal{O}(k^{-2}), \quad (4.1.4)$$

et surtout que

$$\sum_{k \in \mathbb{N}} k^2 \|\nabla F(x_k)\|^2 < +\infty. \quad (4.1.5)$$

On retrouve donc les propriétés de convergence des trajectoires de (DIN-AVD) avec notamment la décroissance rapide du gradient de  $F$  le long des itérés. Notons qu'en pratique, l'hypothèse  $\beta \in (0, 2\sqrt{s})$  est restrictive puisqu'elle force  $\beta < \frac{2}{\sqrt{L}}$  où  $L$  est généralement grand. Ainsi, (IGAHD) peut être vu comme une version perturbée du schéma de gradient accéléré de Nesterov où le terme perturbatif calibré par  $\beta$  doit être contrôlé pour obtenir des propriétés de convergence.

On peut préciser que (IGAHD) n'est pas le seul schéma dérivé de (DIN-AVD). On pourra consulter [14, 15] pour s'informer sur ces autres méthodes.

## 4.2 Résultats de convergence sous conditions de croissance

Le but de ce travail est d'analyser le comportement des trajectoires de (DIN-AVD) sous l'hypothèse que  $F$  vérifie des conditions de croissance. En particulier, une des questions que l'on se pose est : est-ce que les propriétés d'intégrabilité du gradient démontrées dans le cas convexe (4.1.1) peuvent être renforcées sous des hypothèses de croissance ?

En effet, on sait que les trajectoires de (AVD) sont particulièrement enclines à osciller autour de l'ensemble des minimiseurs de la fonction  $F$  lorsque celle-ci vérifie des hypothèses de croissance, ralentissant ainsi leur convergence. Si ce problème d'oscillation peut être atténué par l'ajout du terme de friction lié à la hessienne, alors les schémas issus de (DIN-AVD) pourraient être des solutions pertinentes lors de la minimisation de fonctions satisfaisant des hypothèses de croissance. On peut voir cette approche comme une alternative aux schémas de restart qui atténuent les oscillations en réinitialisant l'inertie occasionnellement. Maulen et Peypouquet proposent d'ailleurs de combiner ces deux philosophies dans [95].

Cette étude est dans la lignée de celles proposées par Aujol, Dossal et Rondepierre dans [27, 26] pour les trajectoires de (AVD).

### 4.2.1 Fonctions à croissance quadratique

On considère d'abord une fonction  $F$  de classe  $C^2$  admettant un unique minimiseur  $x^*$ . On suppose également que  $F$  vérifie la condition de croissance quadratique globale  $\mathcal{G}_\mu^2$  pour  $\mu > 0$ , c'est-à-dire :

$$\forall x \in \mathbb{R}^N, \quad \frac{\mu}{2} \|x - x^*\|^2 \leq F(x) - F^*, \quad (4.2.1)$$

et la condition de platitude  $\mathcal{F}_\gamma$  avec  $\gamma \geq 1$  :

$$\forall x \in \mathbb{R}^N, \quad F(x) - F^* \leq \frac{1}{\gamma} \langle \nabla F(x), x - x^* \rangle. \quad (4.2.2)$$

On s'intéresse donc aux fonctions ayant une croissance quadratique ce qui inclut les fonctions fortement convexes. Sous de telles hypothèses, il est montré dans [27, Théorème 4.2] que si  $\gamma \leq 2$  et  $\alpha > 1 + \frac{2}{\gamma}$ , alors les trajectoires de (AVD) vérifient :

$$F(x(t)) - F^* = \mathcal{O}\left(t^{-\frac{2\alpha\gamma}{\gamma+2}}\right). \quad (4.2.3)$$

Dans le même cadre, Aujol, Dossal et Rondepierre donnent une borne supérieure sur l'erreur en temps fini [26, Théorème 5] qui correspond asymptotiquement au résultat ci-dessus. De plus, la borne sur  $F(x(t)) - F^*$  calculée par les auteurs est proportionnelle à  $\alpha^{\frac{2\alpha\gamma}{\gamma+2}}$ . Cela suggère que choisir  $\alpha$  aussi grand que possible pour obtenir les meilleures vitesses n'est pas toujours pertinent.

Les résultats démontrés avec mes collaborateurs sur la convergence des trajectoires de (DIN-AVD) sont résumés ci-dessous. Un schéma de preuve est donné en suivant et la preuve complète se trouve en Section 4.A.1.

**Théorème 4.2.1.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  admettant un unique minimiseur  $x^* \in \mathbb{R}^N$ . Supposons que  $F$  vérifie  $\mathcal{G}_\mu^2$  et  $\mathcal{F}_\gamma$  pour  $\mu > 0$  et  $\gamma \geq 1$ . On note  $x$  la solution de (DIN-AVD) pour tout  $t \geq t_0$  avec  $t_0 > 0$ ,  $\alpha > 0$  et  $\beta > 0$ . Soit  $\lambda = \frac{2\alpha}{\gamma+2} < \alpha$ .

1. Si  $\alpha > 1 + \frac{2}{\gamma}$ , alors pour tout  $t \geq t_0 + \beta(\alpha - \lambda)$ ,

$$F(x(t)) - F^* \leq \frac{K}{(t + \beta(\lambda - \alpha))^{\frac{2\alpha\gamma}{\gamma+2}}} \quad (4.2.4)$$

où  $K$  dépend de  $t_0$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$  et  $\mu$ . En particulier, si  $t_0 \leq \frac{\alpha r^*}{(\gamma+2)\sqrt{\mu}}$ , alors pour tout  $t \geq \frac{\alpha r^*}{(\gamma+2)\sqrt{\mu}} + \beta(\alpha - \lambda)$ , l'inégalité (4.2.4) est vérifiée pour

$$K = C_1 e^{\frac{2\gamma}{\gamma+2} C_2 (\alpha - 1 - \frac{2}{\gamma})} \left(1 + \frac{\beta\gamma\sqrt{\mu}}{r^*}\right) E_m(t_0) \left(\frac{\alpha r^*}{(\gamma+2)\sqrt{\mu}}\right)^{\frac{2\alpha\gamma}{\gamma+2}}, \quad (4.2.5)$$

2. Si  $\alpha = 1 + \frac{2}{\gamma}$ , alors pour tout  $t \geq t_0 + \beta$ ,

$$F(x(t)) - F^* \leq \left( (t_0 + \beta)^2 + \frac{\lambda^2 + \sqrt{\mu}}{\mu} \right) e^{\frac{\beta}{t_0}} \frac{E_m(t_0)}{t(t - \beta)}, \quad (4.2.6)$$

3. Si  $\alpha \geq 1 + \frac{2}{\gamma}$ , alors

$$\int_{t_0}^{+\infty} u^{\frac{2\alpha\gamma}{\gamma+2}} \|\nabla F(x(u))\|^2 du < +\infty, \quad (4.2.7)$$

où :

- $r^*$  est l'unique racine positive du polynôme

$$r \mapsto r^3 - (1 + C_0)r^2 - 2(1 + \sqrt{2})r - 4,$$

- $C_0 = \beta \frac{\sqrt{\mu}\gamma(\gamma\lambda - 1)}{\gamma\lambda - 2}$ ,
- $C_1 = \left(1 + \frac{2}{r^*}\right)^2$ ,
- $C_2 = \frac{1+C_0}{r^*} + \frac{1+\sqrt{2}}{r^{*2}} + \frac{4}{3r^{*3}}$ ,
- $E_m : t \mapsto \left(1 + \frac{\beta\alpha}{t}\right) (F(x(t)) - F^*) + \frac{1}{2} \|\dot{x}(t) + \beta\nabla F(x(t))\|^2$ .

### Éléments de preuve.

La preuve du Théorème 4.2.1 étant longue et technique, on ne donne ici que le raisonnement qui conduit aux résultats énoncés. La preuve complète est détaillée en Section 4.A.1.

Ce théorème se décompose en trois points. Les deux premiers concernent la décroissance de l'erreur  $F(x(t)) - F^*$  pour différents choix de  $\alpha$  et le troisième est un résultat d'intégrabilité sur le gradient de  $F$  le long de la trajectoire  $x$ . Chaque résultat s'appuie sur une analyse de Lyapunov (voir Section 2.3.3) basée sur l'énergie définie comme suit :

$$\mathcal{E}(t) = (t^2 + t\beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{1}{2} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2, \quad (4.2.8)$$

avec  $\lambda = \frac{2\alpha}{\gamma+2}$ .

La première étape de la preuve consiste à obtenir un contrôle sur l'énergie  $\mathcal{E}$  qui se caractérise ici par une inégalité différentielle de la forme suivante :

$$\forall t \geq T, \quad \mathcal{E}'(t) + \frac{p}{t + \beta(\lambda - \alpha)} \mathcal{E}(t) \leq \tilde{\varphi}(t + \beta(\lambda - \alpha)) \mathcal{E}(t), \quad (4.2.9)$$

pour  $T \geq t_0$ ,  $p \geq 0$  et  $\tilde{\varphi} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  bien choisis. Cette inégalité est suffisante pour prouver les deux premiers points mais on obtient en pratique un résultat légèrement plus fort, c'est-à-dire pour tout  $t \geq T$ ,

$$\mathcal{E}'(t) + \frac{p}{t + \beta(\lambda - \alpha)} \mathcal{E}(t) + \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \leq \tilde{\varphi}(t + \beta(\lambda - \alpha)) \mathcal{E}(t), \quad (4.2.10)$$

où le terme supplémentaire permet de prouver le troisième résultat. En définissant  $\mathcal{H}$  comme

$$\mathcal{H} : t \mapsto \mathcal{E}(t)(t + \beta(\lambda - \alpha))^p e^{-\tilde{\Phi}(t + \beta(\lambda - \alpha))}, \quad (4.2.11)$$

où  $\tilde{\Phi}(t) = -\int_t^{+\infty} \tilde{\varphi}(x) dx$ , on peut déduire de (4.2.9) que  $\mathcal{H}$  est décroissante sur  $[T, +\infty)$ .

Ainsi, comme  $\mathcal{E}(t) \geq (t + \beta(\lambda - \alpha))^2 (F(x(t)) - F^*)$ , on peut écrire que pour tout  $t_1 \geq T$ ,

$$\forall t \geq t_1, \quad F(x(t)) - F^* \leq e^{-\tilde{\Phi}(t + \beta(\lambda - \alpha))} \frac{\mathcal{H}(t_1)}{(t + \beta(\lambda - \alpha))^{p+2}}. \quad (4.2.12)$$

Cette inégalité permet d'aboutir aux deux premiers résultats :

- l'équation (4.2.4) donnée dans le premier point provient d'une simplification de (4.2.12). La borne spécifiée dans (4.2.5) repose sur une optimisation du paramètre  $t_1$  dans le but d'obtenir le contrôle le plus précis sur l'erreur  $F(x(t)) - F^*$ . En ce sens, on choisit  $t_1$  comme le minimiseur de

$$t \mapsto (t + \beta(\lambda - \alpha))^p e^{-\tilde{\Phi}(t + \beta(\lambda - \alpha))}.$$

En développant (4.2.5) pour cette valeur de  $t_1$ , on trouve la borne voulue.

- le second point est une réécriture de (4.2.12) dans le cas  $\alpha = 1 + \frac{2}{\gamma}$  pour  $t_1 = T$ .

La démonstration du troisième résultat s'appuie sur (4.2.10). En définissant  $\mathcal{G}$  comme

$$\mathcal{G} : t \mapsto \mathcal{H}(t) + \beta \int_T^t u(u + \beta(\lambda - \alpha))^{p+1} e^{-\tilde{\Phi}(u + \beta(\lambda - \alpha))} \|\nabla F(x(u))\|^2 du, \quad (4.2.13)$$

l'inégalité (4.2.10) implique que  $\mathcal{G}$  décroît sur  $[T, +\infty)$  et comme  $\mathcal{H}$  est une fonction positive :

$$\forall t \geq T, \quad \beta \int_T^t u(u + \beta(\lambda - \alpha))^{p+1} e^{-\tilde{\Phi}(u + \beta(\lambda - \alpha))} \|\nabla F(x(u))\|^2 du \leq \mathcal{G}(T). \quad (4.2.14)$$



Quelques calculs supplémentaires amènent au résultat voulu. □

Le premier point du Théorème 4.2.1 garantit que les trajectoires de (DIN-AVD) assurent une décroissance asymptotique de l'erreur similaire à celle obtenue pour les trajectoires de (AVD) si  $\alpha > 1 + \frac{2}{\gamma}$ , puisque  $F(x(t)) - F^* = \mathcal{O}\left(t^{-\frac{2\alpha\gamma}{\gamma+2}}\right)$  dans les deux cas. Le deuxième point permet d'étendre ce résultat au cas  $\alpha = 1 + \frac{2}{\gamma}$ . Notons que toute fonction  $\mu$ -fortement convexe vérifie  $\mathcal{G}_\mu^2$  et  $\mathcal{F}_1$  (convexité) et en conséquence, on retrouve le résultat démontré par Attouch, Peypouquet et Redont [21] pour cette classe de fonctions (si les hypothèses additionnelles sont également vérifiées) :

$$F(x(t)) - F^* = \mathcal{O}\left(t^{-\frac{2\alpha}{3}}\right). \tag{4.2.15}$$

Notre résultat est plus fort puisqu'il tient pour des hypothèses plus faibles, c'est-à-dire les fonctions convexes à croissance quadratique.

De plus, (4.2.5) fournit une borne non-asymptotique fine sur  $F(x(t)) - F^*$  qui souligne l'influence de  $\alpha$  et  $\beta$  sur la convergence des trajectoires. Lorsque  $\beta$  tend vers 0, on retrouve la même borne que celle donnée dans [26, Théorème 5] pour les trajectoires de (AVD). On peut également remarquer que la borne supérieure sur l'erreur est proportionnelle à  $\alpha^{\frac{2\alpha\gamma}{\gamma+2}}$  comme pour les trajectoires de (AVD). Cela souligne que choisir  $\alpha$  trop grand pourrait être contre-productif. En optimisant les bornes (4.2.5) et (4.2.6) par rapport à  $\beta$ , on trouve que la valeur optimale est 0. L'ajout du terme de friction lié à la hessienne ne permet donc pas d'obtenir des meilleurs résultats de convergence en suivant nos schémas de preuve. Cependant, la présence de ce terme permet d'obtenir une propriété d'intégrabilité du gradient qui n'est pas valable pour (AVD). En effet, le troisième point est directement lié à la friction géométrique liée à la hessienne et cette convergence rapide du gradient de  $F$  le long de la trajectoire garantit une atténuation des oscillations. Il s'agit également d'une amélioration du résultat d'intégrabilité prouvé par Attouch, Peypouquet et Redont [21] pour (DIN-AVD) dans le cas convexe :

$$\int_{t_0}^{+\infty} u^2 \|\nabla F(x(u))\|^2 du < +\infty. \tag{4.2.16}$$

On considère ici des fonctions convexes qui vérifient  $\mathcal{G}_\mu^2$  pour  $\mu > 0$  et on peut donc appliquer la Proposition 2.1.16 qui relie l'hypothèse de croissance quadratique à la propriété de Łojasiewicz d'exposant  $\frac{1}{2}$ . Comme on fait l'hypothèse additionnelle que  $F$  est de classe  $C^2$ , cela signifie que pour tout  $x \in \mathbb{R}^N$  :

$$\frac{\mu}{2}(F(x) - F^*) \leq \|\nabla F(x)\|^2. \tag{4.2.17}$$

Ainsi, notre résultat d'intégrabilité sur le gradient implique automatiquement une propriété similaire sur l'erreur. On exploite cet argument pour établir le théorème ci-dessous dont la preuve se trouve en Section 4.A.2.

**Théorème 4.2.2.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  admettant un unique minimiseur  $x^* \in \mathbb{R}^N$  et vérifiant  $\mathcal{G}_\mu^2$  pour  $\mu > 0$ . On note  $x$  la solution de (DIN-AVD) pour tout  $t \geq t_0$  avec  $t_0 > 0$ ,  $\alpha > 0$  et  $\beta > 0$ . Pour tout  $\varepsilon \in (0, 1)$ ,

$$F(x(t)) - F^* = \mathcal{O}(t^{-\alpha+\varepsilon}), \quad (4.2.18)$$

et

$$\int_{t_0}^{+\infty} u^{\alpha-\varepsilon} (F(x(u)) - F^*) du < +\infty. \quad (4.2.19)$$

Le Théorème 4.2.2 donne un meilleur résultat asymptotique que celui démontré pour les fonctions fortement convexes dans [21] ( $\mathcal{O}(t^{-\frac{2\alpha}{3}})$ ) tout en étant valable sous des hypothèses plus faibles. La propriété d'intégrabilité énoncée dans le deuxième point permet d'obtenir les résultats ci-dessous. Une preuve est donnée en Section 4.A.3.

**Corollaire 4.2.3.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  admettant un unique minimiseur  $x^* \in \mathbb{R}^N$  et vérifiant  $\mathcal{G}_\mu^2$  pour  $\mu > 0$ . On note  $x$  la solution de (DIN-AVD) pour tout  $t \geq t_0$  avec  $t_0 > 0$ ,  $\alpha > 0$  et  $\beta > 0$ . Pour tout  $\varepsilon \in (0, 1)$ , quand  $t$  tend vers  $+\infty$ ,

1.

$$F(z(t)) - F^* = o(t^{-\alpha-1+\varepsilon}), \quad (4.2.20)$$

$$\text{où } z : t \mapsto \frac{\int_{t/2}^t u^{\alpha-\varepsilon} x(u) du}{\int_{t/2}^t u^{\alpha-\varepsilon} du}.$$

2.

$$\inf_{u \in [t/2, t]} (F(x(u)) - F^*) = o(t^{-\alpha-1+\varepsilon}). \quad (4.2.21)$$

3.

$$\liminf_{t \rightarrow +\infty} t^{\alpha+1-\varepsilon} \log(t)(F(x(t)) - F^*) = 0, \quad (4.2.22)$$

$$\text{où } \liminf_{t \rightarrow +\infty} f(t) = \lim_{t \rightarrow +\infty} \left[ \inf_{\tau \geq t} f(\tau) \right].$$

Notons que le Théorème 4.2.2 s'appuie sur le Lemme 4.A.4 qui indique que toute fonction convexe  $C^2$  satisfait automatiquement  $\mathcal{F}_{2-\delta}$  localement pour tout  $\delta \in (0, 1]$ . L'hypothèse  $C^2$  est utilisée ici pour que la hessienne soit bien définie mais il est possible de dériver des schémas numériques de (DIN-AVD) sans avoir besoin de cette hypothèse. Ainsi, les résultats énoncés dans ce théorème sont forts mais risquent de ne pas être transférables dans le cadre discret sans cette hypothèse  $C^2$ .

## 4.2.2 Fonctions plates

Dans cette section, on s'intéresse aux fonctions qui vérifient  $\mathcal{F}_{\gamma_1}$  et  $\mathcal{G}^{\gamma_2}$  pour  $\gamma_1 \geq \gamma_2 > 2$ . De telles fonctions sont dites à géométrie plate puisqu'elles se comportent comme  $\|x - x^*\|^{\gamma_1}$  autour de leur minimiseur avec  $\gamma_1 > 2$ . Plus  $\gamma_1$  est grand, plus la fonction est plate alors que pour des valeurs proches de 2, on se rapproche d'une

fonction à croissance quadratique. Dans [27], Aujol, Dossal et Rondepierre étudient la convergence des trajectoires de (AVD) pour cette géométrie et démontrent que si  $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$ , alors

$$F(x(t)) - F^* = \mathcal{O}\left(t^{-\frac{2\gamma_2}{\gamma_2-2}}\right). \quad (4.2.23)$$

On énonce un résultat de convergence similaire pour les trajectoires de (DIN-AVD) ainsi qu'une propriété d'intégrabilité de  $\nabla F$ . On donne un schéma de preuve au-dessous et la preuve complète se trouve en Section 4.A.4.

**Théorème 4.2.4.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  admettant un unique minimiseur  $x^* \in \mathbb{R}^N$ . Supposons que  $F$  vérifie  $\mathcal{F}_{\gamma_1}$  et  $\mathcal{G}^{\gamma_2}$  pour  $\gamma_1 > 2$  et  $\gamma_2 > 2$  tels que  $\gamma_1 \geq \gamma_2$ . On note  $x$  la solution de (DIN-AVD) pour tout  $t \geq t_0$  où  $t_0 > 0$ ,  $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$  et  $\beta > 0$ . Quand  $t$  tend vers  $+\infty$ ,*

$$F(x(t)) - F^* = \mathcal{O}\left(t^{-\frac{2\gamma_1}{\gamma_1-2}}\right). \quad (4.2.24)$$

De plus,

$$\int_{t_0}^{+\infty} u^{\frac{2\gamma_1}{\gamma_1-2}} \|\nabla F(x(u))\|^2 du < +\infty. \quad (4.2.25)$$

**Éléments de preuve.** La preuve du Théorème 4.2.4 s'appuie sur une analyse de Lyapunov et en particulier sur deux énergies  $\mathcal{E}$  et  $\mathcal{H}$  définies comme suit :

$$\begin{aligned} \mathcal{E}(t) &= (t^2 + t\beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{\xi}{2} \|x(t) - x^*\|^2 \\ &\quad + \frac{1}{2} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2, \end{aligned} \quad (4.2.26)$$

et

$$\mathcal{H}(t) = t^p \mathcal{E}(t), \quad (4.2.27)$$

où  $x^*$  est l'unique minimiseur de  $F$ ,  $\lambda \in \mathbb{R}$ ,  $\xi \in \mathbb{R}$  et  $p > 0$ .

La première étape est de montrer que pour un jeu de paramètres bien choisi  $(\lambda, \xi, p)$  on a :

$$\forall t \geq t_1, \mathcal{H}'(t) + \beta t^{p+1} (t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \leq t^p \beta C_1 (F(x(t)) - F^*), \quad (4.2.28)$$

avec  $t_1 > 0$  et  $C_1 > 0$ . Contrairement à ce que l'on aimerait obtenir pour contrôler  $\mathcal{H}$ , le terme de droite est strictement positif.

En conséquence, la deuxième étape est d'étudier l'énergie  $\mathcal{G}$  définie par :

$$\mathcal{G} : t \mapsto \mathcal{H}(t) - \beta C_1 \int_{t_1}^t u^p (F(x(u)) - F^*) du, \quad (4.2.29)$$

avec  $t_1 > T$  pour  $T > 0$  bien choisi. En fait, on définit ici  $\mathcal{G}$  de manière à ce que (4.2.28) implique que  $\mathcal{G}'(t) \leq 0$ . En exploitant la décroissance de  $\mathcal{G}$ , on peut montrer que :

$$F(x(t)) - F^* = \mathcal{O}\left(\frac{1}{t^{p+2}}\right). \quad (4.2.30)$$

Pour cela, on introduit la fonction  $v$  définie par :

$$v : t \mapsto t(t + \beta(\lambda - \alpha))^{p+1}(F(x(t)) - F^*), \quad (4.2.31)$$

et on montre qu'elle est bornée en utilisant les hypothèses de géométrie vérifiées par  $F$ .

Pour démontrer le troisième point, c'est-à-dire que :

$$\int_{t_0}^{+\infty} u^{\frac{2\gamma_1}{\gamma_1-2}} \|\nabla F(x(u))\|^2 du < +\infty, \quad (4.2.32)$$

on s'appuie sur une nouvelle fonction  $\mathcal{F}$  définie par

$$\mathcal{F} : t \mapsto \mathcal{H}(t) - \beta C_1 \int_{t_1}^t u^{p-1} a(u) du + \beta \int_{t_1}^t u^{p+1} (u + \beta(\lambda - \alpha)) \|\nabla F(x(u))\|^2 du. \quad (4.2.33)$$

Cette fonction est décroissante sur  $[t_1, +\infty)$  d'après (4.2.28) et en utilisant le fait que  $v$  est bornée, on peut arriver au résultat espéré.  $\square$

Le résultat asymptotique (4.2.24) est légèrement moins bon que (4.2.23) démontré dans [27] pour les trajectoires de (AVD) car on fait l'hypothèse que  $\gamma_1 \geq \gamma_2$ . On donne cependant un résultat d'intégrabilité sur le gradient qui n'est a priori pas valable pour (AVD) et qui est plus fort que celui donné par Attouch, Peypouquet et Redont [21] sous une simple hypothèse de convexité.

Comme l'hypothèse de croissance quadratique, l'hypothèse  $\mathcal{G}^{\gamma_2}$  peut être reliée à l'inégalité de Łojasiewicz d'exposant  $1 - \frac{1}{\gamma_2}$ . En appliquant la Proposition 2.1.18, on obtient que si  $F$  est de classe  $C^1$  et vérifie  $\mathcal{G}^{\gamma_2}$  pour  $\gamma_2 > 2$ , alors il existe  $c > 0$  tel que pour tout  $x \in \mathbb{R}^N$

$$c(F(x) - F^*) \leq \|\nabla F(x)\|^{\frac{\gamma_2}{\gamma_2-1}}. \quad (4.2.34)$$

Cette propriété combinée à (4.2.25) permet d'écrire que sous les hypothèses du Théorème 4.2.4, on a :

$$\int_{t_0}^{+\infty} u^{\frac{2\gamma_1}{\gamma_1-2}} (F(x(u)) - F^*)^{\frac{\gamma_2-1}{2\gamma_2}} du < +\infty. \quad (4.2.35)$$

Ce résultat implique notamment le corollaire suivant dont la preuve se trouve en Section 4.A.5.

**Corollaire 4.2.5.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  admettant un unique minimiseur  $x^* \in \mathbb{R}^N$ . Supposons que  $F$  vérifie  $\mathcal{F}_{\gamma_1}$  et  $\mathcal{G}^{\gamma_2}$  pour  $\gamma_1 > 2$  et  $\gamma_2 > 2$  tels que  $\gamma_1 \geq \gamma_2$ . On note  $x$  la solution de (DIN-AVD) pour tout  $t \geq t_0$  où  $t_0 > 0$ ,  $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$  et  $\beta > 0$ . Quand  $t$  tend vers  $+\infty$ ,*

$$\inf_{u \in [t/2, t]} F(x(u)) - F^* = o\left(t^{-\frac{(3\gamma_1-2)\gamma_2}{2(\gamma_1-2)(\gamma_2-1)}}\right). \quad (4.2.36)$$

On peut constater que pour ces hypothèses de géométrie sur  $F$ , l'inégalité de Łojasiewicz ne permet pas d'obtenir des résultats aussi fort que dans le Corollaire 4.2.3. Toutefois, pour  $\gamma_1 > 2$  et  $\gamma_2 \in \left(2, \frac{4\gamma_1}{\gamma_1+2}\right)$ , on a  $\frac{(3\gamma_1-2)\gamma_2}{2(\gamma_1-2)(\gamma_2-1)} > \frac{2\gamma_1}{\gamma_1-2}$ . Ainsi, dans ce cadre, le taux de convergence donné par le Corollaire 4.2.5 sur  $\inf_{u \in [t/2, t]} F(x(u)) - F^*$  est meilleur que celui prouvé dans le Théorème 4.2.4 pour  $F(x(t)) - F^*$ .

### 4.3 Observations numériques

#### 4.3.1 Atténuation des oscillations sur un problème de moindres carrés

Après avoir donné des résultats théoriques, je présente ici quelques observations sur des expériences numériques. On a vu dans la section précédente que l'ajout du terme lié à la hessienne permet d'obtenir de nouvelles garanties théoriques avec notamment l'intégrabilité du gradient que l'on peut relier à une atténuation du comportement oscillatoire des trajectoires. On illustre ici cette propriété sur un problème classique de moindres carrés :

$$\min_{x \in \mathbb{R}^n} F(x) := \|Ax - b\|_2^2, \tag{4.3.1}$$

où  $A \in \mathcal{M}_{N \times N}(\mathbb{R})$  and  $b \in \mathbb{R}^N$ . La fonction  $F$  ainsi définie est convexe,  $C^2$  et vérifie l'hypothèse  $\mathcal{G}_\mu^2$  où  $\mu > 0$  est la plus petite valeur propre de  $A^T A$ . En fait, si le noyau de  $A$  est réduit à  $\{0\}$ ,  $F$  est  $\mu$ -fortement convexe. Il est important de préciser que ce problème ne présente pas un intérêt majeur en pratique puisqu'on peut calculer sa solution directement avec l'expression  $x^* = (A^T A)^{-1} A^T b$ . Il nous permet toutefois d'analyser le comportement des itérés de schémas numériques.

On applique (IGAHD), c'est-à-dire l'algorithme introduit dans [14, 15] défini par :

$$\begin{cases} x_k = y_{k-1} - s \nabla F(y_{k-1}), \\ y_k = x_k + \alpha_k(x_k - x_{k-1}) - \beta \sqrt{s}(\nabla F(x_k) - \nabla F(x_{k-1})) - \frac{\beta \sqrt{s}}{k} \nabla F(x_{k-1}), \end{cases} \tag{4.3.2}$$

où  $\alpha_k = \frac{k-1}{k+\alpha-1}$ ,  $\alpha > 0$ ,  $\beta \geq 0$  et  $s > 0$ . Pour  $\beta = 0$  et  $\alpha = 3$ , on retrouve le schéma de gradient accéléré de Nesterov. Ce schéma numérique est dérivé de l'EDO ci-dessous qui est une légère variation de (DIN-AVD) :

$$\forall t \geq t_0, \ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \beta H_F(x(t)) \dot{x}(t) + \left(1 + \frac{\beta}{t}\right) \nabla F(x(t)) = 0. \tag{4.3.3}$$

L'ajout du terme évanescent rattaché au gradient de  $F$  conserve la structure de la dynamique tout en facilitant les calculs pour dériver un schéma. Nous n'avons pas travaillé sur la convergence théorique des itérés de (IGAHD) donc je conseille aux lecteurs de consulter [14, 15] pour plus de détails. Comme précisé en Section 4.1.3,

les auteurs Attouch, Chbani, Fadili et Riahi montrent notamment que les itérés de (IGAHD) vérifient

$$F(x_k) - F^* = \mathcal{O}(k^{-2}), \quad (4.3.4)$$

si  $\alpha \geq 3$ ,  $s \leq \frac{1}{L}$  et  $\beta \leq 2\sqrt{s}$

On compare dans la Figure 4.1 la convergence des itérés de (IGAHD) pour plusieurs valeurs de  $\beta$  avec celle des itérés de la méthode de gradient accéléré de Nesterov pour observer l'influence du terme lié à la hessienne. On peut constater que celui-ci a une influence non-négligeable sur le comportement des itérés et en particulier sur leur oscillation. En effet, en augmentant la valeur de  $\beta$ , on atténue de plus en plus les oscillations en commençant par celles de haute fréquence. Pour cette expérience numérique, on a généré  $A$  et  $b$  aléatoirement avec  $N = 500$ .

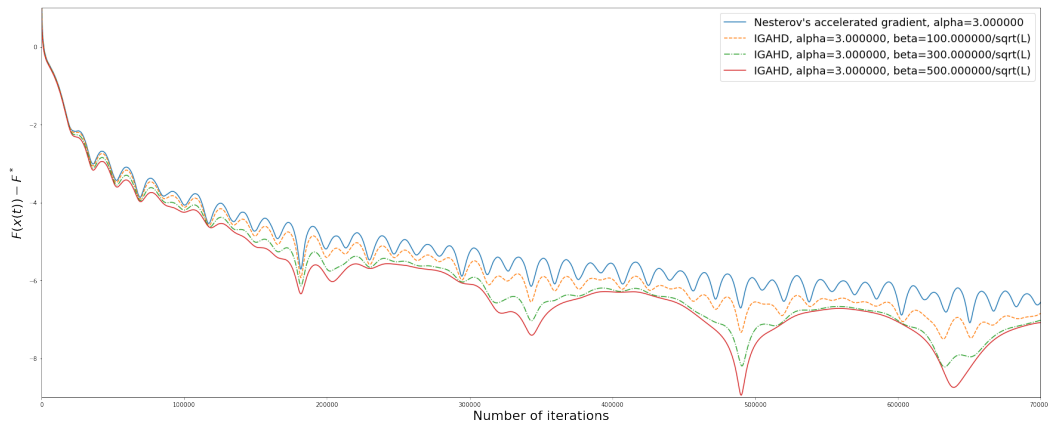


FIGURE 4.1 – Convergence des itérés de (IGAHD) pour différentes valeurs de  $\beta$  sur un problème de moindres carrés ( $N = 500$ ).

### 4.3.2 Étude de la valeur critique de $\beta$

Pour les tests présentés dans la Figure 4.1, on a pris des valeurs de  $\beta$  très grandes ( $\beta \geq 100/\sqrt{L}$ ) et les propriétés de convergence de (IGAHD) démontrées dans [14, 15] ne sont donc pas valables (on devrait avoir  $\beta \leq 2/\sqrt{L}$ ). Dans ce cas, j'ai observé qu'augmenter  $\beta$  est toujours bénéfique jusqu'à une valeur critique pour laquelle les itérés ne convergent plus. La borne théorique  $2/\sqrt{L}$  a donc tendance à être nettement sous-optimale numériquement, ce qui peut être observé sur la Figure 4.2. Cette observation a donc motivé l'étude de la valeur critique du paramètre  $\beta$ . Précisons que cette étude ne porte que sur le cas très précis du problème des moindres carrés.

#### 4.3.2.1 Un schéma de backtracking sur $\beta$

Dans un premier temps, le constat que les plus grandes valeurs de  $\beta$  améliore la convergence de (IGAHD) m'a inspiré l'élaboration d'une stratégie de backtracking sur ce paramètre. Cette méthode est heuristique et aucun résultat théorique n'est

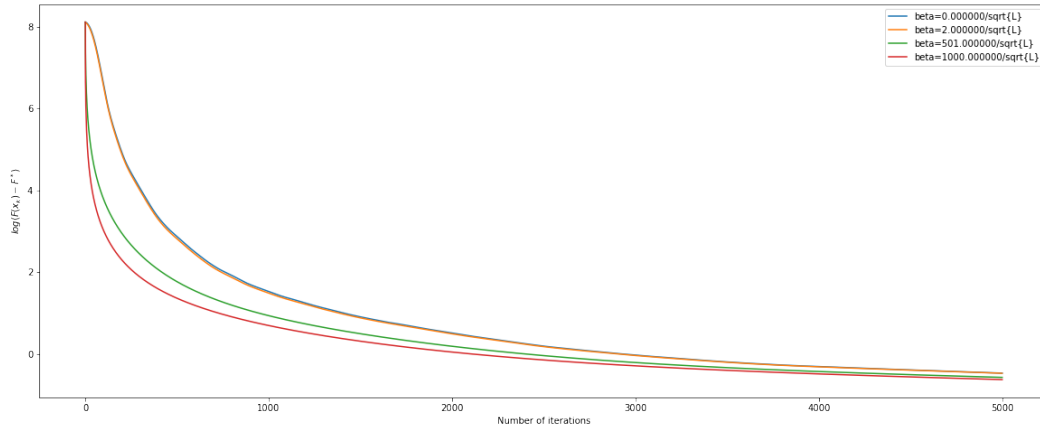


FIGURE 4.2 – Convergence des itérés de (IGAHD) pour différentes valeurs de  $\beta$  sur un problème de moindres carrés ( $N = 1000$ ).

fourni. On écrit d'abord (IGAHD) sous la forme ci-dessous où  $EC$  désigne une condition de sortie.

---

**Algorithme 12** : IGAHD( $x_0, \alpha, \beta, s, EC$ )

---

**Entrée:**  $x_0 \in \mathbb{R}^N$ ,  $\alpha > 0$ ,  $\beta > 0$ ,  $s > 0$ ,  $EC$ .

$k = 0$ ,  $y_0 = x_0$ ;

**répéter**

$k = k + 1$ ;

$x_k = y_k - s \nabla F(y_k)$ ;

$\alpha_k = \frac{k-1}{k+\alpha-1}$ ;

$y_k = x_k + \alpha_k(x_k - x_{k-1}) - \beta \sqrt{s}(\nabla F(x_k) - \nabla F(x_{k-1})) - \frac{\beta \sqrt{s}}{k} \nabla F(x_{k-1})$ ;

**jusqu'à**  $EC$ .

**retourne**  $x_k$ .

---

Le schéma heuristique que je propose s'écrit comme suit où on définit la condition de sortie  $EC$  comme :

$$(\mathcal{EC}) \quad \left\{ \begin{array}{l} (\langle \nabla F(y_{k-1}), x_k - x_{k-1} \rangle > 0 \text{ and } k < k_{max}) \\ \text{ou} \\ \|\nabla F(x_k)\|^2 \leq \varepsilon, \end{array} \right.$$

avec  $k_{max} > 0$ .

---

**Algorithme 13** : IGAHD avec backtracking sur  $\beta$

---

**Entrée:**  $x_0 \in \mathbb{R}^N$ ,  $\alpha > 0$ ,  $\beta_0 > 0$ ,  $s > 0$ ,  $k_{max} > 0$ ,  $\gamma \in (0, 1)$ .

$j = 0$ ;

**répéter**

$\tilde{x} = \text{IGAHD}(x_0, \alpha, \beta_j, s, \mathcal{EC})$ ;

$j = j + 1$ ;

$\beta_j = \gamma\beta_{j-1}$ ;

**jusqu'à**  $\|\nabla F(\tilde{x})\|^2 \leq \varepsilon$ .

**retourne**  $\tilde{x}$ .

---

La stratégie est la suivante : à partir d'un point initial  $x_0$  et une première valeur  $\beta_0 \gg 2/\sqrt{L}$ , on lance (IGAHD) pour vérifier si les premières itérations ont un comportement satisfaisant, c'est-à-dire si la valeur de  $F$  diminue. Si c'est le cas, on conserve cette valeur de  $\beta$  jusqu'à atteindre la précision voulue, sinon on la diminue via un paramètre  $\gamma \in (0, 1)$  ( $\beta_1 = \gamma\beta_0$ ).

Pour vérifier que les premiers itérés de (IGAHD) ne divergent pas, on utilise la condition

$$\langle \nabla F(y_{k-1}), x_k - x_{k-1} \rangle > 0, \quad (4.3.5)$$

qui est une des conditions de restart heuristiques proposées par O'Donoghue et Candès dans [108]. Cette condition est évaluée seulement sur les  $k_{max}$  premières itérations car plusieurs expériences ont montré que les comportements problématiques liés à des valeurs trop grandes de  $\beta$  apparaissent dès les premiers itérés.

La Figure 4.3 permet d'observer la convergence des itérés de ce schéma en comparaison avec celle de (IGAHD). On choisit  $k_{max} = 4$  et  $\beta_0 = 5000/\sqrt{L}$  pour les deux occurrences de l'Algorithme 13. On peut constater un net gain apporté par le backtracking sur  $\beta$  qui permet d'extraire le maximum du terme lié à la hessienne. On doit par contre préciser que la consistance de ces résultats n'est absolument pas garantie puisque la stratégie a été construite autour d'un exemple précis et le choix des paramètres également.

#### 4.3.2.2 Analyse du $\beta$ critique

Concrètement, l'Algorithme 13 cherche la valeur de  $\beta$  la plus élevée assurant la convergence de (IGAHD). Ainsi, on peut estimer la valeur du  $\beta$  critique à partir duquel (IGAHD) ne converge plus en utilisant la stratégie du schéma.

C'est ce qui a été fait pour plusieurs occurrences du problème de moindres carrés (4.3.1) en faisant varier la taille du problème  $N$ , la valeur de  $\alpha$  et la *seed* du générateur aléatoire. De cette manière, on peut analyser la dépendance de cette valeur critique  $\hat{\beta}$  en fonction de différents paramètres. Par la suite, on analysera principalement  $\xi$  défini tel que  $\hat{\beta} = \xi/\sqrt{L}$ .

Ma première intuition était que  $\xi$  allait être corrélée au conditionnement  $\frac{L}{\mu}$  d'une certaine façon. La Figure 4.4 démontre le contraire puisqu'aucune corrélation claire n'apparaît entre  $\xi$  et  $\sqrt{\frac{L}{\mu}}$ .



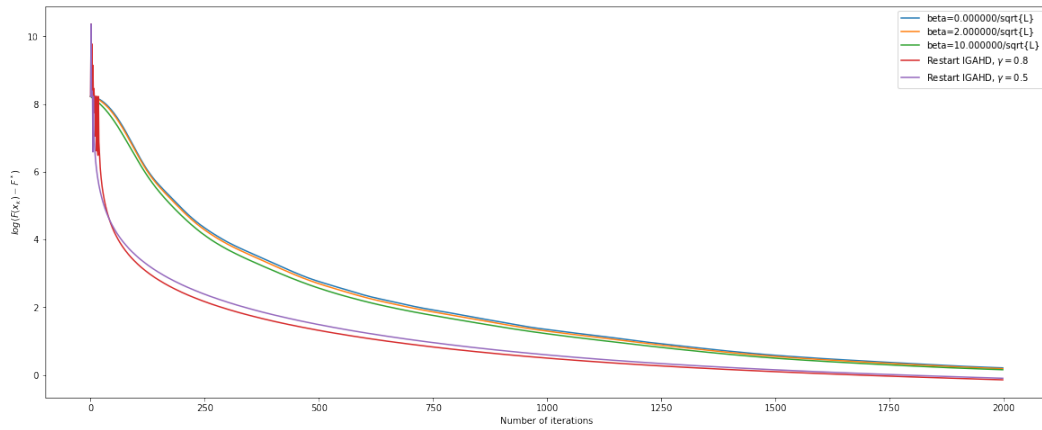


FIGURE 4.3 – Convergence des itérés de (IGAMD) et de l’Algorithme 13 pour différentes valeurs de  $\beta$  sur un problème de moindres carrés ( $N = 1000$ ).

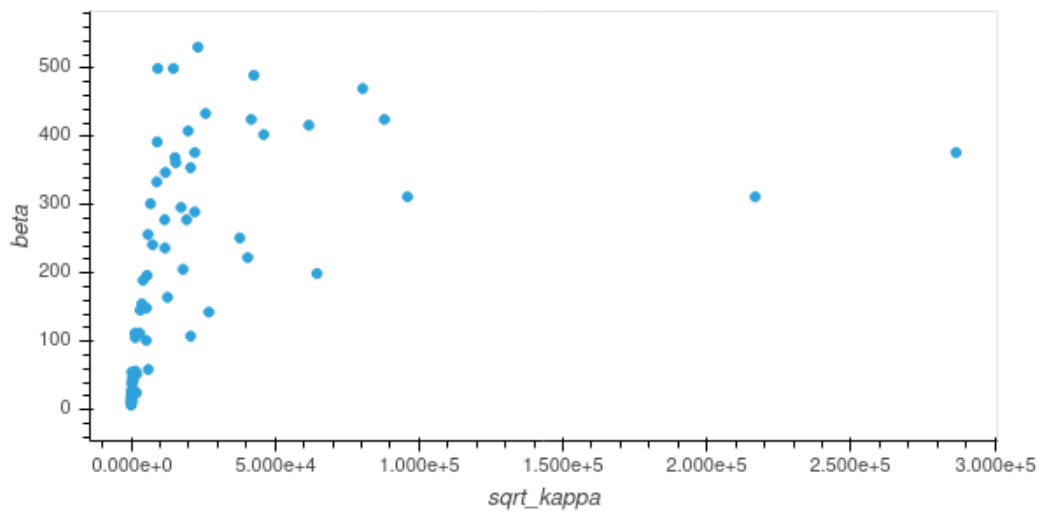


FIGURE 4.4 – Évolution de l’estimation de  $\xi$  en fonction de  $\sqrt{\frac{L}{\mu}}$  pour  $\alpha = 3$ .

De manière surprenante, la Figure 4.5 semble indiquer que  $\xi$  est proportionnel à la dimension  $N$ . On peut également observer que  $\xi$  et  $\alpha$  semblent inversement proportionnels ce qui est renforcé par la Figure 4.6.

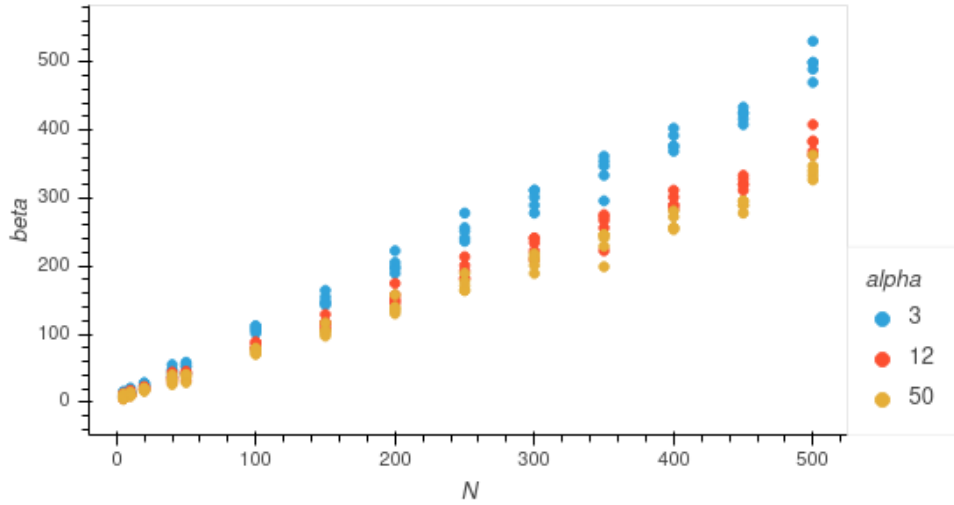


FIGURE 4.5 – Évolution de l'estimation de  $\xi$  en fonction de  $N$  pour  $\alpha = 3$ ,  $\alpha = 12$  et  $\alpha = 50$ .

