

Stage ingénieur  
présenté par

**Baptiste GUILBERY**

Elève Ingénieur de l'INSA Rennes

Spécialité MA

Année scolaire 2022-2023

## **Etude de méthodes du premier ordre accélérées en optimisation convexe**

**Lieu du Projet de Fin d'Études**

**IRMAR**

**Tuteur du Projet de Fin d'Études**

**Mounir HADDOU**

**Correspondant pédagogique INSA**

**Aziz BELMILOUDI**



# Rapport de stage

## Optimisation non différentiable

Baptiste GUILBERY

4 octobre 2024

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Optimisation non différentiable</b>	<b>2</b>
1.1 Sans contrainte . . . . .	2
1.2 Avec une contrainte d'inégalité . . . . .	4
1.2.1 Convergence . . . . .	5
<b>2 Étude de l'équation différentielle ordinaire iDINAM2</b>	<b>6</b>
2.1 iDINAM . . . . .	6
2.2 iDINAM d'ordre 2 . . . . .	7
2.2.1 Existence des solutions . . . . .	7
2.2.2 Comportement asymptotique de iDINAM2 . . . . .	8
<b>3 Étude de l'algorithme iDINAM2-split</b>	<b>11</b>
3.1 iDINAM2-split . . . . .	11
3.2 Etude numérique . . . . .	13
3.3 Première Analyse . . . . .	14
3.4 Comparaison de iDINAM-split et de iDINAM2-split . . . . .	15
3.5 iDINAMN-split . . . . .	18
<b>4 Conclusion</b>	<b>20</b>
<b>Bibliographie</b>	<b>21</b>
<b>5 Annexes</b>	
5.1 Preuve du théorème 1 . . . . .	
5.2 Preuve du théorème 2 . . . . .	
5.3 Preuve Théorème 3 . . . . .	
5.4 Preuve du théorème 4 . . . . .	
5.5 Etude numérique . . . . .	
<b>6 Abstract</b>	

# Introduction

J'ai effectué mon stage à l'Institut de Recherche Mathématique de Rennes (IRMAR). Il est composé de 230 membres répartis entre plusieurs structures (INSA, ENS Rennes, CNRS...) et fait partie des plus grands instituts de recherche en mathématiques de France. L'institut est divisé en trois pôles : Analyse, Géométrie et Aléatoire. Au sein de ceux ci, il y a 10 équipes au total qui travaillent sur des sujets plus spécialisés. J'étais rattaché à l'équipe analyse numérique du pôle Analyse.

L'objectif de mon stage était double. Je souhaitais principalement découvrir le monde de la recherche. Le sujet initial d'étude d'algorithmes d'optimisation non-différentiable a été redéfini pendant celui-ci. Le second objectif était de développer un nouvel algorithme nommé iDINAM2-split. Cet algorithme est dérivé de iDINAM-split présenté par van Nam VO lors de sa thèse [VO22]. iDINAM2-split converge plus rapidement que iDINAM-split sous certaines conditions. Pour cela, j'ai travaillé principalement seul, mais aidé par Mr Haddou mon maître de stage et Mr Ley principalement pour les démonstrations des différents théorèmes.

Ce rapport est découpé en trois parties. Dans la première, je présente le problème d'optimisation non différentiable proposé comme sujet de stage et une rapide étude numérique de quelques algorithmes que j'ai découverts lors de recherches documentaires. Dans la seconde partie, suite à la modification du périmètre, j'étudie théoriquement une nouvelle équation différentielle ordinaire dérivée de iDINAM également introduite dans la thèse de van Nam VO. Et ensuite, dans le troisième chapitre, je propose un algorithme basé sur cette équation différentielle ordinaire que l'étudie théoriquement et numériquement.

# Chapitre 1

## Optimisation non différentiable

Lors de mon stage, j'ai commencé par considérer le problème d'optimisation :

$$\min_{x \in \mathbb{R}^n} R(x) + f(x) \quad (1.1)$$

ou  $R \in \Gamma_0$  ( $\Gamma_0$  est l'ensemble des fonctions convexes propres et semi-continues inférieurement) et  $f(x)$  est convexe, coercive et dérivable. Ce problème est général car il permet de minimiser des fonctions différentiables et des fonctions non différentiables en même temps. Ces dernières ne sont pas utilisables avec les algorithmes classiques d'optimisation (descente de gradient, méthode du gradient projeté).

J'ai commencé par étudier le cas sans contrainte puis dans un second temps en ajoutant une contrainte. Mes implémentations de tous les algorithmes présentées dans cette partie sont disponibles sur <https://github.com/LeProblemeEDMN/StageOND>.

### 1.1 Sans contrainte

Dans cette première partie, je vais présenter les trois algorithmes permettant de résoudre ce problème que j'ai étudié. J'ai commencé par lire le cours d'optimisation différentiable [M23]. Celui-ci, m'a permis d'implémenter en python l'algorithme du sous-gradient qui peut être vu comme une version pour une fonction non-différentiable de l'algorithme du gradient. Pour résoudre (1.1), j'ai utilisé le sous gradient de  $R(x) + f(x)$  qui vaut  $\partial R(x) + \nabla f(x)$ .

J'ai également implémenté la méthode du gradient proximal à l'aide de ce cours. Celle-ci est inspirée de la méthode du gradient projeté et remplace la projection par l'opérateur proximal défini comme :

$$\text{prox}_{\gamma H}(x) = \arg \min_{y \in \mathbb{R}^n} \gamma R(y) + \frac{1}{2} \|x - y\|^2$$

J'ai également lu le papier [FISTA], celui-ci introduit un nouvel algorithme (FISTA) qui permet de résoudre (1.1). FISTA peut être vu comme une modification de l'algorithme du gradient proximal utilisant un terme inertiel (c'est à dire en utilisant  $x_{k-2}$ ).

J'ai en suite comparé l'efficacité des trois algorithmes. Pour cela j'ai réalisé des profils de performance avec ces trois algorithmes. Lors de l'analyse, j'ai remarqué que pour le problème

considéré, l'algorithme du sous gradient ne convergeait pas la majorité du temps. J'ai donc comparé les deux autres algorithmes entre eux.

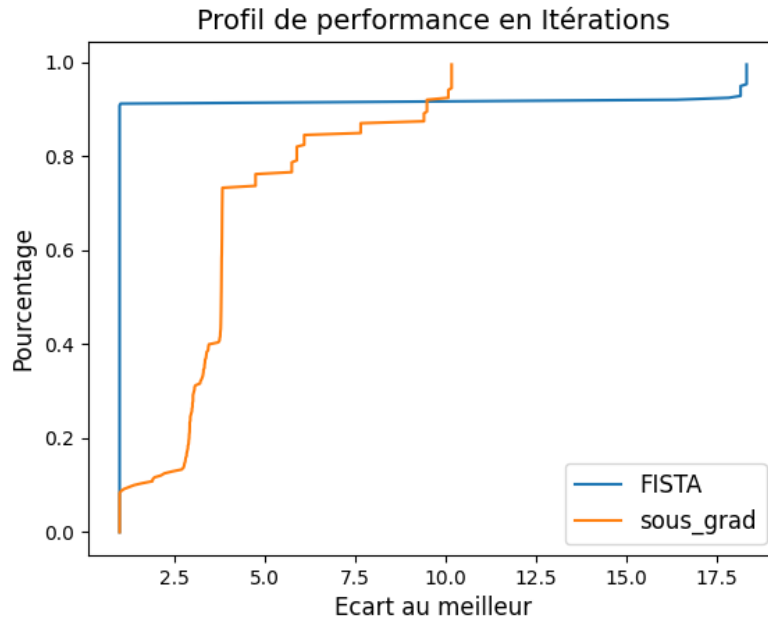


FIGURE 1.1 – Profil de performance en itération de FISTA et du gradient proximal

Ce profil de performance indique que FISTA est toujours plus rapide que l'algorithme du sous gradient. Lors des test, il n'a pas convergé dans environ 10% des cas. J'ai pu conclure de cette analyse que FISTA était plus rapide mais moins robuste que l'algorithme du sous gradient.

Afin de pallier à ce problème il existe une version de FISTA avec des redémarrages qui doivent améliorer la convergence de l'algorithme. Je n'ai pas eu le temps de l'étudier pendant mon stage ayant choisi de me concentrer sur l'étude du nouvel algorithme (contenu des chapitres 2 et 3).

J'ai travaillé également le cas particulier de (1.1) suivant :

$$\min_{x \in \mathbb{R}^n} \|Mx - b\|^2 + \gamma|x|$$

Ce problème peut être résolu avec l'algorithme FISTA. Ce problème est un problème de régression linéaire quand  $\gamma = 0$ . L'introduction du second terme favorise l'apparition de zéros dans les coefficients de la régression dans l'objectif de réduire sa taille de stockage. Plus  $\gamma$  sera grand, plus il y aura de zéros. J'ai réalisé un graphique représentant à gauche le nombre de zéros (sur 40 possible) et à droite l'erreur entre la solution optimale ( $\gamma = 0$ ) pour plusieurs rapport ( $\frac{1}{\gamma}$ ).

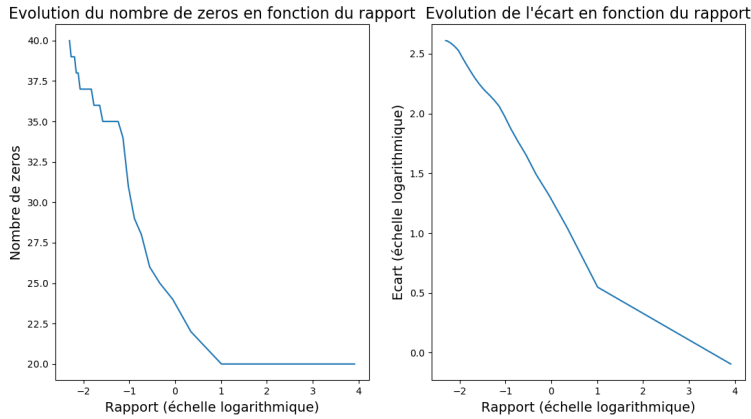


FIGURE 1.2 – Evolution du nombre de 0 en fonction du rapport

Sur la figure 1.2, plus le rapport est petit plus il y a de zéros dans les coefficients de la régression, pour  $\gamma > 7$ , tous les coefficients valent 0 et l'erreur est à son maximum. Pour un coefficient  $\gamma < 0.36$ , le nombre de zéros reste stable à environ 20 (le nombre de zéros présent dans la régression simple).

## 1.2 Avec une contrainte d'inégalité

J'ai ensuite essayé d'ajouter une contrainte à 1.1. Le nouveau problème s'écrit comme ceci :

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & R(x) + f(x) \\ \text{s.t.} \quad & \phi(x) \leq 0 \end{aligned} \quad (1.2)$$

Afin de le résoudre, j'ai utilisé une version simplifiée (car il y a une unique contrainte) l'algorithme d'Uzawa. Celui-ci utilise le Lagrangien défini par :

$$L(x, \lambda) = R(x) + \lambda \phi(x)$$

Voici l'algorithme d'Uzawa : Le problème de résolution de  $\arg \min_{x \in \mathbb{R}^n} L(x, \lambda_k)$  est dans mon cas un

---

### Algorithm 1 Algorithme d'Uzawa

---

**Require:**  $\epsilon > 0, \rho \geq 0, x_0 \in \mathbb{R}^n$

$\lambda_0 \leftarrow 0$

**while**  $\|x_k - x_{k-1}\| > \epsilon$  **do**

$x_{k+1} \in \arg \min_{x \in \mathbb{R}^n} L(x, \lambda_k)$

$\lambda_{k+1} \leftarrow \max(0, \lambda_k + \rho \phi(x_{k+1}))$

**end while**

---

problème d'optimisation composé et peut être résolu par des algorithmes présentés dans la partie précédente.

### 1.2.1 Convergence

J'ai ensuite cherché à démontrer la convergence de cette version simplifiée de l'algorithme d'Uzawa (la démonstration est en annexe 5.1).

**Theorème 1 (Condition de convergence)** *Si l'on a  $R \in \Gamma_0$ ,  $\phi(x)$  convexe, coercive et que le problème (1.1) admet au moins une solution. Si  $\{\lambda_k\}$  et  $\{x_k\}$  sont les suites obtenues avec 1 alors :*

1. *La suite  $\{\lambda_k\}$  converge vers  $\lambda^*$*
2. *La suite  $\{x_k\}$  converge vers  $x^*$*
3.  *$(x^*, \lambda^*)$  est un point selle de  $L(x, \lambda)$  et  $x^*$  est une solution de 1.1*

Pour cela, j'ai cherché à démontrer que la suite  $(x_k, \lambda_k)$  convergeait vers  $(x^*, \lambda^*)$  et qu'il était un point selle du Lagrangien et donc que  $x^*$  était une solution optimale.

J'ai ensuite pu implémenter cet algorithme en python. Je n'ai pas eu le temps d'étudier numériquement ce dernier car j'ai priorisé la création et l'étude de l'algorithme du chapitre 3 en accord avec Mr Haddou. Au début de mon stage, j'ai découvert un nouveau problème 1.1 et implémenté trois nouveaux algorithmes. J'ai ensuite pu utiliser ceux-ci pour résoudre un problème d'optimisation non différentiable avec une unique contrainte.



## Chapitre 2

# Étude de l'équation différentielle ordinaire iDINAM2

Mon problème initial (1.1) peut se reformuler comme  $0 \in \nabla f(x) + \partial R(x)$ . Cette seconde forme est un sous cas de :

$$B(x) + \nabla f(x) = 0 \quad (2.1)$$

Dans ce chapitre, je vais présenter une équation différentielle ordinaire (EDO) ayant pour objectif de trouver un  $x$  pour  $B$  un opérateur  $\lambda$ -cocoercif (voir définition plus bas) et  $f$  une fonction convexe différentiable. A partir de celle-ci, j'ai pu créer un algorithme permettant de résoudre (2.1).

Ce problème générique peut également servir à résoudre des systèmes non linéaires.

Un opérateur  $B$  est dit  $\lambda$ -cocoercif si :

$$\forall x, y \langle B(x) - B(y), x - y \rangle \geq \lambda \|B(x) - B(y)\|^2$$

Cette propriété implique que  $B$  est  $\frac{1}{\lambda}$ -Lipschitzien.

J'ai commencé par lire la thèse de van Nam VO [VO22] j'y ai découvert une première EDO et son algorithme associé pouvant résoudre ce problème et les travaux déjà réalisés sur cette équation. J'ai ensuite pu proposer une amélioration de celle-ci en créant l'EDO iDINAM2 et l'algorithme iDINAM2-split. J'ai ensuite cherché à démontrer l'existence d'une solution puis les propriétés asymptotiques de celle-ci.

### 2.1 iDINAM

Afin de résoudre (2.1), Attouch H. et Maingé P.E. ont étudié l'EDO suivante dans [Att10] :

$$\ddot{x}(t) + \gamma \dot{x}(t) + B(x(t)) + \nabla f(x(t)) = 0 \quad (2.2)$$

Ils ont pu démontrer que cette EDO convergeait bien vers une solution de (2.1) si  $\lambda\gamma^2 > 1$ . Par la suite, van Nam VO a étudié l'EDO suivante dans [VO22] :

$$\ddot{x}(t) + \gamma \dot{x}(t) + B(x(t) + \beta_b \dot{x}(t)) + \nabla f(x(t) + \beta_f \dot{x}(t)) = 0 \quad (\text{iDINAM})$$

Cette équation est très similaire à celle présentée par Attouch H. et Maingé P.E., avec l'ajout de la dérivée seconde dans  $B$  et  $\nabla f$ . Il a également pu démontrer sa convergence faible vers une solution de (2.1) sous quelques hypothèses sur  $\beta_b$  et  $\beta_f$ .

Ces deux EDO utilisent uniquement  $\nabla f$  et  $B$ , cela est très intéressant pour les schémas qui seront créés à partir de ces EDO, ne nécessitant pas de termes tel que  $\nabla^2 f$  et  $\nabla B$  pas toujours disponibles soit du fait de fonctions non différentiables soit en raison des calculs des opérateurs trop coûteux.

## 2.2 iDINAM d'ordre 2

Lors de mon stage j'ai cherché une nouvelle EDO avec les mêmes propriétés que (iDINAM). C'est à dire une EDO ayant pour limite une solution de (2.1) et utilisant uniquement les opérateurs  $B$  et  $\nabla f$ . J'ai considéré dans (iDINAM) les termes  $B(x(t) + \beta_b \dot{x}(t))$  (resp.  $\nabla f(x(t) + \beta_f \dot{x}(t))$ ) comme des approximations de  $B(x(t + \beta_b))$  (resp.  $\nabla f(x(t + \beta_f))$ ) obtenues via un développement de Taylor d'ordre 1 en  $t + \beta_b$  (resp.  $t + \beta_f$ ). J'ai augmenté l'ordre de celui-ci pour obtenir iDINAM2.

$$\ddot{x}(t) + \gamma \dot{x}(t) + B\left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) + \nabla f\left(x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\right) = 0 \quad (\text{iDINAM2})$$

Les paramètres  $\beta_f, \beta_b, \gamma$  seront tous choisis strictement positifs. De plus, par construction, la deuxième propriété est bien vérifiée.

### 2.2.1 Existence des solutions

L'objectif de cette partie va être de démontrer l'existence d'une solution globale à (iDINAM2).

Cela est simple pour (iDINAM), en effet, cette EDO peut simplement être réécrite sous forme résolue. Comme  $B$  et  $\nabla f$  sont lipschitziens, van Nam VO applique le théorème de Cauchy-Lipschitz global pour avoir une unique solution globale sans hypothèse supplémentaire.

J'ai pu obtenir le résultat suivant pour (iDINAM2) (la démonstration est en annexe 5.2) :

**Théorème 2 (Existence)** Soit  $\beta_b, \beta_f, \gamma > 0$ ,  $B$  un opérateur  $\lambda$ -cocercif et  $f$  une fonction convexe,  $C^1$  tel que son gradient soit lipschitzien. Si  $\beta_f \gamma \in ]1; 2[$  et  $\lambda \geq \beta_f \frac{\beta_f^2 - \beta_b^2}{2}$ , que  $(\nabla f + B)^{-1}(0) \neq \emptyset$  et  $\forall t > 0, \int_0^t g(s)ds < \int_0^t \left(a\beta_f - \frac{\rho h + rc}{2}\right) \|X(s)\|^2 + \mu(\|Y(s)\|^2 + \|\dot{x}(s)\|^2)ds$  (les différents termes de cette dernière formule sont ceux de la démonstration). Alors le problème de Cauchy ayant pour équation (iDINAM2) avec les conditions initiales  $x(0) = x_0, \dot{x}(0) = x_1$  admet une solution globale.

Ce résultat est plus faible car il n'y a plus l'unicité de la solution. En effet, (iDINAM2) ne peut plus être écrit sous forme résolue. Il a donc fallu trouver un nouvel argument afin d'avoir un résultat. Pour cela, j'ai ajouté un terme de viscosité pour obtenir :

$$\epsilon > 0, \epsilon \ddot{x}(t) + \ddot{x}(t) + \gamma \dot{x}(t) + B\left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) + \nabla f\left(x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\right) = 0 \quad (\text{iDINAM2e})$$

L'objectif étant de faire tendre  $\epsilon$  vers 0 pour retrouver (iDINAM2). Cette équation à bien une unique solution globale par le théorème de Cauchy-Lipschitz global. J'ai ensuite majoré  $\|x(t)\|$  et  $\|\dot{x}(t)\|$  pour utiliser le théorème d'Ascoli. Celui-ci m'a permis de démontrer le théorème.