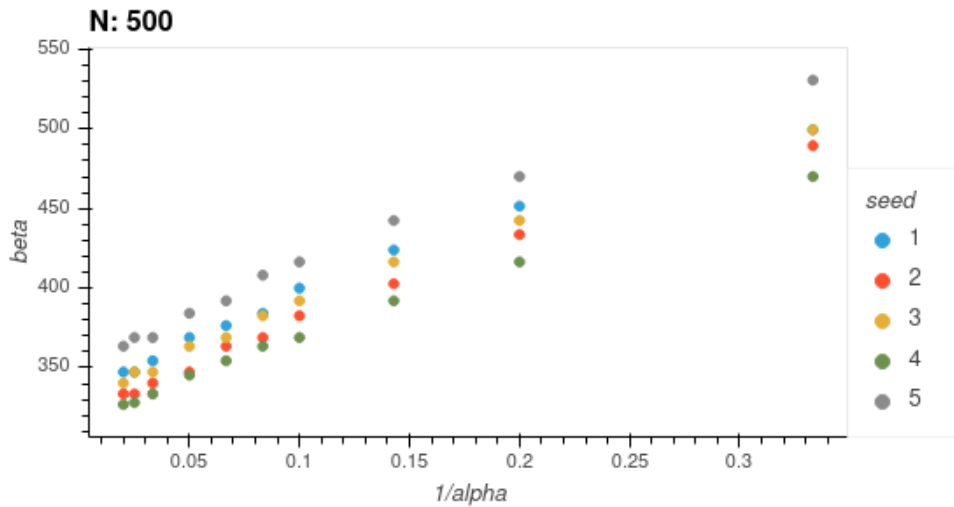


FIGURE 4.6 – Évolution de l'estimation de  $\xi$  en fonction de  $1/\alpha$  pour cinq exemples avec  $N = 500$ .

Ces observations pourraient motiver une analyse théorique de ce phénomène et des tests numériques sur d'autres types de problème. Je n'ai cependant pas poursuivi mes travaux dans cette direction pour la raison suivante : dans [26], Aujol, Dossal et Rondepierre démontrent que FISTA (dans sa forme générale avec  $\alpha > 0$ ) peut être adapté à la géométrie de  $F$  pour atteindre des vitesses comparables à celle d'une exponentielle rapide. Cela consiste à calibrer  $\alpha$  par rapport à la précision voulue ou par rapport à la géométrie de  $F$  ce qui donne généralement des valeurs élevées pour le paramètre. Mes expériences montrent que l'ajout du terme lié à la hessienne a

de moins en moins d'impact quand  $\alpha$  augmente. Ainsi, il semblerait que pour les valeurs les plus pertinentes de  $\alpha$ , le gain apporté par les termes supplémentaires dans (IGAHD) soit négligeable. De plus, mes expériences suggèrent que les performances de (IGAHD) sont meilleures pour un  $\alpha$  bien choisi et  $\beta = 0$  que pour  $\alpha$  mal calibré et  $\beta$  proche de sa valeur critique.

Je pense toutefois qu'une étude plus approfondie de ces questions serait pertinent.

## 4.A Preuves

### 4.A.1 Preuve du Théorème 4.2.1

Soient  $\alpha \geq 1 + \frac{2}{\gamma}$  et  $\lambda = \frac{2\alpha}{\gamma+2}$ . On étudie l'énergie de Lyapunov définie pour  $t \geq t_0$  telle que :

$$\mathcal{E}(t) = (t^2 + t\beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{1}{2} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2.$$

Premièrement, on peut dériver cette énergie par rapport à  $t$  ce qui donne le lemme suivant dont la preuve est donnée en Section 4.A.6.1.

**Lemme 4.A.1.** *Pour tout  $t \geq t_0$  :*

$$\begin{aligned} \mathcal{E}'(t) &= (2 - \gamma\lambda)(t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) - \beta(\lambda - \alpha) (F(x(t)) - F^*) \\ &\quad - \lambda(t + \beta(\lambda - \alpha)) [-\gamma(F(x(t)) - F^*) + \langle \nabla F(x(t)), x(t) - x^* \rangle] \\ &\quad + \frac{\lambda(\lambda + 1 - \alpha)}{t} \langle x(t) - x^*, t(\dot{x}(t) + \beta\nabla F(x(t))) \rangle \\ &\quad + \frac{\lambda + 1 - \alpha}{t} \|t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 - \beta t(t + \beta(\lambda + \alpha)) \|\nabla F(x(t))\|^2. \end{aligned} \tag{4.A.1}$$

En appliquant le Lemme 4.A.1 et en exploitant les hypothèses vérifiées par  $F$ , on peut déduire le lemme ci-dessous. On donne une preuve en Section 4.A.6.2.

**Lemme 4.A.2.** *Soit  $t \geq \max\{t_0, \beta(\alpha - \lambda)\}$ . On a :*

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t} \mathcal{E}(t) &\leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta\nabla F(x(t)) \rangle \right) \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) - \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2, \end{aligned} \tag{4.A.2}$$

où  $K(\alpha) = \frac{2\alpha\gamma}{(\gamma+2)^2} \left( \alpha - 1 - \frac{2}{\gamma} \right) = \frac{(\gamma\lambda - 2)\lambda}{2}$ . En particulier, si  $\alpha = 1 + \frac{2}{\gamma}$ , alors

$$\mathcal{E}'(t) + \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \leq \beta(\alpha - \lambda)(F(x(t)) - F^*). \tag{4.A.3}$$

**Cas  $\alpha > 1 + \frac{2}{\gamma}$ .** Pour cette hypothèse sur  $\alpha$ , on peut réécrire le résultat du Lemme 4.A.2 comme suit. La preuve se trouve en Section 4.A.6.3.

**Lemme 4.A.3.** Soit  $t > \max\{t_0, \beta(\alpha - \lambda)\}$ . Si  $\alpha > 1 + \frac{2}{\gamma}$ , alors on a

$$\mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t + \beta(\lambda - \alpha)}\mathcal{E}(t) + \beta t(t + \beta(\lambda - \alpha))\|\nabla F(x(t))\|^2 \leq \tilde{\varphi}(t + \beta(\lambda - \alpha))\mathcal{E}(t), \quad (4.A.4)$$

où

$$\tilde{\varphi} : t \mapsto \frac{K(\alpha)}{\mu t^2} \left( \sqrt{\mu}(1 + C_0) + \frac{2\alpha}{(\gamma + 2)t}(1 + \sqrt{2}) + \frac{4\alpha^2}{(\gamma + 2)^2\sqrt{\mu}t^2} \right),$$

$$\text{et } C_0 = \beta \frac{\sqrt{\mu}\gamma(\gamma\lambda - 1)}{\gamma\lambda - 2}.$$

On définit  $\mathcal{H}$  pour tout  $t > \max\{t_0, \beta(\alpha - \lambda)\}$  de la manière suivante

$$\mathcal{H} : t \mapsto \mathcal{E}(t)(t + \beta(\lambda - \alpha))^{\gamma\lambda - 2} e^{-\tilde{\Phi}(t + \beta(\lambda - \alpha))},$$

où  $\tilde{\Phi}(t) = -\int_t^{+\infty} \tilde{\varphi}(x)dx$ . Le résultat du Lemme 4.A.3 implique que  $\mathcal{H}'(t) \leq 0$  pour tout  $t > \max\{t_0, \beta(\alpha - \lambda)\}$ . Ainsi, pour tout  $t_1 > \max\{t_0, \beta(\alpha - \lambda)\}$  et  $t \geq t_1$ ,  $\mathcal{H}(t) \leq \mathcal{H}(t_1)$ , et donc,

$$\mathcal{E}(t) \leq \mathcal{E}(t_1) \left( \frac{t_1 + \beta(\lambda - \alpha)}{t + \beta(\lambda - \alpha)} \right)^{\lambda\gamma - 2} e^{-\tilde{\Phi}(t_1 + \beta(\lambda - \alpha)) + \tilde{\Phi}(t + \beta(\lambda - \alpha))}. \quad (4.A.5)$$

Si on choisit  $t_1 = t_0 + \beta(\alpha - \lambda)$ , cette inégalité assure que pour tout  $t \geq t_0 + \beta(\alpha - \lambda)$ ,

$$\mathcal{E}(t) \leq \mathcal{E}(t_0 + \beta(\alpha - \lambda)) \left( \frac{t_0}{t + \beta(\lambda - \alpha)} \right)^{\lambda\gamma - 2} e^{-\tilde{\Phi}(t_0) + \tilde{\Phi}(t + \beta(\lambda - \alpha))}. \quad (4.A.6)$$

On peut observer que la primitive  $\tilde{\Phi}(t) = -\int_t^{+\infty} \tilde{\varphi}(x)dx$  de  $\tilde{\varphi}$  a l'expression suivante

$$\tilde{\Phi}(t) = -\frac{K(\alpha)}{\mu} \left( \frac{\sqrt{\mu}(1 + C_0)}{t} + \frac{\alpha(1 + \sqrt{2})}{(\gamma + 2)t^2} + \frac{4\alpha^2}{3(\gamma + 2)^2\sqrt{\mu}t^3} \right), \quad (4.A.7)$$

ce qui assure que  $\tilde{\Phi}$  est négative. Ainsi, pour tout  $t \geq t_0 + \beta(\alpha - \lambda)$ ,

$$F(x(t)) - F^* \leq e^{-\tilde{\Phi}(t_0)}\mathcal{E}(t_0 + \beta(\alpha - \lambda)) \frac{t_0^{\lambda\gamma - 2}}{(t + \beta(\lambda - \alpha))^{\frac{2\alpha\gamma}{\gamma + 2}}}, \quad (4.A.8)$$

et le premier point du théorème est démontré.

On peut choisir de paramétrer  $t_1$  de manière à avoir un contrôle plus fin sur  $\mathcal{E}(t)$  dans l'équation (4.A.5). On fait ici le choix de prendre  $t_1$  comme le minimiseur de la fonction ci-dessous

$$u \mapsto (u + \beta(\lambda - \alpha))^{\lambda\gamma - 2} e^{-\Phi(u + \beta(\lambda - \alpha))}.$$

On trouve donc que  $u = t_1 + \beta(\lambda - \alpha)$  vérifie

$$\frac{\gamma\lambda - 2}{u} = \tilde{\varphi}(u). \quad (4.A.9)$$

De plus, comme  $\lambda\gamma - 2 = \frac{\gamma+2}{\alpha}K(\alpha)$ , l'égalité (4.A.9) peut être réécrite :

$$\frac{\gamma + 2}{\alpha u} = \frac{1}{\mu u^2} \left( \sqrt{\mu}(1 + C_0) + \frac{2\alpha}{(\gamma + 2)u}(1 + \sqrt{2}) + \frac{4\alpha^2}{(\gamma + 2)^2 \sqrt{\mu} u^2} \right) \quad (4.A.10)$$

En définissant  $r = (\gamma + 2)\frac{\sqrt{\mu}}{\alpha}u$ , cette égalité est équivalente à

$$r^3 - (1 + C_0)r^2 - 2(1 + \sqrt{2})r - 4 = 0. \quad (4.A.11)$$

Pour tout  $C_0 > 0$ , le polynôme  $r \mapsto r^3 - (1 + C_0)r^2 - 2(1 + \sqrt{2})r - 4$  a une unique racine positive que l'on note  $r^*$ . En choisissant  $t_1 = \frac{\alpha}{(\gamma+2)\sqrt{\mu}}r^* + \beta(\alpha - \lambda)$ , on a que si  $t_1 > \max\{t_0, \beta(\alpha - \lambda)\}$  (ce qui est vrai pour  $t_1 \geq t_0 + \beta(\alpha - \lambda)$ ), alors on peut contrôler  $\mathcal{E}$  de la manière suivante

$$\forall t \geq t_1, \mathcal{E}(t) \leq \mathcal{E}(t_1) \left( \frac{\alpha r^*}{(t + \beta(\lambda - \alpha))(\gamma + 2)\sqrt{\mu}} \right)^{\lambda\gamma - 2} e^{-\tilde{\Phi}\left(\frac{\alpha}{(\gamma+2)\sqrt{\mu}}r^*\right) + \tilde{\Phi}(t + \beta(\lambda - \alpha))} \quad (4.A.12)$$

On introduit  $E_m$  qui pour tout  $t \geq t_0$  vérifie

$$E_m(t) = \left(1 + \frac{\beta\alpha}{t}\right) (F(x(t)) - F^*) + \frac{1}{2} \|\dot{x}(t) + \beta\nabla F(x(t))\|^2. \quad (4.A.13)$$

On peut relever que cette énergie est décroissante puisque pour tout  $t \geq t_0$ ,

$$E'_m(t) = -\frac{\beta\alpha}{t^2} (F(x(t)) - F^*) - \frac{\alpha}{t} \|\dot{x}(t)\|^2 - \beta \|\nabla F(x(t))\|^2 \leq 0, \quad (4.A.14)$$

et par conséquent  $E_m$  est bornée uniformément sur  $[t_0, +\infty[$ . On a donc que

$$\begin{aligned} \mathcal{E}(t_1) &= (t_1^2 + t_1\beta(\lambda - \alpha)) (F(x(t_1)) - F^*) \\ &\quad + \frac{1}{2} \|\lambda(x(t_1) - x^*) + t_1(\dot{x}(t_1) + \beta\nabla F(x(t_1)))\|^2 \\ &= (t_1^2 + t_1\beta(\lambda - \alpha)) (F(x(t_1)) - F^*) + \frac{t_1^2}{2} \|\dot{x}(t_1) + \beta\nabla F(x(t_1))\|^2 \\ &\quad + \frac{\lambda^2}{2} \|x(t_1) - x^*\|^2 + \lambda t_1 \langle (x(t_1) - x^*), \dot{x}(t_1) + \beta\nabla F(x(t_1)) \rangle \\ &\leq (t_1^2 + t_1\beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{\lambda}{2} (\lambda + t_1\sqrt{\mu}) \|x(t_1) - x^*\|^2 \\ &\quad + \frac{t_1}{2} \left(t_1 + \frac{\lambda}{\sqrt{\mu}}\right) \|\dot{x}(t) + \beta\nabla F(x(t))\|^2, \end{aligned} \quad (4.A.15)$$

en utilisant l'équation

$$|\langle x(t) - x^*, \dot{x}(t) + \beta\nabla F(x(t)) \rangle| \leq \frac{\sqrt{\mu}}{2} \|x(t) - x^*\|^2 + \frac{1}{2\sqrt{\mu}} \|\dot{x}(t) + \beta\nabla F(x(t))\|^2. \quad (4.A.16)$$

Comme la fonction  $F$  vérifie  $\mathcal{G}_\mu^2$  et  $\frac{\lambda}{\sqrt{\mu}} = \frac{2}{r^*}(t_1 + \beta(\lambda - \alpha))$ , on obtient que

$$\begin{aligned}
\mathcal{E}(t_1) &\leq \left( t_1^2 + t_1\beta(\lambda - \alpha) + \frac{\lambda^2}{\mu} + t_1 \frac{\lambda}{\sqrt{\mu}} \right) (F(x(t_1)) - F^*) \\
&\quad + \frac{t_1}{2} \left( t_1 + \frac{\lambda}{\sqrt{\mu}} \right) \|\dot{x}(t_1) + \beta \nabla F(x(t_1))\|^2 \\
&\leq \left( 1 + \frac{2}{r^*} \right)^2 (t_1^2 + t_1\beta(\lambda - \alpha)) (F(x(t_1)) - F^*) \\
&\quad + \frac{1}{2} \left( \left( 1 + \frac{2}{r^*} \right) t_1^2 + \frac{2}{r^*} t_1(t_1 + \beta(\lambda - \alpha)) \right) \|\dot{x}(t_1) + \beta \nabla F(x(t_1))\|^2 \\
&\leq \left( 1 + \frac{2}{r^*} \right)^2 t_1^2 \left( F(x(t_1)) - F^* + \frac{1}{2} \|\dot{x}(t_1) + \beta \nabla F(x(t_1))\|^2 \right) \\
&\leq \left( 1 + \frac{2}{r^*} \right)^2 t_1^2 E_m(t_1) \leq \left( 1 + \frac{2}{r^*} \right)^2 t_1^2 E_m(t_0).
\end{aligned} \tag{4.A.17}$$

La fonction  $\tilde{\Phi}$  définie par (4.A.7) est négative pour tout  $t \geq 0$  et comme  $t_1 = \frac{\alpha}{(\gamma+2)\sqrt{\mu}} r^* + \beta(\alpha - \lambda)$ , on a :

$$\tilde{\Phi}(t_1 + \beta(\lambda - \alpha)) = -\frac{\gamma+2}{\alpha} K(\alpha) \left( \frac{1+C_0}{r^*} + \frac{1+\sqrt{2}}{r^{*2}} + \frac{4}{3r^{*3}} \right).$$

On peut donc écrire que pour tout  $t \geq t_1$  :

$$F(x(t)) - F^* \leq C_1 e^{\frac{2\gamma}{\gamma+2} C_2 (\alpha - 1 - \frac{2}{\gamma})} \left( 1 + \frac{\beta(\alpha - \lambda)(\gamma+2)\sqrt{\mu}}{\alpha r^*} \right) E_m(t_0) \left( \frac{\alpha r^*}{(\gamma+2)\sqrt{\mu}(t + \beta(\lambda - \alpha))} \right)^{\frac{2\alpha\gamma}{\gamma+2}}, \tag{4.A.18}$$

où

$$C_1 = \left( 1 + \frac{2}{r^*} \right)^2, \quad C_2 = \frac{1+C_0}{r^*} + \frac{1+\sqrt{2}}{r^{*2}} + \frac{4}{3r^{*3}}. \tag{4.A.19}$$

On définit  $\mathcal{G}$  comme suit

$$\mathcal{G} : t \mapsto \mathcal{H}(t) + \beta \int_{t_0 + \beta(\alpha - \lambda)}^t u(u + \beta(\lambda - \alpha))^{\gamma\lambda - 1} e^{-\tilde{\Phi}(u + \beta(\lambda - \alpha))} \|\nabla F(x(u))\|^2 du, \tag{4.A.20}$$

Le Lemme 4.A.3 assure que  $\mathcal{G}'(t) \leq 0$  pour tout  $t \geq t_0 + \beta(\alpha - \lambda)$ . Par conséquent, pour tout  $t \geq t_0 + \beta(\alpha - \lambda)$ ,

$$\mathcal{G}(t) \leq \mathcal{G}(t_0 + \beta(\alpha - \lambda)),$$

et comme  $\mathcal{H}$  est une fonction positive :

$$\beta \int_{t_0 + \beta(\alpha - \lambda)}^t u(u + \beta(\lambda - \alpha))^{\gamma\lambda - 1} e^{-\tilde{\Phi}(u + \beta(\lambda - \alpha))} \|\nabla F(x(u))\|^2 du \leq \mathcal{G}(t_0 + \beta(\alpha - \lambda)).$$

De plus,  $\tilde{\Phi}$  est une fonction négative donc :

$$\beta \int_{t_0 + \beta(\alpha - \lambda)}^t u(u + \beta(\lambda - \alpha))^{\gamma\lambda - 1} \|\nabla F(x(u))\|^2 du \leq \mathcal{G}(t_0 + \beta(\alpha - \lambda)).$$

On peut en déduire que :

$$\int_{t_0+\beta(\alpha-\lambda)}^{+\infty} u(u+\beta(\lambda-\alpha))^{\gamma\lambda-1} \|\nabla F(x(u))\|^2 du < +\infty.$$

Comme  $u \mapsto (1 + \beta \frac{\lambda-\alpha}{u})^{\lambda\gamma-1}$  est une fonction décroissante sur  $(t_0 + \beta(\alpha - \lambda), +\infty)$ , on obtient que :

$$\begin{aligned} \int_{t_0+\beta(\alpha-\lambda)}^{+\infty} u(u+\beta(\lambda-\alpha))^{\gamma\lambda-1} \|\nabla F(x(u))\|^2 du \geq \\ \left(1 + \beta \frac{\lambda-\alpha}{t_0+\beta(\alpha-\lambda)}\right)^{\lambda\gamma-1} \int_{t_0+\beta(\alpha-\lambda)}^{+\infty} u^{\frac{2\alpha\gamma}{\gamma+2}} \|\nabla F(x(u))\|^2 du. \end{aligned} \quad (4.A.21)$$

D'autre part, la fonction  $u \mapsto u^{\frac{2\alpha\gamma}{\gamma+2}} \|\nabla F(x(u))\|^2$  est bornée sur  $(t_0, t_0 + \beta(\alpha - \lambda))$  et on a donc :

$$\int_{t_0}^{+\infty} u^{\frac{2\alpha\gamma}{\gamma+2}} \|\nabla F(x(u))\|^2 du < +\infty. \quad (4.A.22)$$

**Cas**  $\alpha = 1 + \frac{2}{\gamma}$ .

Le Lemme 4.A.2 indique que pour tout  $t > \max\{t_0, \beta\}$ ,

$$\mathcal{E}'(t) + \beta t(t-\beta) \|\nabla F(x(t))\|^2 \leq \frac{\beta}{t(t-\beta)} \mathcal{E}(t), \quad (4.A.23)$$

en utilisant notamment la relation  $\alpha - \lambda = 1$ . Cette inégalité implique que  $t \mapsto \mathcal{E}(t)e^{\frac{\beta}{t-\beta}}$  est une fonction décroissante sur  $(t_0 + \beta, +\infty)$ . Ainsi, pour tout  $t \geq t_0 + \beta$ ,

$$\mathcal{E}(t) \leq \mathcal{E}(t_0 + \beta) e^{-\frac{\beta}{t-\beta} + \frac{\beta}{t_0}} \leq \mathcal{E}(t_0 + \beta) e^{\frac{\beta}{t_0}}.$$

De plus,

$$\begin{aligned} \mathcal{E}(t_0 + \beta) &= t_0(t_0 + \beta) (F(x(t_0 + \beta)) - F^*) \\ &+ \frac{1}{2} \|\lambda(x(t_0 + \beta) - x^*) + t(\dot{x}(t_0 + \beta) + \beta \nabla F(x(t_0 + \beta)))\|^2 \\ &\leq \left( (t_0 + \beta)^2 + \frac{\lambda^2 + \sqrt{\mu}}{\mu} \right) (F(x(t_0 + \beta)) - F^*) \\ &+ \frac{(t_0 + \beta)^2 + \frac{1}{\sqrt{\mu}}}{2} \|\dot{x}(t_0 + \beta) + \beta \nabla F(x(t_0 + \beta))\|^2 \\ &\leq \left( (t_0 + \beta)^2 + \frac{\lambda^2 + \sqrt{\mu}}{\mu} \right) E_m(t_0 + \beta) \leq \left( (t_0 + \beta)^2 + \frac{\lambda^2 + \sqrt{\mu}}{\mu} \right) E_m(t_0), \end{aligned}$$

en utilisant l'inégalité (4.A.16). En conséquence, pour tout  $t \geq t_0 + \beta$ ,

$$F(x(t)) - F^* \leq \left( (t_0 + \beta)^2 + \frac{\lambda^2 + \sqrt{\mu}}{\mu} \right) e^{\frac{\beta}{t_0}} \frac{E_m(t_0)}{t(t-\beta)}. \quad (4.A.24)$$

L'inégalité (4.A.23) assure également que

$$t \mapsto \mathcal{E}(t)e^{\frac{\beta}{t-\beta}} + \int_{t_0+\beta}^t \beta u(u-\beta)e^{\frac{\beta}{u-\beta}} \|\nabla F(x(u))\|^2 du, \quad (4.A.25)$$

est une fonction bornée sur  $(t_0 + \beta, +\infty)$ . Étant donné que  $\mathcal{E}(t)e^{\frac{\beta}{t-\beta}}$  est une fonction positive pour tout  $t \geq t_0 + \beta$ , on en déduit qu'il existe  $M > 0$  tel que pour tout  $t \geq t_0 + \beta$ ,

$$\int_{t_0+\beta}^t (u-\beta)^2 \|\nabla F(x(u))\|^2 du \leq \int_{t_0+\beta}^t u(u-\beta)e^{\frac{\beta}{u-\beta}} \|\nabla F(x(u))\|^2 du < M,$$

et ainsi,

$$\int_{t_0+\beta}^{+\infty} (u-\beta)^2 \|\nabla F(x(u))\|^2 du < +\infty. \quad (4.A.26)$$

En utilisant les mêmes arguments que dans le cas  $\alpha > 1 + \frac{2}{\gamma}$ , on trouve que :

$$\int_{t_0}^{+\infty} u^2 \|\nabla F(x(u))\|^2 du < +\infty. \quad (4.A.27)$$

□

#### 4.A.2 Preuve du Théorème 4.2.2

Soit  $F$  une fonction convexe  $C^2$  vérifiant  $\mathcal{G}_\mu^2$  pour  $\mu > 0$ . De par sa convexité,  $F$  vérifie automatiquement  $\mathcal{F}_1$  et le lemme ci-dessous indique que  $F$  satisfait également  $\mathcal{F}_{2-\delta}$  localement pour tout  $\delta \in (0, 1]$ . On donne une preuve en Section 4.A.6.4.

**Lemme 4.A.4.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  ayant un ensemble de minimiseurs  $X^*$  non vide. Alors, pour tout  $\delta \in (0, 1]$ , la fonction  $F$  vérifie  $\mathcal{F}_{2-\delta}$  localement, c'est-à-dire que pour tout  $x^* \in X^*$ , il existe  $\nu > 0$  tel que*

$$\forall x \in B(x^*, \nu), \quad F(x) - F^* \leq \frac{1}{2-\delta} \langle \nabla F(x), x - x^* \rangle. \quad (4.A.28)$$

Soient  $\alpha \geq 3$ ,  $\beta > 0$  et  $\varepsilon \in (0, 1)$ . Comme  $F$  vérifie  $\mathcal{F}_1$ , les deux premiers points du Théorème 4.2.1 indiquent qu'il existe une fonction décroissante  $\phi$  telle que :

$$\forall t \geq t_0 + \frac{\beta\alpha}{3}, \quad F(x(t)) - F^* \leq \phi(t),$$

où  $\phi(t) \rightarrow 0$  quand  $t \rightarrow +\infty$ . Comme  $F$  satisfait  $\mathcal{G}_\mu^2$ , pour tout  $\nu > 0$ , il existe  $T \geq t_0 + \frac{\beta\alpha}{3}$  tel que  $t \geq T$ ,  $x(t) \in B(x^*, \nu)$ . En conséquence, pour tout  $\delta \in (0, 1]$ , il existe  $T \geq t_0 + \frac{\beta\alpha}{3}$  tel que pour tout  $t \geq T$ ,

$$F(x(t)) - F^* \leq \frac{1}{2-\delta} \langle x(t) - x^*, \nabla F(x(t)) \rangle.$$

Soit  $\delta = \frac{4\varepsilon}{\alpha+\varepsilon}$ . Comme  $\alpha \geq 3$  et  $\varepsilon \in (0, 1)$ , la condition  $\alpha \geq 1 + \frac{2}{2-\delta}$  est vérifiée. En prenant  $T$  comme temps initial de (DIN-AVD), le premier point du Théorème 4.2.1 donne le premier résultat voulu et le troisième point garantit qu'il existe  $M > 0$  tel que

$$\int_T^{+\infty} u^{\alpha-\varepsilon} \|\nabla F(x(u))\|^2 du < M.$$



Comme la fonction  $F$  vérifie  $\mathcal{G}_\mu^2$ , la Proposition 2.1.16 indique que  $F$  satisfait l'inégalité de Łojasiewicz d'exposant  $\frac{1}{2}$  ce qui implique que

$$\int_T^{+\infty} u^{\alpha-\varepsilon} (F(x(u)) - F^*) du < \frac{2M}{\mu}.$$

D'autre part, la fonction  $u \mapsto u^{\alpha-\varepsilon}(F(x(u)) - F^*)$  est bornée sur  $(t_0, T)$  et on a donc que

$$\int_{t_0}^T u^{\alpha-\varepsilon} (F(x(u)) - F^*) du < +\infty,$$

et ainsi,

$$\int_{t_0}^{+\infty} u^{\alpha-\varepsilon} (F(x(u)) - F^*) du < +\infty.$$

□

### 4.A.3 Preuve du Corollaire 4.2.3

Le premier résultat est obtenu en appliquant le lemme ci-dessous au Théorème 4.2.2. Une preuve du lemme est donnée en Section 4.A.6.5.

**Lemme 4.A.5.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe ayant un ensemble de minimiseurs non vide. Supposons que pour un certain  $t_1 > 0$  et un certain  $\delta > 0$ ,  $F$  vérifie :*

$$\int_{t_1}^{+\infty} u^\delta (F(x(u)) - F^*) du < +\infty.$$

*Soit  $z : t \mapsto \frac{\int_{t/2}^t u^\delta x(u) du}{\int_{t/2}^t u^\delta du}$ . Alors, quand  $t \rightarrow +\infty$ , on a*

$$F(z(t)) - F^* = o\left(t^{-\delta-1}\right). \tag{4.A.29}$$

Les deux autres résultats viennent de l'application du Lemme 4.A.6 à  $\phi : x \mapsto F(x) - F^*$ . On donne une preuve de ce lemme en Section 4.A.6.6.

**Lemme 4.A.6.** *Soit  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}^+$  une fonction telle que pour un certain  $t_1 > 0$  et un certain  $\delta > 0$ , on a :*

$$\int_{t_1}^{+\infty} u^\delta \phi(x(u)) du < +\infty.$$

*Alors, quand  $t \rightarrow +\infty$ ,*

$$\inf_{u \in [t/2, t]} \phi(x(u)) = o\left(t^{-\delta-1}\right) \quad \text{et} \quad \liminf_{t \rightarrow +\infty} t^{\delta+1} \log(t) \phi(x(t)) = 0. \tag{4.A.30}$$

## 4.A.4 Preuve du Théorème 4.2.4

On définit l'énergie de Lyapunov  $\mathcal{E}$  de la manière suivante :

$$\begin{aligned} \mathcal{E}(t) &= (t^2 + t\beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{\xi}{2} \|x(t) - x^*\|^2 \\ &\quad + \frac{1}{2} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2, \end{aligned}$$

où  $x^*$  est l'unique minimiseur de  $F$ ,  $\lambda \in \mathbb{R}$  et  $\xi \in \mathbb{R}$ . On définit également  $\mathcal{H}$  ci-dessous

$$\mathcal{H}(t) = t^p \mathcal{E}(t),$$

où  $p > 0$ . En utilisant les notations ci-dessous

$$\begin{aligned} a(t) &= t(F(x(t)) - F^*), \quad b(t) = \frac{1}{2t} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2, \\ c(t) &= \frac{1}{2t} \|x(t) - x^*\|^2, \end{aligned}$$

on obtient que

$$\mathcal{E}(t) = (t + \beta(\lambda - \alpha))a(t) + t(b(t) + \xi c(t)).$$

On choisit  $p = \frac{4}{\gamma_1 - 2}$ ,  $\lambda = \frac{2}{\gamma_1 - 2}$  et  $\xi = \lambda(\lambda + 1 - \alpha)$ . L'énergie  $\mathcal{E}$  est dérivable et le lemme ci-dessous donne un contrôle sur sa dérivée  $\mathcal{E}'$ . La preuve se trouve en Section 4.A.6.9.

**Lemme 4.A.7.** *Soient  $p = \frac{4}{\gamma_1 - 2}$ ,  $\lambda = \frac{2}{\gamma_1 - 2}$  et  $\xi = \lambda(\lambda + 1 - \alpha)$ . Pour tout  $t \geq \max(t_0, \beta(\alpha - \lambda), \beta(2(\alpha - \lambda) - 1))$ , on a*

$$\begin{aligned} \mathcal{E}'(t) &\leq ((2 - \lambda\gamma_1)t + \beta(\lambda - \alpha - \lambda\gamma_1(2(\lambda - \alpha) + 1))) (F(x(t)) - F^*) \\ &\quad + 2(\lambda + 1 - \alpha)b(t) - 2\lambda^2(\lambda + 1 - \alpha)c(t) - \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2. \end{aligned} \tag{4.A.31}$$

En appliquant ce lemme, on a que pour tout  $t \geq \max(t_0, \beta(\alpha - \lambda), \beta(2(\alpha - \lambda) - 1))$  :

$$\begin{aligned} \mathcal{H}'(t) &= t^{p-1} (p\mathcal{E}(t) + t\mathcal{E}'(t)) \leq t^{p-1} [p\mathcal{E}(t) + 2t(\lambda + 1 - \alpha)b(t) \\ &\quad + t((2 - \lambda\gamma_1)t + \beta(\lambda - \alpha - \lambda\gamma_1(2(\lambda - \alpha) + 1))) (F(x(t)) - F^*) \\ &\quad - 2t(\lambda^2(\lambda + 1 - \alpha))c(t) - \beta t^2(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2] \\ &\leq t^p ((2 - \gamma_1\lambda + p)a(t) + 2(\lambda - \alpha + 1) + p)b(t) + \lambda(\lambda + 1 - \alpha)(p - 2\lambda)c(t) \\ &\quad + t^{p-1} \beta((p + 1)(\lambda - \alpha) - \lambda\gamma_1(2(\lambda - \alpha) + 1))a(t) \\ &\quad - \beta t^{p+1}(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2. \end{aligned}$$

Comme  $p = \frac{4}{\gamma_1 - 2}$  et  $\lambda = \frac{2}{\gamma_1 - 2}$  cela garantit que

$$\mathcal{H}'(t) \leq 2t^p \left( \frac{\gamma_1 + 2}{\gamma_1 - 2} - \alpha \right) b(t) + t^{p-1} \beta C_1 a(t) - \beta t^{p+1}(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2, \tag{4.A.32}$$

où  $C_1 = (p+1)(\lambda - \alpha) - \lambda\gamma_1(2(\lambda - \alpha) + 1)$ . Sous l'hypothèse que  $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$ , la constante  $C_1$  est strictement positive et (4.A.32) assure que

$$\mathcal{H}'(t) \leq t^{p-1}\beta C_1 a(t) - \beta t^{p+1}(t + \beta(\lambda - \alpha))\|\nabla F(x(t))\|^2. \quad (4.A.33)$$

On définit la fonction  $\mathcal{G}$  ci-dessous

$$\mathcal{G} : t \mapsto \mathcal{H}(t) - \beta C_1 \int_{t_1}^t u^{p-1} a(u) du,$$

où  $t_1 > \max\{t_0, \beta(2(\alpha - \lambda) - 1), t_m\}$  et  $t_m > \beta(\alpha - \lambda)$  satisfait

$$\frac{t_m^p}{(t_m + \beta(\lambda - \alpha))^{p+1}} \beta C_1 \leq \frac{1}{2}. \quad (4.A.34)$$

Comme  $t \mapsto \frac{t^p}{(t + \beta(\lambda - \alpha))^{p+1}}$  est une fonction décroissante sur  $(\beta(\alpha - \lambda), +\infty)$  et tend vers 0,  $t_m$  est bien défini. L'équation (4.A.33) implique que  $\mathcal{G}'(t) \leq 0$  pour tout  $t \geq t_1$  et il existe donc une constante  $A \in \mathbb{R}$  telle que  $\mathcal{G}(t) \leq A$  pour tout  $t \geq t_1$ , et pour tout  $t \geq t_1$ ,

$$\begin{aligned} \mathcal{G}(t) &= t^p((t + \beta(\lambda - \alpha))a(t) + tb(t) + t\xi c(t)) - \beta C_1 \int_{t_1}^t u^{p-1} a(u) du \\ &\geq t^p((t + \beta(\lambda - \alpha))a(t) + t\xi c(t)) - \beta C_1 \int_{t_1}^t u^{p-1} a(u) du \\ &\geq (t + \beta(\lambda - \alpha))^{p+1} a(t) + t^{p+1} \xi c(t) - \beta C_1 \int_{t_1}^t u^{p-1} a(u) du. \end{aligned}$$

De plus,  $-\beta C_1 \int_{t_1}^t u^{p-1} a(u) du \geq -\left(\frac{t_1}{t_1 + \beta(\lambda - \alpha)}\right)^{p+1} \beta C_1 \int_{t_1}^t \frac{(u + \beta(\lambda - \alpha))^{p+1} a(u)}{u^2} du$ .

Rappelons que comme  $F$  vérifie  $\mathcal{G}^{\gamma_2}$ , il existe  $K > 0$  tel que

$$\forall x \in \mathbb{R}^N, \quad Kd(x, X^*)^{\gamma_2} \leq F(x) - F^*,$$

et par conséquent,

$$t^{p+1} \xi c(t) = t^p \frac{\xi}{2} \|x(t) - x^*\|^2 = t^p \frac{\xi}{2K^{\frac{2}{\gamma_2}}} (K\|x(t) - x^*\|^{\gamma_2})^{\frac{2}{\gamma_2}}.$$

Comme  $\xi < 0$  et  $F$  a un unique minimiseur  $x^*$ ,

$$\begin{aligned} t^{p+1} \xi c(t) &= t^p \frac{\xi}{2K^{\frac{2}{\gamma_2}}} (Kd(x(t), X^*)^{\gamma_2})^{\frac{2}{\gamma_2}} \geq t^p \frac{\xi}{2K^{\frac{2}{\gamma_2}}} (F(x(t)) - F^*)^{\frac{2}{\gamma_2}} \\ &\geq t^{p-\frac{2}{\gamma_2}} \frac{\xi}{2K^{\frac{2}{\gamma_2}}} a(t)^{\frac{2}{\gamma_2}} \geq t^{p-\frac{2}{\gamma_2}-(p+1)\frac{2}{\gamma_2}} \frac{\xi}{2K^{\frac{2}{\gamma_2}}} (t^{p+1} a(t))^{\frac{2}{\gamma_2}} \\ &\geq t^{p-\frac{2}{\gamma_2}-(p+1)\frac{2}{\gamma_2}} \left(\frac{t_1}{t_1 + \beta(\lambda - \alpha)}\right)^{\frac{2(p+1)}{\gamma_2}} \frac{\xi}{2K^{\frac{2}{\gamma_2}}} ((t + \beta(\alpha - \lambda))^{p+1} a(t))^{\frac{2}{\gamma_2}}. \end{aligned}$$

Rappelons que  $p = \frac{4}{\gamma_1 - 2}$ , et donc  $p - \frac{2}{\gamma_2} - (p+1)\frac{2}{\gamma_2} = \frac{4(\gamma_2 - \gamma_1)}{\gamma_2(\gamma_1 - 2)} \leq 0$  puisque  $\gamma_1 \geq \gamma_2$ . Comme  $t \geq t_1$ , on trouve que

$$t^{p+1}\xi c(t) \geq \frac{t_1^{p-\frac{2}{\gamma_2}}}{(t_1 + \beta(\lambda - \alpha))^{\frac{2(p+1)}{\gamma_2}}} \frac{\xi}{2K^{\frac{2}{\gamma_2}}} ((t + \beta(\alpha - \lambda))^{p+1} a(t))^{\frac{2}{\gamma_2}}.$$

On définit  $v : t \mapsto (t + \beta(\lambda - \alpha))^{p+1} a(t)$  sur  $[t_1, +\infty)$ . Pour tout  $t \geq t_1$ , on a

$$\mathcal{G}(t) \geq v(t) - C_2 v(t)^{\frac{2}{\gamma_2}} - \left( \frac{t_1}{t_1 + \beta(\lambda - \alpha)} \right)^{p+1} \beta C_1 \int_{t_1}^t \frac{v(u)}{u^2} du, \quad (4.A.35)$$

où  $C_2 = -\frac{t_1^{p-\frac{2}{\gamma_2}}}{(t_1 + \beta(\lambda - \alpha))^{\frac{2(p+1)}{\gamma_2}}} \frac{\xi}{2K^{\frac{2}{\gamma_2}}} > 0$ .