La plus grande partie de la démonstration a pour objectif de majorer  $\|x(t)\|$  et  $\|\dot{x}(t)\|$ . Pour cela, je me suis appuyé sur les travaux de van Nam VO [VO22] qui dans sa démonstration sur les propriétés asymptotiques de (iDINAM) obtient les mêmes résultats. Les deux équations étant très similaires, j'ai adapté cette dernière. J'ai étudié cette fonction :

$$\begin{aligned} \varepsilon_p(t) := a \left( f \left( x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right) - f(p) - \langle \nabla f(p), x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) - p \rangle \right) + \\ \frac{1}{2} \|x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) - p\|^2 + \frac{d}{2} \|x - p\|^2 \end{aligned}$$

Celle-ci est positive et l'objectif est de démontrer qu'elle est strictement décroissante (en choisissant bien  $a$  et  $d$ ), et de ce fait bornée, ce qui m'a permis d'affirmer que  $\|x - p\|$  et  $\|x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) - p\|$  sont bornées. Pour cela, de nouveaux termes dus à l'ajout de  $\frac{\beta_f^2}{2} \ddot{x}(t)$  sont apparus. Afin de les contrôler, j'ai du rajouter l'hypothèse :  $\forall t > 0, \int_0^t g(s) ds < \int_0^t \left( a\beta_f - \frac{\rho h + rc}{2} \right) \|X(s)\|^2 + \mu(\|Y(s)\|^2 + \|\dot{x}(s)\|^2) ds$ . Et j'ai ensuite démontré des majorations sur  $\|x(t)\|$  et  $\|\dot{x}(t)\|$  à partir de ces résultats.

Dans le théorème, précédent l'hypothèse  $\forall t > 0, \int_0^t g(s) ds < \int_0^t \left( a\beta_f - \frac{\rho h + rc}{2} \right) \|X(s)\|^2 + \mu(\|Y(s)\|^2 + \|\dot{x}(s)\|^2) ds$  implique que dans le cas général il faut calculer une trajectoire avant de savoir si le théorème est applicable. Mais en réalité, il est simple d'observer la convergence de la trajectoire une fois celle-ci calculée. Cette hypothèse, rend alors le théorème inutile en l'état pour statuer à l'avance la convergence de l'algorithme.

Cette condition peut également être reformulée en utilisant des formes quadratiques (avant d'obtenir (5.14)) avec  $X, Y, \ddot{x}, \dot{x}$ . Après des calculs, il est possible de déterminer des nouvelles conditions sur les paramètres et cette hypothèse :  $\forall t > 0, \int_0^t \langle \ddot{x}(s), x(s) - p \rangle ds < \int_0^t \mu_1 \|X(s)\|^2 + \mu_2 \|Y(s)\|^2 + \mu_3 \|\dot{x}(s)\|^2 + \mu_4 \|\ddot{x}(s)\|^2 ds$ . Je n'ai pas choisi de présenter cette méthode dans ma preuve car elle ajouterait de nouvelles contraintes sur les paramètres sans permettre d'avoir une hypothèse simplifiée d'utilisation. Cependant, cette nouvelle formulation pourrait être plus facile à démontrer.

De plus, en considérant  $f(x) = x^2$  et  $B(x) = 0$ , il s'avère que pour  $\beta_f = \frac{1}{2}, \gamma = 1$  iDINAM2 possède une solution globale alors même que les conditions du théorème ne sont pas satisfaites ( $\gamma\beta_f < 1$ ). Celui-ci sert donc à prouver qu'il existe des paramètres permettant d'obtenir l'existence d'une solution sans pouvoir déterminer à l'avance si des valeurs spécifiques fourniront un résultat.

### 2.2.2 Comportement asymptotique de iDINAM2

Dans cette partie, une fois démontré qu'il existe sous certaines condition un solution globale à iDINAM2, l'objectif est de prouver qu'elle converge bien vers une solution de (2.1). J'ai pu obtenir le résultat suivant (la démonstration est en annexe 5.3) :

**Théorème 3 (Convergence)** Soit  $\beta_b, \beta_f, \gamma > 0$ ,  $B$  un opérateur  $\lambda$ -cocercif et  $f$  une fonction convexe,  $\mathcal{C}^1$  tel que son gradient soit lipschitzien. Si  $\beta_f \gamma \in ]1; 2[$  et  $\lambda \geq \beta_f \frac{\beta_f^2 - \beta_b^2}{2}$ ,  $\forall p \in$

$(\nabla f + B)^{-1}(0)$ , que  $\int_0^\infty \|\ddot{x}(s)\|^2 ds < \infty$  et  $(\nabla f + B)^{-1}(0) \neq \emptyset$  et  $\forall t > 0$ ,  $\int_0^t g(s) ds < \int_0^t \left(a\beta_f - \frac{\rho h + rc}{2}\right) \|X(s)\|^2 + \mu(\|Y(s)\|^2 + \|\dot{x}(s)\|^2) ds$  (les différents termes de cette dernière formule sont ceux de la démonstration). Nous avons les résultats suivants :

$$x(t) \text{ converge faiblement vers } x^*, B(x^*) + \nabla f(x^*) = 0$$

$$\lim_{t \rightarrow +\infty} \|\dot{x}(t)\| = 0, \lim_{t \rightarrow +\infty} \|\ddot{x}(t)\| = 0$$

$$\lim_{t \rightarrow +\infty} \|B(x(t)) - B(p)\| = 0, \lim_{t \rightarrow +\infty} \|\nabla f(x(t)) - \nabla f(p)\| = 0$$

Et les majorations suivantes :

$$\int_0^\infty \|\dot{x}(s)\|^2 ds < \infty, \int_0^\infty \|\ddot{x}(s)\|^2 ds < \infty$$

$$\int_0^\infty \left\| \nabla f \left( x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right) - \nabla f(p) \right\|^2 dt < \infty, \int_0^\infty \left\| B \left( x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t) \right) - B(p) \right\|^2 dt < \infty$$

$$\int_0^\infty \left\| \frac{d}{dt} \nabla f \left( x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right) \right\|^2 dt < \infty, \int_0^\infty \left\| \frac{d}{dt} B \left( x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t) \right) \right\|^2 dt < \infty$$

Pour le prouver, je me suis appuyé sur la démonstration du théorème 2 (en prenant  $\epsilon = 0$ ) afin d'obtenir des majorations pour  $\|x(t)\|$  et  $\|\dot{x}(t)\|$ . Cela implique que les hypothèses faites au théorème 2 doivent toujours être validées. Pour la suite, je me suis appuyé sur la thèse de van Nam VO [VO22], pour cela l'argument utilise la propriété d'Opial (voir [Pey10]) :

**Lemme 1 (Lemme d'Opial)** Soit  $x_n$  une suite dans un espace de Hilbert  $H$  et  $F \in X$ , si l'on a :

$$\|x_n - p\| \text{ converge pour tout } p \in F$$

L'ensemble des valeurs vers lequel  $x_n$  converge faiblement est inclus dans  $F$ .  
Alors  $x_n$  converge faiblement vers  $x^* \in F$

Pour prouver ces deux points, j'ai cherché à démontrer en premier lieu les quatre limites du théorème. Pour cela, j'ai utilisé le lemme de Barlabat :

**Lemme 2 (Lemme de Barlabat)** Si  $p \geq 1, q > 1$  et que  $f \in L^p(0, \infty), \dot{f} \in L^q(0, \infty)$  alors  $\lim_{t \rightarrow \infty} f(t) = 0$ .

Afin d'établir que  $\int_0^\infty \left\| B \left( x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t) \right) - B(p) \right\|^2 dt < \infty$  il faut démontrer que :  $\int_0^\infty \|\ddot{x}(s)\|^2 ds < \infty$ .

Je n'ai pas réussi à le faire, je l'ai mis en tant qu'hypothèse afin de pouvoir continuer la démonstration.

Une fois les quatre limites démontrées, la seconde assumption du lemme d'Opial est facilement vérifiable. Pour la première, j'ai utilisé le lemme suivant (la preuve peut être trouvée dans [Att10]) :

**Lemme 3** Soit  $w \in C^2(0, \infty)$  une fonction minorée pour laquelle il existe  $g \in L^1(0, \infty)$  et  $\omega > 0$  :

$\ddot{x}(t) + \omega \dot{x}(t) \leq g(t)$  Alors  $w$  converge quand  $t \rightarrow \infty$

En l'appliquant avec  $w(t) = \frac{1}{2}\|x(t) - p\|^2$  j'ai pu démontrer le théorème 3.

Comme pour le théorème 1, il apparaît que le théorème 2 n'est pas utilisable avant d'avoir trouvé une solution à (iDINAM2). Pour y remédier, il faut démontrer que  $\int_0^\infty \|\ddot{x}(s)\|^2 ds < \infty$ . L'intérêt de ce théorème est d'affirmer qu'il existe des paramètres ou (iDINAM2) converge mais sans pouvoir vérifier à l'avance si les valeurs choisies fourniront un résultat.

Lors de mon stage, j'ai pu apporter les premiers éléments de réponse à l'existence et la convergence de iDINAM2. Afin de compléter les preuves il est nécessaire de démontrer deux majorations. Lors de mes simulations numériques, je n'ai pas pu trouver de contre-exemple. Cela pourrait faire l'objet d'une prochaine recherche.

## Chapitre 3

# Étude de l'algorithme iDINAM2-split

Dans ce chapitre je vais présenter mon travail de conception d'un schéma basé sur (iDINAM2) ainsi qu'une preuve de la convergence et l'étude numérique. En effet, même si les démonstrations du chapitre précédent ne permettent pas d'obtenir des résultats sur la convergence de iDINAM2, il est intéressant de créer un schéma basé sur celui-ci afin d'obtenir des résultats numériques.