Prenons  $t_2 > t_1$  et  $t^* = \operatorname{argmax}_{t \in [t_1, t_2]} v(t)$ . On obtient alors,

$$\begin{aligned} \mathcal{G}(t^*) &\geq v(t^*) - C_2 v(t^*)^{\frac{2}{\gamma_2}} - \left( \frac{t_1}{t_1 + \beta(\lambda - \alpha)} \right)^{p+1} \beta C_1 \int_{t_1}^{t^*} \frac{v(u)}{u^2} du \\ &\geq v(t^*) - C_2 v(t^*)^{\frac{2}{\gamma_2}} - \left( \frac{t_m}{t_m + \beta(\lambda - \alpha)} \right)^{p+1} \beta C_1 \int_{t_m}^{+\infty} \frac{v(t^*)}{u^2} du \\ &\geq v(t^*) - C_2 v(t^*)^{\frac{2}{\gamma_2}} - \frac{t_m^p}{(t_m + \beta(\lambda - \alpha))^{p+1}} \beta C_1 v(t^*) \\ &\geq \frac{1}{2} v(t^*) - C_2 v(t^*)^{\frac{2}{\gamma_2}}. \end{aligned}$$

Comme  $\mathcal{G}(t) \leq A$  pour tout  $t \geq t_1$ , on l'inégalité suivante t :

$$v(t^*) - 2C_2 v(t^*)^{\frac{2}{\gamma_2}} \leq 2A,$$

et donc

$$v(t^*)^{\frac{2}{\gamma_2}} \left( v(t^*)^{1-\frac{2}{\gamma_2}} - 2C_2 \right) \leq 2A. \quad (4.A.36)$$

**Lemme 4.A.8.** Soient  $x \in \mathbb{R}^+$ ,  $\delta \in (0, 1)$ ,  $K_1 > 0$  et  $K_2 > 0$ . On a

$$x^\delta (x^{1-\delta} - K_1) \leq K_2 \quad \implies \quad x \leq \left( K_2^{1-\delta} + K_1 \right)^{\frac{1}{1-\delta}}.$$

En appliquant le Lemme 4.A.8 à (4.A.36), on trouve que

$$v(t^*) \leq \left( (2A)^{1-\frac{2}{\gamma_2}} + 2C_2 \right)^{\frac{\gamma_2}{\gamma_2-2}}, \quad (4.A.37)$$

et en conséquence, pour tout  $t \in [t_1, t_2]$

$$v(t) \leq \left( (2A)^{1-\frac{2}{\gamma_2}} + 2C_2 \right)^{\frac{\gamma_2}{\gamma_2-2}}. \quad (4.A.38)$$

Cette borne ne dépend pas de  $t_2$  donc on peut en déduire que la fonction  $v$  est bornée sur  $[t_1, +\infty)$ . Cela signifie qu'il existe une constante  $M > 0$  telle que pour tout  $t \geq t_1$  :

$$v(t) = (t + \beta(\lambda - \alpha))^{p+1} a(t) = t(t + \beta(\lambda - \alpha))^{p+1} (F(x(t)) - F^*) \leq M,$$

ce qui implique que

$$F(x(t)) - F^* \leq \frac{M}{t(t + \beta(\lambda - \alpha))^{p+1}} \leq \left( \frac{t_1}{t_1 + \beta(\lambda - \alpha)} \right)^{p+1} \frac{M}{t^{p+2}}, \quad (4.A.39)$$

c'est-à-dire, quand  $t \rightarrow +\infty$

$$F(x(t)) - F^* = \mathcal{O} \left( t^{-\frac{2\gamma_1}{\gamma_1-2}} \right). \quad (4.A.40)$$

On définit maintenant  $\mathcal{F}$  pour démontrer le deuxième point :

$$\mathcal{F} : t \mapsto \mathcal{H}(t) - \beta C_1 \int_{t_1}^t u^{p-1} a(u) du + \beta \int_{t_1}^t u^{p+1} (u + \beta(\lambda - \alpha)) \|\nabla F(x(u))\|^2 du.$$

L'équation (4.A.33) donne que  $\mathcal{F}'(t) \leq 0$  pour tout  $t \geq t_1$  donc il existe une constante  $B \in \mathbb{R}$  telle que  $\mathcal{F}(t) \leq B$  pour tout  $t \geq t_1$ . En appliquant (4.A.35), on peut montrer que pour tout  $t \geq t_1$

$$\begin{aligned} \mathcal{F}(t) &\geq v(t) - C_2 v(t)^{\frac{2}{\gamma_2}} - \left( \frac{t_1}{t_1 + \beta(\lambda - \alpha)} \right)^{p+1} \beta C_1 \int_{t_1}^t \frac{v(u)}{u^2} du \\ &\quad + \beta \int_{t_1}^t u^{p+1} (u + \beta(\lambda - \alpha)) \|\nabla F(x(u))\|^2 du. \end{aligned}$$

Nous avons montré qu'il existe une constante  $M > 0$  telle que pour tout  $t \geq t_1$ ,  $v(t) \leq M$ . En conséquence,

$$- \left( \frac{t_1}{t_1 + \beta(\lambda - \alpha)} \right)^{p+1} \beta C_1 \int_{t_1}^t \frac{v(u)}{u^2} du \geq -M \beta C_1 \frac{t_1^p}{(t_1 + \beta(\lambda - \alpha))^{p+1}}.$$

**Lemme 4.A.9.** Soit  $g : x \mapsto x - Kx^\delta$  avec  $K > 0$  et  $\delta \in (0, 1)$ . Pour tout  $x \geq 0$ , on a

$$g(x) \geq K(\delta - 1)(\delta K)^{\frac{\delta}{1-\delta}}.$$

Le Lemme 4.A.9 garantit que pour tout  $t \geq t_1$

$$v(t) - C_2 v(t)^{\frac{2}{\gamma_2}} \geq -C_2 \left( 1 - \frac{2}{\gamma_2} \right) \left( \frac{2C_2}{\gamma_2} \right)^{\frac{2}{\gamma_2-2}}. \quad (4.A.41)$$

Par conséquent,

$$\begin{aligned} \mathcal{F}(t) &\geq -C_2 \left( 1 - \frac{2}{\gamma_2} \right) \left( \frac{2C_2}{\gamma_2} \right)^{\frac{2}{\gamma_2-2}} - M \beta C_1 \frac{t_1^p}{(t_1 + \beta(\lambda - \alpha))^{p+1}} \\ &\quad + \beta \int_{t_1}^t u^{p+1} (u + \beta(\lambda - \alpha)) \|\nabla F(x(u))\|^2 du. \end{aligned}$$

Comme il existe une constante  $B \in \mathbb{R}$  telle que  $\mathcal{F}(t) \leq B$  pour tout  $t \geq t_1$ , on peut écrire que

$$\begin{aligned} \beta \int_{t_1}^t u^{p+1} (u + \beta(\lambda - \alpha)) \|\nabla F(x(u))\|^2 du &\leq B + C_2 \left(1 - \frac{2}{\gamma_2}\right) \left(\frac{2C_2}{\gamma_2}\right)^{\frac{2}{\gamma_2-2}} \\ &\quad + M\beta C_1 \frac{t_1^p}{(t_1 + \beta(\lambda - \alpha))^{p+1}}, \end{aligned}$$

et donc

$$\int_{t_1}^{+\infty} (u + \beta(\lambda - \alpha))^{\frac{2\gamma_1}{\gamma_1-2}} \|\nabla F(x(u))\|^2 du < +\infty. \quad (4.A.42)$$

En utilisant les mêmes arguments que dans la preuve du Théorème 4.2.1 (en Section 4.A.1) et le fait que la fonction  $u \mapsto (u + \beta(\lambda - \alpha))^{\frac{2\gamma_1}{\gamma_1-2}} \|\nabla F(x(u))\|^2$  est bornée sur  $(t_0, t_1)$ , on peut conclure que

$$\int_{t_0}^{+\infty} u^{\frac{2\gamma_1}{\gamma_1-2}} \|\nabla F(x(u))\|^2 du < +\infty. \quad (4.A.43)$$

□

#### 4.A.5 Preuve du Corollaire 4.2.5

Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  ayant un unique minimiseur  $x^*$ . Supposons que  $F$  vérifie  $\mathcal{F}_{\gamma_1}$  et  $\mathcal{G}^{\gamma_2}$  avec  $\gamma_1 > 2$  et  $\gamma_2 > 2$  où  $\gamma_1 \geq \gamma_2$ . Soit  $x$  une solution de (DIN-AVD) pour tout  $t \geq t_0$  où  $t_0 > 0$ ,  $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$  et  $\beta > 0$ . Le Théorème 4.2.4 assure que

$$\int_{t_0}^{+\infty} u^{\frac{2\gamma_1}{\gamma_1-2}} \|\nabla F(x(u))\|^2 du < +\infty.$$

De plus, comme  $F$  satisfait l'hypothèse  $\mathcal{G}^{\gamma_2}$  pour un certain  $\gamma_2 > 2$ , la Proposition 2.1.18 indique que  $F$  vérifie l'inégalité de Łojasiewicz d'exposant  $1 - \frac{1}{\gamma_2}$  ce qui implique que

$$\int_{t_0}^{+\infty} u^{\frac{2\gamma_1}{\gamma_1-2}} (F(x(u)) - F^*)^{\frac{2(\gamma_2-1)}{\gamma_2}} du < +\infty. \quad (4.A.44)$$

En appliquant le Lemme 4.A.6 à  $\phi : x \mapsto (F(x) - F^*)^{\frac{2(\gamma_2-1)}{\gamma_2}}$ , on trouve que lorsque  $t$  tend vers  $+\infty$ ,

$$\inf_{u \in [t/2, t]} (F(x(u)) - F^*)^{\frac{2(\gamma_2-1)}{\gamma_2}} = o\left(t^{-\frac{3\gamma_1-2}{\gamma_1-2}}\right).$$

On peut donc conclure que,

$$\inf_{u \in [t/2, t]} F(x(u)) - F^* = o\left(t^{-\frac{(3\gamma_1-2)\gamma_2}{2(\gamma_1-2)(\gamma_2-1)}}\right).$$

#### 4.A.6 Autres preuves

##### 4.A.6.1 Preuve du Lemme 4.A.1

Rappelons que l'énergie  $\mathcal{E}$  est définie pour tout  $t \geq t_0$  comme :

$$\mathcal{E}(t) = (t^2 + t\beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{1}{2} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2,$$

Cette énergie est différentiable et sa dérivée  $\mathcal{E}'$  vérifie

$$\begin{aligned} \mathcal{E}'(t) &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + (t^2 + t\beta(\lambda - \alpha)) \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ &\quad + \lambda(\lambda + 1 - \alpha) \langle x(t) - x^*, \dot{x}(t) \rangle + \lambda(\beta - t) \langle \nabla F(x(t)), x(t) - x^* \rangle \\ &\quad + t(\lambda + 1 - \alpha) \|\dot{x}(t)\|^2 + t(\beta - t) \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ &\quad + t\beta(\lambda + 1 - \alpha) \langle \nabla F(x(t)), \dot{x}(t) \rangle + t\beta(\beta - t) \|\nabla F(x(t))\|^2 \\ &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + 2t\beta(\lambda + 1 - \alpha) \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ &\quad + \lambda(\lambda + 1 - \alpha) \langle x(t) - x^*, \dot{x}(t) \rangle + \lambda(\beta - t) \langle \nabla F(x(t)), x(t) - x^* \rangle \\ &\quad + t(\lambda + 1 - \alpha) \|\dot{x}(t)\|^2 + t\beta(\beta - t) \|\nabla F(x(t))\|^2. \end{aligned}$$

En réorganisant les termes, on obtient que :

$$\begin{aligned} \mathcal{E}'(t) &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) \\ &\quad + \frac{\lambda + 1 - \alpha}{t} [t^2 \|\dot{x}(t)\|^2 + t^2 \beta^2 \|\nabla F(x(t))\|^2 + 2t^2 \beta \langle \nabla F(x(t)), \dot{x}(t) \rangle] \\ &\quad + \frac{\lambda(\lambda + 1 - \alpha)}{t} [\langle x(t) - x^*, t\dot{x}(t) \rangle + \langle x(t) - x^*, t\beta\nabla F(x(t)) \rangle] \\ &\quad - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 - \lambda(t + \beta(\lambda - \alpha)) \langle \nabla F(x(t)), x(t) - x^* \rangle \\ &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{\lambda + 1 - \alpha}{t} \|t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 \\ &\quad + \frac{\lambda(\lambda + 1 - \alpha)}{t} \langle x(t) - x^*, t(\dot{x}(t) + \beta\nabla F(x(t))) \rangle \\ &\quad - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 - \lambda(t + \beta(\lambda - \alpha)) \langle \nabla F(x(t)), x(t) - x^* \rangle. \end{aligned}$$

Enfin, on peut réécrire  $\mathcal{E}'$  de manière à faire apparaître certains termes qui seront gérés via nos hypothèses sur  $F$  :

$$\begin{aligned} \mathcal{E}'(t) &= (2 - \gamma\lambda)(t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) - \beta(\lambda - \alpha) (F(x(t)) - F^*) \\ &\quad - \lambda(t + \beta(\lambda - \alpha)) [-\gamma(F(x(t)) - F^*) + \langle \nabla F(x(t)), x(t) - x^* \rangle] \\ &\quad + \frac{\lambda(\lambda + 1 - \alpha)}{t} \langle x(t) - x^*, t(\dot{x}(t) + \beta\nabla F(x(t))) \rangle \\ &\quad + \frac{\lambda + 1 - \alpha}{t} \|t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2. \end{aligned}$$

## 4.A.6.2 Preuve du Lemme 4.A.2

Notons d'abord que pour tout  $t \geq t_0$ ,

$$\begin{aligned} \frac{\gamma\lambda - 2}{t}\mathcal{E}(t) &= - (2 - \gamma\lambda)(t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) \\ &\quad + \frac{1}{2} \frac{\gamma\lambda - 2}{t} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2. \end{aligned} \quad (4.A.45)$$

Alors, en appliquant le Lemme 4.A.2, on trouve :

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t}\mathcal{E}(t) &= \frac{1}{2} \frac{\gamma\lambda - 2}{t} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 \\ &\quad - \lambda(t + \beta(\lambda - \alpha)) [-\gamma(F(x(t)) - F^*) + \langle \nabla F(x(t)), x - x^* \rangle] \\ &\quad + \frac{\lambda(\lambda + 1 - \alpha)}{t} \langle x(t) - x^*, t(\dot{x}(t) + \beta\nabla F(x(t))) \rangle \\ &\quad + \frac{\lambda + 1 - \alpha}{t} \|t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*). \end{aligned} \quad (4.A.46)$$

De plus, la fonction  $F$  vérifie  $\mathcal{F}_\gamma$ , pour tout  $t \geq \max\{t_0, \beta(\alpha - \lambda)\}$

$$\lambda(t + \beta(\lambda - \alpha)) [-\gamma(F(x(t)) - F^*) + \langle \nabla F(x(t)), x - x^* \rangle] \geq 0,$$

et par conséquent,

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t}\mathcal{E}(t) &\leq \frac{1}{2} \frac{\gamma\lambda - 2}{t} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 \\ &\quad + \frac{\lambda(\lambda + 1 - \alpha)}{t} \langle x(t) - x^*, t(\dot{x}(t) + \beta\nabla F(x(t))) \rangle \\ &\quad + \frac{\lambda + 1 - \alpha}{t} \|t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\ &\leq \frac{\gamma\lambda - 2}{2t} \|\lambda(x(t) - x^*)\|^2 \\ &\quad + \left( \frac{\lambda + 1 - \alpha}{t} + \frac{\gamma\lambda - 2}{2t} \right) \|t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2 \\ &\quad + \left( \frac{\lambda + 1 - \alpha}{t} + \frac{\gamma\lambda - 2}{t} \right) \langle \lambda(x(t) - x^*), t(\dot{x}(t) + \beta\nabla F(x(t))) \rangle \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \end{aligned}$$

Comme  $2(\lambda - \alpha) + \gamma\lambda = 0$ , on a donc :

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t}\mathcal{E}(t) &\leq \frac{\gamma\lambda - 2}{2t} \|\lambda(x(t) - x^*)\|^2 \\ &\quad + \frac{\lambda + \gamma\lambda - \alpha - 1}{t} \langle \lambda(x(t) - x^*), t(\dot{x}(t) + \beta\nabla F(x(t))) \rangle \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2. \end{aligned}$$



Ainsi,

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t} \mathcal{E}(t) &\leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \right) \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2, \end{aligned} \quad (4.A.47)$$

$$\text{avec } K(\alpha) = \frac{2\alpha\gamma}{(\gamma+2)^2} \left( \alpha - 1 - \frac{2}{\gamma} \right).$$

#### 4.A.6.3 Preuve du Lemme 4.A.3

Le Lemme 4.A.2 garantit que pour tout  $t > \beta(\alpha - \lambda)$ , on a

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t} \mathcal{E}(t) &\leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \right) \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2. \end{aligned}$$

En ajoutant le terme  $\left( \frac{\gamma\lambda - 2}{t + \beta(\lambda - \alpha)} - \frac{\gamma\lambda - 2}{t} \right) \mathcal{E}(t)$  aux deux membres de l'inégalité, on obtient

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t + \beta(\lambda - \alpha)} \mathcal{E}(t) &\leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \right) \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) + \left( \frac{\gamma\lambda - 2}{t + \beta(\lambda - \alpha)} - \frac{\gamma\lambda - 2}{t} \right) \mathcal{E}(t) \\ &\quad - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\ &\leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \right) \\ &\quad + \beta(\alpha - \lambda) (F(x(t)) - F^*) + \frac{\beta(\alpha - \lambda)(\gamma\lambda - 2)}{t(t + \beta(\lambda - \alpha))} \mathcal{E}(t) \\ &\quad - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \end{aligned}$$

Rappelons que pour tout  $t > \beta(\alpha - \lambda)$ ,

$$F(x(t)) - F^* \leq \frac{\mathcal{E}(t)}{t(t + \beta(\lambda - \alpha))}, \quad (4.A.48)$$

ce qui implique que

$$\begin{aligned} \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t + \beta(\lambda - \alpha)} \mathcal{E}(t) &\leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \right) \\ &\quad + \left( \frac{\beta(\alpha - \lambda)}{t(t + \beta(\lambda - \alpha))} + \frac{\beta(\alpha - \lambda)(\gamma\lambda - 2)}{t(t + \beta(\lambda - \alpha))} \right) \mathcal{E}(t) \\ &\quad - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\ &\leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \right) \\ &\quad + \frac{\beta(\alpha - \lambda)(\gamma\lambda - 1)}{t^2 + t\beta(\lambda - \alpha)} \mathcal{E}(t) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2. \end{aligned} \quad (4.A.49)$$

L'étape suivante consiste à trouver une borne supérieure de

$$\frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle,$$

dépendant de  $\mathcal{E}(t)$ . On y parvient en appliquant les inégalités données dans le lemme ci-dessous dont la preuve se trouve en Section 4.A.6.7.

**Lemme 4.A.10.** *Soient  $u \in \mathbb{R}^N$ ,  $v \in \mathbb{R}^N$  et  $a > 0$ . On a que*

$$|\langle u, v \rangle| \leq \frac{a}{2} \|u\|^2 + \frac{1}{2a} \|v\|^2, \quad (4.A.50)$$

et

$$\|u\|^2 \leq (1+a)\|u+v\|^2 + \left(1 + \frac{1}{a}\right) \|v\|^2. \quad (4.A.51)$$

Ce lemme indique que pour tout  $t > \beta(\alpha - \lambda)$  et  $\theta > 0$ ,

$$|\langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle| \leq \frac{\sqrt{\mu}}{2} \|x(t) - x^*\|^2 + \frac{1}{2\sqrt{\mu}} \|\dot{x}(t) + \beta \nabla F(x(t))\|^2, \quad (4.A.52)$$

et de plus,

$$\begin{aligned} t^2 \|\dot{x}(t) + \beta \nabla F(x(t))\|^2 &\leq \left(1 + \theta \frac{\alpha}{t\sqrt{\mu}}\right) \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 \\ &\quad + \lambda^2 \left(1 + \frac{t\sqrt{\mu}}{\theta\alpha}\right) \|x(t) - x^*\|^2. \end{aligned} \quad (4.A.53)$$

En conséquence, pour tout  $\theta > 0$ ,

$$\begin{aligned} &\frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \\ &\leq \left(\frac{\lambda}{t} + \frac{\sqrt{\mu}}{2}\right) \|x(t) - x^*\|^2 + \frac{1}{2\sqrt{\mu}} \|\dot{x}(t) + \beta \nabla F(x(t))\|^2 \\ &\leq \left(\frac{\lambda^2}{2\sqrt{\mu}t^2} + \frac{\lambda}{t} \left(1 + \frac{\lambda}{2\theta\alpha}\right) + \frac{\sqrt{\mu}}{2}\right) \|x(t) - x^*\|^2 \\ &\quad + \left(\frac{\theta\alpha}{2\mu t^3} + \frac{1}{2\sqrt{\mu}t^2}\right) \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 \\ &\leq \left(\frac{\lambda^2}{\mu^{3/2}t^2} + \frac{2\lambda}{\mu t} \left(1 + \frac{\lambda}{2\theta\alpha}\right) + \frac{1}{\sqrt{\mu}}\right) (F(x(t)) - F^*) \\ &\quad + \left(\frac{\theta\alpha}{2\mu t^3} + \frac{1}{2\sqrt{\mu}t^2}\right) \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2, \end{aligned} \quad (4.A.54)$$

puisque  $F$  satisfait  $\mathcal{G}_\mu^2$  et possède un unique minimiseur.

Comme  $\alpha > \lambda$  On a  $\frac{1}{t} < \frac{1}{t+\beta(\lambda-\alpha)}$  et ainsi,

$$\begin{aligned}
 & \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \\
 & \leq \left( \frac{\lambda^2}{\mu^{3/2}(t+\beta(\lambda-\alpha))^2} + \frac{2\lambda}{\mu(t+\beta(\lambda-\alpha))} \left(1 + \frac{\lambda}{2\theta\alpha}\right) + \frac{1}{\sqrt{\mu}} \right) (F(x(t)) - F^*) \\
 & + \left( \frac{\theta\alpha}{2\mu(t+\beta(\lambda-\alpha))^3} + \frac{1}{2\sqrt{\mu}(t+\beta(\lambda-\alpha))^2} \right) \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 \\
 & \leq \left( \frac{\lambda^2}{\mu^{3/2}(t+\beta(\lambda-\alpha))^4} + \frac{2\lambda}{\mu(t+\beta(\lambda-\alpha))^3} \left(1 + \frac{\lambda}{2\theta\alpha}\right) + \frac{1}{\sqrt{\mu}(t+\beta(\lambda-\alpha))^2} \right) t(t + \beta(\lambda - \alpha))(F(x(t)) - F^*) \\
 & + \left( \frac{\theta\alpha}{\mu(t+\beta(\lambda-\alpha))^3} + \frac{1}{\sqrt{\mu}(t+\beta(\lambda-\alpha))^2} \right) \frac{1}{2} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2.
 \end{aligned} \tag{4.A.55}$$

On choisit le paramètre  $\theta$  pour assurer que  $\frac{2\lambda}{\mu(t+\beta(\lambda-\alpha))^3} \left(1 + \frac{\lambda}{2\theta\alpha}\right) = \frac{\theta\alpha}{\mu(t+\beta(\lambda-\alpha))^3}$ .  
Cela conduit au choix  $\theta = \frac{2}{\gamma+2}(1 + \sqrt{2})$  qui mène aux inégalités ci-dessous :

$$\begin{aligned}
 & \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \\
 & \leq \frac{1}{\mu(t+\beta(\lambda-\alpha))^2} \left( \frac{\lambda^2}{\sqrt{\mu}(t+\beta(\lambda-\alpha))^2} + \frac{\lambda}{t+\beta(\lambda-\alpha)} \left(1 + \sqrt{2}\right) + \sqrt{\mu} \right) t(t + \beta(\lambda - \alpha))(F(x(t)) - F^*) \\
 & + \frac{1}{\mu(t+\beta(\lambda-\alpha))^2} \left( \frac{\lambda}{t+\beta(\lambda-\alpha)} \left(1 + \sqrt{2}\right) + \sqrt{\mu} \right) \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 \\
 & \leq \frac{1}{\mu(t+\beta(\lambda-\alpha))^2} \left( \frac{\lambda^2}{\sqrt{\mu}(t+\beta(\lambda-\alpha))^2} + \frac{\lambda}{t+\beta(\lambda-\alpha)} \left(1 + \sqrt{2}\right) + \sqrt{\mu} \right) \mathcal{E}(t).
 \end{aligned} \tag{4.A.56}$$

En réinjectant ce résultat à (4.A.49), on obtient que,

$$\begin{aligned}
 \mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t + \beta(\lambda - \alpha)} \mathcal{E}(t) & \leq K(\alpha) \left( \frac{\lambda}{t} \|x(t) - x^*\|^2 + \langle x(t) - x^*, \dot{x}(t) + \beta \nabla F(x(t)) \rangle \right) \\
 & + \frac{\beta(\alpha - \lambda)(\gamma\lambda - 1)}{t(t + \beta(\lambda - \alpha))} \mathcal{E}(t) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\
 & \leq \frac{K(\alpha)}{\mu(t+\beta(\lambda-\alpha))^2} \left( \frac{\lambda^2}{\sqrt{\mu}(t+\beta(\lambda-\alpha))^2} + \frac{\lambda}{t+\beta(\lambda-\alpha)} \left(1 + \sqrt{2}\right) + \sqrt{\mu} \right) \mathcal{E}(t) \\
 & + \frac{\beta(\alpha - \lambda)(\gamma\lambda - 1)}{(t + \beta(\lambda - \alpha))^2} \mathcal{E}(t) - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2.
 \end{aligned} \tag{4.A.57}$$

On définit  $C_0 = \frac{\beta\sqrt{\mu}(\alpha-\lambda)(\gamma\lambda-1)}{K(\alpha)}$  ce qui permet d'écrire que

$$\begin{aligned}
 \mathcal{E}'(t) + \frac{\gamma\lambda-2}{t+\beta(\lambda-\alpha)} \mathcal{E}(t) & \leq \frac{K(\alpha)}{\mu(t+\beta(\lambda-\alpha))^2} \left( \frac{\lambda^2}{\sqrt{\mu}(t+\beta(\lambda-\alpha))^2} + \frac{\lambda}{t+\beta(\lambda-\alpha)} \left(1 + \sqrt{2}\right) + \sqrt{\mu}(1 + C_0) \right) \mathcal{E}(t) \\
 & - t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2.
 \end{aligned} \tag{4.A.58}$$

et donc,

$$\mathcal{E}'(t) + \frac{\gamma\lambda - 2}{t + \beta(\lambda - \alpha)} \mathcal{E}(t) + t\beta(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \leq \varphi(t + \beta(\lambda - \alpha)) \mathcal{E}(t) \tag{4.A.59}$$

où

$$\varphi : t \mapsto \frac{K(\alpha)}{\mu t^2} \left( \sqrt{\mu}(1 + C_0) + \frac{2\alpha}{(\gamma + 2)t} (1 + \sqrt{2}) + \frac{4\alpha^2}{(\gamma + 2)^2 \sqrt{\mu} t^2} \right). \tag{4.A.60}$$

## 4.A.6.4 Preuve du Lemme 4.A.4

Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$  ayant un ensemble de minimiseurs  $X^*$  non vide. Soient  $\delta \in (0, 1]$  et  $x^* \in X^*$ .

On introduit le lemme ci-dessous qui est prouvé en Section 4.A.6.8.

**Lemme 4.A.11.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$ . Pour tout  $x \in \mathbb{R}^N$  et  $\varepsilon > 0$ , il existe  $\nu > 0$  tel que pour tout  $y \in B(x, \nu)$  :*

$$(1 - \varepsilon)(y - x)^T H_F(x)(y - x) \leq (y - x)^T H_F(y)(y - x) \leq (1 + \varepsilon)(y - x)^T H_F(x)(y - x). \quad (4.A.61)$$

Comme  $F$  est une fonction  $C^2$ , le Lemme 4.A.11 assure qu'il existe  $\nu > 0$  tel que pour tout  $x \in B(x^*, \nu)$  :

$$\left(1 - \frac{\delta}{4 - \delta}\right) K(x) \leq (x - x^*)^T H_F(x)(x - x^*) \leq \left(1 + \frac{\delta}{4 - \delta}\right) K(x), \quad (4.A.62)$$

où  $K(x) = (x - x^*)^T H_F(x^*)(x - x^*)$ .

Soit  $\phi_{x,x^*}$  défini par

$$\begin{aligned} \phi_{x,x^*} : [0, 1] &\rightarrow \mathbb{R} \\ t &\mapsto F(tx + (1 - t)x^*), \end{aligned} \quad (4.A.63)$$

avec  $x \in B(x^*, \nu)$ . La fonction  $\phi_{x,x^*}$  est deux fois dérivable et pour tout  $t \in [0, 1]$ , on a

$$\begin{aligned} \phi'_{x,x^*}(t) &= (x - x^*)^T \nabla F(tx + (1 - t)x^*), \\ \phi''_{x,x^*}(t) &= (x - x^*)^T H_F(tx + (1 - t)x^*)(x - x^*). \end{aligned}$$

En réécrivant (4.A.62) au point  $tx + (1 - t)x^*$  pour un certain  $t \in [0, 1]$  on trouve

$$\left(1 - \frac{\delta}{4 - \delta}\right) \phi''_{x,x^*}(0) \leq \phi''_{x,x^*}(t) \leq \left(1 + \frac{\delta}{4 - \delta}\right) \phi''_{x,x^*}(0). \quad (4.A.64)$$

En intégrant le membre de gauche de (4.A.64) et comme  $\phi'_{x,x^*}(0) = 0$  (car  $\nabla F(x^*) = 0$ ), on obtient que

$$\forall t \in [0, 1], \quad \left(1 - \frac{\delta}{4 - \delta}\right) \phi''_{x,x^*}(0)t \leq \phi'_{x,x^*}(t).$$

En intégrant le membre de droite de (4.A.64), on a

$$\forall t \in [0, 1], \quad \phi_{x,x^*}(t) - \phi_{x,x^*}(0) \leq \left(1 + \frac{\delta}{4 - \delta}\right) \phi''_{x,x^*}(0) \frac{t^2}{2},$$

et par conséquent,

$$\forall t \in [0, 1], \quad \phi_{x,x^*}(t) - \phi_{x,x^*}(0) \leq \frac{1}{2 - \delta} t \phi'_{x,x^*}(t).$$

En prenant  $t = 1$  et en exprimant  $\phi_{x,x^*}$  et  $\phi'_{x,x^*}$  on peut déduire que

$$F(x) - F^* \leq \frac{1}{2 - \delta} \langle \nabla F(x), x - x^* \rangle. \quad (4.A.65)$$

**4.A.6.5 Preuve du Lemme 4.A.5**

Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe ayant un ensemble de minimiseurs  $X^*$  non vide. Supposons que pour un certain  $t_1 > 0$  et un certain  $\delta > 0$ ,  $F$  vérifie :

$$\int_{t_1}^{+\infty} u^\delta (F(x(u)) - F^*) du < +\infty. \tag{4.A.66}$$

Soit  $\varepsilon > 0$ . Cette hypothèse assure qu'il existe  $t_2 \geq 2t_1$  tel que :

$$\forall t \geq t_2, \quad \int_{t/2}^t u^\delta (F(x(u)) - F^*) du < \varepsilon.$$

On définit  $z$  de la manière suivante :

$$z : t \mapsto \frac{\int_{t/2}^t u^\delta x(u) du}{\int_{t/2}^t u^\delta du}.$$

Soit  $t \geq t_2$ . On définit  $\nu$  comme :

$$\begin{aligned} \nu : \mathcal{B}([t/2, t]) &\rightarrow [0, 1] \\ A &\mapsto \frac{\int_A u^\delta du}{\int_{t/2}^t u^\delta du}, \end{aligned}$$

où  $\mathcal{B}([t/2, t])$  la tribu borélienne sur  $[t/2, t]$ . On peut alors écrire que  $z(t) = \int_{t/2}^t x(u) d\nu(u)$ . Comme  $\nu([t/2, t]) = 1$  et  $F$  est une fonction convexe, on peut utiliser l'inégalité de Jensen pour écrire que :

$$\begin{aligned} F(z(t)) - F^* &= F\left(\int_{t/2}^t x(u) d\nu(u)\right) - F^* \\ &\leq \int_{t/2}^t F(x(u)) d\nu(u) - F^* \\ &\leq \int_{t/2}^t (F(x(u)) - F^*) d\nu(u) \\ &\leq \frac{\varepsilon}{\int_{t/2}^t u^\delta du} \end{aligned}$$

Ainsi, quand  $t$  tend vers  $+\infty$ ,  $F(z(t)) - F^* = o(t^{-\delta-1})$ .

**4.A.6.6 Preuve du Lemme 4.A.6**

Soit une fonction  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}^+$  telle que pour un certain  $t_1 > 0$  et un certain  $\delta > 0$ ,  $\phi$  vérifie :

$$\int_{t_1}^{+\infty} u^\delta \phi(x(u)) du < +\infty. \tag{4.A.67}$$

Soit  $\varepsilon > 0$ . Cette hypothèse assure qu'il existe  $t_2 \geq 2t_1$  tel que

$$\forall t \geq t_2, \quad \int_{t/2}^t u^\delta \phi(x(u)) du < \varepsilon.$$

Ainsi, pour tout  $t \geq t_2$ ,

$$\inf_{u \in [t/2, t]} \phi(x(u)) \int_{t/2}^t u^\delta du < \varepsilon,$$

et

$$\inf_{u \in [t/2, t]} \phi(x(u)) < \frac{\varepsilon(\delta + 1)}{t^{\delta+1} - \left(\frac{t}{2}\right)^{\delta+1}}.$$

Par conséquent, quand  $t \rightarrow +\infty$ ,

$$\inf_{u \in [t/2, t]} \phi(x(u)) = o\left(t^{-\delta-1}\right). \quad (4.A.68)$$

On rappelle que  $\liminf_{t \rightarrow +\infty} f(t) = \lim_{t \rightarrow +\infty} \left[ \inf_{\tau \geq t} f(\tau) \right]$ . Comme  $\phi$  est une fonction positive, on trouve que

$$\liminf_{t \rightarrow +\infty} t^{\delta+1} \log(t) \phi(x(t)) = l \geq 0.$$

Supposons que  $l > 0$ . Alors il existe  $\hat{t} > t_1$  tel que :

$$\forall t \geq \hat{t}, \quad t^{\delta+1} \log(t) \phi(x(t)) \geq \frac{l}{2},$$

et donc,

$$\forall t \geq \hat{t}, \quad t^\delta \phi(x(t)) \geq \frac{l}{2t \log(t)}.$$

Ceci est impossible si l'équation (4.A.67) est vérifiée. On obtient donc que  $l = 0$ .

#### 4.A.6.7 Preuve du Lemme 4.A.10

Soient  $u \in \mathbb{R}^N$ ,  $v \in \mathbb{R}^N$  et  $a > 0$ . La première inégalité provient des calculs suivants

$$\langle u, v \rangle = \frac{1}{2} \left\| \sqrt{a}u - \frac{v}{\sqrt{a}} \right\|^2 - \frac{a}{2} \|u\|^2 - \frac{1}{2a} \|v\|^2 \geq -\frac{a}{2} \|u\|^2 - \frac{1}{2a} \|v\|^2,$$

et

$$\langle u, v \rangle = \frac{a}{2} \|u\|^2 + \frac{1}{2a} \|v\|^2 - \frac{1}{2} \left\| \sqrt{a}u + \frac{v}{\sqrt{a}} \right\|^2 \leq \frac{a}{2} \|u\|^2 + \frac{1}{2a} \|v\|^2.$$

On obtient la deuxième inégalité en écrivant  $\|u\|^2$  de la manière suivante :

$$\|u\|^2 = \|u + v\|^2 + \|v\|^2 - 2\langle u + v, v \rangle,$$

et en appliquant la première inégalité à  $\langle u + v, v \rangle$ .

4.A.6.8 Preuve du Lemme 4.A.11

Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe  $C^2$ . Les dérivées partielles de second ordre de  $F$  sont notées  $\partial_{ij}F = \frac{\partial^2 F}{\partial x_i \partial x_j}$  pour tout  $(i, j) \in \llbracket 1, n \rrbracket^2$ .

Soient  $x \in \mathbb{R}^N$  et  $\varepsilon > 0$ . Pour tout  $(i, j) \in \llbracket 1, n \rrbracket^2$ ,  $\partial_{ij}F$  est une fonction continue sur  $\mathbb{R}^N$  et donc

$$\exists \tilde{\nu} > 0, \forall y \in B(x, \tilde{\nu}), (1 - \varepsilon)\partial_{ij}F(x) \leq \partial_{ij}F(y) \leq (1 + \varepsilon)\partial_{ij}F(x).$$

En prenant la valeur minimale de  $\tilde{\nu}$  sur tout les  $(i, j) \in \llbracket 1, n \rrbracket^2$ , on peut déduire qu'il existe  $\tilde{\nu} > 0$  tel que :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \forall y \in B(x, \tilde{\nu}), (1 - \varepsilon)\partial_{ij}F(x) \leq \partial_{ij}F(y) \leq (1 + \varepsilon)\partial_{ij}F(x). \quad (4.A.69)$$

Soit  $\nu = \min \left\{ \tilde{\nu}, \left( n \max_{(i,j) \in \llbracket 1, n \rrbracket^2} |\partial_{ij}F(x)| \right)^{-\frac{1}{2}} \right\}$ ,  $y \in B(x, \nu)$  et  $h = y - x$ . L'équation (4.A.69) assure que pour tout  $(i, j) \in \llbracket 1, n \rrbracket^2$  :

$$\partial_{ij}F(x)h_i h_j - \varepsilon |\partial_{ij}F(x)h_i h_j| \leq \partial_{ij}F(y)h_i h_j \leq \partial_{ij}F(x)h_i h_j + \varepsilon |\partial_{ij}F(x)h_i h_j|. \quad (4.A.70)$$

Rappelons que pour tout  $(i, j) \in \llbracket 1, n \rrbracket^2$ ,  $(H_F(x))_{i,j} = \partial_{ij}F(x)$  et donc

$$\forall (x, h) \in \mathbb{R}^N \times \mathbb{R}^N, h^T H_F(x) h = \sum_{i=1}^n \sum_{j=1}^n \partial_{ij}F(x) h_i h_j. \quad (4.A.71)$$

En sommant (4.A.70) sur l'ensemble des  $(i, j) \in \llbracket 1, n \rrbracket^2$ , on obtient que

$$h^T H_F(x) h - \varepsilon \sum_{i=1}^n \sum_{j=1}^n |\partial_{ij}F(x)h_i h_j| \leq h^T H_F(y) h \leq h^T H_F(x) h + \varepsilon \sum_{i=1}^n \sum_{j=1}^n |\partial_{ij}F(x)h_i h_j|.$$

Notons que  $|h_i h_j| \leq \frac{1}{2} (h_i^2 + h_j^2)$  pour tout  $(i, j) \in \llbracket 1, n \rrbracket^2$ , ce qui donne que

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n |\partial_{ij}F(x)h_i h_j| &\leq \max_{(i,j) \in \llbracket 1, n \rrbracket^2} |\partial_{ij}F(x)| \sum_{i=1}^n \sum_{j=1}^n |h_i h_j| \\ &\leq n \max_{(i,j) \in \llbracket 1, n \rrbracket^2} |\partial_{ij}F(x)| \|h\|^2 \\ &\leq n \max_{(i,j) \in \llbracket 1, n \rrbracket^2} |\partial_{ij}F(x)| \nu^2 \\ &\leq 1. \end{aligned}$$

Par conséquent,

$$(1 - \varepsilon)h^T H_F(x) h \leq h^T H_F(y) h \leq (1 + \varepsilon)h^T H_F(x) h.$$

## 4.A.6.9 Preuve du Lemme 4.A.7

On considère l'énergie  $\mathcal{E}$  définie pour tout  $t \geq t_0$  par

$$\begin{aligned} \mathcal{E}(t) &= (t^2 + t\beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{\xi}{2} \|x(t) - x^*\|^2 \\ &\quad + \frac{1}{2} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))\|^2. \end{aligned}$$

Soit  $v : t \mapsto \lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t)))$ . La fonction  $v$  est dérivable et on a :

$$\begin{aligned} v'(t) &= \lambda\dot{x}(t) + t\ddot{x}(t) + \dot{x}(t) + \beta\nabla F(x(t)) + t\beta\nabla^2 F(x(t))\dot{x}(t) \\ &= (\lambda + 1)\dot{x}(t) + (-\alpha\dot{x}(t) - t\beta\nabla^2 F(x(t))\dot{x}(t) - t\nabla F(x(t))) + \beta\nabla F(x(t)) \\ &\quad + t\beta\nabla^2 F(x(t))\dot{x}(t) \\ &= (\lambda + 1 - \alpha)\dot{x}(t) + (\beta - t)\nabla F(x(t)). \end{aligned}$$

En dérivant  $\mathcal{E}$ , on obtient que :

$$\begin{aligned} \mathcal{E}'(t) &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + (t^2 + t\beta(\lambda - \alpha)) \langle \nabla F(x(t)), \dot{x}(t) \rangle + \langle v(t), v'(t) \rangle \\ &\quad + \xi \langle x(t) - x^*, \dot{x}(t) \rangle. \end{aligned}$$

Quelques développements donnent que

$$\begin{aligned} \langle v(t), v'(t) \rangle &= \langle \lambda(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla F(x(t))), (\lambda + 1 - \alpha)\dot{x}(t) + (\beta - t)\nabla F(x(t)) \rangle \\ &= \lambda(\lambda + 1 - \alpha) \langle x(t) - x^*, \dot{x}(t) \rangle + \lambda(\beta - t) \langle \nabla F(x(t)), x(t) - x^* \rangle \\ &\quad + t(\lambda + 1 - \alpha) \|\dot{x}(t)\|^2 + t(\beta - t) \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ &\quad + t\beta(\lambda + 1 - \alpha) \langle \nabla F(x(t)), \dot{x}(t) \rangle + t\beta(\beta - t) \|\nabla F(x(t))\|^2. \end{aligned}$$

Par conséquent,

$$\begin{aligned} \mathcal{E}'(t) &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + (t^2 + t\beta(\lambda - \alpha)) \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ &\quad + \lambda(\lambda + 1 - \alpha) \langle x(t) - x^*, \dot{x}(t) \rangle + \lambda(\beta - t) \langle \nabla F(x(t)), x(t) - x^* \rangle \\ &\quad + t(\lambda + 1 - \alpha) \|\dot{x}(t)\|^2 + t(\beta - t) \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ &\quad + t\beta(\lambda + 1 - \alpha) \langle \nabla F(x(t)), \dot{x}(t) \rangle + t\beta(\beta - t) \|\nabla F(x(t))\|^2 \\ &\quad + \xi \langle x(t) - x^*, \dot{x}(t) \rangle \\ &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + 2t\beta(\lambda + 1 - \alpha) \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ &\quad + \lambda(\lambda + 1 - \alpha) \langle x(t) - x^*, \dot{x}(t) \rangle + \lambda(\beta - t) \langle \nabla F(x(t)), x(t) - x^* \rangle \\ &\quad + t(\lambda + 1 - \alpha) \|\dot{x}(t)\|^2 + t\beta(\beta - t) \|\nabla F(x(t))\|^2 \\ &\quad + \xi \langle x(t) - x^*, \dot{x}(t) \rangle. \end{aligned}$$



Alors, en réarrangeant les termes on a :

$$\begin{aligned}
 \mathcal{E}'(t) &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) \\
 &+ \frac{\lambda + 1 - \alpha}{t} [t^2 \|\dot{x}(t)\|^2 + 2t^2 \beta \langle \nabla F(x(t)), \dot{x}(t) \rangle + t^2 \beta^2 \|\nabla F(x(t))\|^2] \\
 &+ \frac{\lambda(\lambda + 1 - \alpha)}{t} [\langle x(t) - x^*, t\dot{x}(t) \rangle + \langle x(t) - x^*, t\beta \nabla F(x(t)) \rangle] \\
 &- \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\
 &- \lambda(t + \beta(\lambda - \alpha)) \langle \nabla F(x(t)), x(t) - x^* \rangle \\
 &+ \xi \langle x(t) - x^*, \dot{x}(t) \rangle \\
 &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{\lambda + 1 - \alpha}{t} \|t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 \\
 &+ \frac{\lambda(\lambda + 1 - \alpha)}{t} \langle x(t) - x^*, t(\dot{x}(t) + \beta \nabla F(x(t))) \rangle \\
 &- \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\
 &- \lambda(t + \beta(\lambda - \alpha)) \langle \nabla F(x(t)), x(t) - x^* \rangle \\
 &+ \xi \langle x(t) - x^*, \dot{x}(t) \rangle \\
 &= (2t + \beta(\lambda - \alpha)) (F(x(t)) - F^*) + \frac{\lambda + 1 - \alpha}{t} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 \\
 &- \frac{\lambda^2(\lambda + 1 - \alpha)}{t} \|x(t) - x^*\|^2 - \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 \\
 &- \lambda(t + \beta(2(\lambda - \alpha) + 1)) \langle \nabla F(x(t)), x(t) - x^* \rangle + (\xi - \lambda(\lambda + 1 - \alpha)) \langle x(t) - x^*, \dot{x}(t) \rangle.
 \end{aligned}$$

Cette expression permet d'écrire que

$$\begin{aligned}
 \mathcal{E}'(t) &= ((2 - \lambda\gamma_1)t + \beta(\lambda - \alpha - \lambda\gamma_1(2(\lambda - \alpha) + 1))) (F(x(t)) - F^*) \\
 &+ \lambda(t + \beta(2(\lambda - \alpha) + 1)) [\gamma_1 (F(x(t)) - F^*) - \langle \nabla F(x(t)), x(t) - x^* \rangle] \\
 &+ \frac{\lambda + 1 - \alpha}{t} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 - \frac{\lambda^2(\lambda + 1 - \alpha)}{t} \|x(t) - x^*\|^2 \\
 &- \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2 + (\xi - \lambda(\lambda + 1 - \alpha)) \langle x(t) - x^*, \dot{x}(t) \rangle.
 \end{aligned}$$

Comme  $F$  vérifie  $\mathcal{F}_{\gamma_1}$ , pour tout  $t \geq \beta(2(\alpha - \lambda) - 1)$ , on a

$$\lambda(t + \beta(2(\lambda - \alpha) + 1)) [\gamma_1 (F(x(t)) - F^*) - \langle \nabla F(x(t)), x(t) - x^* \rangle] \leq 0.$$

Enfin, on trouve que

$$\begin{aligned}
 \mathcal{E}'(t) &\leq ((2 - \lambda\gamma_1)t + \beta(\lambda - \alpha - \lambda\gamma_1(2(\lambda - \alpha) + 1))) (F(x(t)) - F^*) \\
 &+ \frac{\lambda + 1 - \alpha}{t} \|\lambda(x(t) - x^*) + t(\dot{x}(t) + \beta \nabla F(x(t)))\|^2 \\
 &- \frac{\lambda^2(\lambda + 1 - \alpha)}{t} \|x(t) - x^*\|^2 - \beta t(t + \beta(\lambda - \alpha)) \|\nabla F(x(t))\|^2.
 \end{aligned}$$

# Schémas inertiels et non unicité des minimiseurs

---

Ce chapitre traite de l'hypothèse d'unicité du minimiseur dans l'analyse de convergence de méthodes inertielles et de systèmes dynamiques associés. Dans la littérature, les résultats de convergence les plus forts pour des systèmes dynamiques et des schémas numériques lorsque  $F$  vérifie une condition de croissance s'appuient, sauf exception, sur l'hypothèse que  $F$  a un unique minimiseur. Les questions posées dans ce chapitre sont donc : **à quoi est due cette hypothèse ? Est-il possible d'obtenir des résultats de convergence rapide en s'en affranchissant ?**

Après avoir détaillé les résultats existants, je donne une stratégie pour contourner cette hypothèse dans le cas continu : j'introduis une hypothèse de géométrie sur l'ensemble de minimiseurs  $X^*$  permettant d'**étendre plusieurs résultats de convergence sans unicité du minimiseur**. En particulier, les théorèmes 5.3.2 et 5.3.3 assurent que la décroissance de l'erreur le long des trajectoires du système de Heavy-Ball avec friction (HBF) est exponentielle rapide si  $F$  vérifie des hypothèses plus faibles que la forte convexité et **sans hypothèse d'unicité du minimiseur**. L'hypothèse de géométrie sur  $X^*$  permet également d'obtenir des résultats de convergence améliorés pour les trajectoires de (AVD) (système associé au schéma de gradient accéléré de Nesterov) lorsque  $F$  vérifie des hypothèses de croissance (théorèmes 5.3.4 et 5.3.5).

Dans un deuxième temps, je démontre que cette hypothèse d'unicité du minimiseur n'est pas nécessaire à l'obtention de résultats de convergence rapide dans le cas discret en étudiant V-FISTA [31] et FISTA [32]. Les théorèmes 5.4.1 et 5.4.2 assurent que la décroissance de l'erreur pour les itérés de V-FISTA avec un choix d'inertie pertinent est **exponentielle rapide** si on considère une fonction à croissance quadratique. Les résultats obtenus ne s'appuient sur **aucune hypothèse sur l'ensemble des minimiseurs et des garanties sont données pour un choix de paramètre sous-optimal** (Corollaire 5.4.4). Enfin, une analyse de FISTA est proposée sous des hypothèses de croissance plus faibles que la forte convexité. Le résultat principal est donné dans le Théorème 5.4.6 qui assure que **les itérés de FISTA convergent fortement dès que  $F$  vérifie une hypothèse de croissance faible  $\mathcal{G}^\gamma$  avec  $\gamma > 2$  pour un terme inertielle  $\alpha$  suffisamment grand**.

## Contents

---

|            |                                |            |
|------------|--------------------------------|------------|
| <b>5.1</b> | <b>Problématique</b>           | <b>134</b> |
| <b>5.2</b> | <b>Revue de la littérature</b> | <b>135</b> |

|            |  |            |
|------------|--|------------|
| 5.2.1      | Systèmes dynamiques . . . . .  | 135        |
| 5.2.2      | Schémas numériques . . . . .   | 138        |
| <b>5.3</b> | <b>Non unicité dans le cas continu . . . . .</b>                     | <b>142</b> |
| 5.3.1      | Stratégie . . . . .  | 142        |
| 5.3.2      | Résultats de convergence sans unicité du minimiseur . . . . .        | 146        |
| <b>5.4</b> | <b>Impact de la non unicité sur les schémas numériques . . . . .</b> | <b>148</b> |
| 5.4.1      | Convergence de V-FISTA . . . . .                                     | 149        |
| 5.4.2      | Convergence de FISTA . . . . .                                       | 152        |
| 5.4.3      | Preuves . . . . .  | 154        |
| <b>5.A</b> | <b>Preuves techniques . . . . .</b>                                  | <b>170</b> |
| 5.A.1      | Preuve du Lemme 5.3.1 . . . . .                                      | 170        |
| 5.A.2      | Preuve du Théorème 5.3.2 . . . . .                                   | 170        |
| 5.A.3      | Preuve du Théorème 5.3.3 . . . . .                                   | 171        |
| 5.A.4      | Preuve du Théorème 5.3.4 . . . . .                                   | 172        |
| 5.A.5      | Preuve du Théorème 5.3.5 . . . . .                                   | 173        |
| 5.A.6      | Preuve du Lemme 5.4.7 . . . . .                                      | 175        |
| 5.A.7      | Preuve du Lemme 5.4.8 . . . . .                                      | 175        |
| 5.A.8      | Preuve du Lemme 5.4.9 . . . . .                                      | 176        |
| 5.A.9      | Preuve du Lemme 5.4.10 . . . . .                                     | 176        |
| 5.A.10     | Preuve du Corollaire 5.4.4 . . . . .                                 | 177        |
| 5.A.11     | Preuve du Lemme 5.4.11 . . . . .                                     | 177        |
| 5.A.12     | Preuve du Lemme 5.4.13 . . . . .                                     | 179        |
| 5.A.13     | Preuve du Lemme 5.4.14 . . . . .                                     | 180        |

## 5.1 Problématique

Lorsque l'on s'intéresse à la convergence de méthodes inertiels de premier ordre ou de systèmes dynamiques liés à ces méthodes sous des conditions plus fortes que la convexité, il existe peu de résultats de convergence rapide ne s'appuyant pas sur l'unicité des minimiseurs de la fonction  $F$ . La plupart des preuves utilisent le fait que le minimum de  $F$  ne soit atteint qu'en un unique point  $x^*$ . Dans le cas contraire, les vitesses de convergence démontrées sont plus faibles, ou des hypothèses additionnelles doivent être faites.

Cette hypothèse d'unicité du minimiseur peut être restrictive en pratique. Par exemple, les fonctions de la forme

$$F : x \mapsto \phi(Ax) + \|x\|_1, \quad (5.1.1)$$

avec  $\phi : \mathbb{R}^P \rightarrow \mathbb{R}$ ,  $A : \mathbb{R}^N \rightarrow \mathbb{R}^P$  et  $P > 0$  peuvent avoir un ensemble de minimiseurs qui ne se réduit pas à un point si  $\text{Ker} A \neq \{0\}$  tout en vérifiant une condition de croissance plus forte que la convexité. C'est le cas de la fonctionnelle associée au

problème LASSO [130] qui vérifie l'hypothèse de croissance quadratique  $\mathcal{G}_\mu^2$  mais qui peut avoir une infinité de minimiseurs :

$$F : x \mapsto \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1. \quad (5.1.2)$$

La question de l'unicité des minimiseurs de cette fonctionnelle a d'ailleurs été étudiée par Tibshirani dans [131] et par Bello-Cruz, Li et Nghia dans [37]. Les auteurs de ce dernier papier étudient également les fonctions avec régularisation  $\ell_1$  de manière plus générale.

Le fait d'obtenir des résultats de convergence rapide sous cette condition d'unicité du minimiseur peut donc être un frein à l'application de méthodes inertielles lorsque l'on s'intéresse à un problème ayant possiblement une infinité de solutions.

Cette hypothèse d'unicité du minimiseur peut être demandée de manière implicite ou explicite. La condition de croissance la plus étudiée dans la littérature est la forte convexité ( $\mathcal{S}_\mu$ ) qui ne peut être vérifiée que pour des fonction ayant un unique minimiseur. Dans ce cas, l'unicité du minimiseur est requise de manière implicite. Des hypothèses de géométrie moins restrictives comme la quasi-forte convexité ( $\mathcal{Q}_\mu$ ) ou la condition de croissance quadratique ( $\mathcal{G}_\mu^2$ ) n'impliquent pas l'unicité du minimiseur. Cependant, les résultats de convergence les plus rapides démontrés dans la littérature pour ces classes de fonctions s'appuient généralement sur l'hypothèse additionnelle que  $X^* = \{x^*\}$ .

Les questions que je me suis posées avec mes collaborateurs sont donc les suivantes : cette hypothèse d'unicité du minimiseur est-elle absolument nécessaire à l'obtention de résultats de convergence rapide ou est-ce une hypothèse liée aux techniques de preuve ? Peut-on trouver un moyen d'obtenir des vitesses rapides en s'en affranchissant ?

On présente dans les sections suivantes les résultats de convergence obtenus dans la littérature sous condition de croissance pour des systèmes dynamiques liés à l'optimisation et des schémas de premier ordre.

## 5.2 Revue de la littérature

### 5.2.1 Systèmes dynamiques

Comme précisé en Section 2.3.1, Polyak s'appuie en 1964 [112] sur le système de Heavy-Ball avec friction pour dériver une méthode d'optimisation de premier ordre, appelée méthode de Heavy-Ball de Polyak dans ce manuscrit. On rappelle la définition de ce système dynamique ci-après :

$$\forall t \geq t_0, \quad \ddot{x}(t) + \alpha \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{HBF})$$

où  $\alpha > 0$  et  $t_0 \geq 0$ . L'étude de Polyak est proposée pour les fonctions vérifiant une hypothèse de forte convexité ( $\mathcal{S}_\mu$ ) et dérivables deux fois ( $C^2$ ). Sous ces conditions, Polyak démontre la décroissance exponentielle rapide de l'erreur le long de la trajectoire de (HBF) :  $F(x(t)) - F^* = \mathcal{O}(e^{-2\sqrt{\mu}t})$ . Comme dans le cas discret (voir la

discussion en Section 3.1.1), on fait la distinction entre décroissance exponentielle rapide (taux de la forme  $e^{-K\sqrt{\mu}t}$  avec  $K > 0$ ) et décroissance exponentielle lente (taux de la forme  $e^{-K\mu t}$  avec  $K > 0$ ) puisque le paramètre de croissance  $\mu$  est généralement petit.

Par la suite, Siegel [126] prouve que sous une simple hypothèse de forte convexité, les trajectoires de (HBF) peuvent assurer une décroissance de l'erreur de  $\mathcal{O}(e^{-\sqrt{\mu}t})$ . Aujol, Dossal et Rondepierre proposent dans [24] et [25] des analyses de convergence des trajectoires de (HBF) respectivement pour des fonctions quasi-fortement convexes ( $\mathcal{Q}_\mu$ ) et à croissance quadratique ( $\mathcal{G}_\mu^2$ ). Pour chaque hypothèse, les auteurs démontrent une décroissance exponentielle rapide de l'erreur (respectivement  $\mathcal{O}(e^{-\sqrt{2\mu}t})$  et  $\mathcal{O}(e^{-(2-\sqrt{2})\sqrt{\mu}t})$ ) à condition que la fonction  $F$  ait un unique minimiseur  $x^*$ . Les preuves de convergence, comme celle de Siegel, s'appuient sur l'analyse d'énergies de Lyapunov.

Dans [36], Bégout, Bolte et Jendoubi étudient (HBF) sous hypothèse de Kurdyka-Łojasiewicz à partir de l'analyse de fonctions désingularisantes. Cette approche leur permet de prouver que les solutions de (HBF) peuvent garantir une décroissance exponentielle lente de l'erreur  $\mathcal{O}(e^{-\mu t})$  si  $F$  est  $C^2$  et vérifie une hypothèse de croissance quadratique  $\mathcal{G}_\mu^2$ . L'unicité du minimiseur de  $F$  n'est pas requise ici mais le taux de convergence est exponentiel lent.

Dans [114], Polyak et Shcherbakov proposent une analyse de Lyapunov qui permet de s'affranchir de l'unicité du minimiseur pour les fonctions  $C^2$  vérifiant la condition de Polyak-Łojasiewicz à gradient  $L$ -Lipschitz. Cette condition de croissance peut notamment être vérifiée par des fonctions non convexes et correspond à une hypothèse de croissance quadratique dans le cas convexe. Le taux de convergence obtenu est exponentiel mais les constantes ne sont pas précisées par les auteurs. En développant leurs calculs on peut trouver une décroissance en  $\mathcal{O}\left(e^{-2\frac{\mu^{\frac{3}{2}}}{(\sqrt{2}+1)\mu+L}t}\right)$  qui n'est pas exponentielle rapide puisque  $L$  est plus grand que  $\mu$  et en général,  $L \gg \mu$ . Apidopoulos, Ginatta et Villa analysent la convergence des trajectoires de (HBF) sous des conditions similaires sur  $F$  dans [8]. En considérant que  $F$  n'est qu'une fois différentiable, les auteurs démontrent une décroissance de l'erreur de l'ordre de  $\mathcal{O}\left(e^{-2\left(\sqrt{\frac{L}{\mu}}-\sqrt{\frac{L-\mu}{\mu}}\right)\sqrt{\mu}t}\right)$  le long des trajectoires. Le taux de décroissance démontré ne repose pas sur l'hypothèse que le minimiseur  $x^*$  est unique mais il n'est pas exponentiel rapide.

Ainsi, aucune analyse dans la littérature ne permet, à notre connaissance, d'obtenir un taux de décroissance exponentiel rapide de l'erreur le long de la trajectoire de (HBF) sans faire l'hypothèse de l'unicité du minimiseur.

Ces résultats de convergence sont résumés dans le Tableau 5.1.

| Référence                       | Hypothèse sur $F$                                    | Coefficient $K$ dans le taux exponentiel de $F(x(t)) - F^*$               |
|---------------------------------|--|---|
| Polyak [112]                    | $\mathcal{S}_\mu$ et $C^2$                           | $2\sqrt{\mu}$   |
| Siegel [126]                    | $\mathcal{S}_\mu$                                    | $\sqrt{\mu}$  |
| Aujol, Dossal, Rondepierre [24] | $\mathcal{Q}_\mu$<br>Unicité du minimiseur           | $\sqrt{2\mu}$   |
| Aujol, Dossal, Rondepierre [25] | $\mathcal{G}_\mu^2$<br>Unicité du minimiseur         | $(2 - \sqrt{2})\sqrt{\mu}$  |
| Bégout, Bolte, Jendoubi [36]    | $\mathcal{G}_\mu^2$ et $C^2$                         | $\mu$   |
| Polyak, Shcherbakov [114]       | Polyak-Łojasiewicz, $C^2$ et gradient $L$ -Lipschitz | $2\frac{\mu^{\frac{3}{2}}}{(\sqrt{2}+1)\mu+L}$                            |
| Apidopoulos, Ginatta, Villa [8] | Polyak-Łojasiewicz et gradient $L$ -Lipschitz        | $2\left(\sqrt{\frac{L}{\mu}} - \sqrt{\frac{L-\mu}{\mu}}\right)\sqrt{\mu}$ |

TABLEAU 5.1 – Taux de convergence obtenus pour  $F(x(t)) - F^*$  où  $x$  est solution de (HBF) pour le meilleur choix de  $\alpha$  (qui dépend de l'hypothèse faite sur  $F$ ). La constante  $K$  est définie telle que  $F(x(t)) - F^* = \mathcal{O}(e^{-Kt})$ .

Dans [127], Su, Boyd et Candès montrent le lien entre le système dynamique avec friction asymptotiquement évanescence (AVD) et le schéma de gradient accéléré de Nesterov [103] :

$$\forall t \geq t_0, \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{AVD})$$

où  $\alpha > 0$  et  $t_0 \geq 0$ .

Leur analyse porte principalement sur le cas où  $F$  est seulement convexe (on a alors  $F(x(t)) - F^* = \mathcal{O}(t^{-2})$  si  $\alpha \geq 3$ ) mais les auteurs démontrent également que si  $F$  est fortement convexe et  $\alpha \geq \frac{9}{2}$ , les trajectoires de (AVD) assurent une décroissance de l'erreur de l'ordre de  $\mathcal{O}(t^{-3})$ . Ce résultat est ensuite amélioré par Attouch, Chbani, Peypouquet et Redont dans [16] qui montrent que pour tout  $\alpha > 3$  et  $F$  fortement convexe, la décroissance de l'erreur est de l'ordre de  $\mathcal{O}(t^{-\frac{2\alpha}{3}})$ .

Dans [27, 26], Aujol, Dossal et Rondepierre analysent la convergence des trajectoires de (AVD) sous des hypothèses de croissance plus faibles que la forte convexité. Les auteurs démontrent que si  $F$  vérifie l'hypothèse de croissance quadratique  $\mathcal{G}_\mu^2$  et  $\mathcal{F}_\gamma$  pour  $\gamma \geq 1$ , alors la décroissance de l'erreur le long des trajectoires de (AVD) est de  $\mathcal{O}(t^{-\frac{2\alpha\gamma}{\gamma+2}})$  pour tout  $\alpha > 1 + \frac{2}{\gamma}$  à condition que  $F$  ait un unique minimiseur. Ce résultat généralise celui démontré dans [16]. Les auteurs de [27] démontrent également que si  $F$  est coercive et vérifie  $\mathcal{F}_{\gamma_1}$  et  $\mathcal{G}^{\gamma_2}$  pour  $\gamma_2 \geq \gamma_1 > 2$ , alors pour tout  $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$  la décroissance de l'erreur est de  $\mathcal{O}(t^{-\frac{2\gamma_2}{\gamma_2-2}})$ . On peut remarquer que l'unicité du minimiseur n'est pas nécessaire ici et est remplacée par une hypothèse de coercivité. Ce dernier résultat est complété par Luo et Xiao dans [91] en rem-

plaçant l'hypothèse de coercivité par une hypothèse d'unicité du minimiseur. Les auteurs montrent que pour  $\alpha \in \left(\frac{\gamma_1+2}{\gamma_1}, \frac{\gamma_1+2}{\gamma_1} \cdot \frac{\gamma_2}{\gamma_2-2}\right)$ , la décroissance de l'erreur est de l'ordre de  $\mathcal{O}\left(t^{-\frac{2\alpha\gamma_1}{\gamma_1+2}}\right)$ .

Ces résultats sont récapitulés dans le Tableau 5.2.

| Référence                                | Hypothèse sur $F$  | Taux de convergence de $F(x(t)) - F^*$   |
|--|--|--|
| Su, Boyd, Candès [127]                   | $\mathcal{S}_\mu$  | $\mathcal{O}(t^{-3}), \alpha \geq \frac{9}{2}$   |
| Attouch, Chbani, Peypouquet, Redont [16] | $\mathcal{S}_\mu$  | $\mathcal{O}\left(t^{-\frac{2\alpha}{3}}\right), \alpha > 3$   |
| Aujol, Dossal, Rondepierre [27, 26]      | $\mathcal{F}_\gamma$ et $\mathcal{G}_\mu^2$<br>Unicité du minimiseur   | $\mathcal{O}\left(t^{-\frac{2\alpha\gamma}{\gamma+2}}\right), \alpha > 1 + \frac{2}{\gamma}$   |
| Aujol, Dossal, Rondepierre [27]          | $\mathcal{F}_{\gamma_1}$ et $\mathcal{G}^{\gamma_2}$ avec<br>$\gamma_2 \geq \gamma_1 > 2$<br>$F$ coercive          | $\mathcal{O}\left(t^{-\frac{2\gamma_2}{\gamma_2-2}}\right), \alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$   |
| Luo, Xiao [91]                           | $\mathcal{F}_{\gamma_1}$ et $\mathcal{G}^{\gamma_2}$ avec<br>$\gamma_2 \geq \gamma_1 > 2$<br>Unicité du minimiseur | $\mathcal{O}\left(t^{-\frac{2\alpha\gamma_1}{\gamma_1+2}}\right),$<br>$\alpha \in \left(\frac{\gamma_1+2}{\gamma_1}, \frac{\gamma_1+2}{\gamma_1} \cdot \frac{\gamma_2}{\gamma_2-2}\right)$ |

TABLEAU 5.2 – Taux de convergence obtenus pour  $F(x(t)) - F^*$  où  $x$  est solution de (AVD).

On peut également remarquer que le système (DIN-AVD) a été étudié sous hypothèse de forte convexité par Attouch, Peypouquet et Redont dans [21] et sous hypothèses de croissance (Chapitre 4) et l'unicité du minimiseur de  $F$  est requise pour chacun des résultats.

Ces résultats soulèvent une question naturelle : peut-on s'affranchir de cette hypothèse d'unicité du minimiseur et obtenir des taux de convergence rapide ? On donne une approche permettant de répondre favorablement à cette question dans la Section 5.3.

### 5.2.2 Schémas numériques

Le premier schéma inertiel de premier ordre est la méthode de Heavy-Ball introduite par Polyak dans [112] dans le cadre de la minimisation de fonctions  $C^2$  fortement convexes. Le taux de convergence de l'erreur garanti par ce schéma est exponentiel rapide puisque  $F(x_k) - F^* = \mathcal{O}\left(e^{-4\sqrt{\kappa}k}\right)$  où  $\kappa = \frac{\mu}{L}$  est l'inverse du conditionnement de  $F$ ,  $\mu$  est son paramètre de croissance et  $L$  est la constante de Lipschitz de  $\nabla F$ .

Nesterov introduit dans [105, 106] un schéma de type Heavy-Ball qui est adapté aux fonctions différentiables fortement convexes et qui assure une décroissance exponentielle rapide de l'erreur sans l'hypothèse  $C^2$ . Le taux de convergence démontré

est de l'ordre de  $\mathcal{O}\left(e^{-\sqrt{\kappa}k}\right)$ . Necoara, Nesterov et Glineur montrent dans [102] que ce taux est valable pour des hypothèses plus faibles sur  $F$ . En effet, si  $F$  est quasi-fortement convexe et si la projection des itérés du schéma sur  $X^*$  est unique, alors la vitesse de convergence prouvée dans le cadre fortement convexe est toujours valide. Cela indique que l'hypothèse d'unicité du minimiseur n'est pas nécessaire ici pour avoir un taux exponentiel rapide mais l'hypothèse de projection unique sur  $X^*$  reste restrictive.

Plus récemment, plusieurs schémas ont été proposés pour optimiser des fonctions fortement convexes. Bubeck, Lee et Singh [45] introduisent une méthode de descente géométrique pour les fonctions vérifiant  $\mathcal{S}_\mu$  qui assure une décroissance de l'erreur de  $\mathcal{O}\left(e^{-\sqrt{\kappa}k}\right)$ . Le schéma est accompagné d'une interprétation géométrique et est défini pour les fonctions différentiables. Chen, Ma et Liu [52] proposent une extension de cette méthode pour les fonctions composites et démontrent la même vitesse de convergence. Dans [58], Drusvyatskiy, Fazel et Roy introduisent une méthode basée sur le moyennage quadratique des itérés qui a le même comportement que la méthode de descente géométrique définie dans [45]. Van Scoy, Freeman et Lynch proposent dans [133] la méthode du triple élan (*Triple momentum method*) pour l'optimisation de fonctions différentiables fortement convexes. Ce schéma atteint la meilleure vitesse pour cette classe de fonctions dans la littérature avec une décroissance de l'erreur de  $\mathcal{O}\left(e^{-2\sqrt{\kappa}k}\right)$ . Cette vitesse est également atteinte par ITEM (*Information-Theoretic Exact Method*) de Taylor et Drori [128] sous les mêmes hypothèses. Ce dernier schéma a la particularité d'atteindre le taux  $\mathcal{O}(k^{-2})$  si la fonction  $F$  ne vérifie qu'une hypothèse de convexité.

Ghadimi, Feysmahdavian et Johansson démontrent dans [65] que la méthode Heavy-Ball de Polyak pour un terme de friction quelconque  $\alpha \in [0, 1)$  et une certaine condition sur le pas converge linéairement lorsque  $F$  est fortement convexe et de classe  $C^1$ . En retirant l'hypothèse  $C^2$ , le taux de convergence obtenu est exponentiel lent  $\mathcal{O}\left(e^{-\kappa k}\right)$ .

Les méthodes présentées ci-dessus sont dédiées au cas où  $F$  est différentiable. Plusieurs schémas ont également été proposés pour la minimisation de fonctions composites vérifiant des hypothèses de croissance.

Dans [126], Siegel introduit un schéma dérivé du système de Heavy-Ball (HBF) qui assure une vitesse de décroissance de  $\mathcal{O}\left(e^{-\sqrt{\kappa}k}\right)$  pour les fonctions fortement convexes. Cette vitesse est aussi atteinte sous les mêmes hypothèses par V-FISTA introduit par Beck [31] qui est une version proximale de la méthode de gradient accéléré de Nesterov en fortement convexe. Ochs, Brox et Pock proposent iPiasco dans [109] qui est un schéma adapté aux fonctions composites de la forme  $F = f + h$  où  $f$  ou  $h$  fortement convexe. Les auteurs démontrent la convergence linéaire de leur algorithme et plus précisément de la quantité  $\|\langle x_k - x^*, x_{k+1} - x^* \rangle\|$ . La vitesse de décroissance de cette quantité est exponentielle rapide.

Dans [24], Aujol, Dossal et Rondepierre dérivent un schéma à partir du système (HBF) pour l'optimisation de fonctions différentiables fortement convexes qui est



adaptable pour les fonctions composites. Pour cette classe de fonctions, les auteurs prouvent que ce schéma assure une décroissance exponentielle rapide de l'erreur  $\mathcal{O}\left(e^{(-\sqrt{2\kappa}+\mathcal{O}(\kappa))k}\right)$ . Un schéma adapté aux fonctions composites à croissance quadratique est proposé par les mêmes auteurs dans [25]. Ce schéma, également dérivé de (HBF), garantit un taux de convergence exponentiel rapide  $\mathcal{O}\left(e^{-(2-\sqrt{2})\sqrt{\kappa}+\mathcal{O}(\kappa))k}\right)$  à condition que  $F$  ait un unique minimiseur. Si cette hypothèse n'est pas vérifiée, la vitesse obtenue devient exponentielle lente  $\mathcal{O}\left(e^{(-\kappa+\varepsilon+\mathcal{O}(\kappa))k}\right)$ .

Ainsi, l'hypothèse d'unicité du minimiseur semble nécessaire pour obtenir une décroissance exponentielle rapide de l'erreur si  $F$  vérifie une hypothèse de croissance, à moins de faire une hypothèse d'unicité de la projection sur  $X^*$  qui peut être restrictive.

Les résultats énoncés ci-dessus sont résumés dans le Tableau 5.3.

Le schéma de gradient accéléré de Nesterov [103] et ses extensions aux fonctions composites parmi lesquelles FISTA [32] de Beck et Teboulle sont des méthodes particulièrement pertinentes pour l'optimisation de fonctions convexes. Pour des hypothèses de géométrie plus fortes que la convexité, il s'avère que les méthodes type Heavy-Ball (ou les autres méthodes présentées au-dessus) sont plus rapides à condition de bien paramétrer le paramètre de friction.

En effet, si  $F$  est fortement convexe, Su, Boyd et Candès démontrent dans [127] que FISTA pour  $\alpha \geq \frac{9}{2}$  assure une vitesse de convergence polynomiale  $\mathcal{O}(k^{-3})$  qui est donc moins bonne que les vitesses de décroissance exponentielles rapides garanties par les schémas présentés précédemment. Rappelons que l'on appelle FISTA le schéma ci-dessous :

$$\forall k \in \mathbb{N}, \begin{cases} y_k = x_k + \frac{k}{k+\alpha}(x_k - x_{k-1}), \\ x_{k+1} = \text{prox}_{sh}(y_k - s\nabla f(y_k)) \end{cases} \quad (\text{FISTA})$$

où  $y_0 = x_0$ ,  $s > 0$  et  $\alpha > 0$ . Ce taux de convergence est amélioré par Attouch et Cabot [12] sous des hypothèses plus faibles. Les auteurs montrent que si  $F$  vérifie  $\mathcal{G}_\mu^2$  et a un unique minimiseur, alors pour tout  $\alpha > 0$  on a  $F(x_k) - F^* = \mathcal{O}\left(k^{-\frac{2\alpha}{3}}\right)$ . Dans [7], Apidopoulos, Aujol, Dossal et Rondepierre complètent ce résultat en montrant que si  $F$  vérifie en plus  $\mathcal{F}_\gamma$  pour  $\gamma \geq 1$ , alors pour tout  $\alpha \geq 1 + \frac{2}{\gamma}$  la vitesse de convergence de l'erreur est de  $\mathcal{O}\left(k^{-\frac{2\alpha\gamma}{\gamma+2}}\right)$ . Aujol, Dossal et Rondepierre proposent une analyse en temps fini de FISTA pour ces hypothèses dans [26] et montrent que des vitesses comparables à celles d'exponentielles rapides peuvent être obtenues à précision fixée pour un bon choix de  $\alpha$ . Les auteurs de [7] proposent également une analyse de convergence de FISTA pour des géométries plates. Si  $F$  vérifie  $\mathcal{F}_{\gamma_1}$  et  $\mathcal{G}^{\gamma_2}$  avec  $\gamma_2 \geq \gamma_1 > 2$  et a un unique minimiseur, il est montré que pour  $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$  on a  $F(x_k) - F^* = \mathcal{O}\left(k^{-\frac{2\gamma_2}{\gamma_2-2}}\right)$ .

Comme pour les schémas présentés auparavant, les résultats de convergence améliorés pour FISTA sous condition de croissance reposent sur une hypothèse d'unicité du minimiseur.

| Algorithme   | Référence   | Hypothèse sur $F$   | Taux de convergence de $F(x_k) - F^*$   |
|--|---|---|---|
| Heavy-Ball de Polyak   | Polyak [112]  | $\mathcal{S}_\mu$ et $C^2$  | $\mathcal{O}\left(e^{-4\sqrt{\kappa}k}\right)$                                    |
| Méthode de gradient accéléré en fortement convexe (schéma type Heavy-Ball) | Nesterov[105, 106]<br>Necoara, Nesterov, Glineur[102] | $\mathcal{S}_\mu$ et $C^1$<br>$\mathcal{Q}_\mu$ , $C^1$ et projection unique sur $X^*$ des itérés | $\mathcal{O}\left(e^{-\sqrt{\kappa}k}\right)$                                     |
| Méthode de descente géométrique  | Bubeck, Lee, Singh [45]<br>Chen, Ma, Liu [52]         | $\mathcal{S}_\mu$<br>Adapté au cas composite  | $\mathcal{O}\left(e^{-\sqrt{\kappa}k}\right)$                                     |
| Méthode du triple élan ( <i>triple momentum</i> )                          | Van Scoy, Freeman, Lynch [133]                        | $\mathcal{S}_\mu$ et $C^1$  | $\mathcal{O}\left(e^{-2\sqrt{\kappa}k}\right)$                                    |
| ITEM   | Taylor, Drori [128]                                   | $\mathcal{S}_\mu$ et $C^1$  | $\mathcal{O}\left(e^{-2\sqrt{\kappa}k}\right)$                                    |
| Méthode type Heavy-Ball avec friction quelconque                           | Ghadimi, Feysmahdavian, Johansson [65]                | $\mathcal{S}_\mu$ et $C^1$  | $\mathcal{O}\left(e^{-\kappa k}\right)$   |
| Méthode type Heavy-Ball  | Siegel [126]  | $\mathcal{S}_\mu$ et $C^1$<br>Adapté au cas composite   | $\mathcal{O}\left(e^{-\sqrt{\kappa}k}\right)$                                     |
| V-FISTA (Méthode type Heavy-Ball)  | Beck [31]   | $\mathcal{S}_\mu$<br>Adapté au cas composite  | $\mathcal{O}\left(e^{-\sqrt{\kappa}k}\right)$                                     |
| Méthode type Heavy-Ball  | Aujol, Dossal, Rondepierre [24]                       | $\mathcal{S}_\mu$<br>Adapté au cas composite  | $\mathcal{O}\left(e^{(-\sqrt{2\kappa} + \mathcal{O}(\kappa))k}\right)$            |
| Méthode type Heavy-Ball  | Aujol, Dossal, Rondepierre [25]                       | $\mathcal{G}_\mu^2$ et unicité du minimiseur<br>Adapté au cas composite                           | $\mathcal{O}\left(e^{(-(2-\sqrt{2})\sqrt{\kappa} + \mathcal{O}(\kappa))k}\right)$ |
| Méthode type Heavy-Ball  | Aujol, Dossal, Rondepierre [25]                       | $\mathcal{G}_\mu^2$<br>Adapté au cas composite  | $\mathcal{O}\left(e^{(-\kappa + \varepsilon + \mathcal{O}(\kappa))k}\right)$      |

TABLEAU 5.3 – Taux de convergence obtenus pour  $F(x_k) - F^*$  pour différents schémas de premier ordre sous condition de croissance.

Les résultats présentés ci-dessus sont regroupés dans le Tableau 5.4.

On peut noter qu'un moyen d'obtenir des taux de convergence exponentiels rapides sans hypothèse d'unicité du minimiseur est de redémarrer FISTA (voir Chapitre 3). Cependant, on cherche ici à déterminer si les schémas classiques souffrent de la non-unicité des minimiseurs sans stratégie additionnelle.

| Référence  | Hypothèse sur $F$   | Gamme de paramètres                         | Taux de convergence de $F(x_k) - F^*$                         |
|--|---|---|---|
| Su, Boyd, Candès [127]   | $\mathcal{S}_\mu$   | $\alpha \geq \frac{9}{2}$                   | $\mathcal{O}(k^{-3})$   |
| Attouch, Cabot [12]  | $\mathcal{G}_\mu^2$ et unicité du minimiseur  | $\alpha > 0$                                | $\mathcal{O}\left(k^{-\frac{2\alpha}{3}}\right)$              |
| Apidopoulos, Aujol, Dossal, Rondepierre [7]<br>Aujol, Dossal, Rondepierre [26] | $\mathcal{F}_\gamma$ et $\mathcal{G}_\mu^2$ , $\gamma \geq 1$<br>Unicité du minimiseur                          | $\alpha \geq 1 + \frac{2}{\gamma}$          | $\mathcal{O}\left(k^{-\frac{2\alpha\gamma}{\gamma+2}}\right)$ |
| Apidopoulos, Aujol, Dossal, Rondepierre [7]                                    | $\mathcal{F}_{\gamma_1}$ et $\mathcal{G}^{\gamma_2}$ avec $\gamma_2 \geq \gamma_1 > 2$<br>Unicité du minimiseur | $\alpha \geq \frac{\gamma_1+2}{\gamma_1-2}$ | $\mathcal{O}\left(k^{-\frac{2\gamma_2}{\gamma_2-2}}\right)$   |

TABLEAU 5.4 – Taux de convergence obtenus pour  $F(x_k) - F^*$  pour FISTA sous condition de croissance.

### 5.3 Non unicité dans le cas continu

Dans cette section, je propose un cadre permettant d'étendre des résultats de convergence connus pour les trajectoires de (HBF) et (AVD) sous hypothèse de croissance sans avoir besoin d'hypothèse d'unicité du minimiseur. Cette approche s'appuie sur une hypothèse de géométrie sur l'ensemble de minimiseurs  $X^*$ . Notons que comme  $F$  est supposée convexe et continue, l'ensemble  $X^*$  est convexe et fermé.

#### 5.3.1 Stratégie

Les deux systèmes dynamiques auxquels on s'intéresse sont le système de Heavy-Ball avec friction (HBF) :

$$\forall t \geq t_0, \ddot{x}(t) + \alpha \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{HBF})$$

avec  $\alpha > 0$ ,  $t_0 \geq 0$  et  $(x(t_0), \dot{x}(t_0)) = (x_0, \dot{x}_0)$ , et le système dynamique avec friction asymptotiquement évanescence (AVD) :

$$\forall t \geq t_0, \ddot{x}(t) + \frac{\alpha}{t} \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{AVD})$$

avec  $\alpha > 0$ ,  $t_0 \geq 0$  et  $(x(t_0), \dot{x}(t_0)) = (x_0, \dot{x}_0)$ .

Une approche classique pour analyser la convergence des trajectoires de ces systèmes est d'utiliser une analyse de Lyapunov (voir 2.3.3). En étudiant une énergie de Lyapunov bien choisie, il est possible de déduire des propriétés de convergence le

long des trajectoires solutions en exploitant de potentielles hypothèses de géométrie sur  $F$ .

Dans la littérature, les énergies de Lyapunov permettant de déterminer des vitesses de convergence rapides ont une forme similaire en fonction du système dynamique considéré. Pour analyser la convergence des trajectoires de (HBF) sous hypothèses de croissance, Siegel [126] et Aujol, Dossal et Rondepierre [24, 25] s'appuient sur l'énergie définie ci-dessous :

$$\mathcal{E}_{\text{HBF}}(t) = F(x(t)) - F^* + \frac{1}{2}\|\lambda(x(t) - x^*) + \dot{x}(t)\|^2 + \frac{\xi}{2}\|x(t) - x^*\|^2, \quad (5.3.1)$$

avec différents choix de  $\lambda > 0$  et  $\xi \in \mathbb{R}$ .

D'autre part, le système (AVD) est étudié par Su, Boyd et Candès [127], Attouch, Chbani, Peypouquet et Redont [16] et Aujol, Dossal et Rondepierre [27, 26] à partir de l'énergie de Lyapunov ci-dessous :

$$\forall t \geq t_0, \mathcal{E}_{\text{AVD}}(t) = t^2 (F(x(t)) - F^*) + \frac{1}{2}\|\lambda(x(t) - x^*) + t\dot{x}(t)\|^2, \quad (5.3.2)$$

avec différentes valeurs de  $\lambda > 0$ . Ces deux énergies dépendent d'un minimiseur  $x^*$  de  $F$ , fixe en fonction du temps.

L'hypothèse d'unicité du minimiseur n'est pas nécessaire pour obtenir les résultats démontrés par Su, Boyd et Candès [127] et Attouch, Chbani, Peypouquet et Redont [16] dans le cas convexe pour (AVD). En effet, il est montré que pour tout  $x^* \in X^*$ ,  $\mathcal{E}_{\text{AVD}}$  est décroissante ce qui suffit à déduire une borne sur  $F(x(t)) - F^*$ .

En revanche, si on considère que  $F$  vérifie une hypothèse de croissance supplémentaire comme  $\mathcal{G}_\mu^2$  (croissance quadratique) ou  $\mathcal{Q}_\mu$  (quasi-forte convexité), le fait que ces énergies dépendent d'un  $x^* \in X^*$  fixe est limitant pour déterminer des vitesses améliorées. En effet, si  $F$  vérifie  $\mathcal{G}_\mu^2$  pour  $\mu > 0$ , alors

$$\forall x \in \mathbb{R}^N, \frac{\mu}{2}d(x, X^*)^2 \leq F(x) - F^*, \quad (5.3.3)$$

ce qui signifie que l'on ne peut écrire l'inégalité  $\|x(t) - x^*\|^2 \leq \frac{2}{\mu}(F(x(t)) - F^*)$  qu'à condition que  $x^*$  soit l'unique projection de  $x(t)$  sur  $X^*$  pour tout  $t \geq t_0$ . De même, l'hypothèse de quasi-forte convexité assure l'inégalité suivante pour tout  $x \in \mathbb{R}^N$  :

$$F^* \geq F(x) + \langle \nabla F(x), x^* - x \rangle + \frac{\mu}{2}\|x^* - x\|^2, \quad (5.3.4)$$

où  $x^*$  est la projection de  $x$  sur  $X^*$ . Ces inégalités ne peuvent donc pas être exploitées à moins de faire l'hypothèse que l'ensemble de minimiseurs de  $F$  est réduit à un point (ou que la projection de  $x(t)$  sur  $X^*$  soit constante en fonction de  $t$ ).

C'est à partir de ce constat que je propose de légèrement modifier ces énergies de Lyapunov, le but étant d'exploiter ces inégalités liées aux hypothèses de croissance sans avoir à faire d'hypothèse d'unicité du minimiseur. Les nouvelles énergies de Lyapunov sont définies comme suit pour tout  $t \geq t_0$  :

$$\mathcal{E}_{\text{HBF}}^*(t) = F(x(t)) - F^* + \frac{1}{2}\|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2 + \frac{\xi}{2}\|x(t) - x^*(t)\|^2, \quad (5.3.5)$$

$$\mathcal{E}_{\text{AVD}}^*(t) = t^2 (F(x(t)) - F^*) + \frac{1}{2}\|\lambda(x(t) - x^*(t)) + t\dot{x}(t)\|^2, \quad (5.3.6)$$

où  $\lambda > 0$ ,  $\xi \in \mathbb{R}$  et  $x^*(t)$  est défini pour tout  $t \geq t_0$  comme la projection de  $x(t)$  sur  $X^*$ . Rappelons que l'on considère que  $F$  est continue et convexe donc  $X^*$  est un ensemble convexe fermé et la projection sur  $X^*$  est bien définie. Ainsi, on a pour tout  $t \geq t_0$  que

$$x^*(t) = \arg \min_{x^* \in X^*} \|x(t) - x^*\|^2, \quad (5.3.7)$$

et notamment

$$\|x(t) - x^*(t)\| = d(x(t), X^*). \quad (5.3.8)$$

Cette légère modification conduit à une question importante pour poursuivre une analyse de Lyapunov classique et dériver ces énergies par rapport à  $t$  : **à quelle condition la fonction  $t \mapsto x^*(t)$  est-elle dérivable ?**

On note  $P_{X^*}$  l'opérateur de projection sur  $X^*$ . On a donc que pour tout  $t \geq t_0$ ,  $x^*(t) = P_{X^*}(x(t))$ . La régularité de  $t \mapsto x^*(t)$  est donc étroitement liée à celle de l'opérateur  $P_{X^*}$  et en particulier, si  $P_{X^*}$  est directionnellement dérivable, alors  $t \mapsto x^*(t)$  est dérivable à droite (et à gauche). Quand elle est définie, sa dérivée à droite est égale à  $P'_{X^*}(x(t), \dot{x}(t))$  où  $P'_{X^*}(x, d)$  désigne la dérivée de  $P_{X^*}$  au point  $x \in \mathbb{R}^N$  dans la direction  $d \in \mathbb{R}^N$ .

La différentiabilité de l'opérateur de projection sur un ensemble convexe a été notamment étudiée par Holmes [74], Zarantonello [136], Fitzpatrick et Phelps [63], Hiriart-Urruty [73] ou encore Noll [107]. Ces travaux mettent en évidence le lien entre la régularité de l'opérateur de projection sur un ensemble convexe  $K \subset \mathbb{R}^N$  et la régularité de la frontière de  $K$  que l'on note  $\partial K$ . En particulier, il est montré que  $P_{X^*}$  est directionnellement dérivable si  $\partial K$  est  $C^2$ . Cette condition est vérifiée si pour un certain  $x_0 \in \text{int}(K)$ , la fonctionnelle de Minkowski pour  $K$  par rapport à  $x_0$  définie par

$$\forall x \in \mathbb{R}^N, \mu_{K, x_0}(x) := \inf\{t > 0 : x \in t(K - x_0) + x_0\}, \quad (5.3.9)$$

est deux fois dérivable au voisinage de  $\partial K$ .

Dans [42], Bonnans, Cominetti et Shapiro donnent une expression de la dérivée directionnelle de  $P_K$  quand  $K$  est un ensemble convexe régulier du second ordre. Cette hypothèse de géométrie est technique et détaillée par Shapiro dans [124]. On peut noter que tout ensemble convexe polyédral est régulier du second ordre.

Supposons maintenant que l'ensemble de minimiseurs  $X^*$  est régulier du second ordre. On a alors que  $t \mapsto x^*(t)$  est dérivable à droite et les énergies de Lyapunov  $\mathcal{E}_{\text{HBF}}^*$  et  $\mathcal{E}_{\text{AVD}}$  également. Pour simplifier les notations, les dérivées à droite respectives de ces trois fonctions seront désignées par  $\dot{x}^*$ ,  $\dot{\mathcal{E}}_{\text{HBF}}^*$  et  $\dot{\mathcal{E}}_{\text{AVD}}^*$ . On obtient alors pour tout  $t \geq t_0$  :

$$\dot{\mathcal{E}}_{\text{HBF}}^*(t) = D_{\text{HBF}}(t) - (\lambda^2 + \xi) \langle x(t) - x^*(t), \dot{x}^*(t) \rangle - \lambda \langle \dot{x}(t), \dot{x}^*(t) \rangle, \quad (5.3.10)$$

$$\dot{\mathcal{E}}_{\text{AVD}}^*(t) = D_{\text{AVD}}(t) - \lambda^2 \langle x(t) - x^*(t), \dot{x}^*(t) \rangle - \lambda t \langle \dot{x}(t), \dot{x}^*(t) \rangle, \quad (5.3.11)$$

où

$$D_{\text{HBF}}(t) = \langle \nabla F(x(t)), \dot{x}(t) \rangle + \langle \lambda(x(t) - x^*(t)) + \dot{x}(t), \lambda \dot{x}(t) + \ddot{x}(t) \rangle + \xi \langle x(t) - x^*(t), \dot{x}(t) \rangle,$$

et

$$D_{\text{AVD}}(t) = 2t \langle F(x(t)) - F^* \rangle + t^2 \langle \nabla F(x(t)), \dot{x}(t) \rangle \\ + \langle \lambda(x(t) - x^*(t)) + t\dot{x}(t), (\lambda + 1)\dot{x}(t) + t\ddot{x}(t) \rangle.$$

Les fonctions  $D_{\text{HBF}}$  et  $D_{\text{AVD}}$  ainsi définies correspondent exactement à  $\dot{\mathcal{E}}_{\text{HBF}}$  et  $\dot{\mathcal{E}}_{\text{AVD}}$  où le  $x^*$  fixe est remplacé par  $x^*(t)$  la projection de  $x(t)$  sur  $X^*$ . Cela signifie que les contrôles obtenus sur  $\dot{\mathcal{E}}_{\text{HBF}}$  et  $\dot{\mathcal{E}}_{\text{AVD}}$  par les auteurs de [126, 25, 24, 27, 26] sous les hypothèses  $\mathcal{Q}_\mu$  ou  $\mathcal{G}_\mu^2$  avec l'hypothèse d'unicité du minimiseur sont ici valables sur  $D_{\text{HBF}}$  et  $D_{\text{AVD}}$  mais sans hypothèse d'unicité.

Pour exploiter cette propriété, on veut d'abord contrôler les termes additionnels  $\langle x(t) - x^*(t), \dot{x}^*(t) \rangle$  et  $\langle \dot{x}(t), \dot{x}^*(t) \rangle$ . Pour avoir une intuition sur le comportement de ces termes, on peut se référer à la Figure 5.1.

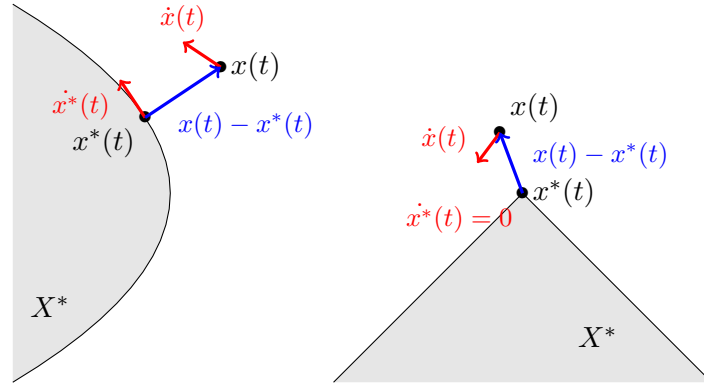


FIGURE 5.1 – Comportement de  $\dot{x}^*$  pour un ensemble de minimiseurs à frontière  $C^2$  (à gauche) et un ensemble de minimiseurs polyédral (à droite).

Dans un premier temps, on peut montrer que  $\langle \dot{x}(t), \dot{x}^*(t) \rangle$  est positif en utilisant l'expression  $\dot{x}^*(t) = \lim_{h \rightarrow 0, h > 0} \frac{x^*(t+h) - x^*(t)}{h}$  et les propriétés de la projection sur un ensemble convexe. En effet, comme  $X^*$  est un ensemble convexe fermé, l'inégalité ci-dessous est vérifiée pour tout  $x \in \mathbb{R}^N$  et  $u \in X^*$  :

$$\langle x - P_{X^*}(x), u - P_{X^*}(x) \rangle \leq 0.$$

Ainsi, pour tout  $t \geq t_0$  et  $h > 0$  on a

$$\langle x(t+h) - x(t), x^*(t+h) - x^*(t) \rangle = \langle x(t+h) - x^*(t+h), x^*(t+h) - x^*(t) \rangle \\ + \|x^*(t+h) - x^*(t)\|^2 \\ + \langle x(t) - x^*(t), x^*(t) - x^*(t+h) \rangle \\ \geq 0.$$

En faisant tendre  $h$  vers 0, on trouve que  $\langle \dot{x}(t), \dot{x}^*(t) \rangle \geq 0$ .