### 3.1 iDINAM2-split

Pour créer l'algorithme iDINAM2-split, j'ai discrétisé (iDINAM2) avec un schéma implicite selon  $B$  et explicite selon  $\nabla f$ , d'où le nom de iDINAM2-split. J'ai obtenu l'équation suivante :

$$\begin{aligned} \frac{1}{h^2}(x_{k+1} - 2x_k + x_{k-1}) + \frac{\gamma}{h}(x_{k+1} - x_k) + B \left( x_{k+1} + \frac{\beta_b}{h}(x_{k+1} - x_k) + \frac{\beta_b^2}{2h^2}(x_{k+1} - 2x_k + x_{k-1}) \right) \\ + \nabla f \left( x_k + \frac{\beta_f}{h}(x_k - x_{k-1}) + \frac{\beta_f^2}{2h^2}(x_k - 2x_{k-1} + x_{k-2}) \right) = 0 \end{aligned} \quad (3.1)$$

En réarrangeant les termes j'ai obtenu :

$$\begin{aligned} hB \left[ x_{k+1} \left( 1 + \frac{\beta_b}{h} + \frac{\beta_b^2}{2h^2} \right) - x_k \left( \frac{\beta_b}{h} + \frac{\beta_b^2}{h^2} \right) + x_{k-1} \frac{\beta_b^2}{2h^2} \right] + x_{k+1} \left( \gamma + \frac{1}{h} \right) = \\ x_k \left( \gamma + \frac{2}{h} \right) - \frac{x_{k-1}}{h} - h \nabla f \left[ x_k \left( 1 + \frac{\beta_f}{h} + \frac{\beta_f^2}{2h^2} \right) - x_{k-1} \left( \frac{\beta_f}{h} + \frac{\beta_f^2}{h^2} \right) + x_{k-2} \frac{\beta_f^2}{2h^2} \right] \end{aligned}$$

En prenant  $\delta_b = 1 + \frac{\beta_b}{h} + \frac{\beta_b^2}{2h^2}$ ,  $\delta_f = 1 + \frac{\beta_f}{h} + \frac{\beta_f^2}{2h^2}$ ,  $\mu_b = \frac{\beta_b}{h} + \frac{\beta_b^2}{h^2}$ ,  $\mu_f = \frac{\beta_f}{h} + \frac{\beta_f^2}{h^2}$  et  $\alpha = \left( \gamma + \frac{1}{h} \right)^{-1} \delta$  j'ai obtenu :

$$\begin{aligned} (Id + \alpha h B) (\delta_b x_{k+1} - \mu_b x_k + (1 + \mu_b - \delta_b) x_{k-1}) = x_k \left( \alpha \gamma + \frac{2\alpha}{h} - \mu_b \right) + \\ x_{k-1} \left( 1 + \mu_b - \delta_b - \frac{\alpha}{h} \right) - \alpha h \nabla f (\delta_f x_k - \mu_b x_{k-1} + (1 + \mu_f - \delta_f) x_{k-2}) \end{aligned}$$

**Algorithm 2** iDINAM2-split**Require:**  $x_0, x_1, x_2 \in \mathbb{R}^n$ ,  $\epsilon > 0$ 

$$\delta_b \leftarrow 1 + \frac{\beta_b}{h} + \frac{\beta_b^2}{2h^2}, \delta_f \leftarrow 1 + \frac{\beta_f}{h} + \frac{\beta_f^2}{2h^2}$$

$$\mu_b \leftarrow \frac{\beta_b}{h} + \frac{\beta_b^2}{h^2}, \mu_f \leftarrow \frac{\beta_f}{h} + \frac{\beta_f^2}{h^2}$$

$$\alpha \leftarrow \left(\gamma + \frac{1}{h}\right)^{-1} \delta$$

**while**  $\|x_k - x_{k-1}\| > \epsilon$  **do**

$$z_k \leftarrow \nabla f(\delta_f x_k - \mu_b x_{k-1} + (1 + \mu_f - \delta_f)x_{k-2})$$

$$y_k \leftarrow x_k \left(\alpha\gamma + \frac{2\alpha}{h} - \mu_b\right) + x_{k-1} \left(1 + \mu_b - \delta_b - \frac{\alpha}{h}\right) - \alpha h z_k$$

$$x_{k+1} \leftarrow \frac{1}{\delta_b} (Id + \alpha h B)^{-1}(y_k) + x_k \frac{\mu_b}{\delta_b} - x_{k-1} \frac{1 + \mu_b - \delta_b}{\delta_b}$$

**end while**

A partir de cette dernière équation, je propose l'algorithme suivant

Mon implémentation de iDINAM2-split en python est disponible sur <https://github.com/LeProblemeEDMN/Stag>

Il est important de noter que iDINAM2-split dépend de  $x_{k-2}$  ce qui n'était pas le cas de iDINAM-split (voir [VO22]). Cette dépendance implique, en général une convergence plus rapide en nombre d'itération que pour les schémas sans celle-ci. En contrepartie, pour chaque itération un plus grand nombre d'étape est nécessaire mais pour iDINAM2-split, la complexité d'une itération est majoritairement dictée par la complexité de la résolution de  $(Id + \alpha h B)^{-1}(y_k)$ . Dans l'algorithme 2, si  $B$  est le sous gradient d'un opérateur (soit  $B(x) \in \partial G(x)$ ), le calcul de  $(Id + \alpha h B)^{-1}(y_k)$  peut être effectué avec  $prox_{\alpha h G}(y_k)$ . En effet :

$$x_k = prox_{\alpha h G}(y_k) = \arg \min_{x \in \mathbb{R}^n} \alpha h G(x) + \frac{1}{2} \|x - y_k\|^2$$

Soit :

$$0 = \alpha h B(x_k) + x_k - y_k \Leftrightarrow y_k = (\alpha h B + Id)(x_k)$$

J'ai pu démontrer le théorème suivant, celui ci démontre que la suite  $x_k$  converge vers une solution de (2.1) (la démonstration est en annexe 5.4).

**Theorème 4** Soit  $\beta_b, \beta_f, \gamma > 0$ ,  $B$  un opérateur  $\lambda$ -cococercif et  $f$  une fonction convexe,  $\mathcal{C}^1$  tel que son gradient soit lipschitzien. Si

$$\beta_f \gamma > 1, \lambda > \frac{(1 - h^2)\beta_f^2 - \beta_b^2}{2}, L > h\beta_f - \frac{h^2\beta_f^2}{2}$$

$$\sum_{k=3}^{\infty} \langle x_k - p, x_k - 3x_{k-1} + 3x_{k-2} + x_{k-3} \rangle < \infty$$

En notant :

$$a = \frac{\delta^2 + \delta\sigma + \delta}{2} - \frac{\beta_f^2 \delta^2 r}{4h^2} - \frac{\beta_f^2}{4h^2 r}, b = \frac{\beta_f^4}{8h^4} - \frac{\beta_f^6 r}{16h^6}$$

$$c = \frac{\beta_f^2}{2h^2}(\sigma + \delta) - \frac{\beta_f^4 r \delta}{4h^4}, d = \frac{\beta_f - \beta_b}{2h}(2 - \beta_b - \beta_f) + h\beta_f \delta$$

$$e = h\beta_f - \frac{\beta_b^2 - \beta_f^2}{2} - h^2\beta_f^2, g = \frac{\beta_f^3}{2h} - \frac{\beta_f^2}{2}, i = \lambda + \frac{\beta_b^2 + (h^2 - 1)\beta_f^2}{2}$$

$$j = L - h\beta_f + \frac{h^2\beta_f^2}{2}, r \in \left[ \frac{2h^2}{\beta_f^2 \delta^2} \left( \delta + \sigma + 1 - \frac{\sqrt{\Delta_1}}{2} \right); \frac{2h^2}{\beta_f^2 \delta^2} \left( \delta + \sigma + 1 + \frac{\sqrt{\Delta_1}}{2} \right) \right]$$

Et que il existe  $(\alpha_1, \alpha_2, \alpha_3, \chi_1, \chi_2, \chi_3, \iota_1, \iota_2, \iota_3, \eta_1, \eta_2, \eta_3) \in \mathbb{R}_+^9$  tel que

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &\leq 1, \quad \chi_1 + \chi_2 + \chi_3 \leq 1 \\ \iota_1 + \iota_2 + \iota_3 &\leq 1, \quad \eta_1 + \eta_2 + \eta_3 \leq 1 \\ 4\alpha_1\chi_1ab - c^2 &> 0, \quad 4\alpha_2\iota_1ai - d^2 > 0 \\ 4\chi_2\iota_2bi - \frac{\beta_f^6}{4h^2} &> 0, \quad 4\alpha_3\eta_1aj - \delta^2(1 - h\beta_f)^2 > 0 \\ 4\iota_3\eta_2ij - e^2 &> 0, \quad 4\chi_3\eta_3bj - g^2 > 0 \end{aligned}$$

Alors  $\forall p \in (\nabla f + B)^{-1}(0)$ , nous avons les résultats suivants :

$$\begin{aligned} x_k &\text{ converge vers } x^*, \quad B(x^*) + \nabla f(x^*) = 0 \\ \sum_{k=0}^{\infty} \|\nabla f(x_k) - \nabla f(p)\|^2 &< \infty, \quad \sum_{k=0}^{\infty} \|B(x_k) - B(p)\|^2 < \infty \\ \lim_{k \rightarrow \infty} \|B(x_k) - B(p)\| &= 0, \quad \lim_{k \rightarrow \infty} \|\nabla f(x_k) - \nabla f(p)\| = 0, \quad \lim_{k \rightarrow \infty} \|x_{k+1} - x_k\| = 0 \end{aligned}$$

Pour prouver ce théorème, je me suis basé sur la démonstration de la convergence de iDINAM-split de van Nam VO dans [VO22]. En considérant la fonction suivante :

$$E_k = \frac{1}{2} \|x_k - p\|^2 + \frac{\beta_f}{h} (x_k - x_{k-1}) + \frac{\beta_f^2}{2h^2} (x_k - 2x_{k-1} + x_{k-2})^2 + \frac{\delta}{2} \|x_k - p\|^2 \quad (3.2)$$

J'ai ensuite cherché à démontrer que cette fonction était bornée. Dans le cas de (iDINAM), van Nam VO obtenait des conditions sur les paramètres qui assuraient qu'une forme quadratique était définie positive ce qui permettait d'obtenir la décroissance de  $E_k$ . Mais dans le cas de (iDINAM2) de nombreux nouveaux termes apparaissent nécessitant l'introduction de six formes quadratiques en vue d'obtenir la décroissance de  $E_k$ . Les conditions pour que toutes les formes quadratiques soient positives sont données par les dix inéquations du théorème.

A ce stade, il reste deux terme. Je n'ai pas réussi à les majorer ce qui m'aurait permis de prouver la décroissance de  $E_k$ . Je me suis donc écarté de la preuve de [VO22] pour chercher à montrer que  $E_k$  était bornée. Pour cela, j'ai prouvé que la série de mes deux termes (positifs) était finie. Le terme  $\|x_k - x_{k-1}\|^2 - \|x_{k+1} - x_k\|^2$  est simple (somme télescopique) car  $\|x_1 - x_0\|^2 < +\infty$ . Pour le second, j'ai du ajouter l'hypothèse  $\sum_{k=3}^{\infty} \langle x_k - p, x_k - 3x_{k-1} + 3x_{k-2} + x_{k-3} \rangle < \infty$ . A partir de ce résultat j'ai pu démontrer le théorème en utilisant le lemme d'Opial.