D'autre part, si on se réfère à l'expression de la dérivée directionnelle de la projection sur un ensemble régulier du second ordre donnée dans [42, Théorème 7.2], on a que pour tout  $t \geq t_0$ ,  $\langle \dot{x}(t), x(t) - x^*(t) \rangle = 0$ .

Par conséquent, on trouve que pour tout  $t \geq t_0$  et pour tout jeu de paramètre  $(\lambda, \xi)$ ,  $\dot{\mathcal{E}}_{\text{HBF}}^*(t) \leq D_{\text{HBF}}(t)$  et  $\dot{\mathcal{E}}_{\text{AVD}}^*(t) \leq D_{\text{AVD}}(t)$ . Les analyses proposées dans [126, 25, 24, 27, 26] peuvent donc être étendues à ce nouveau cadre : les inégalités différentielles vérifiées par  $\dot{\mathcal{E}}_{\text{HBF}}$  et  $\dot{\mathcal{E}}_{\text{AVD}}$  sont aussi vérifiées par  $\dot{\mathcal{E}}_{\text{HBF}}^*$  et  $\dot{\mathcal{E}}_{\text{AVD}}^*$ . Comme  $\dot{\mathcal{E}}_{\text{HBF}}^*$  et  $\dot{\mathcal{E}}_{\text{AVD}}^*$  sont des dérivées à droite, on a besoin du lemme ci-dessous pour obtenir les résultats de convergence désirés. La preuve est donnée en Section 5.A.1.

**Lemme 5.3.1.** *Soit  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  une fonction continue dérivable à droite et à gauche. Supposons que*

$$\forall t \geq t_0, \phi_+(t) \leq \psi(t), \quad (5.3.12)$$

où  $\phi_+(t) = \lim_{h \rightarrow 0, h > 0} \frac{\phi(t+h) - \phi(t)}{h}$  désigne la dérivée à droite de  $\phi$  pour un certain  $t$ .

Alors,

$$\forall t \geq t_0, \phi(t) \leq \phi(t_0) + \int_{t_0}^t \psi(u) du. \quad (5.3.13)$$

Les résultats de convergence que l'on peut généraliser sans unicité du minimiseur avec l'hypothèse de régularité de second ordre sur  $X^*$  sont présentés dans la section suivante.

## 5.3.2 Résultats de convergence sans unicité du minimiseur

### 5.3.2.1 Propriétés de convergence des trajectoires de (HBF)

On énonce ci-dessous deux théorèmes portant sur la convergence des trajectoires de (HBF) lorsque  $F$  vérifie  $\mathcal{G}_\mu^2$  ou  $\mathcal{Q}_\mu$  pour  $\mu > 0$ .

**Théorème 5.3.2.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe différentiable ayant un ensemble de minimiseurs  $X^*$  non vide. Supposons que  $F$  vérifie  $\mathcal{G}_\mu^2$  pour  $\mu > 0$  et que  $X^*$  est un ensemble régulier du second ordre. Alors, la solution  $x$  de (HBF) pour tout  $t_0 \geq 0$  et  $\alpha = \left(2 - \frac{\sqrt{2}}{2}\right) \sqrt{\mu}$  vérifie :*

$$\forall t \geq t_0, F(x(t)) - F^* \leq \left(\frac{11}{2} - 2\sqrt{2}\right) M_0 e^{-(2-\sqrt{2})\sqrt{\mu}(t-t_0)}, \quad (5.3.14)$$

où  $M_0 = F(x(t_0)) - F^* + \frac{1}{2} \|\dot{x}(t_0)\|^2$ . De plus,

$$\|\dot{x}(t)\| = \mathcal{O}\left(e^{-\left(1-\frac{\sqrt{2}}{2}\right)\sqrt{\mu}t}\right). \quad (5.3.15)$$

Le premier point est obtenu à partir des calculs de Aujol, Dossal et Rondepierre dans [25, Théorème 1]. En appliquant la démarche présentée dans la section précédente, on peut déduire les mêmes vitesses de convergence en remplaçant l'hypothèse d'unicité du minimiseur par une condition de géométrie sur  $X^*$ . On donne ici le taux de convergence pour le meilleur choix de  $\alpha$  mais le résultat général de [25,

Théorème 1] est également valable. Le deuxième point assure que les trajectoires de (HBF) sont finies. Une preuve pour ce deuxième point est donnée en Section 5.A.2.

Le théorème ci-dessous concerne le cas où  $F$  est  $\mu$ -quasi fortement convexe.

**Théorème 5.3.3.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe différentiable ayant un ensemble de minimiseurs  $X^*$  non vide. Supposons que  $F$  vérifie  $\mathcal{Q}_\mu$  pour  $\mu > 0$  et que  $X^*$  est un ensemble régulier du second ordre. Alors, la solution  $x$  de (HBF) pour tout  $t_0 \geq 0$  et  $\alpha = \frac{3}{\sqrt{2}}\sqrt{\mu}$  vérifie :*

$$\forall t \geq t_0, F(x(t)) - F^* \leq 39M_0 e^{-\sqrt{2\mu}(t-t_0)}, \quad (5.3.16)$$

où  $M_0 = F(x(t_0)) - F^* + \frac{1}{2}\|\dot{x}(t_0)\|^2$ . De plus,

$$\|\dot{x}(t)\| = \mathcal{O}\left(e^{-\frac{\sqrt{2}}{2}\sqrt{\mu}t}\right). \quad (5.3.17)$$

La preuve de ce théorème s'articule autour de la démonstration de [24, Théorème 3.2]. Le premier point est directement obtenu grâce à l'hypothèse sur  $X^*$  et la démarche expliquée en Section 5.3.1 et le deuxième point s'appuie sur quelques calculs supplémentaires précisés en Section 5.A.3.

Dans ces deux théorèmes, les vitesses de décroissance de l'erreur le long de la trajectoire de (HBF) sont exponentielles rapides sans hypothèse d'unicité du minimiseur. Il s'agit des premiers résultats de ce type à ma connaissance. Ces théorèmes demandent cependant une hypothèse de géométrie sur  $X^*$  liée à la dérivabilité de l'opérateur de projection sur cet ensemble.

### 5.3.2.2 Propriétés de convergence des trajectoires de (AVD)

On s'intéresse ici à la convergence des trajectoires de (AVD) sous hypothèses de croissance sur  $F$ . Le théorème ci-dessous concerne les fonctions à croissance quadratique et est une extension de [26, Théorème 5].

**Théorème 5.3.4.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe différentiable ayant un ensemble de minimiseurs  $X^*$  non vide. Supposons que  $F$  vérifie  $\mathcal{G}_\mu^2$  pour  $\mu > 0$ , et que  $X^*$  est un ensemble régulier du second ordre. Alors, pour  $\mu$  suffisamment petit, la solution  $x$  de (AVD) pour tout  $t_0 \geq 0$  et  $\alpha > 3$  vérifie :*

$$\forall t \geq \frac{\alpha r^*}{3\sqrt{\mu}} \geq t_0, F(x(t)) - F^* \leq C_1 e^{\frac{2}{3}C_2(\alpha-3)} M_0 \left(\frac{\alpha r^*}{3t\sqrt{\mu}}\right)^{\frac{2\alpha}{3}}, \quad (5.3.18)$$

où  $M_0 = F(x(t_0)) - F^* + \frac{1}{2}\|\dot{x}(t_0)\|^2$ ,  $r^* \simeq 3$  est l'unique racine positive du polynôme  $r \mapsto r^3 - r^2 - 2(1 + \sqrt{2})r - 4$  et

$$C_1 = 1 + \frac{2}{r^*} + \frac{4}{r^{*2}}, \quad C_2 = \frac{1}{r^*} + \frac{1 + \sqrt{2}}{r^{*2}} + \frac{4}{3r^{*3}}.$$

De plus, on a :

$$\|\dot{x}(t)\| = \mathcal{O}\left(t^{-\frac{\alpha}{3}}\right). \quad (5.3.19)$$



Ce théorème s'applique aux fonctions convexes vérifiant  $\mathcal{G}_\mu^2$  mais le résultat démontré dans [26, Théorème 5] pour les fonctions vérifiant  $\mathcal{F}_\gamma$  et  $\mathcal{G}_\mu^2$  est aussi valable. On écrit ici le cas particulier  $\gamma = 1$  pour clarifier l'écriture.

Le premier point du Théorème 5.3.4 est obtenu grâce à l'hypothèse de géométrie faite sur  $X^*$  et la stratégie expliquée en Section 5.3.1. Une preuve du deuxième résultat est donnée en Section 5.A.4.

Le théorème suivant n'a pas d'équivalent exact dans la littérature contrairement aux trois précédents. On fait l'hypothèse que  $F$  satisfait l'hypothèse de croissance  $\mathcal{G}^\gamma$  pour  $\gamma > 2$  qui est une hypothèse de croissance plus faible que  $\mathcal{G}_\mu^2$ . Aujol, Dossal et Rondepierre prouvent dans [27] que si  $F$  vérifie en plus l'hypothèse  $\mathcal{F}_r$  où  $r \in (2, \gamma]$  et est coercive, alors pour tout  $\alpha \geq \frac{r+2}{r-2}$  les trajectoires de (AVD) vérifient :

$$F(x(t)) - F^* = \mathcal{O}\left(t^{-\frac{2\gamma}{\gamma-2}}\right). \quad (5.3.20)$$

Ce résultat est étendu dans [123] dans le cas perturbé. Notons que ce résultat ne nécessite pas que  $F$  ait un unique minimiseur.

Le théorème ci-dessous énonce un taux de convergence similaire sans nécessiter l'hypothèse  $\mathcal{F}_r$  avec  $r > 2$ . L'hypothèse de coercivité est remplacée par une hypothèse sur la géométrie de  $X^*$ .

**Théorème 5.3.5.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction convexe différentiable ayant un ensemble de minimiseurs  $X^*$  non vide. Supposons que  $F$  vérifie  $\mathcal{G}^\gamma$  pour  $\gamma > 2$ , et que  $X^*$  est un ensemble régulier du second ordre. Alors, la solution  $x$  de (AVD) pour tout  $\alpha > \frac{9}{2} + \frac{6}{\gamma-2}$  satisfait*

$$F(x(t)) - F^* = \mathcal{O}\left(\frac{1}{t^{\frac{2\gamma}{\gamma-2}}}\right), \quad (5.3.21)$$

et

$$\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^{\frac{\gamma}{\gamma-2}}}\right). \quad (5.3.22)$$

La preuve de ce théorème est donnée en Section 5.A.5.

On peut noter que le deuxième point de ce théorème assure que la solution  $x$  de (AVD) est de longueur finie et qu'elle converge donc vers un minimiseur de  $F$ .

## 5.4 Impact de la non unicité sur les schémas numériques

Après avoir étudié la convergence des trajectoires de (HBF) et (AVD) sans unicité du minimiseur et sous des hypothèses de croissance sur  $F$ , on se concentre ici sur des résultats analogues pour des schémas numériques dérivés, à savoir V-FISTA et FISTA. Lors de l'analyse des systèmes dynamiques, on a introduit une hypothèse de géométrie sur  $X^*$  pour pouvoir dériver nos énergies de Lyapunov contenant le terme  $x^*(t)$ . On utilise la même philosophie que dans l'analyse en continu pour retirer l'hypothèse d'unicité du minimiseur mais aucune hypothèse sur  $X^*$  n'est nécessaire pour obtenir les résultats démontrés dans cette section.

### 5.4.1 Convergence de V-FISTA

On s'intéresse d'abord à la convergence de V-FISTA introduit par Beck dans [31, Section 10.7.7] pour les fonctions composites fortement convexes. Ce schéma est une adaptation du schéma de gradient accéléré de Nesterov en fortement convexe introduit dans [105] pour les fonctions composites. Il s'agit donc d'une méthode de type Heavy-Ball. Pour un vecteur initial  $x_0 \in \mathbb{R}^N$ , cet algorithme est défini par :

$$\forall k \in \mathbb{N}, \begin{cases} y_k = x_k + \alpha(x_k - x_{k-1}), \\ x_{k+1} = \text{prox}_{s_h}(y_k - s\nabla f(y_k)), \end{cases} \quad (\text{V-FISTA})$$

avec  $s = \frac{1}{L}$ ,  $x_{-1} = x_0$  et  $\alpha$  est défini égal à  $\frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}}$  où  $\kappa = \frac{\mu}{L}$ .

Contrairement à FISTA, le paramètre d'inertie  $\alpha$  mis en jeu dans V-FISTA est constant dans le temps, d'où le lien avec le système dynamique (HBF). Comme toute méthode de type Heavy-Ball, il est crucial de bien définir  $\alpha$  pour converger rapidement et le meilleur choix est dicté par la quantité  $\kappa = \frac{\mu}{L}$ .

Dans cette section, on donne des résultats de convergence sur la suite définie par V-FISTA lorsque  $F$  vérifie  $\mathcal{G}_\mu^2$  pour  $\mu > 0$ . Les deux théorèmes énoncés sont basés sur deux énergies de Lyapunov légèrement différentes. Le premier théorème donne des résultats de convergence plus faibles mais il est plus lisible, tant au niveau de l'énoncé que de la preuve.

**Théorème 5.4.1.** *Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie par (V-FISTA) avec  $s = \frac{1}{L}$  et  $\alpha = 1 - \frac{5}{3\sqrt{3}}\sqrt{\kappa}$ . Si  $\kappa \leq \frac{1}{3}$ , alors pour tout  $k \in \mathbb{N}$*

$$F(x_k) - F^* \leq \frac{4}{3} \left(1 - \frac{2}{3\sqrt{3}}\sqrt{\kappa}\right)^k (F(x_0) - F^*), \quad (5.4.1)$$

et

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(e^{-\frac{\sqrt{\kappa}}{3\sqrt{3}}k}\right). \quad (5.4.2)$$

Une preuve de ce théorème est donnée en Section 5.4.3.2.

Ce théorème assure que pour une certaine valeur de  $\alpha$ , la décroissance de l'erreur le long des itérés de (V-FISTA) est exponentielle rapide avec un taux au pire cas de  $\mathcal{O}\left(e^{-\frac{2}{3\sqrt{3}}\sqrt{\kappa}k}\right)$ . Cette vitesse de convergence est meilleure que celle assurée par FISTA avec restart périodique optimal ( $\mathcal{O}\left(e^{-\frac{1}{e}\sqrt{\frac{\mu}{L}}k}\right)$ ) sous les mêmes hypothèses. Cela suggère donc que si le paramètre  $\mu$  est connu, le meilleur choix est d'utiliser (V-FISTA) qui fait intervenir un paramètre d'inertie constant. Le deuxième point garantit que la trajectoire des itérés est finie et donc que la suite  $(x_k)_{k \in \mathbb{N}}$  converge fortement vers un minimiseur de  $F$ .

Dans le théorème ci-dessous, on donne un résultat plus fort et plus général sur les itérés de (V-FISTA) pour des hypothèses similaires. On donne une preuve en Section 5.4.3.3.

**Théorème 5.4.2.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie par (V-FISTA) avec  $s = \frac{1}{L}$  et  $\alpha = 1 - \omega\sqrt{\kappa} > 0$ . On a pour tout  $k \in \mathbb{N}$  :

$$F(x_k) - F^* \leq (1 + (\omega - \tau)^2 + (\omega - \tau)\omega\tau\sqrt{\kappa}) (1 - \tau\sqrt{\kappa} + \tau^2\kappa)^k (F(x_0) - F^*), \quad (5.4.3)$$

et

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(e^{-\frac{1}{2}\tau\sqrt{\kappa}(1-\tau\sqrt{\kappa})k}\right), \quad (5.4.4)$$

où  $\tau > 0$  vérifie l'inégalité ci-dessous :

$$(1 - \omega\sqrt{\kappa})\tau^3 - \omega(2 - \omega\sqrt{\kappa})\tau^2 + (\omega^2 + 2)\tau - \omega \leq 0. \quad (5.4.5)$$

La vitesse de convergence donnée dans ce théorème est cachée dans l'inégalité (5.4.5). En effet, en fixant  $\alpha = 1 - \omega\sqrt{\kappa}$  pour  $\omega \in \left(0, \frac{1}{\sqrt{\kappa}}\right)$ , la meilleure constante  $\tau$  est celle qui vérifie

$$(1 - \omega\sqrt{\kappa})\tau^3 - \omega(2 - \omega\sqrt{\kappa})\tau^2 + (\omega^2 + 2)\tau - \omega = 0. \quad (5.4.6)$$

Bien que les solutions de cette équation aient une forme analytique complexe, on donne un graphe affichant  $\tau$  en fonction de  $\omega$  dans la Figure 5.2. Cette figure montre

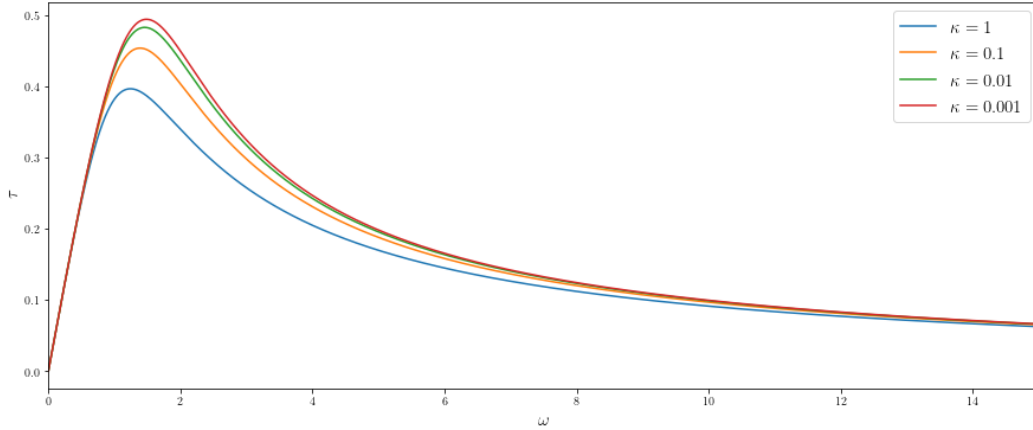


FIGURE 5.2 – Valeur de la meilleure constante  $\tau$  en fonction de  $\omega$  pour des valeurs de  $\kappa$  telles que  $(\tau, \omega, \kappa)$  vérifie (5.4.6).

qu'en fonction de la valeur du conditionnement, il existe une valeur optimale pour  $\omega$  maximisant la plus grande valeur de  $\tau$  admissible. De plus, on peut facilement montrer que si  $(\tau, \omega, \kappa)$  vérifie (5.4.5), alors pour tout  $\hat{\kappa} < \kappa$ ,  $(\tau, \omega, \hat{\kappa})$  vérifie également (5.4.5).

Comme les conséquences du Théorème 5.4.2 sont difficilement lisibles en l'état, on donne plus de détails sur la valeur de la constante  $\tau$  dans le corollaire ci-dessous.

**Corollaire 5.4.3.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie par (V-FISTA) avec  $s = \frac{1}{L}$  et  $\alpha = 1 - \omega\sqrt{\kappa} > 0$ . Si  $\kappa \leq K$ , alors pour tout  $k \in \mathbb{N}$  :

$$\begin{aligned} F(x_k) - F^* &\leq C (1 - \tau\sqrt{\kappa} + \tau^2\kappa)^k (F(x_0) - F^*) \\ &\leq C (1 - \sigma\sqrt{\kappa})^k (F(x_0) - F^*) \end{aligned} \quad (5.4.7)$$

et

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(e^{-\frac{\sigma}{2}\sqrt{\kappa}n}\right), \quad (5.4.8)$$

pour les jeux de paramètres donnés dans le Tableau 5.5.

| $K$           | $\omega$ | $\tau$       | $\sigma$ | $C$   |
|---------------|----------|--------------|----------|-------|
| 1             | 1.2      | <b>0.39</b>  | 0.23     | 2.04  |
| $\frac{1}{3}$ | 1.32     | <b>0.42</b>  | 0.31     | 2.1   |
| $10^{-1}$     | 1.39     | <b>0.45</b>  | 0.38     | 2.07  |
| $10^{-2}$     | 1.46     | <b>0.48</b>  | 0.45     | 2.03  |
| $10^{-3}$     | 1.49     | <b>0.494</b> | 0.486    | 2.02  |
| $10^{-4}$     | 1.495    | <b>0.498</b> | 0.495    | 2.002 |

TABLEAU 5.5 – Jeux de paramètres admissibles.

Ce corollaire souligne que les meilleurs taux de convergence obtenus pour les itérés de (V-FISTA) sont légèrement moins bons que ceux démontrés dans [25] pour un autre schéma type Heavy-Ball lorsque  $F$  a un unique minimiseur (l'erreur décroît comme  $\mathcal{O}(e^{-(2-\sqrt{2})\sqrt{\kappa} + \mathcal{O}(\kappa)})^k$ ). Cependant, le schéma proposé par les auteurs est spécifiquement construit pour obtenir la meilleure vitesse, alors que (V-FISTA) est une méthode plus simple et générique. De plus, la vitesse énoncée n'est plus valable lorsque  $F$  n'a pas un unique minimiseur, les auteurs montrant que le taux de convergence au pire cas est alors légèrement moins bon que  $\mathcal{O}(e^{-\kappa k})$ .

Le corollaire ci-dessous donne un taux de convergence au pire cas de (V-FISTA) dans le cas où  $\alpha$  n'est pas nécessairement calibré par rapport à  $\frac{\mu}{L}$  mais de la forme  $1 - \theta$ . Une preuve est donnée en Section 5.A.10.

**Corollaire 5.4.4.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ . Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie par (V-FISTA) avec  $s = \frac{1}{L}$  et  $\alpha = 1 - \theta$  avec  $\theta \in [\frac{3}{2}\sqrt{\kappa}, 1)$ . Alors, si  $\kappa \leq \frac{1}{10}$ ,

$$F(x_k) - F^* = \mathcal{O}\left(e^{-\tau\kappa k}\right), \quad (5.4.9)$$

et

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(e^{-\frac{\tau}{2}\kappa k}\right) \quad (5.4.10)$$

où  $\tau = \frac{2}{3\theta} \left(1 - \frac{2}{3\theta}\sqrt{\kappa}\right)$ .

Ce corollaire fournit des garanties théoriques sur la convergence de (V-FISTA) pour un choix d'inertie sous-optimal. En effet, on se positionne dans la situation où  $\alpha = 1 - \theta$  avec  $\theta \in [\frac{3}{2}\sqrt{\kappa}, 1)$ , c'est-à-dire que  $\alpha$  est plus petit que sa valeur optimale qui vaut approximativement  $1 - \frac{3}{2}\sqrt{\kappa}$ .

En pratique, ce corollaire permet d'obtenir des garanties de convergence pour différents choix de  $\alpha$  :

- Supposons que l'on choisit  $\alpha = 1 - C\sqrt{\kappa}$  pour  $C \geq \frac{3}{2}$ . On a alors  $\theta = C\sqrt{\kappa}$  et ainsi, le Corollaire 5.4.4 garantit que la suite donnée par (V-FISTA) pour ce choix de paramètre assure une décroissance de l'erreur de l'ordre de  $\mathcal{O}\left(e^{-\tau\sqrt{\kappa}k}\right)$  où  $\tau = \frac{2(3C-2)}{9C^2}$ .
- Supposons que l'on choisit arbitrairement  $\alpha = \frac{9}{10}$  sans connaître la valeur théorique de  $\kappa$ . On a donc  $\theta = \frac{1}{10}$ . Si  $\theta \geq \frac{3}{2}\sqrt{\kappa}$  (c'est-à-dire si  $\kappa < \frac{1}{225}$ ), alors le Corollaire 5.4.4 assure que l'erreur le long des itérés de (V-FISTA) pour ce choix de  $\alpha$  décroît au moins aussi vite que  $e^{-\tau\kappa k}$  avec  $\tau = \frac{20}{3}\left(1 - \frac{20}{3}\sqrt{\kappa}\right)$  qui peut être majoré par  $\frac{100}{27} \approx 3.7$ . On peut conclure que si  $F$  est suffisamment mal conditionnée, les garanties de (V-FISTA) avec ce choix de  $\alpha$  sont meilleures que celles de Forward-Backward qui assure  $F(x_k) - F^* = \mathcal{O}\left(e^{-\kappa k}\right)$  [64].

Le fait d'obtenir des garanties de convergence lorsque  $\alpha$  n'est pas optimal est notable puisque la plupart des algorithmes de type Heavy-Ball proposés dans la littérature sont accompagnés de théorèmes de convergence valables uniquement pour le meilleur choix d'inertie. Grâce au Corollaire 5.4.4, on peut déduire des vitesses de convergence pour (V-FISTA) même si  $\kappa$  n'est pas connu exactement.

### 5.4.2 Convergence de FISTA

Dans cette section, on étudie les propriétés de convergence de FISTA introduit par Beck et Teboulle [32]. En particulier, on s'intéresse à la formulation proposée par Su, Boyd et Candès [127] et Chambolle et Dossal [51] :

$$\forall k \in \mathbb{N}, \begin{cases} y_k = x_k + \frac{k}{k+\alpha}(x_k - x_{k-1}), \\ x_{k+1} = \text{prox}_{sh}(y_k - s\nabla f(y_k)), \end{cases} \quad (\text{FISTA})$$

où  $x_0 \in \mathbb{R}^N$ ,  $s \leq \frac{1}{L}$ ,  $\alpha \geq 3$  et  $x_{-1} = x_0$ .

Si  $F$  est une fonction convexe vérifiant  $\mathcal{H}_L$  pour  $L > 0$ , (FISTA) garantit une décroissance de l'erreur de l'ordre de  $k^{-2}$  pour tout  $\alpha \geq 3$ , ce qui est l'ordre optimal pour les méthodes de premier ordre [105]. Pour des hypothèses plus fortes sur la fonction  $F$ , les résultats démontrés dans la littérature nécessitent que  $F$  ait un unique minimiseur. Dans le théorème ci-dessous, on généralise le résultat démontré dans [26, Théorème 6] pour les fonctions vérifiant  $\mathcal{G}_\mu^2$  et qui s'appuie sur une hypothèse d'unicité du minimiseur. La preuve est une adaptation de l'originale à ce cas relaxé et est située en Section 5.4.3.4.

**Théorème 5.4.5.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}_\mu^2$  pour  $L > 0$  et  $\mu > 0$ .

Soient  $\alpha \geq 3 + \frac{3}{\sqrt{2}}$  et  $\kappa = \frac{\mu}{L}$ . Alors il existe  $\kappa_0 > 0$  tel que si  $\kappa \in (0, \kappa_0]$ , la suite  $(x_k)_{k \in \mathbb{N}}$  définie par (FISTA) avec  $s = \frac{1}{L}$  vérifie :

$$\forall k \geq \frac{3\alpha}{\sqrt{\kappa}}, \quad F(x_k) - F^* \leq \frac{9}{4} e^{-2} M_0 \left( \frac{8e}{3\sqrt{\kappa}} \alpha \right)^{\frac{2\alpha}{3}} k^{-\frac{2\alpha}{3}}. \quad (5.4.11)$$

où  $M_0 = F(x_0) - F^*$  correspond à l'énergie potentielle initiale du système. De plus,

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(k^{-\frac{\alpha}{3}}\right). \quad (5.4.12)$$

Ce théorème permet d'étendre le taux de convergence  $\mathcal{O}\left(k^{-\frac{2\alpha}{3}}\right)$  connu pour les fonctions fortement convexes ou vérifiant  $\mathcal{G}_\mu^2$  avec unicité du minimiseur dans un cas plus général. De plus, le deuxième point garantit la convergence forte des itérés vers un minimiseur de  $F$ .

Comme la borne donnée sur  $F(x_k) - F^*$  est valable pour un nombre fini d'itérations, il est possible d'optimiser le paramètre  $\alpha$  pour obtenir la meilleure vitesse en fonction de la précision voulue. Une analyse sur la façon de paramétrer  $\alpha$  est détaillée par Aujol, Dossal et Rondepierre dans [26] et est extensible ici puisque la borne est identique. En particulier, le [26, Théorème 3] est toujours valable sans hypothèse d'unicité du minimiseur. Ainsi, si  $F$  vérifie les hypothèses du Théorème 5.4.5, pour  $\varepsilon > 0$  et  $\alpha = 3 \ln\left(\frac{3\sqrt{LM_0}}{e\sqrt{2\varepsilon}}\right)$ , il existe  $\kappa_0 > 0$  tel que pour tout  $0 < \kappa \leq \kappa_0$ , les itérés de (FISTA) vérifient  $\|g(x_k)\| \leq \varepsilon$  pour tout  $k \geq k^{FISTA}$  où

$$k^{FISTA} := \frac{8e^2}{\sqrt{\kappa}} \ln\left(\frac{3\sqrt{LM_0}}{e\sqrt{2\varepsilon}}\right), \quad (5.4.13)$$

et  $g$  désigne le gradient composite. Un tel calibrage permet d'obtenir des vitesses comparables à celles de stratégies de restart et ne nécessite pas de connaître  $\mu$ . Une estimation de  $M_0 = F(x_0) - F^*$  est toutefois nécessaire.

Quand on considère une hypothèse plus faible comme  $\mathcal{G}^\gamma$  pour  $\gamma > 2$ , Apidopoulos, Aujol, Dossal et Rondepierre démontrent dans [7] que la décroissance de l'erreur est de  $\mathcal{O}\left(k^{-\frac{2\gamma}{\gamma-2}}\right)$  pour  $\alpha \geq \frac{r+2}{r-2}$  à condition que  $F$  vérifie également l'hypothèse de platitude  $\mathcal{F}_r$  avec  $r \in (2, \gamma]$  et ait un unique minimiseur. On donne ci-dessous un théorème plus général ne s'appuyant pas sur une hypothèse  $\mathcal{F}_r$  ni sur l'unicité du minimiseur de  $F$ . La preuve est donnée en Section 5.4.3.5.

**Théorème 5.4.6.** Soit  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction qui satisfait  $\mathcal{H}_L$  et  $\mathcal{G}^\gamma$  pour  $L > 0$  et  $\gamma > 2$ . Alors, pour  $\alpha > 5 + \frac{8}{\gamma-2}$ , la suite  $(x_k)_{k \in \mathbb{N}}$  définie par (FISTA) avec  $s = \frac{1}{L}$  vérifie :

$$F(x_k) - F^* = \mathcal{O}\left(k^{-\frac{2\gamma}{\gamma-2}}\right), \quad (5.4.14)$$

et

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(k^{-\frac{\gamma}{\gamma-2}}\right). \quad (5.4.15)$$

Ces résultats garantissent que si  $F$  vérifie une hypothèse de croissance faible  $\mathcal{G}^\gamma$  pour  $\gamma > 2$ , le taux de convergence asymptotique original  $\mathcal{O}(k^{-2})$  peut être amélioré en choisissant  $\alpha$  suffisamment grand. De plus, on peut écrire grâce au deuxième point que pour un tel choix de  $\alpha$ ,  $\sum_{k=0}^{+\infty} \|x_k - x_{k-1}\| < +\infty$  et donc que la suite  $(x_k)_{k \in \mathbb{N}}$  converge fortement vers un minimiseur de  $F$ .

Notons que l'hypothèse  $\mathcal{G}^\gamma$  est définie globalement dans ce manuscrit mais un tel résultat peut être étendu si l'hypothèse est locale. Ainsi, **les itérés de (FISTA) convergent fortement vers un minimiseur de  $F$  dès que  $F$  vérifie une hypothèse de croissance locale à condition que  $\alpha$  soit suffisamment grand.**

### 5.4.3 Preuves

#### 5.4.3.1 Notations et résultats préliminaires

Les démonstrations situées dans cette section sont construites autour d'énergies de Lyapunov discrètes qui font intervenir les mêmes termes. Dans chacune des preuves, on utilisera les notations suivantes :

$$\begin{aligned} w_k &= \frac{2}{L}(F(x_k) - F^*), \quad h_k = \|x_k - x_k^*\|^2, \\ \delta_k &= \|x_k - x_{k-1}\|^2, \quad \gamma_k^* = \|x_k^* - x_{k-1}^*\|^2. \end{aligned} \quad (5.4.16)$$

Dans le même esprit que pour notre analyse en continu, on fait intervenir la suite  $(x_k^*)_{k \in \mathbb{N}}$  où pour tout  $k \geq 0$ ,  $x_k^*$  représente la projection de  $x_k$  sur  $X^*$ . Pour gérer les termes additionnels que cette suite peut induire, on introduit le lemme ci-dessous. La preuve est donnée en Section 5.A.6.

**Lemme 5.4.7.** *Pour tout  $k \in \mathbb{N}^*$ , on a :*

1.  $\langle x_k - x_k^*, x_k - x_{k-1} \rangle = \frac{1}{2}(h_k - h_{k-1} + \delta_k - \gamma_k^*) + \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle.$
2.  $\langle x_{k-1} - x_{k-1}^*, x_k - x_{k-1} \rangle = \frac{1}{2}(h_k - h_{k-1} - \delta_k + \gamma_k^*) + \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle,$

On donne également le lemme ci-dessous qui énonce des propriétés sur les suites générées par (FISTA) et (V-FISTA). La preuve est basée sur le lemme de descente prouvé par Chambolle et Dossal dans [51] et se situe en Section 5.A.7.

**Lemme 5.4.8.** *Soit  $(x_k)_{k \in \mathbb{N}}$  la suite générée par (FISTA) ou (V-FISTA) avec  $s = \frac{1}{L}$ . Alors, pour tout  $k \in \mathbb{N}^*$ ,*

$$w_{k+1} - w_k \leq \alpha_k^2 \delta_k - \delta_{k+1}, \quad (5.4.17)$$

and

$$\begin{aligned} w_{k+1} &\leq (1 + \alpha_k)h_k + (\alpha_k^2 + \alpha_k)\delta_k - \alpha_k h_{k-1} - h_{k+1} - \gamma_{k+1}^* - \alpha_k \gamma_k^* \\ &\quad + 2\alpha_k \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - 2\langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle, \end{aligned} \quad (5.4.18)$$

où  $\alpha_k$  vaut  $\frac{k}{k+\alpha}$  pour (FISTA) et  $\alpha$  pour (V-FISTA).

Dans les preuves qui suivent, plusieurs majorations sont liées aux propriétés de la projection sur un ensemble convexe. En effet, si  $C$  est un ensemble convexe fermé tel que  $C \subset \mathbb{R}^N$ , alors pour tout  $x \in \mathbb{R}^N$  et  $y \in C$ ,

$$\langle x - p, y - p \rangle \leq 0,$$

où  $p$  désigne la projection de  $x$  sur  $C$ . Cette propriété garantit que l'on a pour tout  $k \geq 1$

$$\langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle \geq 0,$$

ou encore

$$\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \leq 0.$$

D'autre part, l'ensemble  $X^*$  est assuré d'être convexe et fermé car la fonction  $F$  est convexe et semi-continue inférieurement (tout ensemble de niveau de  $F$  est fermé).

#### 5.4.3.2 Preuve du Théorème 5.4.1

Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie par (V-FISTA) pour  $s = \frac{1}{L}$  et  $\alpha > 0$  à définir. On va analyser l'énergie de Lyapunov discrète ci-dessous :

$$\mathcal{E}_k = \frac{2}{L}(F(x_k) - F^*) + \|x_k - x_{k-1} + \lambda(x_{k-1} - x_{k-1}^*)\|^2, \quad (5.4.19)$$

où  $\lambda > 0$  et  $x_k^*$  désigne la projection de  $x_k$  sur  $X^*$  pour tout  $k \in \mathbb{N}$ . Pour le choix  $\lambda = \sqrt{\kappa}$  et  $x^* k = x^*$  pour tout  $k$ , on retrouve l'énergie considérée par Beck dans [31].

Le but de cette preuve est de déterminer  $\tau > 0$  le plus grand possible tel que pour un jeu de paramètres bien choisi,

$$\mathcal{E}_{k+1} - (1 - \tau\sqrt{\kappa})\mathcal{E}_k \leq 0. \quad (5.4.20)$$

La preuve est divisée en trois parties. Dans un premier temps, on s'appuie sur les notations et lemmes énoncés dans la Section 5.4.3.1 pour gérer les termes liés à la non-unicité des minimiseurs de  $F$ . Ensuite, on donne un jeu de paramètres permettant d'exploiter les propriétés de croissance de  $F$  et obtenir l'inégalité souhaitée (5.4.20). Enfin, on démontre la convergence forte de la suite  $(x_k)_{k \in \mathbb{N}}$  en contrôlant  $\|x_k - x_{k_1}\|$ .

**Travail préliminaire :** Rappelons que l'on utilise les notations introduites en Section (5.4.16). En réécrivant (5.4.19) et en utilisant le deuxième point du Lemme 5.4.7, on trouve que

$$\mathcal{E}_k = w_k + (1 - \lambda)\delta_k + \lambda(h_k - h_{k-1}) + \lambda^2 h_{k-1} + \lambda\gamma_k^* + 2\lambda\langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle. \quad (5.4.21)$$



Le Lemme 5.4.8 assure que pour  $\lambda \leq 1$  :

$$\begin{aligned} w_{k+1} - (1 - \lambda)w_k &\leq \alpha(\alpha + \lambda)\delta_k - (1 - \lambda)\delta_{k+1} + \alpha\lambda(h_k - h_{k-1}) + \lambda(h_k - h_{k+1}) \\ &\quad - \lambda\gamma_{k+1}^* - \alpha\lambda\gamma_k^* + 2\alpha\lambda\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \\ &\quad - 2\lambda\langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle. \end{aligned} \quad (5.4.22)$$

Cette inégalité combinée à (5.4.21) donne que

$$\mathcal{E}_{k+1} - (1 - \lambda)\mathcal{E}_k \leq a_1\delta_k + a_2(h_k - h_{k-1}) + a_3h_k + \mathcal{X}_k^*, \quad (5.4.23)$$

où :

$$\begin{aligned} - a_1 &= \alpha(\alpha + \lambda) - (1 - \lambda)^2, \\ - a_2 &= \alpha\lambda - \lambda(1 - \lambda) + (1 - \lambda)\lambda^2, \\ - a_3 &= \lambda^3, \end{aligned}$$

et  $\mathcal{X}_k^*$  est défini par

$$\begin{aligned} \mathcal{X}_k^* &= -\lambda(1 - \lambda + \alpha)\gamma_k^* + 2\alpha\lambda\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \\ &\quad - 2\lambda(1 - \lambda)\langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle. \end{aligned}$$

Les propriétés du projeté sur un ensemble convexe garantissent que pour tout  $k \in \mathbb{N}$  :

$$\begin{cases} \gamma_k^* \geq 0, \\ \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \leq 0, \\ \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle \geq 0. \end{cases}$$

Par conséquent  $\mathcal{X}_k^* \leq 0$  et on a donc :

$$\mathcal{E}_{k+1} - (1 - \lambda)\mathcal{E}_k \leq a_1\delta_k + a_2(h_k - h_{k-1}) + a_3h_k. \quad (5.4.24)$$

**Décroissance de  $F(x_k) - F^*$  :** Rappelons que l'on cherche  $\tau > 0$  tel que :

$$\mathcal{E}_{k+1} - (1 - \tau\sqrt{\kappa})\mathcal{E}_k \leq 0.$$

On choisit le jeu de paramètres suivant :

$$\tau = \frac{2}{3\sqrt{3}}, \quad \lambda = \frac{1}{\sqrt{3}}\sqrt{\kappa}, \quad \alpha = 1 - \frac{5}{3\sqrt{3}}\sqrt{\kappa} = 1 - \frac{5}{3}\lambda.$$

On trouve alors que :

$$\begin{aligned} \mathcal{E}_{k+1} - (1 - \tau\sqrt{\kappa})\mathcal{E}_k &= \mathcal{E}_{k+1} - (1 - \lambda)\mathcal{E}_k + \left(\frac{2}{3\sqrt{3}} - \frac{1}{\sqrt{3}}\right)\sqrt{\kappa}\mathcal{E}_k \\ &\leq a_1\delta_k + a_2(h_k - h_{k-1}) + a_3h_k - \frac{1}{3\sqrt{3}}\sqrt{\kappa}\mathcal{E}_k, \end{aligned} \quad (5.4.25)$$

avec :

$$\begin{aligned} - a_1 &= -\frac{\lambda}{3}\left(1 - \frac{\lambda}{3}\right) = \frac{\sqrt{\kappa}}{27}(\sqrt{\kappa} - 3\sqrt{3}), \\ - a_2 &= \lambda^2\left(\frac{1}{3} - \lambda\right) = \frac{\kappa}{3\sqrt{3}}\left(\frac{1}{\sqrt{3}} - \sqrt{\kappa}\right), \end{aligned}$$

$$- a_3 = \lambda^3 = \frac{\kappa^{\frac{3}{2}}}{3\sqrt{3}}.$$

Si  $\kappa \leq \frac{1}{3}$ , alors  $a_1 \leq 0$  et donc

$$\mathcal{E}_{k+1} - \left(1 - \frac{2}{3\sqrt{3}}\sqrt{\kappa}\right) \mathcal{E}_k \leq \frac{\kappa}{3\sqrt{3}} \left(\frac{1}{\sqrt{3}} - \sqrt{\kappa}\right) (h_k - h_{k-1}) + \frac{\kappa^{\frac{3}{2}}}{3\sqrt{3}} h_k - \frac{1}{3\sqrt{3}} \sqrt{\kappa} \mathcal{E}_k. \quad (5.4.26)$$

De plus, comme la condition  $\kappa \leq \frac{1}{3}$  assure que  $a_2 \geq 0$ , on peut appliquer le lemme ci-dessous qui est prouvé en Section 5.A.8.

**Lemme 5.4.9.** *Soit  $(x_k)_{k \in \mathbb{N}}$  la suite donnée par (V-FISTA) et  $\lambda = \frac{1}{\sqrt{3}}\sqrt{\kappa}$ . Pour tout  $k \in \mathbb{N}$  :*

$$h_k - h_{k-1} \leq \frac{\sqrt{3}}{\sqrt{\kappa}} (\mathcal{E}_k - w_k). \quad (5.4.27)$$

Le Lemme 5.4.9 assure que si  $\kappa \leq \frac{1}{3}$ , alors

$$\mathcal{E}_{k+1} - \left(1 - \frac{2}{3\sqrt{3}}\sqrt{\kappa}\right) \mathcal{E}_k \leq \frac{\kappa^{\frac{3}{2}}}{3\sqrt{3}} h_k - \frac{\sqrt{\kappa}}{3} \left(\frac{1}{\sqrt{3}} - \sqrt{\kappa}\right) w_k - \frac{\kappa}{3} \mathcal{E}_k. \quad (5.4.28)$$

De plus, comme  $F$  vérifie  $\mathcal{G}_\mu^2$ , on a  $h_k \leq \frac{w_k}{\kappa}$  et donc

$$\begin{aligned} \mathcal{E}_{k+1} - \left(1 - \frac{2}{3\sqrt{3}}\sqrt{\kappa}\right) \mathcal{E}_k &\leq \left(\frac{\sqrt{\kappa}}{3\sqrt{3}} - \frac{\sqrt{\kappa}}{3} \left(\frac{1}{\sqrt{3}} - \sqrt{\kappa}\right)\right) w_k - \frac{\kappa}{3} \mathcal{E}_k \\ &\leq \frac{\kappa}{3} w_k - \frac{\kappa}{3} \mathcal{E}_k. \end{aligned} \quad (5.4.29)$$

En remarquant que  $w_k \leq \mathcal{E}_k$ , on peut déduire que

$$\mathcal{E}_{k+1} - \left(1 - \frac{2}{3\sqrt{3}}\sqrt{\kappa}\right) \mathcal{E}_k \leq 0. \quad (5.4.30)$$

Ainsi,

$$\mathcal{E}_k \leq \left(1 - \frac{2}{3\sqrt{3}}\sqrt{\kappa}\right)^k \mathcal{E}_0. \quad (5.4.31)$$

Comme on considère que  $x_{-1} = x_0$ , on trouve que  $\mathcal{E}_0 = w_0 + \lambda^2 h_0$ . Par conséquent, la condition de croissance  $\mathcal{G}_\mu^2$  assure que  $\mathcal{E}_0 \leq \frac{4}{3} w_0$  et

$$F(x_k) - F^* \leq \frac{4}{3} \left(1 - \frac{2}{3\sqrt{3}}\sqrt{\kappa}\right)^k (F(x_0) - F^*). \quad (5.4.32)$$

**Convergence des trajectoires :** Soit  $b_k = \|x_k - x_{k-1} + \lambda(x_{k-1} - x_{k-1}^*)\|^2$ . En appliquant l'inégalité  $\|u\|^2 = 2\|u + v\|^2 + 2\|v\|^2$ , on trouve que :

$$\delta_k \leq 2b_k + \frac{2}{\lambda^2} h_{k-1}. \quad (5.4.33)$$

En revenant à la définition de  $\mathcal{E}$  et grâce à la géométrie de  $F$ , on a

$$\delta_k \leq 2\mathcal{E}_k + \frac{2}{\lambda^2 \kappa} w_{k-1}. \quad (5.4.34)$$

Alors, en appliquant l'inégalité (5.4.30), on obtient que :

$$\delta_k \leq \left( 2 \left( 1 - \frac{2}{3\sqrt{3}} \sqrt{\kappa} \right) + \frac{2}{\lambda^2 \kappa} \right) \mathcal{E}_{k-1}, \quad (5.4.35)$$

et en conséquence,

$$\delta_k \leq \left( 2 \left( 1 - \frac{2}{3\sqrt{3}} \sqrt{\kappa} \right) + \frac{2}{\lambda^2 \kappa} \right)^{k-1} \mathcal{E}_0. \quad (5.4.36)$$

Ainsi,

$$\|x_k - x_{k-1}\| = \mathcal{O} \left( e^{-\frac{1}{3\sqrt{3}} \sqrt{\kappa} n} \right). \quad (5.4.37)$$

### 5.4.3.3 Preuve du Théorème 5.4.2

**Structure de la preuve :** La preuve du Théorème 5.4.2 est construite autour de l'analyse proposée dans [25] pour étudier la convergence des trajectoires de (HBF), c'est-à-dire l'EDO suivante :

$$\ddot{x}(t) + \alpha_c \dot{x}(t) + \nabla F(x(t)) = 0, \quad (\text{HBF})$$

pour  $\alpha_c > 0$ .

Comme la suite définie par (V-FISTA) peut être vue comme une discrétisation de (HBF), on adapte la stratégie proposée en continu à ce cadre discret. Il est important de noter que le paramètre  $\alpha_c$  dans (HBF) ne joue pas le même rôle que  $\alpha$  dans (V-FISTA). En effet,  $\alpha_c$  se comporte plutôt comme  $\sqrt{L}(1 - \alpha)$ .

La preuve du Théorème 5.4.2 se construit autour de l'énergie ci-dessous :

$$\forall k \in \mathbb{N}, \quad \mathcal{E}_k = \frac{2}{L} (F(x_k) - F^*) + \alpha \|x_k - x_{k-1} + \lambda(x_k - x_k^*)\|^2 + \lambda(1 - \alpha)^2 \|x_k - x_k^*\|^2. \quad (5.4.38)$$

La stratégie de la preuve est relativement simple : on cherche un jeu de paramètres  $(\alpha, \lambda, \nu) \in (\mathbb{R}^+)^3$  tel que pour tout  $k \in \mathbb{N}$ ,

$$\mathcal{E}_{k+1} - \mathcal{E}_k + \nu \mathcal{E}_{k+1} \leq 0. \quad (5.4.39)$$

Une fois cette inégalité démontrée, des calculs simples donnent que

$$\forall k \in \mathbb{N}, \quad \mathcal{E}_k \leq (1 - \nu + \nu^2)^k \mathcal{E}_0, \quad (5.4.40)$$

ce qui permet de conclure.

**Preuve :** Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie dans (V-FISTA). En utilisant les notations définies dans (5.4.16), on peut réécrire notre énergie de la manière suivante

$$\forall k \in \mathbb{N}, \quad \mathcal{E}_k = w_k + \lambda(\lambda\alpha + (1 - \alpha)^2)h_k + \alpha\delta_k + 2\alpha\lambda \langle x_k - x_k^*, x_k - x_{k-1} \rangle. \quad (5.4.41)$$

D'après le premier point du Lemme 5.4.7,

$$\begin{aligned} \mathcal{E}_k = & w_k + \lambda(\lambda\alpha + (1 - \alpha)^2)h_k + \lambda\alpha(h_k - h_{k-1}) + (1 + \lambda)\alpha\delta_k \\ & - \lambda\alpha\gamma_k^* + 2\lambda\alpha \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \quad (5.4.42)$$

Les propriétés de la projection sur un ensemble convexe donnent que pour tout  $k \in \mathbb{N}$  :

$$\mathcal{E}_k \leq w_k + \lambda(\lambda\alpha + (1 - \alpha)^2)h_k + \lambda\alpha(h_k - h_{k-1}) + (1 + \lambda)\alpha\delta_k. \quad (5.4.43)$$

En utilisant l'expression (5.4.42), on peut démontrer le lemme ci-dessous. La preuve est donnée en Section 5.A.9.

**Lemme 5.4.10.** *Pour tout  $k \in \mathbb{N}$ , on a*

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &\leq -\lambda w_{k+1} + \lambda\alpha(\lambda + \alpha - 1)(h_{k+1} - h_k) \\ &\quad + (\lambda\alpha + \alpha - 1)\delta_{k+1} - \alpha(1 - \alpha - \lambda\alpha)\delta_k. \end{aligned} \quad (5.4.44)$$

Cette inégalité combinée avec (5.4.43) garantit que pour tout  $\nu > 0$ ,

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k + \nu\mathcal{E}_{k+1} &\leq (\nu - \lambda)w_{k+1} + \lambda\alpha(\lambda + \alpha - 1 + \nu)(h_{k+1} - h_k) \\ &\quad + ((1 + \lambda)\alpha(1 + \nu) - 1)\delta_{k+1} - \alpha(1 - \alpha - \lambda\alpha)\delta_k \\ &\quad + \nu\lambda(\lambda\alpha + (1 - \alpha)^2)h_{k+1}. \end{aligned} \quad (5.4.45)$$

On prend le jeu de paramètres suivant :

$$\alpha = 1 - \omega\sqrt{\kappa}, \quad \nu = \tau\sqrt{\kappa}, \quad \lambda = 1 - \alpha - \nu = (\omega - \tau)\sqrt{\kappa}. \quad (5.4.46)$$

Pour ce choix de paramètres, on peut écrire l'inégalité suivante

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k + \tau\sqrt{\kappa}\mathcal{E}_{k+1} &\leq (2\tau - \omega)\sqrt{\kappa}w_{k+1} - (\omega\tau + (\omega - \tau)^2 + \omega\tau(\omega - \tau)\sqrt{\kappa})\kappa\delta_{k+1} \\ &\quad - (1 - \omega\sqrt{\kappa})(\tau + (\omega - \tau)\sqrt{\kappa})\sqrt{\kappa}\delta_k \\ &\quad + \tau(\omega - \tau)(\omega - \tau + \omega\tau\sqrt{\kappa})\kappa^{\frac{3}{2}}h_{k+1}. \end{aligned} \quad (5.4.47)$$

Notons que l'on veut que  $\alpha > 0$  ce qui implique que  $1 - \omega\sqrt{\kappa} > 0$ . Supposons en plus que  $\omega > 2\tau$ . Alors, on obtient que pour tout  $k \in \mathbb{N}$ ,

$$\mathcal{E}_{k+1} - \mathcal{E}_k + \tau\sqrt{\kappa}\mathcal{E}_{k+1} \leq -(\omega - 2\tau)\sqrt{\kappa}w_{k+1} + \tau(\omega - \tau)(\omega - \tau + \omega\tau\sqrt{\kappa})\kappa^{\frac{3}{2}}h_{k+1}. \quad (5.4.48)$$

Comme  $F$  satisfait  $\mathcal{G}_\mu^2$ , on peut écrire que  $\kappa h_{k+1} \leq w_{k+1}$  et ainsi,

$$\mathcal{E}_{k+1} - \mathcal{E}_k + \tau\sqrt{\kappa}\mathcal{E}_{k+1} \leq (\tau(\omega - \tau)(\omega - \tau + \omega\tau\sqrt{\kappa}) - \omega + 2\tau)\kappa^{\frac{3}{2}}h_{k+1}. \quad (5.4.49)$$

Par conséquent, si

$$(1 - \omega\sqrt{\kappa})\tau^3 - \omega(2 - \omega\sqrt{\kappa})\tau^2 + (\omega^2 + 2)\tau - \omega \leq 0, \quad (5.4.50)$$

alors pour tout  $k \in \mathbb{N}$ ,

$$\mathcal{E}_{k+1} - \mathcal{E}_k + \tau\sqrt{\kappa} \leq 0. \quad (5.4.51)$$

Notons que les solutions de (5.4.50) vérifient directement  $\omega > 2\tau$ .

Des calculs simples montrent que (5.4.51) implique

$$\mathcal{E}_k \leq (1 - \tau\sqrt{\kappa} + \tau^2\kappa)^k \mathcal{E}_0. \quad (5.4.52)$$

Comme  $x_{-1} = x_0$ ,

$$\mathcal{E}_0 = w_0 + \lambda (\lambda\alpha + (1 - \alpha)^2) h_0 = w_0 + ((\omega - \tau)^2 + (\omega - \tau)\omega\tau\sqrt{\kappa}) \kappa h_0, \quad (5.4.53)$$

et en utilisant l'hypothèse  $\mathcal{G}_\mu^2$ ,

$$\mathcal{E}_0 \leq (1 + (\omega - \tau)^2 + (\omega - \tau)\omega\tau\sqrt{\kappa}) w_0. \quad (5.4.54)$$

De plus, pour tout  $k \in \mathbb{N}$ ,  $w_k \leq \mathcal{E}_k$ . Donc si (5.4.50) est vérifiée, alors pour tout  $k \in \mathbb{N}$  :

$$F(x_k) - F^* \leq (1 + (\omega - \tau)^2 + (\omega - \tau)\omega\tau\sqrt{\kappa}) (1 - \tau\sqrt{\kappa} + \tau^2\kappa)^k (F(x_0) - F^*). \quad (5.4.55)$$

En appliquant l'inégalité  $\|u\|^2 = 2\|u + v\|^2 + 2\|v\|^2$ , on déduit que :

$$\delta_k \leq 2b_k + \frac{2}{\lambda^2} h_k, \quad (5.4.56)$$

où  $b_k = \|\lambda(x_k - x_k^*) + x_k - x_{k-1}\|^2$ . L'hypothèse  $\mathcal{G}_\mu^2$  donne que

$$\delta_k \leq 2b_k + \frac{2}{\lambda^2\kappa} w_k, \quad (5.4.57)$$

et étant donné la définition de  $\mathcal{E}_k$  on a

$$\delta_k \leq \left( \frac{2}{\alpha} + \frac{2}{\lambda^2\kappa} \right) \mathcal{E}_k. \quad (5.4.58)$$

En combinant l'inégalité ci-dessus à (5.4.52), on montre finalement que si (5.4.50) est vérifiée, alors

$$\|x_k - x_{k-1}\| = \mathcal{O} \left( e^{-\frac{1}{2}\tau\sqrt{\kappa}(1-\tau\sqrt{\kappa})n} \right). \quad (5.4.59)$$

#### 5.4.3.4 Preuve du Théorème 5.4.5

La preuve du Théorème 5.4.5 est une adaptation de la preuve de [26, Théorème 6] en retirant l'hypothèse que  $F$  a un unique minimiseur. La structure est proche avec des termes additionnels liés à la non-unicité des minimiseurs. L'aspect technique réside dans la façon de contrôler ces termes.

On analyse l'énergie de Lyapunov discrète définie ci-dessous :

$$E_k = \frac{2k^2}{L} (F(x_k) - F^*) + \|\lambda(x_{k-1} - x_{k-1}^*) + k(x_k - x_{k-1})\|^2, \quad (5.4.60)$$

avec  $\lambda = \frac{2\alpha}{3}$ . On utilise les notations introduites dans (5.4.16) et  $\alpha_k$  désigne  $\frac{k}{k+\alpha}$ . Ainsi, on peut réécrire l'énergie de Lyapunov de la façon suivante :

$$E_k = k^2 w_k + \lambda^2 h_{k-1} + k^2 \delta_k + 2\lambda k \langle x_{k-1} - x_{k-1}^*, x_k - x_{k-1} \rangle,$$

et en appliquant le second point du Lemme 5.4.7, on a

$$E_k = k^2 w_k + \lambda k h_k + (\lambda^2 - \lambda k) h_{k-1} + (k^2 - \lambda k) \delta_k + \lambda k \gamma_k^* + 2\lambda k \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle.$$

Pour tout  $k \in \mathbb{N}$ , on a :

$$\begin{aligned}
E_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) E_k &= (k+1)^2 w_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) k^2 w_k \\
&+ ((k+1)^2 - \lambda(k+1)) \delta_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) (k^2 - \lambda k) \delta_k \\
&+ \left(\lambda^2 - \lambda(k+1) - \lambda k \left(1 - \frac{\lambda - 2}{k}\right)\right) h_k + \lambda(k+1) h_{k+1} \\
&- (\lambda^2 - \lambda k) \left(1 - \frac{\lambda - 2}{k}\right) h_{k-1} \\
&+ \lambda(k+1) \gamma_{k+1}^* + 2\lambda(k+1) \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle \\
&- \lambda(k - \lambda + 2) \gamma_k^* - 2\lambda(k - \lambda + 2) \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle.
\end{aligned}$$

Des calculs simples donnent que

$$(k+1)^2 w_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) k^2 w_k = n(k - \lambda + 2)(w_{k+1} - w_k) + (\lambda k + 1)w_{k+1}.$$

Par conséquent, le Lemme 5.4.8 garantit que pour tout  $k \in \mathbb{N}^*$  :

$$\begin{aligned}
&(k+1)^2 w_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) k^2 w_k \\
&\leq n(k - \lambda + 2) (\alpha_k^2 \delta_k - \delta_{k+1}) \\
&+ (\lambda k + 1) ((1 + \alpha_k) h_k + (\alpha_k^2 + \alpha_k) \delta_k - \alpha_k h_{k-1} - h_{k+1} - \gamma_{k+1}^* - \alpha_k \gamma_k^*) \\
&+ 2(\lambda k + 1) (\alpha_k \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle).
\end{aligned}$$

Ainsi,

$$\begin{aligned}
E_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) E_k &\leq A_1(k, \alpha) \delta_k + A_2(k, \alpha) \delta_{k+1} + B_1(k, \alpha) h_{k-1} \\
&+ B_2(k, \alpha) h_k + B_3(k, \alpha) h_{k+1} + D_1(k, \alpha) \gamma_{k+1}^* \\
&+ D_2(k, \alpha) \gamma_k^* + D_3(k, \alpha) \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle \\
&+ D_4(k, \alpha) \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle \\
&+ D_5(k, \alpha) \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle,
\end{aligned} \tag{5.4.61}$$

où :

$$\begin{aligned}
- A_1(k, \alpha) &= \frac{17\alpha^2}{9} - \frac{8\alpha}{3} + 2 - \alpha \frac{(10\alpha^2 - 18\alpha + 9)k + 7\alpha^3 - 12\alpha^2 + 6\alpha}{3(k+\alpha)^2}, \\
- A_2(k, \alpha) &= 1 - \frac{2\alpha}{3}, \\
- B_1(k, \alpha) &= -\frac{2\alpha^2}{9} + \frac{4\alpha}{3} - 1 + \frac{3\alpha - 2\alpha^3}{3(k+\alpha)} + \frac{8\alpha^3 - 24\alpha^2}{27n}, \\
- B_2(k, \alpha) &= \frac{2\alpha^2}{9} - 2\alpha + 2 - \frac{3\alpha - 2\alpha^3}{3(k+\alpha)}, \\
- B_3(k, \alpha) &= \frac{2\alpha}{3} - 1, \\
- D_1(k, \alpha) &= \frac{2\alpha}{3} - 1, \\
- D_2(k, \alpha) &= -\frac{4\alpha}{3}k - 1 - \frac{4\alpha}{3} + \frac{10\alpha^2}{9} - \frac{\alpha(2\alpha^2 - 3)}{3(k+\alpha)}, \\
- D_3(k, \alpha) &= \frac{4\alpha}{3} - 2, \\
- D_4(k, \alpha) &= -\frac{4\alpha}{3}k - \frac{8\alpha}{3} + \frac{8\alpha^2}{9}, \\
- D_5(k, \alpha) &= \frac{4\alpha}{3}k + 2 - \frac{4\alpha^2}{3} + \frac{\alpha(4\alpha^2 - 6)}{3(k+\alpha)}.
\end{aligned}$$

Comme  $B_3(k, \alpha) = -A_2(k, \alpha) = D_1(k, \alpha) = \frac{1}{2}D_3(k, \alpha)$ , et

$$B_1(k, \alpha) + B_2(k, \alpha) + B_3(k, \alpha) = \frac{8\alpha^2}{27} \frac{\alpha - 3}{k} = \frac{4\alpha K(\alpha)}{3k},$$

où  $K(\alpha) = \frac{2\alpha(\alpha-3)}{9}$ , on peut écrire que

$$\begin{aligned} E_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) E_k &\leq \frac{4\alpha K(\alpha)}{3k} h_k + A_1(k, \alpha) \delta_k + B_1(k, \alpha) (h_{k-1} - h_k) \\ &\quad + B_3(k, \alpha) (h_{k+1} - h_k - \delta_{k+1}) + B_3(k, \alpha) \gamma_{k+1}^* \\ &\quad + D_2(k, \alpha) \gamma_k^* + 2B_3(k, \alpha) \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle \\ &\quad + D_4(k, \alpha) \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle \\ &\quad + D_5(k, \alpha) \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \quad (5.4.62)$$

Par la suite, on a besoin du lemme ci-dessous qui est une adaptation de [26, Lemme 4]. La preuve est donnée en Section 5.A.11.

**Lemme 5.4.11.** *Pour tout  $k \geq 1$  et  $(A, B) \in \mathbb{R}^2$ ,*

$$\begin{aligned} A\delta_k + B(h_{k-1} - h_k) &\leq \left(2|A + B| + \frac{\sqrt{2}|B|}{\sqrt{\kappa}}\right) \left(1 + \frac{4\alpha^2}{9\kappa k^2}\right) \frac{E_k}{\left(k - \frac{2\alpha}{3}\right)^2} \\ &\quad - B\gamma_k^* + 2B \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \quad (5.4.63)$$

Cette inégalité permet d'écrire que :

$$\begin{aligned} \frac{4\alpha K(\alpha)}{3} \frac{h_k}{k} + A_1(k, \alpha) \delta_k + B_1(k, \alpha) (h_{k-1} - h_k) &\leq \frac{\hat{C}_1(k, \alpha, \kappa) E_k}{(k - \lambda)^2} \\ &\quad - B_1(k, \alpha) \gamma_k^* + 2B_1(k, \alpha) \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle, \end{aligned}$$

et

$$\begin{aligned} B_3(k, \alpha) (h_{k+1} - h_k - \delta_{k+1}) &\leq \frac{\hat{C}_2(k, \alpha, \kappa) E_{k+1}}{(k + 1 - \lambda)^2} - 2B_3(k, \alpha) \langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle \\ &\quad + B_3(k, \alpha) \gamma_{k+1}^*, \end{aligned}$$

où

$$\hat{C}_1(k, \alpha, \kappa) = 2 \left| \frac{5}{3} \alpha^2 - \frac{4\alpha}{3} + 1 + R(k, \alpha) \right| + \sqrt{2} \left( \frac{\left| -\frac{2\alpha^2}{9} + \frac{4\alpha}{3} - 1 + Q(k, \alpha) \right|}{\sqrt{\kappa}} \right) \left( 1 + \frac{4\alpha^2}{9\kappa k^2} \right) + \frac{4\alpha K(\alpha)}{3\kappa k}$$

avec :

$$\begin{aligned} |R(\alpha, k)| &= \left| A_1(k, \alpha) + B_1(k, \alpha) - \left( \frac{5}{3} \alpha^2 - \frac{4\alpha}{3} + 1 \right) \right| \leq \frac{8\alpha^3}{k} \\ |Q(\alpha, k)| &= \frac{\alpha^3}{3k} \left| k \frac{3 - 2\alpha^2}{\alpha^2(k + \alpha)} + 8 \frac{\alpha - 3}{9\alpha} \right| \leq \frac{\alpha^3}{k}, \end{aligned}$$

et  $\hat{C}_2(k, \alpha, \kappa) = \left(\frac{2\alpha}{3} - 1\right) \left(4 + \frac{\sqrt{2}}{\sqrt{\kappa}}\right) \left(1 + \frac{4\alpha^2}{9\kappa(k+1)^2}\right)$ . En revenant à l'inégalité (5.4.62), on trouve que

$$\begin{aligned} E_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) E_k &\leq \frac{\hat{C}_1(k, \alpha, \kappa) E_k}{(k - \lambda)^2} + \frac{\hat{C}_2(k, \alpha, \kappa) E_{k+1}}{(k + 1 - \lambda)^2} \\ &\quad + 2B_3(k, \alpha) \gamma_{k+1}^* + 2B_3(k, \alpha) \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle \\ &\quad - 2B_3(k, \alpha) \langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle + (D_2(k, \alpha) - B_1(k, \alpha)) \gamma_k^* \\ &\quad + D_4(k, \alpha) \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle \\ &\quad + (D_5(k, \alpha) + 2B_1(k, \alpha)) \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned}$$

Comme pour tout  $k \in \mathbb{N}^*$ ,

$$\gamma_k^* + \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle - \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle = \langle x_k - x_{k-1}, x_k^* - x_{k-1}^* \rangle,$$

on a :

$$2B_3(k, \alpha) (\gamma_{k+1}^* + \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle - \langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle) = 2(\lambda - 1) \langle x_{k+1} - x_k, x_{k+1}^* - x_k^* \rangle.$$

De plus, pour tout  $k \geq \frac{2\alpha}{3} - 2$ ,

$$D_4(k, \alpha) \leq D_2(k, \alpha) - B_1(k, \alpha) \leq -(D_5(k, \alpha) + 2B_1(k, \alpha)) \leq -\frac{4\alpha}{3} \left(k - 2 \left(\frac{2\alpha}{3} - 1\right)\right). \quad (5.4.64)$$

Comme la suite  $(x_k)_{k \in \mathbb{N}}$  vérifie les inégalités suivantes,

$$\begin{cases} \gamma_k^* \geq 0, \\ \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle \geq 0, \\ \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \leq 0, \end{cases}$$

on trouve que

$$\begin{aligned} (D_2(k, \alpha) - B_1(k, \alpha)) \gamma_k^* + D_4(k, \alpha) \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle + (D_5(k, \alpha) + 2B_1(k, \alpha)) \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \\ \leq -\frac{4\alpha}{3} \left(k - 2 \left(\frac{2\alpha}{3} - 1\right)\right) \langle x_k - x_{k-1}, x_k^* - x_{k-1}^* \rangle, \end{aligned}$$

qui est un terme positif si  $k \geq \frac{4\alpha}{3} - 2$ . Par conséquent, si cette condition est vérifiée on a

$$\begin{aligned} E_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) E_k &\leq \frac{\hat{C}_1(k, \alpha, \kappa) E_k}{(k - \lambda)^2} + \frac{\hat{C}_2(k, \alpha, \kappa) E_{k+1}}{(k + 1 - \lambda)^2} \\ &\quad + 2(\lambda - 1) \langle x_{k+1} - x_k, x_{k+1}^* - x_k^* \rangle. \end{aligned} \quad (5.4.65)$$

Comme  $\langle x_{k+1} - x_k, x_{k+1}^* - x_k^* \rangle \leq \delta_{k+1}$ , pour tout  $k \geq \frac{4\alpha}{3} - 2$ ,

$$E_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) E_k \leq \frac{\hat{C}_1(k, \alpha, \kappa) E_k}{(k - \lambda)^2} + \frac{\hat{C}_2(k, \alpha, \kappa) E_{k+1}}{(k + 1 - \lambda)^2} + 2(\lambda - 1) \delta_{k+1}.$$

Alors d'après (5.A.34) et l'inégalité  $h_k \leq \frac{E_k}{\kappa k^2}$  assurée par l'hypothèse  $\mathcal{G}_\mu^2$ ,

$$\delta_{k+1} \leq \frac{2}{(k + 1 - \lambda)^2} b_{k+1} + \frac{8\alpha^2}{9(k + 1 - \lambda)^2} h_{k+1} \leq \frac{2}{(k + 1 - \lambda)^2} \left(1 + \frac{4\alpha^2}{9\kappa(k + 1)^2}\right) E_{k+1}.$$



Ainsi,

$$E_{k+1} - \left(1 - \frac{\lambda - 2}{k}\right) E_k \leq \frac{\tilde{C}_1(k, \alpha, \kappa) E_k}{(k - \lambda)^2} + \frac{\tilde{C}_2(k, \alpha, \kappa) E_{k+1}}{(k + 1 - \lambda)^2}, \quad (5.4.66)$$

où  $\tilde{C}_1(k, \alpha, \kappa) = \hat{C}_1(k, \alpha, \kappa)$  and

$$\begin{aligned} \tilde{C}_2(k, \alpha, \kappa) &= \hat{C}_2(k, \alpha, \kappa) + 4(\lambda - 1) \left(1 + \frac{4\alpha^2}{9\kappa(k+1)^2}\right) \\ &= \left(\frac{2\alpha}{3} - 1\right) \left(8 + \frac{\sqrt{2}}{\sqrt{\kappa}}\right) \left(1 + \frac{4\alpha^2}{9\kappa(k+1)^2}\right). \end{aligned}$$

Comme  $\kappa \in (0, 1]$ , pour tout  $k \geq \frac{4\alpha}{3\sqrt{\kappa}}$ , on a que

$$\frac{1}{k - \frac{2\alpha}{3}} \leq \frac{1}{k} (1 + \sqrt{\kappa}),$$

et donc pour tout  $k \geq \frac{4\alpha}{3\sqrt{\kappa}}$ ,

$$E_{k+1} - \left(1 - \frac{\frac{2\alpha}{3} - 2}{k}\right) E_k \leq (1 + \sqrt{\kappa})^2 \left(\tilde{C}_1(k, \alpha, \kappa) \frac{E_k}{k^2} + \tilde{C}_2(k, \alpha, \kappa) \frac{E_{k+1}}{(k+1)^2}\right). \quad (5.4.67)$$

On peut noter que cette inégalité est identique à celle obtenue dans le cas où  $F$  a un unique minimiseur à la différence près que le coefficient  $\tilde{C}_2(k, \alpha, \kappa)$  est légèrement plus grand ici (sous hypothèse d'unicité du minimiseur,  $\tilde{C}_2(k, \alpha, \kappa)$  est égal à  $\hat{C}_2(k, \alpha, \kappa)$ ). Par conséquent, les bornes uniformes sur  $\tilde{C}_1(k, \alpha, \kappa)$  calculées dans [26] sont toujours valides dans ce cas. Ainsi, on peut montrer qu'il existe deux constantes réelles  $\tilde{c}_1$  et  $\tilde{c}_2$  telles que pour tout  $\alpha \geq 3 + \frac{3}{\sqrt{2}}$  et  $k \geq \frac{4\alpha}{3\sqrt{\kappa}}$ ,

$$\tilde{C}_1(k, \alpha, \kappa) \leq \frac{5}{4} \sqrt{\frac{2}{\kappa}} P(\alpha) (1 + \tilde{c}_1 \sqrt{\kappa} + \tilde{c}_2 \kappa), \quad (5.4.68)$$

où  $P : \alpha \mapsto \frac{2}{9}(\alpha - 3) \left(\frac{8}{5}\alpha - 3\right) - 1$ . De plus, on a pour tout  $k \geq \frac{4\alpha}{3\sqrt{\kappa}}$  et  $\alpha \geq 3$ ,

$$\begin{aligned} \tilde{C}_2(k, \alpha, \kappa) &= \left(\frac{2\alpha}{3} - 1\right) \left(8 + \frac{\sqrt{2}}{\sqrt{\kappa}}\right) \left(1 + \frac{4\alpha^2}{9\kappa(k+1)^2}\right) \\ &\leq \frac{5}{4} \sqrt{\frac{2}{\kappa}} \left(\frac{2\alpha}{3} - 1\right) (1 + 4\sqrt{2\kappa}). \end{aligned}$$

Ainsi, pour tout  $\alpha \geq 3 + \frac{3}{\sqrt{2}}$  :

$$\forall k \geq \frac{4\alpha}{3\sqrt{\kappa}}, \quad E_{k+1} - \left(1 - \frac{\frac{2\alpha}{3} - 2}{k}\right) E_k \leq \frac{\mathbf{C}_1(\alpha, \kappa) E_k}{k^2} + \frac{\mathbf{C}_2(\alpha, \kappa) E_{k+1}}{(k+1)^2}, \quad (5.4.69)$$

où :

$$- \mathbf{C}_1(\alpha, \kappa) = \frac{5}{4} \sqrt{\frac{2}{\kappa}} \left[\frac{2}{9}(\alpha - 3) \left(\frac{8}{5}\alpha - 3\right) - 1\right] (1 + \sqrt{\kappa})^2 (1 + \tilde{c}_1 \sqrt{\kappa} + \tilde{c}_2 \kappa),$$

$$- \mathbf{C}_2(\alpha, \kappa) = \frac{5}{4} \sqrt{\frac{2}{\kappa}} \left( \frac{2\alpha}{3} - 1 \right) (1 + \sqrt{\kappa})^2 (1 + 4\sqrt{2\kappa}).$$

À partir de ce point la preuve suit exactement les calculs de [26]. On intègre d'abord l'inégalité grâce au lemme suivant qui est une version légèrement modifiée de [26, Lemme 2].

**Lemme 5.4.12.** *Soit  $\alpha \geq 3$  and  $k_0 \geq \frac{4\alpha}{3\sqrt{\kappa}}$ . Si l'énergie  $E_k$  vérifie (5.4.69), alors :*

$$\forall k \geq k_0, E_k \leq E_{k_0} \left( \frac{k}{k_0} \right)^{-(\frac{2\alpha}{3}-2)} e^{\phi(k_0)}, \quad (5.4.70)$$

où  $\phi(k_0) = \frac{5}{6k_0} \sqrt{\frac{2}{\kappa}} (\alpha - 3) \left( \frac{16}{15}\alpha - 1 \right) \left( 1 + c\kappa^{\frac{1}{4}} \right)$  et  $c > 0$  est indépendant de  $\alpha$ .

La preuve de ce lemme est identique à celle de [26, Lemme 2] malgré la légère modification de la constante  $\mathbf{C}_2(\alpha, \kappa)$ .

Comme  $F(x_k) - F^* \leq \frac{L}{2k^2} E_k$ , on trouve que pour tout  $k \geq \frac{4\alpha}{3\sqrt{\kappa}}$ ,

$$F(x_k) - F^* \leq \frac{L}{2} \left( k_0^{\frac{2\alpha}{3}-2} e^{\phi(k_0)} \right) E_{k_0} k^{-\frac{2\alpha}{3}}.$$

On peut alors choisir  $k_0$  de manière à obtenir un contrôle fin sur  $F(x_k) - F^*$ . En suivant le raisonnement donné dans [26], on fait le choix

$$k_0 = \frac{5}{4} \sqrt{\frac{2}{\kappa}} \left( \frac{16}{15}\alpha - 1 \right) \left( 1 + c\kappa^{\frac{1}{4}} \right),$$

et on peut montrer que l'on a pour  $\kappa$  suffisamment petit :

$$\forall k \geq \frac{3\alpha}{\sqrt{\kappa}}, F(x_k) - F^* \leq \frac{9}{4} e^{-2} M_0 \left( \frac{8e}{3\sqrt{\kappa}} \alpha \right)^{\frac{2\alpha}{3}} k^{-\frac{2\alpha}{3}}. \quad (5.4.71)$$

où  $M_0 = F(x_0) - F^*$ .

Démontrons maintenant le deuxième point du Théorème 5.4.5. On rappelle l'inégalité suivante prouvée dans la Section 5.A.11

$$\delta_k \leq \frac{2}{(k - \lambda)^2} b_k + \frac{8\alpha^2}{9(k - \lambda)^2} h_k, \quad (5.4.72)$$

où  $b_k = \|\lambda(x_{k-1} - x_{k-1}^*) + k(x_k - x_{k-1})\|^2$ . En revenant à la définition de l'énergie  $E_k$ , on obtient

$$\delta_k \leq \frac{2}{(k - \lambda)^2} E_k + \frac{8\alpha^2}{9(k - \lambda)^2} h_k, \quad (5.4.73)$$

et comme  $F$  satisfait  $\mathcal{G}_\mu^2$ , on a  $h_k \leq \frac{E_k}{\kappa k^2}$  et donc

$$\delta_k \leq \frac{2}{(k - \lambda)^2} \left( 1 + \frac{4\alpha^2}{9\kappa k^2} \right) E_k. \quad (5.4.74)$$

Ainsi, pour tout  $k \geq \frac{4\alpha}{3\sqrt{\kappa}}$ ,  $\delta_k \leq \frac{5}{2(k-\lambda)^2} E_k$ . En appliquant le Lemme 5.4.12, on trouve que pour un certain  $K > 0$ ,

$$\delta_k \leq \frac{K}{k^{\frac{2\alpha}{3}}}, \quad (5.4.75)$$

et ainsi

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(k^{-\frac{\alpha}{3}}\right). \quad (5.4.76)$$

□

### 5.4.3.5 Preuve du Théorème 5.4.6

**Structure de la preuve** Les résultats du Théorème 5.4.6 sont obtenus à partir de l'analyse de l'énergie discrète ci-dessous

$$\mathcal{E}_k = \frac{2k^2}{L} (F(x_k) - F^*) + \|\lambda(x_k - x_k^*) + k\alpha_k(x_k - x_{k-1})\|^2 + \xi \|x_k - x_k^*\|^2 + \lambda k \alpha_k^2 \|x_k - x_{k-1}\|^2, \quad (5.4.77)$$

où  $\lambda > 0$  et  $\xi < 0$ . En reprenant les notations introduites dans 5.4.16, on a

$$\mathcal{E}_k = k^2 w_k + b_k + \xi h_k + \lambda k \alpha_k^2 \delta_k, \quad (5.4.78)$$

où  $b_k = \|\lambda(x_k - x_k^*) + k\alpha_k(x_k - x_{k-1})\|^2$ .

L'idée de cette preuve est de montrer que cette énergie de Lyapunov se comporte asymptotiquement comme  $k^{-\frac{4}{\gamma-2}}$ . Cela ne suffit cependant pas à conclure puisque  $\mathcal{E}$  fait intervenir un terme négatif paramétré par  $\xi$ . Ainsi, la fin de la preuve consiste à exploiter l'hypothèse  $\mathcal{G}^\gamma$  de manière à obtenir le résultat désiré. On peut remarquer que la démonstration suit les mêmes principes que celle donnée en Section 5.A.5 dans le cas continu.

Pour étudier le comportement asymptotique de  $\mathcal{E}$ , on définit  $\mathcal{H}$  qui pour tout  $k \in \mathbb{N}$  vérifie  $\mathcal{H}_k = k^p \mathcal{E}_k$  avec  $p = 1 + \frac{4}{\gamma-2}$ . On suit les étapes suivantes :

- En utilisant les propriétés des itérés de (FISTA) et la convexité de  $F$ , on montre que pour un jeu de paramètres  $(\alpha, \lambda, \xi)$  bien choisi et  $k$  suffisamment grand,

$$\mathcal{H}_{k+1} - \mathcal{H}_k \leq A(k+1)^{p+1} w_{k+1} + B(k+1)^{p-1} h_{k+1}, \quad (5.4.79)$$

avec  $A < 0$  et  $B > 0$ .

- À partir de l'inégalité précédente et l'hypothèse de croissance vérifiée par  $F$ , on prouve que pour  $k$  suffisamment grand,

$$\mathcal{H}_k \leq Ck, \quad (5.4.80)$$

pour un certain  $C > 0$ . Cette inégalité garantit que  $\mathcal{E}_k$  décroît asymptotiquement comme  $k^{-\frac{4}{\gamma-2}}$ .

- En revenant à la définition de  $\mathcal{H}_k$  et  $\mathcal{E}_k$  et grâce à l'hypothèse  $\mathcal{G}^\gamma$ , on montre que  $k^{p+1} w_k$  et  $\alpha_k^2 k^{p+1} \delta_k$  sont des suites bornées. Ceci permet de conclure.

**Preuve** Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie par (FISTA) et  $(\mathcal{E}_k)_{k \in \mathbb{N}}$  l'énergie définie par (5.4.77). La première étape de la preuve est d'obtenir une borne supérieure sur  $\mathcal{E}_{k+1} - \mathcal{E}_k$ . On donne une telle inégalité dans le lemme ci-dessous dont la preuve se situe en Section 5.A.12.

**Lemme 5.4.13.** *Soit  $\xi = \lambda(\lambda + 1 - \alpha)$ . Pour tout  $k \in \mathbb{N}^*$ ,*

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k \leq & ((2 - \lambda)k + 1) w_{k+1} + B_1(k)b_{k+1} + B_2(k)h_{k+1} + B_3(k)\delta_{k+1} \\ & - B_4(k) (\gamma_{k+1}^* - 2\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle), \end{aligned} \quad (5.4.81)$$

où :

$$\begin{aligned} - B_1(k) &= \frac{2(\lambda+1-\alpha)}{k+1} + \frac{\alpha(2\lambda+2-\alpha)}{(k+1)^2}, \\ - B_2(k) &= -\frac{2\lambda^2(\lambda+1-\alpha)}{k+1} + \frac{\alpha\lambda^2(\alpha-2\lambda-2)}{(k+1)^2}, \\ - B_3(k) &= \alpha(\lambda+2) - (2\lambda^2 + 2\lambda + 1) + \alpha^2 \frac{k(\lambda-2) + \lambda - 2 - 2\alpha}{(k+1-\alpha)^2}, \\ - B_4(k) &= -2\lambda(\lambda+1-\alpha) - \frac{\alpha^2\lambda}{k+1+\alpha}. \end{aligned}$$

On introduit l'énergie  $\mathcal{H}$  définie par  $\mathcal{H}_k = k^p \mathcal{E}_k$  avec  $p = 1 + \frac{4}{\gamma-2}$ . L'étape suivante est de démontrer l'inégalité ci-dessous. On donne la démonstration de ce lemme en Section 5.A.13.

**Lemme 5.4.14.** *Soit  $\xi = \lambda(\lambda + 1 - \alpha)$ . Si  $\lambda \leq \alpha - 1$ , alors pour tout  $k \in \mathbb{N}^*$ ,*

$$\begin{aligned} \mathcal{H}_{k+1} - \mathcal{H}_k \leq & ((2 - \lambda + p)(k + 1)^{p+1} + R_1(k)) w_{k+1} \\ & + ((2(\lambda + 1 - \alpha) + p)(k + 1)^{p-1} + R_2(k)) b_{k+1} \\ & + (\lambda(\lambda + 1 - \alpha)(p - 2\lambda)(k + 1)^{p-1} + R_3(k)) h_{k+1} \\ & - k^p B_4(k) (\gamma_{k+1}^* - 2\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle), \end{aligned} \quad (5.4.82)$$

avec :

$$\begin{aligned} - R_1(k) &= (\lambda - 1 + p(\lambda - 2)) k^p + p(\lambda - 2)k^{p-1}, \\ - R_2(k) &= ((4\lambda + 6 + 2p - \alpha)\alpha + 2\lambda p) (k + 1)^{p-2} + \alpha^2(6\lambda + 8 + p)(k + 1)^{p-3} + \\ & 2\alpha^3(\lambda + 2)(k + 1)^{p-4}, \\ - R_3(k) &= (\lambda^2(\alpha^2 + 2\alpha + 4\lambda p + p + 1) + \lambda(\alpha - 1)(p - 1)) (k + 1)^{p-2} + \\ & \lambda^2\alpha(2p\lambda + 2p + 6\lambda\alpha + 2\alpha) (k + 1)^{p-3} + 2\alpha^3\lambda^2(\lambda + 2)(k + 1)^{p-4}. \end{aligned}$$

Si on choisit  $\lambda = \alpha - 1 - p$ , on trouve que  $\lambda \leq \alpha - 1$  et

$$\begin{cases} C_1 := 2 - \lambda + p = 3 + 2p - \alpha, \\ C_2 := 2(\lambda + 1 - \alpha) + p = -p, \\ C_3 := \lambda(\lambda + 1 - \alpha)(p - 2\lambda) = p(\alpha - 1 - p)(2\alpha - 2 - 3p), \end{cases} \quad (5.4.83)$$

ce qui implique que pour  $\alpha > 3 + 2p = 5 + \frac{8}{\gamma-2}$ ,

$$\begin{cases} C_1 < 0, \\ C_2 < 0, \\ C_3 > 0. \end{cases} \quad (5.4.84)$$

Étant donné l'ordre des termes  $R_1(k)$ ,  $R_2(k)$  et  $R_3(k)$ , ces inégalités assurent que pour  $k$  suffisamment grand :

$$\begin{cases} C_1(k+1)^{p+1} + R_1(k) < \frac{C_1}{2}(k+1)^{p+1}, \\ C_2(k+1)^{p-1} + R_2(k) < 0, \\ C_3(k+1)^{p-1} + R_3(k) < 2C_3(k+1)^{p-1}. \end{cases} \quad (5.4.85)$$

De plus, pour le choix  $\lambda = \alpha - 1 - p$  on a  $B_4(k) = 2p(\alpha - 1 - p) - \alpha^2 \frac{\alpha-1-p}{k+1+\alpha}$  qui est une suite positive lorsque  $\alpha > 5 + \frac{8}{\gamma-2}$  et  $k$  est suffisamment grand. Comme  $\gamma_{k+1}^* \geq 0$  et  $\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle \leq 0$ , on trouve que

$$k^p B_4(k) (\gamma_{k+1}^* - 2\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle) \geq 0. \quad (5.4.86)$$

Ainsi, si  $\alpha > 5 + \frac{8}{\gamma-2}$ , alors pour  $k$  suffisamment grand :

$$\mathcal{H}_{k+1} - \mathcal{H}_k \leq \frac{C_1}{2}(k+1)^{p+1} w_{k+1} + 2C_3(k+1)^{p-1} h_{k+1} \quad (5.4.87)$$

La suite de la preuve s'appuie sur la géométrie de  $F$  et en particulier sur le lemme ci-dessous qui est obtenu à partir de l'hypothèse  $\mathcal{G}^\gamma$ .

**Lemme 5.4.15.** *Soit  $F$  une fonction vérifiant les hypothèses du Théorème 5.4.6. Si  $p = 1 + \frac{4}{\gamma-2}$ ,*

$$k^{p-1} h_k \leq \left( \frac{L}{2K} \right)^{\frac{2}{\gamma}} (k^{p+1} w_k)^{\frac{2}{\gamma}}, \quad (5.4.88)$$

où  $K > 0$  est une constante permettant de vérifier pour tout  $x \in \mathbb{R}^n$ ,

$$Kd(x, X^*)^\gamma \leq F(x) - F^*. \quad (5.4.89)$$

En appliquant ce lemme, on trouve que pour  $k$  suffisamment grand,

$$\mathcal{H}_{k+1} - \mathcal{H}_k \leq \frac{C_1}{2}(k+1)^{p+1} w_{k+1} + 2C_3 \left( \frac{L}{2K} \right)^{\frac{2}{\gamma}} ((k+1)^{p+1} w_{k+1})^{\frac{2}{\gamma}}, \quad (5.4.90)$$

et il existe donc une constante  $M_0 \in \mathbb{R}$  telle que  $\mathcal{H}_{k+1} - \mathcal{H}_k \leq M_0$ .

Ainsi, il existe  $k_0 \in \mathbb{N}$  tel que  $k \geq k_0$ ,  $\mathcal{H}_k \leq nM_0 + M_1$  avec  $M_1 = \mathcal{H}_{k_0} - k_0 M_0$ . Pour  $k$  suffisamment grand, on a donc  $\mathcal{H}_k \leq 2kM_0$ . En revenant à la définition de  $\mathcal{H}$ , cela implique que :

$$k^{p-1} (k^2 w_k + b_k + \xi h_k + \lambda k \alpha_k^2 \delta_k) \leq 2M_0. \quad (5.4.91)$$

Comme  $\xi = \lambda(\lambda + 1 - \alpha) < 0$ , cela garantit que pour  $k$  suffisamment grand,

$$k^{p+1} w_k - |\xi| k^{p-1} h_k \leq 2M_0, \quad (5.4.92)$$

et en appliquant le Lemme 5.4.15, on a

$$k^{p+1}w_k - |\xi| \left( \frac{L}{2K} \right)^{\frac{2}{\gamma}} (k^{p+1}w_k)^{\frac{2}{\gamma}} \leq 2M_0. \quad (5.4.93)$$

Le lemme ci-dessous permet d'écrire que pour  $k$  suffisamment grand, la suite  $k^{p+1}w_k$  est bornée.

**Lemme 5.4.16.** *Soit  $x \in \mathbb{R}^+$ ,  $\delta \in (0, 1)$ ,  $K_1 > 0$  et  $K_2 > 0$ . On a :*

$$x^\delta (x^{1-\delta} - K_1) \leq K_2 \quad \implies \quad x \leq \left( K_2^{1-\delta} + K_1 \right)^{\frac{1}{1-\delta}}. \quad (5.4.94)$$

Par conséquent, il existe une constante  $M_2 > 0$  telle que pour  $k$  suffisamment grand,  $k^{p+1}w_k \leq M_2$  et en remplaçant  $p$  par sa valeur,

$$F(x_k) - F^* \leq \frac{LM_2}{2k^{\frac{2\gamma}{\gamma-2}}}. \quad (5.4.95)$$

Cela prouve le premier point du théorème.