Je n'ai pas réussi lors de mon stage à finir la démonstration de ce théorème car j'ai du ajouter l'hypothèse  $\sum_{k=3}^{\infty} \langle x_k - p, x_k - 3x_{k-1} + 3x_{k-2} + x_{k-3} \rangle < \infty$ . Comme pour les deux autres théorèmes, il ne peut pas être utilisé pour déterminer à l'avance si des paramètres vont permettre la convergence de l'algorithme. Je pense qu'il est possible de démontrer cette hypothèse, en effet, je n'ai pas trouvé de contre-exemple lors des simulations numériques.

## 3.2 Etude numérique

Dans cette partie, je vais présenter mes travaux sur l'étude numérique de iDINAM2-split. J'ai implémenté l'algorithme et réalisé les analyses en Python3. J'ai divisé celles-ci en trois parties. Dans la première j'ai cherché à caractériser l'effet des paramètres sur la convergence de



l'algorithme. Dans la seconde partie, j'ai comparé iDINAM2-split et iDINAM-split (voir la thèse de van Nam VO [VO22]) qui sont deux algorithmes ayant le même objectif. Et dans la dernière partie j'ai introduit un nouvel algorithme se voulant comme une généralisation de iDINAM2-split et l'ai comparé aux deux algorithmes précédents.

Pour réaliser les tests, j'ai considéré le problème de minimisation suivant :

$$\min_{x \in \mathbb{R}^n} f(x) + G(x) \quad (3.3)$$

Et j'ai pris  $B(x)$  appartenant au sous gradient de  $G(x)$ . Pour résoudre  $(I + \alpha h B)^{-1}(y_k)$  j'ai utilisé comme proposé dans la partie précédente l'opérateur proximal de  $G(x)$ .

### 3.3 Première Analyse

Dans cette partie, j'ai réalisé des graphiques des courbes de  $\|x_k - x^*\|$  en échelle logarithmique. J'ai fait varier un unique paramètre et je les ai tracées pour plusieurs valeurs de celui-ci. Cela permet de voir leurs convergences et quelles valeurs du paramètre permettent la convergence la plus rapide.

J'ai utilisé les trois problèmes suivants pour les analyses numériques (n=5 dans les trois cas).

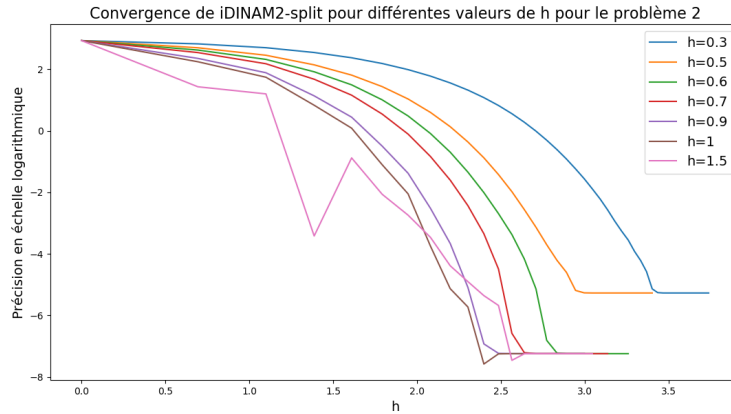
$$B(x) = \partial\|x\|_1, \quad f(x) = \sum_{i=0}^n x_i \ln(x_i) + 1 - x_i, \quad x^{(0)} = x^{(1)} = x^{(2)} = (1, 1, \dots, 1) \quad (\text{Problème 1})$$

$$B(x) = (1 + \ln(x_1), \dots, 1 + \ln(x_n)), \quad f(x) = \|x\|_2^2, \quad x^{(0)} = x^{(1)} = x^{(2)} = (1, 1, \dots, 1) \quad (\text{Problème 2})$$

$$B(x) = \partial\|x\|_1, \quad f(x) = \sum_{k=1}^n kx_k^2 + \left( \sum_{k=1}^n x_k \right)^2, \quad x^{(0)} = x^{(1)} = x^{(2)} = (1, 1, \dots, 1) \quad (\text{Problème 3})$$

Pour le problème 1,  $B(x)$  n'est pas  $\lambda$ -cocoercif et pourtant iDINAM2-split converge bien vers une solution de (2.1).

J'ai obtenu des graphiques comme celui-ci :

FIGURE 3.1 – Trajectoires de iDINAM2-split pour différentes valeurs de  $h$  pour le problème 2

Sur la figure 3.1, il est observable que plus le pas  $h$  est grand plus l'algorithme converge rapidement. De la même manière, j'ai pu déduire des tendances d'évolution du nombre d'itération en fonction des paramètres  $\beta_f$ ,  $\beta_b$  et  $\beta_f\gamma$ . Lors de mes analyses, j'ai eu l'idée d'ajouter un cinquième paramètre  $M$ . Celui-ci transforme  $\nabla f(x)$  et  $B(x)$  par  $M\nabla f(x)$  et  $MB(x)$  dans l'algorithme. Cet ajout implique que  $MB(x)$  est  $\frac{\lambda}{M}$ -cocoercif et que  $M\nabla f(x)$  est  $LM$ -lipschitzienne et que de plus  $\nabla f(x) + B(x) = 0 \Leftrightarrow M\nabla f(x) + MB(x) = 0$  ce qui signifie que les solutions obtenues par cette nouvelle formulation seront identiques. Bien choisi, ce nouveau paramètre permet d'accélérer encore la convergence de l'algorithme.

A partir de ces exemples, suite à l'observation de comportements communs des paramètres sur la vitesse de convergence et la stabilité des trajectoires, j'ai proposé un algorithme permettant de les choisir. Celui-ci commence avec de petites valeurs de  $h, \beta_b, \beta_f$   $M = 1$  et une grande valeur de  $\beta_f\gamma$  (2 par exemple). Puis augmente progressivement  $h, \beta_b, \beta_f$  afin d'accroître la vitesse de convergence. La seconde étape, consiste à réduire  $\gamma\beta_f$  et  $M$  afin de trouver un bon compromis entre la stabilité et la vitesse.

La détermination des paramètres est importante car un mauvais choix est souvent quatre ou cinq fois plus lent. Lors de cette première analyse, j'ai remarqué que les hypothèses du théorème 4 pouvaient ne pas être valides (en général  $\gamma\beta_f < 1$ ) alors que l'algorithme convergeait.

### 3.4 Comparaison de iDINAM-split et de iDINAM2-split

Dans cette partie, j'ai comparé les deux algorithmes avec les mêmes paramètres et ensuite avec des paramètres "optimaux" pour chacun.

L'objectif de la comparaison avec des mêmes paramètres était de déterminer en fonction de ceux-ci l'algorithme le plus rapide et d'observer des points communs et des différences sur l'action de ces paramètres sur ces schémas. Pour cela, j'ai fait un graphique où je faisais varier un unique paramètre. Et pour chaque valeur de celui-ci, je calculais le nombre d'itération  $N$  tel que  $\forall n \geq N, \|x_n - x^*\| < s$  ou  $s$  est un seuil défini à l'avance ( $s = 10^{-10}$  pour le problème 1 et 3 et  $s = 10^{-4}$  pour le problème 2). J'ai obtenu un graphique comme ceci :

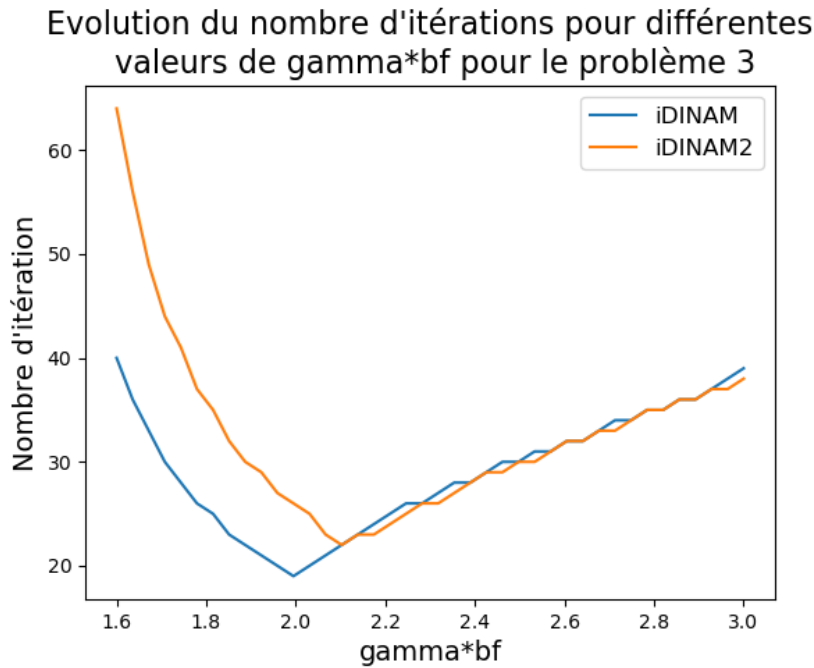


FIGURE 3.2 – Trajectoires de iDINAM2-split et iDINAM2-split pour différents  $\gamma\beta_f$  pour le problème 3

Sur la figure 3.2, il semblerait que si  $\gamma\beta_f$  est petit il soit préférable d'utiliser iDINAM-split. En ayant fait de mêmes avec les quatre autres paramètres, j'ai constaté que iDINAM-split avait une plus grande amplitude de choix. C'est à dire que si iDINAM2-split ne converge pas il est possible que iDINAM-split converge. Cette propriété est très importante, en effet si le résultat a besoin d'être calculé qu'une seule fois il est préférable de prendre iDINAM-split qui convergera plus probablement.