L'inégalité (5.4.91) implique également que

$$k^{p-1} (k^2w_k - |\xi|h_k + b_k) \leq 2M_0. \quad (5.4.96)$$

En appliquant l'inégalité  $\|u\|^2 \leq 2\|u+v\|^2 + 2\|v\|^2$  à  $u = \alpha_k(x_k - x_{k-1})$  et  $v = \lambda(x_k - x_k^*)$ , on obtient

$$b_k \geq \frac{k^2\alpha_k^2}{2}\delta_k - \lambda^2h_k. \quad (5.4.97)$$

En combinant cette inégalité à (5.4.96), on a

$$(k^{p+1}w_k - (\lambda^2 + |\xi|)k^{p-1}h_k) + \frac{\alpha_k^2}{2}k^{p+1}\delta_k \leq 2M_0. \quad (5.4.98)$$

Le Lemme 5.4.15 assure ensuite que

$$k^{p+1}w_k - (\lambda^2 + |\xi|)k^{p-1}h_k \geq k^{p+1}w_k - \left( \frac{L}{2K} \right)^{\frac{2}{\gamma}} (k^{p+1}w_k)^{\frac{2}{\gamma}} \geq M_3, \quad (5.4.99)$$

pour une constante  $M_3 \in \mathbb{R}$ .

Ainsi, pour  $k$  suffisamment grand

$$\delta_k \leq \frac{4M_0 - 2M_3}{\alpha_k^2 k^{p+1}}, \quad (5.4.100)$$

et donc

$$\|x_k - x_{k-1}\| = \mathcal{O}\left(k^{-\frac{\gamma}{\gamma-2}}\right). \quad (5.4.101)$$

## 5.A Preuves techniques

### 5.A.1 Preuve du Lemme 5.3.1

Soit  $\phi'$  la dérivée de  $\phi$  quand elle est bien définie. D'après GC. Young [135], comme  $\phi$  est continue et est dérivable à droite et à gauche, la fonction  $\phi$  est dérivable partout sauf en un ensemble dénombrable de points. Cela implique qu'il existe  $(t_i)_{i \in \llbracket 1, N \rrbracket}$  et  $N \in \mathbb{N}^* \cup \{+\infty\}$  tels que pour tout  $i \in \llbracket 0, N-1 \rrbracket$  et  $t \in (t_i, t_{i+1})$ ,  $\phi'(t)$  est bien défini et vaut  $\phi_+(t)$ . On choisit cette suite de manière à ce qu'elle soit ordonnée c'est-à-dire  $t_0 < t_i < t_{i+1}$  pour tout  $i$  avec  $t_N = +\infty$  quand  $N \neq +\infty$ .

Supposons que  $t \in (t_0, t_1)$ .

- Si  $\phi$  est dérivable en  $t_0$ , alors  $\phi$  est dérivable sur l'intervalle  $[t_0, t_1)$  et  $\phi' = \phi_+$  sur cet intervalle. En conséquence, l'inégalité (5.3.12) garantit que,

$$\phi(t) \leq \phi(t_0) + \int_{t_0}^t \psi(u) du.$$

- Si  $\phi$  n'est pas dérivable en  $t_0$ , alors l'inégalité (5.3.12) assure que pour  $h > 0$  suffisamment petit,

$$\phi(t_0 + h) \leq \phi(t_0) + h\psi(t_0).$$

La discussion du point précédent permet de déduire que  $\phi$  est dérivable sur  $[t_0 + h, t_1)$ . Ainsi, on déduit qu'il existe  $H \in (0, t - t_0)$  tel que pour tout  $h \in (0, H)$  :

$$\phi(t) \leq \phi(t_0 + h) + \int_{t_0 + h}^t \psi(u) du \leq \phi(t_0) + \int_{t_0}^t \psi(u) du + \int_{t_0}^{t_0 + h} (\psi(t_0) - \psi(u)) du.$$

Comme cette inégalité est vérifiée pour tout  $h \in (0, H)$ , on obtient finalement l'inégalité (5.3.13).

On suppose maintenant que  $t = t_1$ . On a prouvé ci-dessus que (5.3.13) est vrai pour tout  $t \in (t_0, t_1)$ . Par conséquent, pour tout  $t \in (t_0, t_1)$ ,

$$\phi(t) \leq \phi(t_0) + \int_{t_0}^t \psi(u) du,$$

et comme  $\phi$  est une fonction continue, l'inégalité est toujours vraie pour  $t = t_1$ .

En utilisant les mêmes arguments, on peut montrer que l'équation (5.3.13) est vraie pour tout  $t > t_1$ . En effet, si  $t > t_1$ , alors  $t \in (t_i, t_{i+1})$  ou  $t = t_i$  pour un certain  $i \in \llbracket 1, N \rrbracket$ . Dans les deux cas, on obtient le résultat voulu en réappliquant le même raisonnement sur les intervalles consécutifs  $(t_j, t_{j+1})$  pour  $0 \leq j \leq i$ .

### 5.A.2 Preuve du Théorème 5.3.2

On s'appuie sur la démonstration de [25, Théorème 1]. Le premier point est une conséquence directe de la discussion de la Section 5.3.1 combinée aux calculs

de [25]. Le deuxième point demande quelques calculs supplémentaires. On choisit  $\alpha = \left(2 - \frac{\sqrt{2}}{2}\right) \sqrt{\mu}$  et on considère l'énergie de Lyapunov suivante :

$$\mathcal{E}(t) = F(x(t)) - F^* + \frac{1}{2} \|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2 + \xi \|x(t) - x^*(t)\|^2, \quad (5.A.1)$$

où  $\lambda = \sqrt{\mu}$  et  $\xi = -\left(1 - \frac{\sqrt{2}}{2}\right) \mu$ .

En suivant le raisonnement énoncé dans la Section 5.3.1 et la preuve de [25, Theorem 1], l'hypothèse vérifiée par  $X^*$  permet d'écrire qu'il existe  $t_1 > t_0$  tel que :

$$\forall t \geq t_1, \mathcal{E}(t) \leq \mathcal{E}(t_1) e^{-(2-\sqrt{2})\sqrt{\mu}(t-t_1)}. \quad (5.A.2)$$

Comme  $F$  vérifie  $\mathcal{G}_\mu^2$ , on obtient que :

$$F(x(t)) - F^* + \xi \|x(t) - x^*(t)\|^2 \geq (\sqrt{2} - 1)(F(x(t)) - F^*), \quad (5.A.3)$$

et donc,

$$\forall t \geq t_1, \mathcal{E}(t) \geq (\sqrt{2} - 1) (F(x(t)) - F^*) + \frac{1}{2} \|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2. \quad (5.A.4)$$

Cette inégalité implique que pour tout  $t \geq t_1$  :

$$F(x(t)) - F^* \leq \frac{1}{\sqrt{2} - 1} \mathcal{E}(t), \quad (5.A.5)$$

et

$$\|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2 \leq 2\mathcal{E}(t). \quad (5.A.6)$$

En appliquant l'inégalité  $\|u\|^2 \leq 2\|u + v\|^2 + 2\|v\|^2$ , on trouve que :

$$\begin{aligned} \|\dot{x}(t)\|^2 &\leq 2\|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2 + 2\mu\|x(t) - x^*(t)\|^2 \\ &\leq 2\|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2 + 4(F(x(t)) - F^*). \end{aligned} \quad (5.A.7)$$

Les inégalités ci-dessus assurent que :

$$\|\dot{x}(t)\|^2 \leq 4 \left(1 + \frac{1}{\sqrt{2} - 1}\right) \mathcal{E}(t). \quad (5.A.8)$$

La borne sur l'énergie donnée dans (5.A.2) amène à la conclusion que :

$$\|\dot{x}(t)\| = \mathcal{O} \left( e^{-\left(1 - \frac{\sqrt{2}}{2}\right)t} \right). \quad (5.A.9)$$

### 5.A.3 Preuve du Théorème 5.3.3

De la même manière que dans la preuve précédente, on s'appuie sur la preuve originale de [24, Théorème 3.2] pour déduire un contrôle sur  $\|\dot{x}(t)\|$ . L'analyse du papier mentionné est construite autour de l'énergie de Lyapunov ci-dessous :

$$\mathcal{E}(t) = F(x(t)) - F^* + \frac{1}{2} \|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2 + \xi \|x(t) - x^*(t)\|^2, \quad (5.A.10)$$



avec  $\lambda = \frac{2\alpha}{3}$  et  $\xi = -\frac{\lambda^2}{2}$ .

La condition de géométrie vérifiée par  $X^*$  permet de déduire les mêmes contrôles sur l'énergie  $\mathcal{E}$  que dans la preuve originale en suivant le raisonnement présenté en Section 5.3.1. En particulier, on peut montrer que pour tout  $t \geq t_0$ ,

$$\mathcal{E}(t) \leq \mathcal{E}(t_0)e^{-\frac{2\alpha}{3}(t-t_0)}, \quad (5.A.11)$$

et de plus,

$$\mathcal{E}(t) \geq \frac{1}{2} \|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2. \quad (5.A.12)$$

Ainsi, on peut évaluer l'inégalité  $\|u\|^2 \leq 2\|u + v\|^2 + 2\|v\|^2$  pour  $u = \dot{x}(t)$  et  $v = \lambda(x(t) - x^*(t))$  pour obtenir

$$\|\dot{x}(t)\|^2 \leq 2\|\lambda(x(t) - x^*(t)) + \dot{x}(t)\|^2 + 2\lambda^2\|x(t) - x^*(t)\|^2, \quad (5.A.13)$$

et donc

$$\frac{1}{4}\|\dot{x}(t)\|^2 \leq \mathcal{E}(t_0)e^{-\frac{2\alpha}{3}(t-t_0)} + \frac{\lambda^2}{2}\|x(t) - x^*(t)\|^2. \quad (5.A.14)$$

En reprenant les calculs de [24], on trouve que pour tout  $t \geq t_0$

$$\frac{\lambda^2}{2}\|x(t) - x^*(t)\|^2 \leq 2 \left( \sqrt{2\mathcal{E}(t_0)} + \frac{\lambda}{2}\|x(t_0) - x^*(t_0)\|e^{-\frac{\lambda}{2}(t-t_0)} \right)^2 e^{-\frac{2\alpha}{3}(t-t_0)}, \quad (5.A.15)$$

ce qui garantit donc que quand  $t$  tend vers  $+\infty$  on a :

$$\|\dot{x}(t)\| = \mathcal{O}\left(e^{-\frac{\alpha}{3}t}\right). \quad (5.A.16)$$

#### 5.A.4 Preuve du Théorème 5.3.4

En suivant la preuve de [26, Theorem 5], on étudie l'énergie de Lyapunov ci-dessous :

$$\mathcal{E}(t) = t^2(F(x(t)) - F^*) + \frac{1}{2}\|\lambda(x(t) - x^*(t)) + t\dot{x}(t)\|^2, \quad (5.A.17)$$

où  $\lambda = \frac{2\alpha}{3}$ . En appliquant le raisonnement décrit en Section 5.3.1 à la preuve de [26, Theorem 5], on trouve qu'il existe  $t_1 \geq t_0$  et  $K > 0$  tels que

$$\forall t \geq t_1, \mathcal{E}(t) \leq Kt^{-\frac{2\alpha}{3}+2}. \quad (5.A.18)$$

En appliquant l'inégalité  $\|u\|^2 \leq 2\|u + v\|^2 + 2\|v\|^2$ , on obtient que

$$\|\dot{x}(t)\|^2 \leq \frac{2}{t^2}\|\lambda(x(t) - x^*(t)) + t\dot{x}(t)\|^2 + \frac{2\lambda^2}{t^2}\|x(t) - x^*(t)\|^2. \quad (5.A.19)$$

Ainsi, l'hypothèse  $\mathcal{G}_\mu^2$  assure que

$$\|\dot{x}(t)\|^2 \leq \frac{4}{t^2} \left( 1 + \frac{\lambda^2}{\mu t^2} \right) \mathcal{E}(t). \quad (5.A.20)$$

L'inégalité (5.A.18) donne que

$$\|\dot{x}(t)\| = \mathcal{O}\left(t^{-\frac{\alpha}{3}}\right). \quad (5.A.21)$$

□

## 5.A.5 Preuve du Théorème 5.3.5

On s'appuie sur l'analyse proposée dans [123]. La preuve est construite de l'énergie définie ci-dessous

$$\mathcal{H}(t) = t^p \left( t^2(F(x(t)) - F^*) + \frac{1}{2} \|\lambda(x(t) - x^*(t)) + t\dot{x}(t)\|^2 + \frac{\xi}{2} \|x(t) - x^*(t)\|^2 \right), \quad (5.A.22)$$

où  $p = 1 + \frac{4}{\gamma-2}$  et  $\lambda > 0$ . On utilisera les notations suivantes :

$$\begin{aligned} a(t) &= t(F(x(t)) - F^*), \quad b(t) = \frac{1}{2t} \|\lambda(x(t) - x^*(t)) + t\dot{x}(t)\|^2 \\ c(t) &= \frac{1}{2t} \|x(t) - x^*(t)\|^2. \end{aligned}$$

On peut donc réécrire  $\mathcal{H}$  comme suit :

$$\mathcal{H}(t) = t^{p+1} (a(t) + b(t) + \xi c(t)).$$

En reprenant la discussion sur la dérivabilité de l'opérateur de projection sur  $X^*$  donnée en Section 5.3.1, on peut écrire que sous l'hypothèse qui est faite sur  $X^*$ , l'énergie  $\mathcal{H}$  est dérivable à droite. Par souci de simplicité, la dérivée à droite de  $\mathcal{H}$  est notée  $\mathcal{H}'$ . En adaptant [123, Lemme 4.4] à notre cas, on trouve que si  $\xi = \lambda(\lambda+1-\alpha)$ , alors

$$\mathcal{H}'(t) \leq t^p ((2+p-\lambda)a(t) + (2(\lambda+1-\alpha)+p)b(t) + \lambda(\lambda+1-\alpha)(p-2\lambda)c(t)).$$

Soit  $\lambda = \alpha - 1 - \frac{p}{2}$ . À condition que  $\alpha > \frac{9}{2} + \frac{6}{\gamma-2}$ , on a

$$\begin{cases} 2+p-\lambda < 0, \\ 2(\lambda+1-\alpha)+p = 0, \\ \lambda(\lambda+1-\alpha)(p-2\lambda) > 0. \end{cases}$$

En conséquence, on obtient que :

$$\mathcal{H}'(t) \leq t^p (Aa(t) + Cc(t)),$$

où  $A = 3 - \alpha + \frac{3p}{2} < 0$  et  $C = p(\alpha - 1 - \frac{p}{2})(\alpha - 1) > 0$ . On peut alors appliquer [123, Lemme 4.5] que l'on rappelle ci-dessous.

**Lemme 5.A.1.** *Si  $F$  vérifie l'hypothèse  $\mathcal{G}^\gamma$  pour  $\gamma > 2$ , alors en définissant  $p_2 = \frac{4}{\gamma-2}$  on a*

$$t^{p_2+1}c(t) \leq \frac{K^{-\frac{2}{\gamma}}}{2} (t^{p_2+1}a(t))^{\frac{2}{\gamma}},$$

où  $K > 0$  est une constante telle que pour tout  $x \in \mathbb{R}^N$  :

$$Kd(x, X^*)^\gamma \leq F(x) - F^*. \quad (5.A.23)$$

Par conséquent, pour tout  $m \in \mathbb{R}$ , il existe  $M \in \mathbb{R}$  tel que pour tout  $t \geq t_0$ ,

$$t^{p_2+1}(mc(t) - a(t)) \leq M.$$

Comme  $p = p_2 + 1$ , ce lemme garantit qu'il existe une constante  $M_1 \in \mathbb{R}$  telle que pour tout  $t \geq t_0$ ,  $\mathcal{H}'(t) \leq M_1$ . Ceci implique qu'il existe  $M_2 \in \mathbb{R}$  qui assure que  $\mathcal{H}(t) \leq M_1 t + M_2$  et donc

$$t^{p+1}(a(t) + \xi c(t)) \leq M_1 t + M_2.$$

Ainsi, pour  $t$  suffisamment grand,

$$t^p a(t) \leq 2M_1 + |\xi| t^p c(t).$$

Le premier point du Lemme 5.A.1 assure qu'il existe une constante  $M_3 > 0$  telle que :

$$t^p a(t) \leq 2M_1 + M_3 (t^p a(t))^{\frac{2}{\gamma}}. \quad (5.A.24)$$

Le Lemme 5.4.94 garantit qu'il existe une constante  $M_4 > 0$  telle que pour  $t$  suffisamment grand,

$$t^p a(t) \leq M_4,$$

et par conséquent,

$$F(x(t)) - F^* \leq \frac{M_4}{t^{p+1}}. \quad (5.A.25)$$

Comme  $p + 1 = \frac{2\gamma}{\gamma-2}$ , le premier point est démontré.

On prouve le deuxième résultat en réutilisant l'inégalité  $\mathcal{H}(t) \leq M_1 t + M_2$ . En appliquant l'inégalité  $\|u\|^2 \leq 2\|u + v\|^2 + 2\|v\|^2$  à  $u = t\dot{x}(t)$  et  $v = \lambda(x(t) - x^*(t))$ , on trouve que

$$b(t) \geq \frac{t}{2} \|\dot{x}(t)\|^2 - \lambda^2 c(t).$$

Donc pour  $t$  suffisamment grand,

$$t^p \left( a(t) - (|\xi| + \lambda^2) c(t) + \frac{t}{2} \|\dot{x}(t)\|^2 \right) \leq 2M_1.$$

Le Lemme 5.A.1 donne donc qu'il existe une constante  $M_5 > 0$  pour laquelle

$$t^p (a(t) - (|\xi| + \lambda^2) c(t)) \geq t^p a(t) - M_5 (t^p a(t))^{\frac{2}{\gamma}}.$$

**Lemme 5.A.2.** Soit  $g : x \mapsto x - Kx^\delta$  pour  $K > 0$  et  $\delta \in (0, 1)$ . Pour tout  $x \geq 0$ ,

$$g(x) \geq K(\delta - 1)(\delta K)^{\frac{\delta}{1-\delta}}. \quad (5.A.26)$$

Le Lemme 5.A.2 assure qu'il existe une constante  $M_6 \in \mathbb{R}$  telle que l'inégalité  $t^p (a(t) - (|\xi| + \lambda^2) c(t)) \geq M_6$  est valable. On a donc que pour  $t$  suffisamment grand,

$$\frac{t^{p+1}}{2} \|\dot{x}(t)\|^2 \leq 2M_1 + M_6,$$

et donc

$$\|\dot{x}(t)\| \leq \frac{M_7}{t^{\frac{p+1}{2}}}, \quad (5.A.27)$$

où  $M_7 = 4M_1 + 2M_6 \geq 0$  et  $\frac{p+1}{2} = \frac{\gamma}{\gamma-2}$ .

### 5.A.6 Preuve du Lemme 5.4.7

Soit  $k \in \mathbb{N}^*$ . En réécrivant

$$x_k - x_k^* = \frac{1}{2} \left( (x_k - x_{k-1}) + (x_{k-1} - x_{k-1}^*) + (x_{k-1}^* - x_k^*) + (x_k - x_k^*) \right),$$

on trouve que

$$\langle x_k - x_k^*, x_k - x_{k-1} \rangle = \frac{1}{2} \delta_k + \frac{1}{2} \langle (x_{k-1} - x_{k-1}^*) + (x_{k-1}^* - x_k^*) + (x_k - x_k^*), x_k - x_{k-1} \rangle.$$

Comme  $x_k - x_{k-1} = (x_k - x_k^*) + (x_k^* - x_{k-1}^*) + (x_{k-1}^* - x_{k-1})$ , on a :

$$\begin{aligned} 2\langle x_k - x_k^*, x_k - x_{k-1} \rangle &= \delta_k + \langle x_{k-1} - x_{k-1}^*, x_k - x_k^* \rangle + \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \\ &\quad - h_{k-1} - \langle x_k^* - x_{k-1}^*, x_k - x_k^* \rangle + \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \\ &\quad - \gamma_k^* + \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - \langle x_{k-1} - x_{k-1}^*, x_k - x_k^* \rangle + h_k \\ &= h_k - h_{k-1} + \delta_k - \gamma_k^* + 2\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned}$$

Le second point est prouvé avec la même approche. Pour tout  $k \in \mathbb{N}^*$ ,

$$x_{k-1} - x_{k-1}^* = \frac{1}{2} \left( (x_{k-1} - x_k) + (x_k - x_k^*) + (x_k^* - x_{k-1}^*) + (x_{k-1}^* - x_{k-1}) \right),$$

et donc :

$$2\langle x_{k-1} - x_{k-1}^*, x_k - x_{k-1} \rangle = -\delta_k + \langle (x_k - x_k^*) + (x_k^* - x_{k-1}^*) + (x_{k-1}^* - x_{k-1}), x_k - x_{k-1} \rangle.$$

En réécrivant  $x_k - x_{k-1}$  comme précédemment, on trouve finalement que :

$$\langle x_{k-1} - x_{k-1}^*, x_k - x_{k-1} \rangle = \frac{1}{2} (h_k - h_{k-1} - \delta_k + \gamma_k^*) + \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle.$$

### 5.A.7 Preuve du Lemme 5.4.8

Le premier point est obtenu directement à partir de [51, Lemme 3.1] qui assure que :

$$F(x_{k+1}) - F(x_k) \leq \frac{L}{2} (\|y_k - x_k\|^2 - \|x_{k+1} - x_k\|^2).$$

En écrivant que  $y_k = x_k + \alpha_k(x_k - x_{k-1})$  et  $\frac{2}{L}(F(x_{k+1}) - F(x_k)) = w_{k+1} - w_k$ , on obtient le résultat souhaité.

En appliquant [51, Lemme 3.1] à deux autres points, on trouve que

$$F(x_{k+1}) - F^* \leq \frac{L}{2} (\|y_k - x_k^*\|^2 - \|x_{k+1} - x_k^*\|^2).$$

Par conséquent :

$$\begin{aligned} w_{k+1} &\leq \|x_k + \alpha_k(x_k - x_{k-1}) - x_k^*\|^2 - \|(x_{k+1} - x_{k+1}^*) + (x_{k+1}^* - x_k^*)\|^2 \\ &\leq h_k + \alpha_k^2 \delta_k - h_{k+1} - \gamma_{k+1}^* + 2\alpha_k \langle x_k - x_k^*, x_k - x_{k-1} \rangle \\ &\quad - 2\langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle. \end{aligned}$$

Rappelons que le premier point du Lemme 5.4.7 assure que :

$$\langle x_k - x_k^*, x_k - x_{k-1} \rangle = \frac{1}{2}(h_k - h_{k-1} + \delta_k - \gamma_k^*) + \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle,$$

et donc :

$$\begin{aligned} w_{k+1} &\leq (1 + \alpha_k)h_k + (\alpha_k^2 + \alpha_k)\delta_k - \alpha_k h_{k-1} - h_{k+1} - \gamma_{k+1}^* - \alpha_k \gamma_k^* \\ &\quad + 2\alpha_k \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - 2\langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle. \end{aligned}$$

### 5.A.8 Preuve du Lemme 5.4.9

Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie dans (V-FISTA) pour  $s = \frac{1}{L}$ . On peut écrire que l'énergie  $(\mathcal{E}_k)_{k \in \mathbb{N}}$  de la façon suivante :

$$\mathcal{E}_k = w_k + (1 - \lambda)\delta_k + \lambda(h_k - h_{k-1}) + \lambda^2 h_{k-1} + \lambda \gamma_k^* + 2\lambda \langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle.$$

Comme  $\langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle \geq 0$ , on a que

$$\mathcal{E}_k \geq w_k + \lambda(h_k - h_{k-1}),$$

ce qui conduit à la conclusion.

### 5.A.9 Preuve du Lemme 5.4.10

Soit  $(x_k)_{k \in \mathbb{N}}$  la suite définie par (V-FISTA). En utilisant l'expression (5.4.42) de  $\mathcal{E}_k$ , on trouve que :

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &= w_{k+1} - w_k + \lambda(\alpha + \lambda\alpha + (1 - \alpha)^2)(h_{k+1} - h_k) + \alpha(1 + \lambda)\delta_{k+1} \\ &\quad - \lambda\alpha(h_k - h_{k-1}) - \alpha(1 + \lambda)\delta_k - \lambda\alpha\gamma_{k+1}^* + \lambda\alpha\gamma_k^* \\ &\quad + 2\lambda\alpha \langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle - 2\lambda\alpha \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \tag{5.A.28}$$

Le premier point du Lemme 5.4.8 combiné à l'inégalité  $-\lambda\alpha\gamma_{k+1}^* + 2\lambda\alpha \langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle \leq 0$  assure que

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &\leq \lambda(\alpha + \lambda\alpha + (1 - \alpha)^2)(h_{k+1} - h_k) + (\alpha + \lambda\alpha - 1)\delta_{k+1} \\ &\quad - \lambda\alpha(h_k - h_{k-1}) - \alpha(1 + \lambda - \alpha)\delta_k + \lambda\alpha\gamma_k^* \\ &\quad - 2\lambda\alpha \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \tag{5.A.29}$$

D'après le deuxième point du Lemme 5.4.8, on a

$$\begin{aligned} 0 &\leq -\lambda w_{k+1} + \lambda(\alpha + \alpha^2)\delta_k + \lambda\alpha(h_k - h_{k-1}) - \lambda(h_{k+1} - h_k) \\ &\quad - \lambda\gamma_{k+1}^* - \lambda\alpha\gamma_k^* + 2\lambda\alpha \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle \\ &\quad - 2\lambda \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle, \end{aligned} \tag{5.A.30}$$

et comme  $\langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle \geq 0$ , on trouve que

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &\leq -\lambda w_{k+1} + \lambda(\alpha + \lambda\alpha + (1 - \alpha)^2 - 1)(h_{k+1} - h_k) \\ &\quad + (\alpha + \lambda\alpha - 1)\delta_{k+1} - \alpha(1 - \alpha - \lambda\alpha)\delta_k. \end{aligned} \tag{5.A.31}$$

En développant  $\alpha + \lambda\alpha + (1 - \alpha)^2 - 1$  on arrive à la conclusion souhaitée.

### 5.A.10 Preuve du Corollaire 5.4.4

Soit  $(x_k)_{k \in \mathbb{N}}$  la suite donnée par (V-FISTA) pour  $\alpha > 0$  et  $s = \frac{1}{L}$ . D'après le Théorème 5.4.2, si  $\alpha = 1 - \omega\sqrt{\kappa}$  pour  $\omega \in \left(0, \frac{1}{\sqrt{\kappa}}\right)$ , alors les équations (5.4.3) et (5.4.4) sont vraies pour tout  $\tau > 0$  satisfaisant :

$$(1 - \omega\sqrt{\kappa}) \tau^3 - \omega(2 - \omega\sqrt{\kappa}) \tau^2 + (\omega^2 + 2)\tau - \omega \leq 0. \quad (5.A.32)$$

On définit

$$P : (\tau, \omega, \kappa) \mapsto (1 - \omega\sqrt{\kappa}) \tau^3 - \omega(2 - \omega\sqrt{\kappa}) \tau^2 + (\omega^2 + 2)\tau - \omega.$$

On peut montrer que si  $\kappa \leq \frac{1}{10}$ , alors pour tout  $\omega \geq \frac{3}{2}$ ,  $P\left(\frac{2}{3\omega}, \omega, \kappa\right) < 0$ . En fait,  $\omega \mapsto \frac{2}{3\omega}$  est une borne inférieure de la plus haute valeur de  $\tau$  qui vérifie (5.4.6) pour  $\omega \geq \frac{3}{2}$  et  $\kappa = \frac{1}{10}$  comme illustré dans la Figure 5.3.

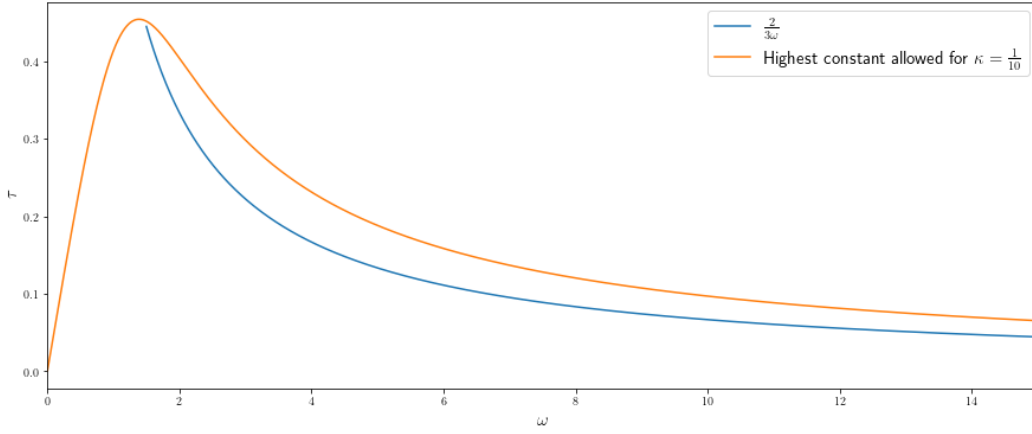


FIGURE 5.3 – Évolution de la plus haute valeur de  $\tau$  vérifiant (5.4.6) pour  $\kappa = \frac{1}{10}$  et de la fonction  $\omega \mapsto \frac{2}{3\omega}$  en fonction de  $\omega$ .

Par conséquent, si  $\kappa \leq \frac{1}{10}$  et  $\omega \in \left(\frac{3}{2}, \frac{1}{\sqrt{\kappa}}\right)$ , alors (5.4.3) et (5.4.4) sont vraies avec  $\tau = \frac{2}{3\omega}$ . De plus, si on considère  $\alpha = 1 - \theta$  pour  $\theta \in \left[\frac{3}{2}\sqrt{\kappa}, 1\right)$ , alors (5.4.3) et (5.4.4) sont vérifiées avec  $\tau = \frac{2}{3\theta}\sqrt{\kappa}$ . Cela mène aux deux points du Corollaire 5.4.4.

### 5.A.11 Preuve du Lemme 5.4.11

Soit  $(A, B) \in \mathbb{R}^2$ . On peut facilement montrer que pour tout  $k \in \mathbb{N}^*$ ,

$$h_{k-1} - h_k - \delta_k = -2\langle x_k - x_{k-1}, x_k - x_k^* \rangle + 2\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - \gamma_k^*.$$

En conséquence, pour tout  $k \in \mathbb{N}^*$ ,

$$\begin{aligned} A\delta_k + B(h_{k-1} - h_k) &= (A + B)\delta_k + B(h_{k-1} - h_k - \delta_k) \\ &= (A + B)\delta_k - 2B\langle x_k - x_{k-1}, x_k - x_k^* \rangle \\ &\quad + 2B\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - B\gamma_k^* \\ &\leq (A + B)\delta_k + 2|B| |\langle x_k - x_{k-1}, x_k - x_k^* \rangle| \\ &\quad + 2B\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - B\gamma_k^*. \end{aligned}$$

De plus, pour tout  $k \in \mathbb{N}^*$  et  $\theta > 0$ ,

$$2|\langle x_k - x_{k-1}, x_k - x_k^* \rangle| \leq \frac{h_k}{\theta} + \theta \delta_k. \quad (5.A.33)$$

Ainsi,

$$A\delta_k + B(h_{k-1} - h_k) \leq (A + B + \theta|B|)\delta_k + \frac{|B|}{\theta}h_k + 2B\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - B\gamma_k^*.$$

On introduit  $b_k := \|\lambda(x_{k-1} - x_{k-1}^*) + k(x_k - x_{k-1})\|^2$ . En développant  $b_k$ , on obtient que :

$$\begin{aligned} b_k &= \|\lambda(x_k - x_k^*) + (k - \lambda)(x_k - x_{k-1}) + \lambda(x_{k-1}^* - x_{k-1}^*)\|^2 \\ &= \|\lambda(x_k - x_k^*) + (k - \lambda)(x_k - x_{k-1})\|^2 + \lambda^2\gamma_k^* \\ &\quad + 2\lambda^2\langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle + 2\lambda(k - \lambda)\langle x_k - x_{k-1}, x_k^* - x_{k-1}^* \rangle. \end{aligned}$$

On peut appliquer l'inégalité ci-dessous à  $u = (k - \lambda)(x_k - x_{k-1})$  et  $v = \lambda(x_k - x_k^*)$  :

$$\|u\|^2 \leq 2\|u + v\|^2 + 2\|v\|^2,$$

ce qui donne :

$$\begin{aligned} (k - \lambda)^2\delta_k &\leq 2\|\lambda(x_k - x_k^*) + (k - \lambda)(x_k - x_{k-1})\|^2 + 2\lambda^2h_k \\ &\leq 2b_k + \frac{8\alpha^2}{9}h_k - \Delta_k^*, \end{aligned}$$

où  $\Delta_k^* = 2(\lambda^2\gamma_k^* + 2\lambda^2\langle x_k - x_k^*, x_k^* - x_{k-1}^* \rangle + 2\lambda(k - \lambda)\langle x_k - x_{k-1}, x_k^* - x_{k-1}^* \rangle)$ . Comme  $\Delta_k^* \geq 0$  on trouve que

$$\delta_k \leq \frac{2}{(k - \lambda)^2}b_k + \frac{8\alpha^2}{9(k - \lambda)^2}h_k, \quad (5.A.34)$$

et

$$\begin{aligned} A\delta_k + B(h_{k-1} - h_k) &\leq (|A + B| + \theta|B|)\frac{2}{(k - \lambda)^2}b_k + \left( (|A + B| + \theta|B|)\frac{8\alpha^2}{9(k - \lambda)^2} + \frac{|B|}{\theta} \right)h_k \\ &\quad + 2B\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - B\gamma_k^*. \end{aligned}$$

Comme  $F$  satisfait  $\mathcal{G}_\mu^2$ , on a  $h_k \leq \frac{w_k}{s\mu}$  et donc,

$$\begin{aligned} A\delta_k + B(h_{k-1} - h_k) &\leq (|A + B| + \theta|B|)\frac{2}{(k - \lambda)^2}b_k + \left( (|A + B| + \theta|B|)\frac{8\alpha^2}{9s\mu(k - \lambda)^2} + \frac{|B|}{s\mu\theta} \right)w_k \\ &\quad + 2B\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - B\gamma_k^*. \end{aligned}$$

Pour le choix  $\theta = \frac{1}{\sqrt{2s\mu}}$ , on trouve donc que

$$\begin{aligned} A\delta_k + B(h_{k-1} - h_k) &\leq (2|A + B| + \frac{\sqrt{2}|B|}{\sqrt{s\mu}})\frac{1}{(k - \lambda)^2}b_k + \left( (2|A + B| + \frac{\sqrt{2}|B|}{\sqrt{s\mu}})\frac{4\alpha^2}{9s\mu(k - \lambda)^2} + \frac{\sqrt{2}|B|}{\sqrt{s\mu}} \right)w_k \\ &\quad + 2B\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - B\gamma_k^*, \end{aligned}$$

et ainsi,

$$\begin{aligned} A\delta_k + B(h_{k-1} - h_k) &\leq \left( 2|A + B| + \frac{\sqrt{2}|B|}{\sqrt{s\mu}} \right) \left( 1 + \frac{4\alpha^2}{9s\mu k^2} \right) \frac{E_k}{(k - \frac{2\alpha}{3})^2} \\ &\quad - B\gamma_k^* + 2B\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned}$$

## 5.A.12 Preuve du Lemme 5.4.13

Rappelons la définition de l'énergie  $\mathcal{E}$  :

$$\mathcal{E}_k = k^2 w_k + b_k + \xi h_k + \lambda k \alpha_k^2 \delta_k. \quad (5.A.35)$$

Notons que pour tout  $k \in \mathbb{N}$ ,

$$b_k = \lambda^2 h_k + k^2 \alpha_k^2 \delta_k + 2\lambda k \alpha_k \langle x_k - x_k^*, x_k - x_{k-1} \rangle, \quad (5.A.36)$$

et en appliquant le deuxième point du Lemme 5.4.7, on trouve donc

$$\begin{aligned} b_k &= \lambda^2 h_k + \lambda k \alpha_k (h_k - h_{k-1}) + n \alpha_k (n \alpha_k + \lambda) \delta_k \\ &\quad - \lambda k \alpha_k \gamma_k^* + 2\lambda k \alpha_k \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \quad (5.A.37)$$

Par conséquent,

$$\begin{aligned} b_{k+1} - b_k &= \lambda(\lambda + (k+1)\alpha_{k+1})(h_{k+1} - h_k) + (k+1)\alpha_{k+1}((k+1)\alpha_{k+1} + \lambda)\delta_{k+1} \\ &\quad - \lambda k \alpha_k (h_k - h_{k-1}) - n \alpha_k (n \alpha_k + \lambda) \delta_k - \lambda(k+1)\alpha_{k+1} \gamma_{k+1}^* + \lambda k \alpha_k \gamma_k^* \\ &\quad + 2\lambda(k+1)\alpha_{k+1} \langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle - 2\lambda k \alpha_k \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \quad (5.A.38)$$

D'autre part, on a

$$(k+1)^2 w_{k+1} - k^2 w_k = k^2 (w_{k+1} - w_k) + (2k+1)w_{k+1}, \quad (5.A.39)$$

et d'après le premier point du Lemme 5.4.8,

$$(k+1)^2 w_{k+1} - k^2 w_k \leq k^2 (\alpha_k^2 \delta_k - \delta_{k+1}) + (2k+1)w_{k+1}, \quad (5.A.40)$$

En combinant (5.A.38) et (5.A.40), on trouve que

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &\leq (2k+1)w_{k+1} - \lambda k (\alpha_k + \alpha_k^2) \delta_k \\ &\quad + ((k+1)\alpha_{k+1}((k+1)\alpha_{k+1} + \lambda) + \lambda(k+1)\alpha_{k+1}^2 - k^2) \delta_{k+1} \\ &\quad + \lambda k \alpha_k (h_{k-1} - h_k) - \lambda \left( \lambda + (k+1)\alpha_{k+1} + \frac{\xi}{\lambda} \right) (h_k - h_{k+1}) \\ &\quad - \lambda(k+1)\alpha_{k+1} \gamma_{k+1}^* + 2\lambda(k+1)\alpha_{k+1} \langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle \\ &\quad + \lambda k \alpha_k \gamma_k^* - 2\lambda k \alpha_k \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \end{aligned} \quad (5.A.41)$$

De plus, la deuxième inégalité du Lemme 5.4.8 assure que

$$\begin{aligned} -\lambda k w_{k+1} + \lambda k (h_k - h_{k+1}) + \lambda k \alpha_k (h_k - h_{k-1}) + \lambda k (\alpha_k + \alpha_k^2) \delta_k - \lambda k \gamma_{k+1}^* - \lambda k \alpha_k \gamma_k^* \\ + 2\lambda k \alpha_k \langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle - 2\lambda k \langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle \leq 0. \end{aligned} \quad (5.A.42)$$

En sommant l'inégalité (5.A.42) à (5.A.41), on obtient

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &\leq ((2-\lambda)k+1)w_{k+1} + ((k+1)\alpha_{k+1}((k+1)\alpha_{k+1} + \lambda) + \lambda(k+1)\alpha_{k+1}^2 - k^2) \delta_{k+1} \\ &\quad - \lambda \left( \lambda + (k+1)\alpha_{k+1} + \frac{\xi}{\lambda} - n \right) (h_k - h_{k+1}) + \mathcal{X}_k^*, \end{aligned} \quad (5.A.43)$$



où

$$\begin{aligned} \mathcal{X}_k^* &= (-\lambda(k+1)\alpha_{k+1} - \lambda k)\gamma_{k+1}^* + 2\lambda(k+1)\alpha_{k+1}\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle \\ &\quad - 2\lambda k\langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle. \end{aligned}$$

Comme  $X^*$  est un ensemble convexe fermé,  $\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle \leq 0$  et  $\langle x_{k+1} - x_{k+1}^*, x_{k+1}^* - x_k^* \rangle \geq 0$  ce qui donne que pour tout  $k \in \mathbb{N}$ ,  $\mathcal{X}_k^* \leq 0$ . En choisissant  $\xi = \lambda(\lambda + 1 - \alpha)$ , on a alors

$$\mathcal{E}_{k+1} - \mathcal{E}_k \leq ((2 - \lambda)k + 1)w_{k+1} + A_1(k)\delta_{k+1} + A_2(k)(h_k - h_{k+1}), \quad (5.A.44)$$

avec :

$$\begin{aligned} - A_1(k) &= 2(\lambda + 1 - \alpha)k + \frac{k^2(3\alpha^2 - 3\alpha\lambda - 2\alpha + 2\lambda + 1) + k(2\alpha^3 - 2\alpha^2\lambda + 2\alpha^2 - 2\alpha\lambda - 2\alpha + 4\lambda + 2) + 1 + 2\lambda + \alpha\lambda}{(k+1+\alpha)^2}, \\ - A_2(k) &= -2\lambda(\lambda + 1 - \alpha) - \frac{\alpha^2\lambda}{k+1+\alpha}. \end{aligned}$$

Notons qu'en réécrivant (5.A.37), on a

$$h_{k-1} - h_k = -\frac{1}{\lambda k \alpha_k} b_k + \frac{\lambda}{k \alpha_k} h_k + \frac{k \alpha_k + \lambda}{\lambda} \delta_k - \gamma_k^* + 2\langle x_{k-1} - x_{k-1}^*, x_k^* - x_{k-1}^* \rangle. \quad (5.A.45)$$

Cela permet d'écrire que

$$\begin{aligned} \mathcal{E}_{k+1} - \mathcal{E}_k &\leq ((2 - \lambda)k + 1)w_{k+1} - \frac{A_2(k)}{\lambda(k+1)\alpha_{k+1}} b_{k+1} + \frac{\lambda A_2(k)}{(k+1)\alpha_{k+1}} h_{k+1} \\ &\quad + \left( A_1(k) + \frac{(k+1)\alpha_{k+1} + \lambda}{\lambda} A_2(k) \right) \delta_{k+1} \\ &\quad - A_2(k) (\gamma_{k+1}^* - 2\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle), \end{aligned} \quad (5.A.46)$$

qui est l'inégalité voulue.

### 5.A.13 Preuve du Lemme 5.4.14

Soit  $\xi = \lambda(\lambda + 1 - \alpha)$  et  $\mathcal{H}_k = k^p \mathcal{E}_k$  avec  $p = 1 + \frac{4}{\gamma-2}$ . On peut écrire que

$$\mathcal{H}_{k+1} - \mathcal{H}_k = k^p (\mathcal{E}_{k+1} - \mathcal{E}_k) + ((k+1)^p - k^p) \mathcal{E}_{k+1}. \quad (5.A.47)$$

De plus, pour tout  $k \in \mathbb{N}$ ,  $(k+1)^p - k^p \in [pk^{p-1}, p(k+1)^{p-1}]$ . Ainsi, en faisant l'hypothèse que  $\lambda \leq \alpha - 1$ , on trouve que  $\xi \leq 0$  et

$$\begin{aligned} ((k+1)^p - k^p) \mathcal{E}_{k+1} &\leq p(k+1)^{p+1} w_{k+1} + p(k+1)^{p-1} b_{k+1} + p\xi k^{p-1} h_{k+1} \\ &\quad + p\lambda(k+1)^p \alpha_{k+1}^2 \delta_{k+1}. \end{aligned} \quad (5.A.48)$$

En appliquant le Lemme 5.4.13 et l'inégalité ci-dessous, on a

$$\begin{aligned} \mathcal{H}_{k+1} - \mathcal{H}_k &\leq (k^p ((2 - \lambda)k + 1) + p(k+1)^{p+1}) w_{k+1} \\ &\quad + (k^p B_1(k) + p(k+1)^{p-1}) b_{k+1} \\ &\quad + (k^p B_2(k) + p\xi k^{p-1}) h_{k+1} \\ &\quad + (k^p B_3(k) + p\lambda(k+1)^p \alpha_{k+1}^2) \delta_{k+1} \\ &\quad - k^p B_4(k) (\gamma_{k+1}^* - 2\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle). \end{aligned} \quad (5.A.49)$$

L'inégalité  $\|u\|^2 \leq 2\|u+v\|^2 + 2\|v\|^2$  appliquée à  $u = \alpha_k(x_k - x_{k-1})$  et  $v = \lambda(x_k - x_k^*)$  donne donc

$$\delta_k \leq \frac{2}{k^2 \alpha_k^2} b_k + \frac{2\lambda^2}{k^2 \alpha_k^2} h_k. \quad (5.A.50)$$

Ainsi,

$$\begin{aligned} \mathcal{H}_{k+1} - \mathcal{H}_k &\leq (k^p ((2 - \lambda)k + 1) + p(k + 1)^{p+1}) w_{k+1} \\ &\quad + \left( k^p B_1(k) + p(k + 1)^{p-1} + 2 \frac{k^p |B_3(k)| + p\lambda(k + 1)^p \alpha_{k+1}^2}{(k + 1)^2 \alpha_{k+1}^2} \right) b_{k+1} \\ &\quad + \left( k^p B_2(k) + p\xi k^{p-1} + 2\lambda^2 \frac{k^p |B_3(k)| + p\lambda(k + 1)^p \alpha_{k+1}^2}{(k + 1)^2 \alpha_{k+1}^2} \right) h_{k+1} \\ &\quad - k^p B_4(k) (\gamma_{k+1}^* - 2\langle x_k - x_k^*, x_{k+1}^* - x_k^* \rangle). \end{aligned} \quad (5.A.51)$$

En remplaçant  $\xi$  par sa valeur et en réorganisant chaque terme, on aboutit à la conclusion.