Pour la comparaison avec des paramètres optimaux j'ai réalisé des profils de performances voir [Dol01]. J'ai implémenté ceux-ci en python. Pour cela j'ai du coder deux scripts. Le premier effectue des mesures du temps d'exécution des algorithmes et de leur nombre d'itérations pour plusieurs instances de six problèmes pouvant être paramétrés (dont les trois problèmes de la première analyse). J'ai stocké les résultats dans un fichier pour l'analyse. Le deuxième script faisait les profils de performances à partir des données.

Voici l'analyse que j'ai pu en tirer.

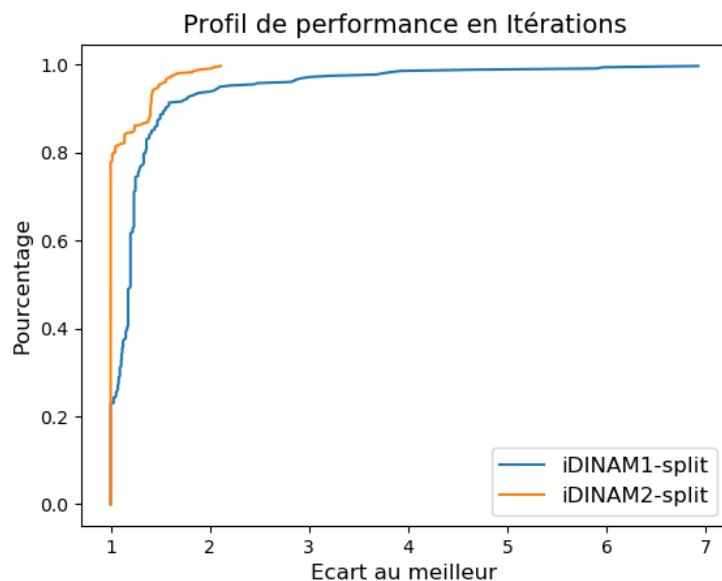


FIGURE 3.3 – Profil de performance du nombre d’itération de iDINAM-split et de iDINAM2-split

Nous pouvons observer sur 3.3 que pour 75% des problèmes, iDINAM2 fait moins d’itérations que iDINAM. Pour les 25% restant, iDINAM2 fait au maximum 2.1 fois plus d’itérations que iDINAM. Alors que iDINAM peut faire jusqu’à 7 fois plus d’itérations que iDINAM2.

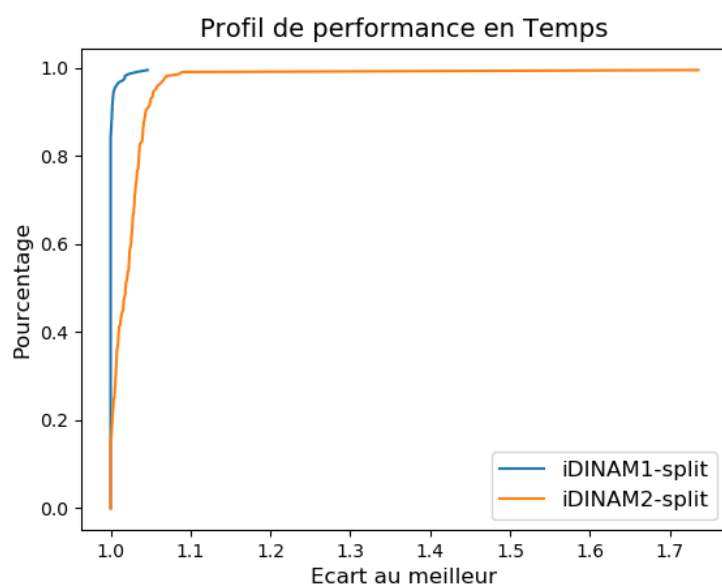


FIGURE 3.4 – Profil de performance du temps d’exécution de iDINAM-split et de iDINAM2-split

Pour générer cette seconde figure 3.4, j'ai pris les tests utilisant seulement les temps d'exécution supérieurs à 0.15 secondes afin d'éliminer les erreurs dues à la mesure du temps. J'ai observé sur 5.13 que iDINAM est toujours plus rapide. Mais cette différence est légère (inférieure à 7% dans 99% des cas) et au maximum 70% de temps d'exécution en plus. Cela est en désaccord avec le résultat précédent qui indiquait que iDINAM2 faisait moins d'itérations que iDINAM et donc devrait être plus rapide. Une hypothèse est que dans les problèmes considérés la résolution de  $(I + \alpha h B)^{-1}$  est simple (pas d'algorithmes itératifs) et par ailleurs iDINAM2-split fait plus d'opérations par itérations. Ces opérations supplémentaires seraient à l'origine de ces temps d'exécution plus longs qui devraient s'inverser si la résolution de  $(I + \alpha h B)^{-1}$  était plus compliquée. Je n'ai pas eu le temps d'approfondir cette observation.

Nous observons donc avec ce profil de performance que iDINAM2 semble être un meilleur algorithme car il est en général plus robuste que iDINAM. Ce résultat m'a étonné car il est en contradiction avec celui de la première analyse où j'ai observé que iDINAM-split est plus robuste que iDINAM2-split pour des mêmes paramètres.

La conclusion de cette comparaison des deux algorithmes est que si l'on ne souhaite pas optimiser les paramètres il vaut mieux utiliser iDINAM-split. Sinon iDINAM2-split converge plus vite en général.

### 3.5 iDINAMN-split

Après avoir étudié iDINAM2-split et l'avoir comparé à iDINAM-split, j'ai cherché à déterminer si il était intéressant de continuer à augmenter le degré du développement limité de  $x(t + \beta_b)$  et  $x(t + \beta_f)$  à l'ordre  $N$ . Pour cela, j'ai créé un nouvel algorithme appelé iDINAMN-split :

---

#### Algorithm 3 iDINAMN split

---

**Require:**  $x_0, \dots, x_N \in \mathbb{R}^n, \epsilon > 0$

$$u_i \leftarrow (-1)^i \frac{\beta_b^i}{i!h^i} \sum_{m=0}^{N-i} \frac{\beta_b^m}{m!h^m}$$

$$v_i \leftarrow (-1)^i \frac{\beta_f^i}{i!h^i} \sum_{m=0}^{N-i} \frac{\beta_f^m}{m!h^m}$$

$$\alpha \leftarrow (1 + \gamma h)^{-1} u_0$$

**while**  $\|x_k - x_{k-1}\| > \epsilon$  **do**

$$z_k \leftarrow \sum_{i=0}^N v_i x_{k-i}$$

$$y_k \leftarrow \alpha(2 + \gamma h)x_k - \alpha x_{k-1} + \sum_{i=1}^N u_i x_{k+1-i} - \alpha h^2 \nabla f(z_k)$$

$$x_{k+1} \leftarrow \frac{1}{u_0} (Id + \alpha h B)^{-1} (y_k) - \sum_{i=1}^N \frac{u_i}{u_0} x_{k+1-i}$$

**end while**

---

En prenant  $N=1$ , je retrouve l'algorithme iDINAM-split et pour  $N=2$  iDINAM2-split.

Afin de comparer les différents algorithmes obtenus, j'ai réalisé un profil de performance pour  $N \in \{1, 2, 3, 4, 6, 10, 16\}$ .

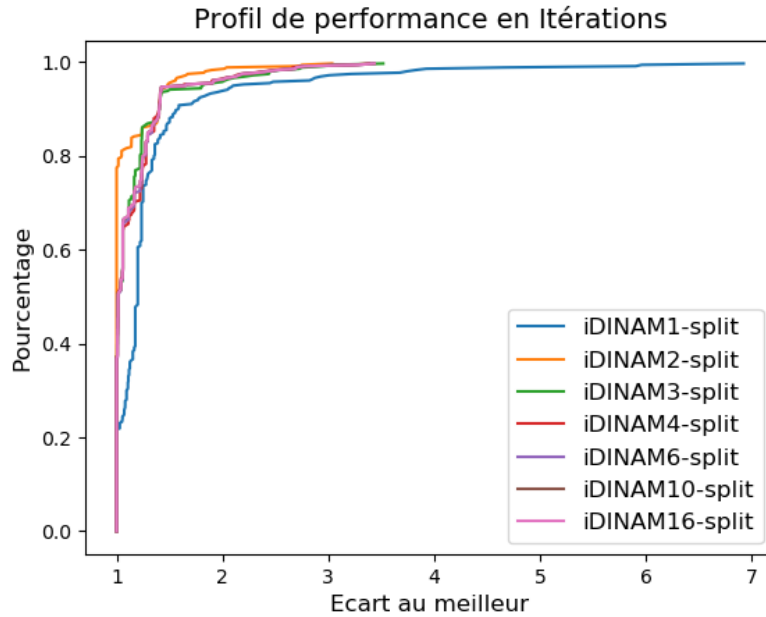


FIGURE 3.5 – Profil de performance du nombre d'itération de iDINAMN-split

J'ai observé que pour  $N > 2$  les algorithmes étaient toujours moins performants que iDINAM2-split. J'en ai donc conclu qu'ils n'étaient pas intéressants et qu'il valait toujours mieux utiliser soit iDINAM-split soit iDINAM2-split.

J'ai donc pu étudier iDINAM2-split et caractériser l'influence des paramètres pour celui-ci. Ce travail est important en pratique car il permet de ne pas prendre au hasard les paramètres. Il a permis d'observer que les hypothèses des théorèmes n'étaient pas forcément nécessaires (comme  $B$   $\lambda$ -cocoercif par exemple) pour que l'algorithme converge. J'ai pu également comparer plusieurs versions de iDINAMN-split afin de trouver la plus adaptée.

## Chapitre 4

# Conclusion

Ce stage, a été l'occasion de faire une première découverte avec le monde de la recherche. J'ai apprécié de pouvoir travailler en autonomie. J'ai trouvé le sujet très passionnant et je suis satisfait des résultats obtenus lors de ce stage même si les théorèmes ne sont pas avec des hypothèses minimales. Cela m'a permis de développer mes compétences rédactionnelles ainsi que ma maîtrise du LATEX. Ce stage a été l'occasion de découvrir le domaine de l'optimisation non différentiable et de nouveaux outils tels que les profils de performance.

A partir des travaux présentés dans la thèse de van Nam VO [VO22], j'ai créé iDINAM2-split. J'ai pu réaliser la démonstration de trois théorèmes concernant iDINAM2-split (existence et convergence de l'équation différentielle ordinaire et convergence de l'algorithme). Mais ces trois preuves sont incomplètes car dans leur forme actuelle elles font intervenir la solution en hypothèse ce qui est une contrainte forte pour leur utilisation. Un objectif de recherche pourrait être de prouver les théorèmes sans ces hypothèses.

J'ai également réalisé une étude numérique des influences des paramètres sur la convergence de iDINAM2-split et l'ai comparé à iDINAM-split dont il dérive. Dans le cas où l'on souhaite choisir les différents paramètres de l'algorithme afin de l'accélérer pour un problème, il est conseillé d'utiliser iDINAM2-split et dans le cas inverse iDINAM-split.

# Bibliographie

- [Dol01] Moré J. DOLAN E. « Benchmarking Optimization Software with Performance Profiles ». In : *Mathematical Programming* (2001).
- [Bec09] Teboulle M. BECK A. « Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems ». In : *IEEE Transactions on Image Processing* (2009).
- [Att10] Maingé P.E. ATTOUCH H. « Asymptotic behavior of second order dissipative evolution equations combining potential with nonpotential effects ». In : *ESAIM Control Optim* (2010).
- [Pey10] Sorin S. PEYPOUQUET J. « Evolution Equations for Maximal Monotone Operators : Asymptotic Analysis in Continuous and Discrete Time ». In : *Journal of Convex Analysis* (2010).
- [VO22] Van Nam VO. « Analyse des dynamiques inertielles et algorithmes associés pour l'optimisation du premier ordre ». 2022.
- [M23] Haddou M. *Notes des cours d'optimisation convexe non différentiable et applications*. 2023.



# Chapitre 5

## Annexes

### 5.1 Preuve du théorème 1

**Preuve :**

Il existe  $\Lambda > 0$  tel que si  $\lambda_0 \in [0; \Lambda]$ ,  $\forall k > 0, \lambda_k \in [0; \Lambda]$

En effet, si  $x_m = \arg \min_{x \in \mathbb{R}^n} \phi(x)$  on a  $\phi(x_m) < 0$  (car le problème admet au moins une solution)

et l'on peut prendre

$$M < \frac{\arg \min_{\phi(x) > 0} R(x) - R(x_m)}{\phi(x_m)} < \infty \text{ car } R \text{ est propre.}$$

Si  $\lambda_k \geq M$  on a  $L(x_m, \lambda_k) < \arg \min_{\phi(x) > 0} L(x, \lambda_k)$  donc  $x_{k+1} \in C = \{x \in \mathbb{R}^n; \phi(x) \leq 0\}$

Et  $\lambda_{k+1} \leq \lambda_k$ . On peut donc prendre

$$\Lambda = \max_{\lambda \in [0; M]} \lambda + \rho \phi(\arg \min_{x \in \mathbb{R}^n} L(x, \lambda)) < \infty$$

Supposons que  $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$  alors comme  $\phi$  est coercive  $\lim_{k \rightarrow \infty} \phi(x_k) = +\infty$  et donc  $\lim_{k \rightarrow \infty} \lambda_k = +\infty$  ce qui est impossible selon le résultat précédent. Donc la suite  $\{x_k\}$  est bornée. Si  $\phi(x_1) \leq 0$  l'algorithme a convergé car  $\forall k, x_k = x_1$  et  $v_k = 0$  et de plus comme  $x_1$  est une solution de 1 elle est un point selle de  $L(x, \lambda)$

Dans le cas  $\phi(x_1) > 0$  nous allons démontrer qu'il existe  $\rho$  tel que :

$$\forall k \geq 0, \phi(x_k) > \phi(x_{k+1}) \geq 0 \quad (5.1)$$

Supposons maintenant que  $\phi(x_k) > 0$ . On note :

$$\delta = L(x_k, \lambda_k) - \min_{\phi(x) < 0} L(x, \lambda_k), \quad \Delta = \phi(x_k) - \min_{x \in \mathbb{R}^n} \phi(x)$$

On a les propriétés suivantes  $\delta \in ]-\infty; 0[, \Delta \in ]0; +\infty[$  car  $R$  est propre et  $\{x_k\}$  et  $\{\lambda_k\}$  sont bornées.

En prenant  $\rho_k \in ]0; \frac{-\delta}{\Delta \phi(x_k)}[$  on obtient,

$$\begin{aligned} \forall x \in C, L(x_k, \lambda_k + \rho_k \phi(x_k)) - L(x, \lambda_k + \rho_k \phi(x_k)) = \\ L(x_k, \lambda_k) - L(x, \lambda_k) + \rho_k \phi(x_k)(\phi(x_k) - \phi(x)) \leq \delta + \rho_k \phi(x_k) \Delta < 0 \\ \Leftrightarrow L(x_k, \lambda_k + \rho_k \phi(x_k)) < L(x, \lambda_k + \rho_k \phi(x_k)) \end{aligned}$$

Soit  $x_{k+1} \notin C$  et  $\phi(x_{k+1}) > 0$  en prenant  $\rho = \min_{k \in \mathbb{N}} \rho_k$  on prouve par récurrence la première inégalité de 5.1 car  $\phi(x_1) > 0$ .

Comme  $\phi(x_k) > 0$  on a  $\lambda_k > \lambda_{k-1}$ .

$$\begin{aligned} L(x_k, \lambda_{k-1}) \leq L(x_{k+1}, \lambda_{k-1}) \Leftrightarrow R(x_k) \leq R(x_{k+1}) + \lambda_{k-1}(\phi(x_{k+1}) - \phi(x_k)) \\ L(x_{k+1}, \lambda_k) \leq L(x_k, \lambda_k) \Leftrightarrow R(x_{k+1}) + v_k(\phi(x_{k+1}) - \phi(x_k)) \leq R(x_k) \end{aligned}$$

En combinant les deux équations on obtient :

$$(v_k - v_{k-1})(\phi(x_{k+1}) - \phi(x_k)) \leq 0$$

Soit  $\phi(x_{k+1}) \leq \phi(x_k)$  par récurrence, la deuxième inégalité de 5.1 est prouvée et donc 5.1 est prouvée.

Donc comme la suite  $\{\phi(x_k)\}$  est minorée et décroissante elle converge. On note  $\phi^*$  sa limite. De plus comme  $\lambda_k < \Lambda$  nous obtenons que  $\phi^* \leq 0$  car sinon comme  $\lambda_{k+1} = \max(0, \lambda_k + \rho \phi(x_k))$ ,  $\lambda_k$  divergerait ce qui est impossible. On en déduit donc que  $\phi^* = 0$ . Ce qui entraîne que la suite  $\{\lambda_k\}$  converge et puis que la suite  $\{x_k\}$  converge. De plus  $\lambda^* = \rho \sum_{k \in \mathbb{N}} \phi(x_k) > 0$ .

Montrons maintenant que  $(x^*, \lambda^*)$  est un point selle de  $L(x, \lambda)$ . Par la définition de  $x_k$  dans l'algorithme 1 nous savons que :

$$\forall x \in \mathbb{R}^n, L(x^*, \lambda^*) \leq L(x, \lambda^*)$$

Si  $\lambda^* = 0$  alors  $\phi(x^*) \leq 0$  car  $0 = \max(0, \rho \phi(x^*)) \Leftrightarrow \phi(x^*) \leq 0$ . On obtient donc que :

$$\forall \lambda > 0, L(x^*, \lambda) = R(x^*) + \lambda \phi(x^*) \leq R(x^*) = L(x^*, \lambda^*)$$

Si  $\lambda^* > 0$  nous savons que  $\phi(x^*) = 0$  on obtient directement que  $\forall \lambda > 0, L(x^*, \lambda) \leq L(x^*, \lambda^*)$ .

On obtient donc en combinant les trois résultats précédents :

$$\forall x \in \mathbb{R}^n, \forall \lambda \geq 0, L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*)$$

Soit  $(x^*, \lambda^*)$  est un point selle de  $L(x, \lambda)$  et donc par le théorème de la dualité forte  $x^*$  est une solution optimale de 1.1.

## 5.2 Preuve du théorème 2

Soit  $p, p' \in (\nabla f + B)^{-1}(0)$  par la cocoercivité de B et la monotonie de  $\nabla f$  :

$$\langle B(p) - B(p'), p - p' \rangle \geq \lambda \|B(p) - B(p')\|^2 \quad (5.2)$$

$$\langle \nabla f(p) - \nabla f(p'), p - p' \rangle \geq 0 \quad (5.3)$$

En additionnant les deux :

$$0 = \langle 0, p - p' \rangle \geq \lambda \|B(p) - B(p')\| \geq 0$$

On en déduit que  $B(p) = B(p')$  et donc que  $\nabla f(p) = \nabla f(p')$ .

Afin de démontrer l'existence et l'unicité du problème associé à (iDINAM2) nous allons étudier cette équation en y ajoutant un terme de viscosité. C'est à dire en rajoutant  $\epsilon \ddot{x}(t)$

$$\epsilon > 0, \epsilon \ddot{x}(t) + \ddot{x}(t) + \gamma \dot{x}(t) + B\left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) + \nabla f\left(x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\right) = 0 \quad (\text{iDINAM2e})$$

On peut écrire (iDINAM2e) sous forme résolue :

$$\begin{pmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} v \\ w \\ U(u, v, w) \end{pmatrix} \quad (5.4)$$

avec  $U(u, v, w) = -w - \gamma v - B\left(u + \beta_b v + \frac{\beta_b^2}{2} w\right) - \nabla f\left(u + \beta_f v + \frac{\beta_f^2}{2} w\right)$ . Comme  $\nabla f$  et  $B$  sont lipschitziens en appliquant le théorème de Cauchy-Lipschitz local nous obtenons l'existence d'une unique solution maximale au problème de Cauchy obtenu à partir de (iDINAM2e).