# Conclusion et perspectives

---

## 6.1 Récapitulatif

Pendant ces trois années, j'ai eu le plaisir de travailler sur des questions variées.

L'étude de méthodes de restart m'a permis de découvrir un moyen simple de gérer l'inertie et a conduit à l'élaboration de deux méthodes visant à s'affranchir de la connaissance des paramètres de géométrie de la fonction à minimiser. En plus d'être efficaces en pratique, des garanties de convergence sont également fournies. De manière générale, les stratégies de restart sont une façon algorithmique pratique de parer les oscillations.

Dans un autre registre, j'ai analysé le comportement du système dynamique (DIN-AVD) faisant intervenir un terme de friction lié à la hessienne de  $F$ . Là où les stratégies de restart réduisent les oscillations en coupant l'inertie "à la main", l'introduction de ce nouveau terme de friction dans le système (AVD) est un moyen plus physique d'atténuer ces oscillations. Cette propriété est de plus en plus forte lorsque l'on considère que  $F$  vérifie des hypothèses de croissance comme démontré dans le Chapitre 4. Cette atténuation du comportement oscillatoire peut être observée chez les schémas dérivés de cette EDO mais les garanties théoriques sur la vitesse de convergence au pire cas de  $F(x_k) - F^*$  ne sont pas meilleures que pour les mêmes schémas sans ce terme de friction.

La dernière partie de ma thèse a été consacrée à la question de l'hypothèse d'unicité du minimiseur de  $F$  dans l'analyse de convergence de schémas inertiels. Les résultats de convergence prouvés dans les cadres continu et discret sans cette hypothèse suggèrent que l'omniprésence de cette hypothèse dans la littérature est soit d'ordre technique (lié à la technique de preuve), soit une conséquence simple de la condition de croissance (comme dans la forte convexité). En aucun cas les méthodes inertielles ne semblent impactées en l'absence de cette hypothèse. Un des résultats notables de ces travaux est la convergence forte des itérés de FISTA lorsque la fonction à minimiser vérifie une hypothèse de croissance.

Ces travaux donnent un éclairage sur la façon dont se comportent les méthodes inertielles quand elles sont appliquées à des fonctions vérifiant des hypothèses situées entre la convexité et la forte convexité. Des alternatives sont proposées pour s'adapter à la géométrie du problème dans le cas où le paramètre de croissance n'est pas connu.

## 6.2 Perspectives

Certaines questions gravitant autour des thématiques de cette thèse sont toujours ouvertes.

Par exemple, les schémas de restart heuristiques introduits par O’Donoghue et Candès dans [108] sont parmi les plus rapides en pratique mais aucune preuve de convergence améliorée n’a encore été démontrée. Une autre interrogation personnelle est : est-il possible de construire un schéma de restart de **V-FISTA** (traitement de l’inertie constant) présentant de meilleures garanties théoriques que les stratégies de redémarrage de **FISTA** ? Les résultats de convergence obtenus dans le Chapitre 5 pourraient donner des éléments de réponse.

D’autre part, je suis très curieux à propos des développements à venir sur les schémas dérivés du système (**DIN-AVD**). À l’heure actuelle, ces méthodes ne donnent pas de garanties significativement meilleures que des schémas classiques et les preuves de convergence s’appuient sur des hypothèses fortes sur certains paramètres. L’atténuation des oscillations offerte par ces schémas reflète cependant un potentiel intéressant.

Concernant mes travaux sur l’hypothèse d’unicité du minimiseur, j’aurais souhaité affaiblir voire supprimer la condition de géométrie faite sur  $X^*$  dans le cas continu pour étendre les résultats connus. Cette hypothèse n’est liée qu’à la dérivabilité de l’opérateur de projection sur un ensemble convexe et elle n’est pas nécessaire pour analyser des méthodes discrètes. Par ailleurs, les vitesses de convergence pour **FISTA** et **V-FISTA** pourraient être légèrement améliorées avec des majorations plus fines.

Limité par le temps, j’aurais également aimé étudier plus en détail l’approche par équations différentielles à haute-résolution proposée par Shi, Du, Jordan et Su [125]. Les auteurs parviennent à construire un objet situé entre le monde continu des systèmes dynamiques et le monde discret des schémas numériques, permettant une meilleure compréhension du comportement de ces schémas et qui pourrait faciliter l’analyse de leur convergence. De même, la stratégie introduite par Drori et Teboulle dans [57] qui consiste à voir l’analyse de performance de schémas de premier ordre comme la résolution d’un problème d’optimisation me semble extrêmement intéressante. Les papiers de Kim et Fessler [81], Taylor, Hendrickx et Glineur [129] et Taylor et Drori [128] me font penser qu’il serait pertinent d’appliquer une telle approche pour étudier la performance de schémas de premier ordre pour des hypothèses situées entre la convexité et la forte convexité.

Enfin, une suite logique à mes travaux serait d’étudier le comportement de méthodes inertielles pour les fonctions non convexes. Dans ce contexte, les résultats de convergence sont peu nombreux et la question de la pertinence de l’inertie est à ma connaissance toujours sans réponse.

# Bibliographie

- [1] Teodoro ALAMO, Pablo KRUPA et Daniel LIMON. “Gradient based restart FISTA”. In : *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2019, p. 3936-3941.
- [2] Teodoro ALAMO, Pablo KRUPA et Daniel LIMON. “Restart of accelerated first order methods with linear convergence under a quadratic functional growth condition”. In : *IEEE Transactions on Automatic Control* (2022).
- [3] Teodoro ALAMO, Daniel LIMON et Pablo KRUPA. “Restart FISTA with global linear convergence”. In : *2019 18th European Control Conference (ECC)*. IEEE. 2019, p. 1969-1974.
- [4] Felipe ALVAREZ. “On the minimizing property of a second order dissipative system in Hilbert spaces”. In : *SIAM Journal on Control and Optimization* 38.4 (2000), p. 1102-1119.
- [5] Felipe ALVAREZ, Hedy ATTOUCH, Jérôme BOLTE et Patrick REDONT. “A second-order gradient-like dissipative dynamical system with hessian-driven damping. : Application to optimization and mechanics”. In : *Journal de mathématiques pures et appliquées* 81.8 (2002), p. 747-779.
- [6] Vassilis APIDOPOULOS, Jean-François AUJOL et Charles DOSSAL. “The Differential Inclusion Modeling FISTA Algorithm and Optimality of Convergence Rate in the Case  $b \leq 3$ ”. In : *SIAM Journal on Optimization* 28.1 (2018), p. 551-574. DOI : [10.1137/17M1128642](https://doi.org/10.1137/17M1128642). eprint : <https://doi.org/10.1137/17M1128642>.
- [7] Vassilis APIDOPOULOS, Jean-François AUJOL, Charles DOSSAL et Aude RONDEPIERRE. “Convergence rates of an inertial gradient descent algorithm under growth and flatness conditions”. In : *Mathematical Programming* 187 (2021), p. 151-193.
- [8] Vassilis APIDOPOULOS, Nicolò GINATTA et Silvia VILLA. “Convergence rates for the heavy-ball continuous dynamics for non-convex optimization, under Polyak–Łojasiewicz condition”. In : *Journal of Global Optimization* 84.3 (2022), p. 563-589.
- [9] Larry ARMIJO. “Minimization of functions having Lipschitz continuous first partial derivatives”. In : *Pacific Journal of mathematics* 16.1 (1966), p. 1-3.
- [10] Hedy ATTOUCH, Al cha BALHAG, Zaki CHBANI et Hassan RIAHI. “Accelerated gradient methods combining Tikhonov regularization with geometric damping driven by the Hessian”. In : *arXiv preprint arXiv :2203.05457* (2022).
- [11] Hedy ATTOUCH et Alexandre CABOT. “Asymptotic stabilization of inertial gradient dynamics with time-dependent viscosity”. In : *Journal of Differential Equations* 263.9 (2017), p. 5412-5458.

- [12] Hedy ATTOUCH et Alexandre CABOT. “Convergence rates of inertial forward-backward algorithms”. In : *SIAM Journal on Optimization* 28.1 (2018), p. 849-874.
- [13] Hedy ATTOUCH, Alexandre CABOT, Zaki CHBANI et Hassan RIAHI. “Rate of convergence of inertial gradient dynamics with time-dependent viscous damping coefficient”. In : *Evolution Equations and Control Theory* 7.3 (2018), p. 353-371.
- [14] Hedy ATTOUCH, Zaki CHBANI, Jalal FADILI et Hassan RIAHI. “Convergence of iterates for first-order optimization algorithms with inertia and Hessian driven damping”. In : *Optimization* (2021), p. 1-40.
- [15] Hedy ATTOUCH, Zaki CHBANI, Jalal FADILI et Hassan RIAHI. “First-order optimization algorithms via inertial systems with Hessian driven damping”. In : *Mathematical Programming* (2022), p. 1-43.
- [16] Hedy ATTOUCH, Zaki CHBANI, Juan PEYPOUQUET et Patrick REDONT. “Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity”. In : *Mathematical Programming* 168 (2018), p. 123-175.
- [17] Hedy ATTOUCH, Zaki CHBANI et Hassan RIAHI. “Rate of convergence of the Nesterov accelerated gradient method in the subcritical case  $\alpha \leq 3$ ”. In : *ESAIM : COCV* 25 (2019), p. 2. DOI : [10.1051/cocv/2017083](https://doi.org/10.1051/cocv/2017083).
- [18] Hedy ATTOUCH, Jalal FADILI et Vyacheslav KUNGURTSEV. “On the effect of perturbations in first-order optimization methods with inertia and Hessian driven damping”. In : *Evolution Equations and Control Theory* 12.1 (2023), p. 71-117. DOI : [10.3934/eect.2022022](https://doi.org/10.3934/eect.2022022).
- [19] Hedy ATTOUCH, Xavier GOUDOU et Patrick REDONT. “The heavy ball with friction method, I. The continuous dynamical system : global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system”. In : *Communications in Contemporary Mathematics* 2.01 (2000), p. 1-34.
- [20] Hedy ATTOUCH et Juan PEYPOUQUET. “The rate of convergence of Nesterov’s accelerated forward-backward method is actually faster than  $1/k^2$ ”. In : *SIAM Journal on Optimization* 26.3 (2016), p. 1824-1834.
- [21] Hedy ATTOUCH, Juan PEYPOUQUET et Patrick REDONT. “Fast convex optimization via inertial dynamics with Hessian driven damping”. In : *Journal of Differential Equations* 261.10 (2016), p. 5734-5783.
- [22] Jean-François AUJOL, Luca CALATRONI, Charles DOSSAL, Hippolyte LABARRIÈRE et Aude RONDEPIERRE. “Parameter-Free FISTA by Adaptive Restart and Backtracking”. working paper or preprint. 2023.
- [23] Jean-François AUJOL, Charles DOSSAL, Văn Hào HOÀNG, Hippolyte LABARRIÈRE et Aude RONDEPIERRE. “Fast convergence of inertial dynamics with Hessian-driven damping under geometry assumptions”. In : *arXiv preprint arXiv :2206.06853* (2022).

- [24] Jean-François AUJOL, Charles DOSSAL et Aude RONDEPIERRE. “Convergence Rates of the Heavy Ball Method for Quasi-strongly Convex Optimization”. In : *SIAM Journal on Optimization* 32.3 (2022), p. 1817-1842.
- [25] Jean-François AUJOL, Charles DOSSAL et Aude RONDEPIERRE. “Convergence rates of the Heavy-Ball method under the Łojasiewicz property”. In : *Mathematical Programming* (2022), p. 1-60.
- [26] Jean-François AUJOL, Charles DOSSAL et Aude RONDEPIERRE. “FISTA is an automatic geometrically optimized algorithm for strongly convex functions”. In : *Mathematical Programming* (2023). ISSN : 1436-4646. DOI : [10.1007/s10107-023-01960-6](https://doi.org/10.1007/s10107-023-01960-6).
- [27] Jean-François AUJOL, Charles DOSSAL et Aude RONDEPIERRE. “Optimal convergence rates for Nesterov acceleration”. In : *SIAM Journal on Optimization* 29.4 (2019), p. 3131-3153.
- [28] Jean-François AUJOL, Charles H DOSSAL, Hippolyte LABARRIÈRE et Aude RONDEPIERRE. “FISTA restart using an automatic estimation of the growth parameter”. working paper or preprint. 2022.
- [29] Mounir BALTI et Ramzi MAY. “Asymptotic for the perturbed heavy ball system with vanishing damping term”. In : *Evolution Equations and Control Theory* 6.2 (2017), p. 177-186. DOI : [10.3934/eect.2017010](https://doi.org/10.3934/eect.2017010).
- [30] Heinz H BAUSCHKE, Jérôme BOLTE et Marc TEBoulLE. “A descent lemma beyond Lipschitz gradient continuity : first-order methods revisited and applications”. In : *Mathematics of Operations Research* 42.2 (2017), p. 330-348.
- [31] Amir BECK. *First-order methods in optimization*. SIAM, 2017.
- [32] Amir BECK et Marc TEBoulLE. “A fast iterative shrinkage-thresholding algorithm for linear inverse problems”. In : *SIAM journal on imaging sciences* 2.1 (2009), p. 183-202.
- [33] Amir BECK et Marc TEBoulLE. “Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems”. In : *IEEE transactions on image processing* 18.11 (2009), p. 2419-2434.
- [34] Stephen BECKER, Jérôme BOBIN et Emmanuel J CANDÈS. “NESTA : A fast and accurate first-order method for sparse recovery”. In : *SIAM Journal on Imaging Sciences* 4.1 (2011), p. 1-39.
- [35] Stephen R BECKER, Emmanuel J CANDÈS et Michael C GRANT. “Templates for convex cone problems with applications to sparse signal recovery”. In : *Mathematical programming computation* 3 (2011), p. 165-218.
- [36] Pascal BÉGOUT, Jérôme BOLTE et Mohamed Ali JENDOUBI. “On damped second-order gradient systems”. In : *Journal of Differential Equations* 259.7 (2015), p. 3115-3143.



- [37] Yunier BELLO-CRUZ, Guoyin LI et Tran Thai An NGHIA. “Quadratic Growth Conditions and Uniqueness of Optimal Solution to Lasso”. In : *Journal of Optimization Theory and Applications* 194.1 (2022), p. 167-190.
- [38] Mario BERTERO, Patrizia BOCCACCI et Valeria RUGGIERO. *Inverse Imaging with Poisson Data*. 2053-2563. IOP, 2018.
- [39] Jérôme BOLTE, Aris DANILIDIS et Adrian LEWIS. “The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems”. In : *SIAM Journal on Optimization* 17.4 (2007), p. 1205-1223.
- [40] Jérôme BOLTE, Aris DANILIDIS, Adrian LEWIS et Masahiro SHIOTA. “Clarke subgradients of stratifiable functions”. In : *SIAM Journal on Optimization* 18.2 (2007), p. 556-572.
- [41] Jérôme BOLTE, Trong Phong NGUYEN, Juan PEYPOUQUET et Bruce W. SUTER. “From error bounds to the complexity of first-order descent methods for convex functions”. In : *Mathematical Programming* 165.2 (2017), p. 471-507. ISSN : 1436-4646. DOI : [10.1007/s10107-016-1091-6](https://doi.org/10.1007/s10107-016-1091-6).
- [42] J Frédéric BONNANS, Roberto COMINETTI et Alexander SHAPIRO. “Sensitivity analysis of optimization problems under second order regular constraints”. In : *Mathematics of Operations Research* 23.4 (1998), p. 806-831.
- [43] Radu Ioan BOȚ, Ernő Robert CSETNEK et Szilárd Csaba LÁSZLÓ. “Tikhonov regularization of a second order dynamical system with Hessian driven damping”. In : *Mathematical Programming* 189 (2021), p. 151-186.
- [44] Stephen BOYD, Lin XIAO et Almir MUTAPCIC. “Subgradient methods”. In : *lecture notes of EE392o, Stanford University, Autumn Quarter 2004* (2003), p. 2004-2005.
- [45] Sébastien BUBECK, Yin Tat LEE et Mohit SINGH. “A geometric alternative to Nesterov’s accelerated gradient descent”. In : *arXiv preprint arXiv :1506.08187* (2015).
- [46] Alexandre CABOT, Hans ENGLER et Sébastien GADAT. “On the long time behavior of second order differential equations with asymptotically small dissipation”. In : *Transactions of the American Mathematical Society* 361.11 (2009), p. 5983-6017.
- [47] Alexandre CABOT, Hans ENGLER et Sébastien GADAT. “Second-order differential equations with asymptotically small dissipation and piecewise flat potentials.” In : *Electronic Journal of Differential Equations (EJDE)[electronic only]* 2009 (2009), p. 33-38.
- [48] Alexandre CABOT et Pierre FRANKEL. “Asymptotics for some semilinear hyperbolic equations with non-autonomous damping”. In : *Journal of Differential Equations* 252.1 (2012), p. 294-322. ISSN : 0022-0396. DOI : <https://doi.org/10.1016/j.jde.2011.09.012>.

- [49] Luca CALATRONI et Antonin CHAMBOLLE. “Backtracking strategies for accelerated descent methods with smooth composite objectives”. In : *SIAM Journal on Optimization* 29.3 (2019), p. 1772-1798.
- [50] Luca CALATRONI, Guillaume GARRIGOS, Lorenzo ROSASCO et Silvia VILLA. “Accelerated iterative regularization via dual diagonal descent”. In : *SIAM Journal on Optimization* 31.1 (2021), p. 754-784.
- [51] Antonin CHAMBOLLE et Ch DOSSAL. “On the convergence of the iterates of the “fast iterative shrinkage/thresholding algorithm””. In : *Journal of Optimization theory and Applications* 166 (2015), p. 968-982.
- [52] Shixiang CHEN, Shiqian MA et Wei LIU. “Geometric descent method for convex composite minimization”. In : *Advances in Neural Information Processing Systems* 30 (2017).
- [53] Patrick L COMBETTES et Jean-Christophe PESQUET. “Proximal splitting methods in signal processing”. In : *Fixed-point algorithms for inverse problems in science and engineering* (2011), p. 185-212.
- [54] Patrick L COMBETTES et Valérie R WAJS. “Signal recovery by proximal forward-backward splitting”. In : *Multiscale modeling & simulation* 4.4 (2005), p. 1168-1200.
- [55] Alexandre D’ASPREMONT. “Smooth optimization with approximate gradient”. In : *SIAM Journal on Optimization* 19.3 (2008), p. 1171-1183.
- [56] Alexandre D’ASPREMONT, Damien SCIEUR, Adrien TAYLOR et al. “Acceleration methods”. In : *Foundations and Trends® in Optimization* 5.1-2 (2021), p. 1-245.
- [57] Yoel DRORI et Marc TEBoulLE. “Performance of first-order methods for smooth convex minimization : a novel approach”. In : *Mathematical Programming* 145.1 (2014), p. 451-482. ISSN : 1436-4646. DOI : [10.1007/s10107-013-0653-0](https://doi.org/10.1007/s10107-013-0653-0).
- [58] Dmitriy DRUSVYATSKIY, Maryam FAZEL et Scott ROY. “An optimal first order method based on optimal quadratic averaging”. In : *SIAM Journal on Optimization* 28.1 (2018), p. 251-271.
- [59] Jonathan ECKSTEIN. “Splitting methods for monotone operators with applications to parallel optimization”. Thèse de doct. Massachusetts Institute of Technology, 1989.
- [60] Jonathan ECKSTEIN et Dimitri P BERTSEKAS. “On the Douglas—Rachford splitting method and the proximal point algorithm for maximal monotone operators”. In : *Mathematical programming* 55 (1992), p. 293-318.
- [61] Olivier FERCOQ et Zheng QU. “Adaptive restart of accelerated gradient methods under local quadratic growth condition”. In : *IMA Journal of Numerical Analysis* 39.4 (2019), p. 2069-2095.

- [62] Olivier FERCOQ et Zheng QU. “Restarting accelerated gradient methods with a rough strong convexity estimate”. In : *arXiv preprint arXiv :1609.07358* (2016).
- [63] Simon FITZPATRICK et R. R. PHELPS. “Differentiability of the Metric Projection in Hilbert Space”. In : *Transactions of the American Mathematical Society* 270.2 (1982), p. 483-501. ISSN : 00029947.
- [64] Guillaume GARRIGOS, Lorenzo ROSASCO et Silvia VILLA. “Convergence of the forward-backward algorithm : beyond the worst-case with the help of geometry”. In : *Mathematical Programming* 198.1 (2023), p. 937-996. ISSN : 1436-4646. DOI : [10.1007/s10107-022-01809-4](https://doi.org/10.1007/s10107-022-01809-4).
- [65] Euhanna GHADIMI, Hamid Reza FEYZMAHDAVIAN et Mikael JOHANSSON. “Global convergence of the heavy-ball method for convex optimization”. In : *2015 European control conference (ECC)*. IEEE. 2015, p. 310-315.
- [66] Pontus GISELSSON et Stephen BOYD. “Monotonicity and restart in fast gradient methods”. In : *53rd IEEE Conference on Decision and Control*. IEEE. 2014, p. 5058-5063.
- [67] Tom GOLDSTEIN, Christoph STUDER et Richard BARANIUK. *FASTA : A Generalized Implementation of Forward-Backward Splitting*. <http://arxiv.org/abs/1501.04979>. 2015.
- [68] Ming GU, Lek-Heng LIM et Cinna Julie WU. “PARNES : A rapidly convergent algorithm for accurate recovery of sparse and approximately sparse signals”. In : *Numerical Algorithms* 64.2 (2013), p. 321-347.
- [69] Osman GÜLER. “New proximal point algorithms for convex minimization”. In : *SIAM Journal on Optimization* 2.4 (1992), p. 649-664.
- [70] Osman GÜLER. “On the convergence of the proximal point algorithm for convex minimization”. In : *SIAM journal on control and optimization* 29.2 (1991), p. 403-419.
- [71] Alain HARAUX et Mohamed-Ali JENDOUBI. “Convergence of solutions of second-order gradient-like systems with analytic nonlinearities”. In : *journal of differential equations* 144.2 (1998), p. 313-320.
- [72] Zachary T. HARMANY, Roummel F. MARCIA et Rebecca M. WILLETT. “This is SPIRAL-TAP : Sparse Poisson Intensity Reconstruction ALgorithms—Theory and Practice”. In : *IEEE Transactions on Image Processing* 21.3 (2012), p. 1084-1096. DOI : [10.1109/TIP.2011.2168410](https://doi.org/10.1109/TIP.2011.2168410).
- [73] Jean-Baptiste HIRIART-URRUTY. “At what points is the projection mapping differentiable?” In : *The American Mathematical Monthly* 89.7 (1982), p. 456-458.
- [74] Richard B HOLMES. “Smoothness of certain metric projections on Hilbert space”. In : *Transactions of the American Mathematical Society* 184 (1973), p. 87-100.

- [75] Franck IUTZELER et Jérôme MALICK. “On the proximal gradient algorithm with alternated inertia”. In : *Journal of Optimization Theory and Applications* 176.3 (2018), p. 688-710.
- [76] Mohamed Ali JENDOUBI et Ramzi MAY. “Asymptotics for a second-order differential equation with nonautonomous damping and an integrable source term”. In : *Applicable Analysis* 94.2 (2015), p. 435-443.
- [77] Hamed KARIMI, Julie NUTINI et Mark SCHMIDT. “Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition”. In : *Machine Learning and Knowledge Discovery in Databases : European Conference, ECML PKDD 2016, Riva del Garda, Italy, September 19-23, 2016, Proceedings, Part I 16*. Springer. 2016, p. 795-811.
- [78] Thomas KERDREUX, Alexandre D’ASPREMONT et Sebastian POKUTTA. “Restarting frank-wolfe”. In : *The 22nd international conference on artificial intelligence and statistics*. PMLR. 2019, p. 1275-1283.
- [79] Donghwan KIM et Jeffrey A FESSLER. “Adaptive restart of the optimized gradient method for convex optimization”. In : *Journal of Optimization Theory and Applications* 178.1 (2018), p. 240-263.
- [80] Donghwan KIM et Jeffrey A FESSLER. “An optimized first-order method for image restoration”. In : *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2015, p. 3675-3679.
- [81] Donghwan KIM et Jeffrey A. FESSLER. “Optimized first-order methods for smooth convex minimization”. In : *Mathematical Programming* 159.1 (2016), p. 81-107. ISSN : 1436-4646. DOI : [10.1007/s10107-015-0949-3](https://doi.org/10.1007/s10107-015-0949-3).
- [82] Krzysztof KURDYKA. “On gradients of functions definable in o-minimal structures”. In : *Annales de l’institut Fourier*. T. 48. 3. 1998, p. 769-783.
- [83] Marta LAZZARETTI, Simone REBEGOLDI, Luca CALATRONI et Claudio ESTATICO. “A Scaled and Adaptive FISTA Algorithm for Signal-Dependent Sparse Image Super-Resolution Problems”. In : *Scale Space and Variational Methods in Computer Vision*. Sous la dir. d’Abderrahim ELMOATAZ, Jalal FADILI, Yvain QUÉAU, Julien RABIN et Loïc SIMON. Cham : Springer International Publishing, 2021, p. 242-253. ISBN : 978-3-030-75549-2.
- [84] Jingwei LIANG, Tao LUO et Carola-Bibiane SCHONLIEB. “Improving “Fast Iterative Shrinkage-Thresholding Algorithm” : Faster, Smarter, and Greedier”. In : *SIAM Journal on Scientific Computing* 44.3 (2022), A1069-A1091.
- [85] Qihang LIN et Lin XIAO. “An adaptive accelerated proximal gradient method and its homotopy continuation for sparse optimization”. In : *Computational Optimization and Applications* 60.3 (2015), p. 633-674. ISSN : 1573-2894. DOI : [10.1007/s10589-014-9694-4](https://doi.org/10.1007/s10589-014-9694-4).
- [86] Pierre-Louis LIONS et Bertrand MERCIER. “Splitting algorithms for the sum of two nonlinear operators”. In : *SIAM Journal on Numerical Analysis* 16.6 (1979), p. 964-979.

- [87] Mingrui LIU et Tianbao YANG. “Adaptive accelerated gradient converging method under  $h\{o\}$  lderian error bound condition”. In : *Advances in Neural Information Processing Systems* 30 (2017).
- [88] Stanislas ŁOJASIEWICZ. “Sur la géométrie semi-et sous-analytique”. In : *Annales de l’institut Fourier*. T. 43. 5. 1993, p. 1575-1595.
- [89] Stanislaw ŁOJASIEWICZ. “Une propriété topologique des sous-ensembles analytiques réels”. In : *Les équations aux dérivées partielles* 117 (1963), p. 87-89.
- [90] Dirk A LORENZ et Thomas POCK. “An inertial forward-backward algorithm for monotone inclusions”. In : *Journal of Mathematical Imaging and Vision* 51 (2015), p. 311-325.
- [91] Jun-Ren LUO et Ti-Jun XIAO. “Optimal Convergence Rates for Damped Inertial Gradient Dynamics with Flat Geometries”. In : *Applied Mathematics & Optimization* 87.3 (2023), p. 53. ISSN : 1432-0606. DOI : [10.1007/s00245-023-09966-4](https://doi.org/10.1007/s00245-023-09966-4).
- [92] Zhi-Quan LUO et Paul TSENG. “Error bounds and convergence analysis of feasible descent methods : a general approach.” In : *Annals of Operations Research* (1993).
- [93] Bernard MARTINET. “Régularisation d’inéquations variationnelles par approximations successives. Rev. Française Informat”. In : *Recherche Opérationnelle* 4 (1970), p. 154-158.
- [94] Mathurin MASSIAS, Alexandre GRAMFORT et Joseph SALMON. “Celer : a fast solver for the lasso with dual extrapolation”. In : *International Conference on Machine Learning*. PMLR. 2018, p. 3315-3324.
- [95] Juan Jose MAULEN et Juan PEYPOUQUET. “A speed restart scheme for a dynamics with Hessian driven damping”. In : *arXiv preprint arXiv :2301.12240* (2023).
- [96] Ramzi MAY. “Asymptotic for a second-order evolution equation with convex potential and vanishing damping term”. In : *Turkish Journal of Mathematics* 41.3 (2017), p. 681-685.
- [97] Ramzi MAY, Chokri MNASRI et Mounir ELLOUMI. “Asymptotic for a second order evolution equation with damping and regularizing terms”. In : *AIMS Mathematics* 6.5 (2021), p. 4901-4914.
- [98] Jean-Jacques MOREAU. “Fonctionnelles convexes”. In : *Séminaire Jean Leray* 2 (1966), p. 1-108.
- [99] Jean-Jacques MOREAU. “Proximité et dualité dans un espace hilbertien”. In : *Bulletin de la Société mathématique de France* 93 (1965), p. 273-299.
- [100] Abdellatif MOUDAFI et M. OLINY. “Convergence of a splitting inertial proximal method for monotone operators”. In : *Journal of Computational and Applied Mathematics* 155.2 (2003), p. 447-454. ISSN : 0377-0427. DOI : [https://doi.org/10.1016/S0377-0427\(02\)00906-8](https://doi.org/10.1016/S0377-0427(02)00906-8).

- [101] Michael MUEHLEBACH et Michael JORDAN. “A dynamical systems perspective on Nesterov acceleration”. In : *International Conference on Machine Learning*. PMLR. 2019, p. 4656-4662.
- [102] Ion NECOARA, Yurii NESTEROV et François GLINEUR. “Linear convergence of first order methods for non-strongly convex optimization”. In : *Mathematical Programming* 175.1 (2019), p. 69-107. ISSN : 1436-4646. DOI : [10.1007/s10107-018-1232-1](https://doi.org/10.1007/s10107-018-1232-1).
- [103] Yurii NESTEROV. “A method for unconstrained convex minimization problem with the rate of convergence  $\mathcal{O}(1/k^2)$ ”. In : *Doklady an ussr*. T. 269. 1983, p. 543-547.
- [104] Yurii NESTEROV. “Gradient methods for minimizing composite functions”. In : *Mathematical programming* 140.1 (2013), p. 125-161.
- [105] Yurii NESTEROV. *Introductory lectures on convex optimization : A basic course*. T. 87. Springer Science & Business Media, 2003.
- [106] Yurii NESTEROV et al. *Lectures on convex optimization*. T. 137. Springer, 2018.
- [107] Dominikus NOLL. “Directional differentiability of the metric projection in Hilbert space”. In : *Pacific Journal of Mathematics* 170.2 (1995), p. 567-592.
- [108] Brendan O’DONOGHUE et Emmanuel CANDÈS. “Adaptive restart for accelerated gradient schemes”. In : *Foundations of computational mathematics* 15 (2015), p. 715-732.
- [109] Peter OCHS, Thomas BROX et Thomas POCK. “iPiasco : Inertial proximal algorithm for strongly convex optimization”. In : *Journal of Mathematical Imaging and Vision* 53 (2015), p. 171-181.
- [110] Gregory B PASSTY. “Ergodic convergence to a zero of the sum of monotone operators in Hilbert space”. In : *Journal of Mathematical Analysis and Applications* 72.2 (1979), p. 383-390. ISSN : 0022-247X. DOI : [https://doi.org/10.1016/0022-247X\(79\)90234-8](https://doi.org/10.1016/0022-247X(79)90234-8).
- [111] B.T. POLYAK. “Gradient methods for the minimisation of functionals”. In : *USSR Computational Mathematics and Mathematical Physics* 3.4 (1963), p. 864-878. ISSN : 0041-5553. DOI : [https://doi.org/10.1016/0041-5553\(63\)90382-3](https://doi.org/10.1016/0041-5553(63)90382-3).
- [112] Boris T POLYAK. “Some methods of speeding up the convergence of iteration methods”. In : *Ussr computational mathematics and mathematical physics* 4.5 (1964), p. 1-17.
- [113] Boris T POLYAK. “Subgradient methods : a survey of Soviet research”. In : *Nonsmooth optimization* 3 (1978), p. 5-29.
- [114] Boris T POLYAK et Pavel S SHCHERBAKOV. “Optimisation and asymptotic stability”. In : *International Journal of Control* 91.11 (2018), p. 2404-2410.



- [115] Michael James David POWELL. “Restart procedures for the conjugate gradient method”. In : *Mathematical programming* 12 (1977), p. 241-254.
- [116] Simone REBEGOLDI et Luca CALATRONI. “Scaled, Inexact, and Adaptive Generalized FISTA for Strongly Convex Optimization”. In : *SIAM Journal on Optimization* 32.3 (2022), p. 2428-2459.
- [117] James RENEGAR et Benjamin GRIMMER. “A Simple Nearly Optimal Restart Scheme For Speeding Up First-Order Methods”. In : *Foundations of Computational Mathematics* 22.1 (2022), p. 211-256. DOI : [10.1007/s10208-021-09502-2](https://doi.org/10.1007/s10208-021-09502-2).
- [118] R Tyrrell ROCKAFELLAR. “Monotone operators and the proximal point algorithm”. In : *SIAM journal on control and optimization* 14.5 (1976), p. 877-898.
- [119] Ralph ROCKAFELLAR. “Characterization of the subdifferentials of convex functions”. In : *Pacific Journal of Mathematics* 17.3 (1966), p. 497-510.
- [120] Vincent ROULET et Alexandre D’ASPREMONT. “Sharpness, restart and acceleration”. In : *Advances in Neural Information Processing Systems* 30 (2017).
- [121] Saverio SALZO et Silvia VILLA. “Inexact and accelerated proximal point algorithms”. In : *Journal of Convex analysis* 19.4 (2012), p. 1167-1192.
- [122] Katya SCHEINBERG, Donald GOLDFARB et Xi BAI. “Fast first-order methods for composite convex optimization with backtracking”. In : *Foundations of Computational Mathematics* 14 (2014), p. 389-417.
- [123] Othmane SEBBOUH, Charles DOSSAL et Aude RONDEPIERRE. “Convergence rates of damped inertial dynamics under geometric conditions and perturbations”. In : *SIAM Journal on Optimization* 30.3 (2020), p. 1850-1877.
- [124] Alexander SHAPIRO. “Differentiability properties of metric projections onto convex sets”. In : *Journal of Optimization Theory and Applications* 169.3 (2016), p. 953-964.
- [125] Bin SHI, Simon S. DU, Michael I. JORDAN et Weijie J. SU. “Understanding the acceleration phenomenon via high-resolution differential equations”. In : *Mathematical Programming* 195.1 (2022), p. 79-148. ISSN : 1436-4646. DOI : [10.1007/s10107-021-01681-8](https://doi.org/10.1007/s10107-021-01681-8).
- [126] Jonathan W. SIEGEL. “Accelerated First-Order Methods : Differential Equations and Lyapunov Functions”. In : *arXiv : Optimization and Control* (2019).
- [127] Weijie SU, Stephen BOYD et Emmanuel CANDÈS. “A differential equation for modeling Nesterov’s accelerated gradient method : theory and insights”. In : *Advances in neural information processing systems* 27 (2014).
- [128] Adrien TAYLOR et Yoel DRORI. “An optimal gradient method for smooth strongly convex minimization”. In : *Mathematical Programming* (2022). ISSN : 1436-4646. DOI : [10.1007/s10107-022-01839-y](https://doi.org/10.1007/s10107-022-01839-y).

- [129] Adrien B. TAYLOR, Julien M. HENDRICKX et François GLINEUR. “Smooth strongly convex interpolation and exact worst-case performance of first-order methods”. In : *Mathematical Programming* 161.1 (2017), p. 307-345. ISSN : 1436-4646. DOI : [10.1007/s10107-016-1009-3](https://doi.org/10.1007/s10107-016-1009-3).
- [130] Robert TIBSHIRANI. “Regression shrinkage and selection via the lasso”. In : *Journal of the Royal Statistical Society : Series B (Methodological)* 58.1 (1996), p. 267-288.
- [131] Ryan J. TIBSHIRANI. “The lasso problem and uniqueness”. In : *Electronic Journal of Statistics* 7.none (2013), p. 1456-1490. DOI : [10.1214/13-EJS815](https://doi.org/10.1214/13-EJS815).
- [132] Paul TSENG. “On accelerated proximal gradient methods for convex-concave optimization”. In : *submitted to SIAM Journal on Optimization* 2.3 (2008).
- [133] Bryan VAN SCOY, Randy A FREEMAN et Kevin M LYNCH. “The fastest known globally convergent first-order method for minimizing strongly convex functions”. In : *IEEE Control Systems Letters* 2.1 (2017), p. 49-54.
- [134] Silvia VILLA, Saverio SALZO, Luca BALDASSARRE et Alessandro VERRI. “Accelerated and inexact forward-backward algorithms”. In : *SIAM Journal on Optimization* 23.3 (2013), p. 1607-1633.
- [135] Grace Chisholm YOUNG. “A note on derivatives and differential coefficients”. In : *Acta mathematica* 37.1 (1914), p. 141-154.
- [136] Eduardo H. ZARANTONELLO. “Projections on Convex Sets in Hilbert Space and Spectral Theory : Part I. Projections on Convex Sets : Part II. Spectral Theory”. In : *Contributions to Nonlinear Functional Analysis*. Academic Press, 1971, p. 237-424. ISBN : 978-0-12-775850-3. DOI : <https://doi.org/10.1016/B978-0-12-775850-3.50013-3>.
- [137] Hui ZHANG. “The restricted strong convexity revisited : analysis of equivalence to error bound and quadratic growth”. In : *Optimization Letters* 11 (2017), p. 817-833.
- [138] Hui ZHANG et Lizhi CHENG. “Restricted strong convexity and its applications to convergence analysis of gradient-type methods in convex optimization”. In : *Optimization Letters* 9.5 (2015), p. 961-979. ISSN : 1862-4480. DOI : [10.1007/s11590-014-0795-x](https://doi.org/10.1007/s11590-014-0795-x).
- [139] Hui ZHANG et Wotao YIN. “Gradient methods for convex minimization : better rates under weaker conditions”. In : *arXiv preprint arXiv :1303.4645* (2013).
- [140] Marcelo VW ZIBETTI, Elias S HELOU, Ravinder R REGATTE et Gabor T HERMAN. “Monotone FISTA with variable acceleration for compressed sensing magnetic resonance imaging”. In : *IEEE transactions on computational imaging* 5.1 (2018), p. 109-119.





## Étude de méthodes inertielles en optimisation et leur comportement sous conditions de géométrie

**Résumé :** Ce manuscrit de thèse est consacré à l'optimisation de fonctions convexes composites dans un cadre déterministe. Dans ce cadre, je me suis concentré sur l'analyse de convergence d'algorithmes d'optimisation sous des hypothèses de croissance.

Lorsque l'on s'intéresse à la minimisation de telles fonctions, les méthodes de premier ordre les plus efficaces sont dites inertielles car elles font apparaître une notion d'élan. Si l'introduction de cette inertie est généralement bénéfique, il est nécessaire de bien paramétrer ce type de méthodes pour obtenir une solution rapidement et éviter des comportements pathologiques. Par exemple, certaines méthodes dites Heavy-Ball nécessitent de connaître la valeur du conditionnement de la fonction à minimiser pour être utilisées.

La question à laquelle cherche à répondre cette thèse est : quelle méthode appliquer en fonction des hypothèses de géométrie vérifiées par  $F$ ? Comme les méthodes étudiées sont inertielles, on peut reformuler cette problématique de la manière suivante : comment maîtriser l'inertie lorsque  $F$  vérifie une hypothèse de croissance plus forte que la convexité? Plusieurs stratégies sont proposées, basées sur des procédés algorithmiques ou sur le parallèle entre schémas numériques et systèmes dynamiques.

Une première partie est consacrée aux méthodes de restart. Les travaux qui y sont présentés font suite au constat suivant : lorsque la fonction à minimiser vérifie une hypothèse de croissance (forte convexité par exemple), les méthodes les plus rapides doivent être calibrées en fonction du paramètre de croissance (ou de manière équivalente en fonction du conditionnement) qui peut être difficile à estimer. Les stratégies de restart visent à répondre à la question : comment profiter des hypothèses de croissance d'une fonction lorsque le paramètre de croissance n'est pas connu? Je propose deux schémas de restart de FISTA permettant d'obtenir des garanties théoriques de convergence rapide en s'affranchissant de la connaissance de certains paramètres de géométrie.

Dans un deuxième temps, on étudie un moyen alternatif permettant de maîtriser l'inertie et atténuer les comportements oscillatoires. Cette analyse porte sur un système dynamique introduit par Attouch, Peyrouquet et Redont qui est proche de celui proposé par Su, Boyd et Candès et fait intervenir un terme de friction lié à la hessienne de  $F$ . L'introduction de ce terme de friction permet d'atténuer les oscillations des trajectoires et des schémas numériques de premier ordre peuvent être dérivés de ce système. On se concentre ici sur les propriétés de convergence des trajectoires de ce système lorsque  $F$  vérifie des hypothèses de croissance. Dans ce contexte, on montre que l'introduction du terme lié à la hessienne permet d'obtenir des propriétés d'intégrabilité sur le gradient de la fonction  $F$  qui sont de plus en plus fortes à mesure que les hypothèses de croissance sont renforcées. Comme les hypothèses considérées sont liées à l'inégalité de Łojasiewicz, ces propriétés donnent des garanties sur la décroissance de l'erreur.

La dernière partie est dédiée à l'hypothèse d'unicité du minimiseur dans l'analyse de convergence de méthodes inertielles. Dans la littérature, les résultats de convergence les plus forts pour des systèmes dynamiques et des schémas numériques lorsque  $F$  vérifie une condition de croissance s'appuient, sauf exception, sur l'hypothèse que  $F$  a un unique minimiseur. Les questions posées dans ce chapitre sont : à quoi est due cette hypothèse? Est-il possible d'obtenir des résultats de convergence rapide en s'en affranchissant? On montre que cette hypothèse est principalement liée aux techniques de preuves et qu'elle n'est pas nécessaire à l'obtention de vitesses rapides. Dans le cadre continu, une hypothèse de géométrie sur  $X^*$  est proposée pour étendre plusieurs résultats connus. On donne enfin des

résultats de convergence pour V-FISTA et FISTA sans hypothèse d'unicité du minimiseur.

## Inertial methods in optimization under growth conditions

**Abstract :** This thesis manuscript is devoted to the optimization of composite convex functions in a deterministic framework. In this context, I have focused on the convergence analysis of optimization algorithms under growth assumptions.

When it comes to minimizing such functions, the most efficient first-order methods are said to be inertial, since they introduce a notion of momentum. While the introduction of this inertia is generally beneficial, it is necessary to parameterize this type of method properly to obtain a solution quickly and avoid pathological behavior. For example, some so-called Heavy-Ball methods require knowledge of the condition number of the function to be minimized in order to be used.

The question this thesis seeks to answer is : which method should be applied given the geometric assumptions verified by  $F$ ? Since the methods studied are inertial, we can rephrase this problem as follows : how can inertia be controlled when  $F$  verifies a growth hypothesis that is stronger than convexity? Several strategies are proposed, based on algorithmic procedures or on the link existing between numerical schemes and dynamical systems.

The first part is devoted to restart methods. The work presented here stems from the following observation : when the function to be minimized satisfies a growth hypothesis (e.g. strong convexity), the fastest methods must be calibrated according to the growth parameter (or, equivalently, according to the condition), which can be difficult to estimate. Restart strategies aim to answer the question : how can we take advantage of the growth assumptions of a function when the growth parameter is not known? Two FISTA restart schemes are introduced that ensure a fast exponential decay of the error without requiring the knowledge of some geometry parameters.

Secondly, we study an alternative approach to controlling inertia and attenuating oscillatory behavior. This analysis concerns a dynamical system introduced by Attouch, Peypouquet and Redont, which is close to that proposed by Su, Boyd and Candès, and involves a friction term driven by the Hessian of  $F$ . The introduction of this friction term dampens trajectory oscillations, and first-order numerical schemes can be derived from this system. We focus on the convergence properties of the trajectories of this system when  $F$  satisfies growth conditions. In this context, we show that the introduction of the Hessian term leads to integrability properties on the gradient of the function  $F$  that become stronger and stronger as the growth assumptions are strengthened. Since the assumptions considered are related to the Łojasiewicz inequality, these properties provide guarantees on error decay.

The last part is devoted to the assumption of uniqueness of the minimizer in the convergence analysis of inertial methods. In the literature, the strongest convergence results for dynamical systems and numerical schemes when  $F$  verifies a growth condition are based, with a few exceptions, on the assumption that  $F$  has a unique minimiser. The questions raised in this chapter are : what is the reason for this assumption? Is it possible to obtain fast convergence results without it? We show that this assumption is mainly related to proof techniques and that it is not necessary to obtain fast convergence. In the continuous framework, a geometry assumption on  $X^*$  is proposed to extend several known results. Finally, convergence results are given for V-FISTA and FISTA without the minimiser uniqueness assumption.