Nous allons maintenant prendre  $p$  tel que  $0 \in B(p) + \nabla f(p)$  et considérer la fonction suivante :

$$\begin{aligned} \varepsilon_p(t) := a \left( f\left(x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\right) - f(p) - \langle \nabla f(p), x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) - p \rangle \right) + \\ \frac{1}{2} \|x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) - p\|^2 + \frac{d}{2} \|x - p\|^2 \end{aligned}$$

La fonction  $\varepsilon_p(t)$  est positive car  $f$  est convexe. En la dérivant, nous obtenons :

$$\begin{aligned} \dot{\varepsilon}_p = a \left\langle \dot{x} + \beta_f \ddot{x} + \frac{\beta_f^2}{2} \ddot{x}, \nabla f\left(x + \beta_f \dot{x} + \frac{\beta_f^2}{2} \ddot{x}(t)\right) - \nabla f(p) \right\rangle + \\ \left\langle \dot{x} + \beta_f \ddot{x} + \frac{\beta_f^2}{2} \ddot{x}, x + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x} - p \right\rangle + d \langle x - p, \dot{x} \rangle \end{aligned}$$

En notant  $X = \nabla f\left(x + \beta_f \dot{x} + \frac{\beta_f^2}{2} \ddot{x}(t)\right) - \nabla f(p)$  et  $Y = B\left(x + \beta_b \dot{x} + \frac{\beta_b^2}{2} \ddot{x}(t)\right) - B(p)$  et en utilisant (iDINAM2) nous obtenons :

$$\begin{aligned} X + Y = \nabla f\left(x + \beta_f \dot{x} + \frac{\beta_f^2}{2} \ddot{x}(t)\right) + B\left(x + \beta_b \dot{x} + \frac{\beta_b^2}{2} \ddot{x}(t)\right) \\ \ddot{x} = -\gamma \dot{x} - X - Y - \epsilon \ddot{x} \end{aligned} \quad (5.5)$$

Puis en réinjectant :

$$\begin{aligned}
 \dot{\epsilon}_p &= a \left\langle \dot{x} + \beta_f(-\gamma\dot{x} - X - Y - \epsilon\ddot{x}) + \frac{\beta_f^2}{2}\ddot{x}, X \right\rangle + \\
 &\quad \left\langle \dot{x} + \beta_f(-\gamma\dot{x} - X - Y - \epsilon\ddot{x}) + \frac{\beta_f^2}{2}\ddot{x}, x + \beta_f\dot{x} + \frac{\beta_f^2}{2}\ddot{x} - p \right\rangle + d\langle x - p, \dot{x} \rangle \\
 &= -a\beta_f\|X\|^2 + \left(a - \frac{\beta_f^2}{2}\right) (1 - \gamma\beta_f)\langle\dot{x}, X\rangle - \frac{\beta_f^2}{2}(1 - \gamma\beta_f)\langle\dot{x}, Y\rangle - a\beta_f\langle X, Y\rangle - \beta_f \\
 &\quad \left\langle x + \beta_f\dot{x} + \frac{\beta_f^2}{2}\ddot{x} - p, X \right\rangle - \beta_f \left\langle x + \beta_f\dot{x} + \frac{\beta_f^2}{2}\ddot{x} - p, Y \right\rangle + (1 - \gamma\beta_f) \left(\beta_f - \gamma\frac{\beta_f^2}{2}\right) \|\dot{x}\|^2 \\
 &\quad + (d + 1 - \gamma\beta_f)\langle x - p, \dot{x} \rangle + \left(\frac{\beta_f^2}{2} - \beta_f\epsilon\right) \left\langle \ddot{x}, aX + x + \dot{x} \left(\beta_f - \gamma\frac{\beta_f^2}{2}\right) - \frac{\beta_f^2}{2}(X + Y) - p \right\rangle \\
 &\quad - \epsilon\frac{\beta_f^2}{2} \langle \ddot{x}, \dot{x}(1 - \gamma\beta_f) - \beta_f(X + Y) \rangle - \left(\frac{\beta_f^2}{2} - \epsilon\beta_f\right) \epsilon\frac{\beta_f^2}{2} \|\ddot{x}\|^2
 \end{aligned}$$

En prenant  $d = \gamma\beta_f - 1 > 0$ , en posant  $g = \left(\frac{\beta_f^2}{2} - \beta_f\epsilon\right) \left\langle \ddot{x}, aX + x + \dot{x} \left(\beta_f - \gamma\frac{\beta_f^2}{2}\right) - \frac{\beta_f^2}{2}(X + Y) - p \right\rangle - \epsilon\frac{\beta_f^2}{2} \langle \ddot{x}, \dot{x}(1 - \gamma\beta_f) - \beta_f(X + Y) \rangle - \left(\frac{\beta_f^2}{2} - \epsilon\beta_f\right) \epsilon\frac{\beta_f^2}{2} \|\ddot{x}\|^2 + \epsilon\beta_f\frac{\beta_f^2 - \beta_b^2}{2} \langle \ddot{x}, Y \rangle$  et en utilisant les inégalités suivantes dues à la cocoercivité de  $B$  et à la convexité de  $f$  :

$$\begin{aligned}
 &\left\langle X, x + \beta_f\dot{x} + \frac{\beta_f^2}{2}\ddot{x} - p \right\rangle \geq 0 \\
 &\left\langle Y, x + \beta_f\dot{x} + \frac{\beta_f^2}{2}\ddot{x} - p \right\rangle \geq \lambda\|Y\|^2 - \left(\frac{\beta_b^2}{2} - \frac{\beta_f^2}{2}\right) \langle Y, -\gamma\dot{x} - X - Y - \epsilon\ddot{x} \rangle - (\beta_b - \beta_f)\langle Y, \dot{x} \rangle
 \end{aligned}$$

On peut donc les insérer afin d'obtenir l'équation suivante :

$$\dot{\epsilon}_p \leq -a\beta_f\|X\|^2 + h\langle\dot{x}, X\rangle + e\|\dot{x}\|^2 - d\langle\dot{x}, Y\rangle - c\langle Y, X\rangle - b\|Y\|^2 + g \quad (5.6)$$

Avec  $b = \lambda - \beta_f\frac{\beta_f^2 - \beta_b^2}{2}$ ,  $c = \beta_f\left(a - \frac{\beta_f^2 - \beta_b^2}{2}\right)$ ,  $d = \beta_f^2\left(\frac{3}{2} - \gamma\beta_f\right) + \beta_f\beta_b\frac{\gamma - \beta_b}{2}$ ,  $e = (1 - \gamma\beta_f)\left(\beta_f - \gamma\frac{\beta_f^2}{2}\right)$  et  $h = \left(a - \frac{\beta_f^2}{2}\right)(1 - \gamma\beta_f)$ .

En prenant  $a > \frac{\beta_f^2}{2}$  nous en déduisons que  $c > 0$ ,  $h < 0$ . Nous avons les inégalités suivantes :

$$\begin{aligned}
 \rho &> 0, \quad h\langle X, \dot{x} \rangle \leq -\frac{\rho h}{2}\|X\|^2 - \frac{h}{2\rho}\|\dot{x}\|^2 \\
 r &> 0, \quad -c\langle X, Y \rangle \leq \frac{rc}{2}\|X\|^2 + \frac{c}{2r}\|Y\|^2
 \end{aligned}$$

En combinant (5.6) avec les équations précédentes nous obtenons :

$$\dot{\epsilon}_p \leq \left(\frac{-\rho h + rc}{2} - a\beta_f\right) \|X\|^2 + \left(e - \frac{h}{2\rho}\right) \|\dot{x}\|^2 + \left(\frac{c}{2r} - b\right) \|Y\|^2 - d\langle\dot{x}, Y\rangle + g \quad (5.7)$$

$$\dot{\epsilon}_p + \left(a\beta_f - \frac{-\rho h + rc}{2}\right) \|X\|^2 + q\langle\dot{x}, Y\rangle - g \leq 0 \quad (5.8)$$

Ou  $q(\dot{x}, Y) = \left(\frac{h}{2\rho} - e\right) \|\dot{x}\|^2 + \left(b - \frac{c}{2r}\right) \|Y\|^2 + d\langle \dot{x}, Y \rangle$ . Nous voulons que  $q$  soit définie positive et que le coefficient devant  $\|X\|^2$  soit positif pour cela il faut que :

$$\left(\frac{h}{2\rho} - e\right) > 0 \quad (5.9)$$

$$\left(b - \frac{c}{2r}\right) > 0 \quad (5.10)$$

$$\left(\frac{h}{2\rho} - e\right) \left(b - \frac{c}{2r}\right) - \frac{d^2}{4} > 0 \quad (5.11)$$

$$a\beta_f - \frac{-\rho h + rc}{2} > 0 \quad (5.12)$$

Nous allons prendre  $\beta_f$  tel que  $e < 0$  comme  $\gamma\beta_f > 1$  il faut que :

$$\beta_f > \gamma \frac{\beta_f^2}{2} \Rightarrow \frac{1}{\gamma} < \beta_f < \frac{2}{\gamma}$$

Nous allons prendre  $\lambda$  tel que  $b > 0$  il faut que :

$$\lambda > \beta_f \frac{\beta_f^2 - \beta_b^2}{2}$$

En prenant,  $r = \frac{\tau c}{2b}$  et  $\rho = \frac{\tau h}{2e}$  avec  $\tau > 1$  les deux premières conditions sont satisfaites (car  $h < 0$ ).

On peut réécrire la troisième comme ceci :

$$-eb \left(1 - \frac{1}{\tau}\right)^2 - \frac{d^2}{4} > 0 \Rightarrow \frac{1}{\tau} < 1 - \sqrt{-\frac{d^2}{4eb}}$$

Et la dernière est vérifiée si :

$$\begin{aligned} \left(1 - \sqrt{-\frac{d^2}{4eb}}\right)^{-1} < \tau \quad 2a\beta_f - \left(\frac{-h^2}{2e} + \frac{c^2}{2b}\right) \tau &= 2a\beta_f - \left(\frac{-h^2}{2e} + \frac{c^2}{2b}\right) \left(1 - \sqrt{-\frac{d^2}{4eb}}\right)^{-1} \\ \Rightarrow a &> \frac{1}{2\beta_f} \left(\frac{-h^2}{2e} + \frac{c^2}{2b}\right) \left(1 - \sqrt{-\frac{d^2}{4eb}}\right)^{-1} > 0 \end{aligned}$$

Car  $\left(\frac{-h^2}{2e} + \frac{c^2}{2b}\right) > 0$ . Il existe donc  $\mu > 0$  tel que  $q(\dot{x}, Y) \geq \mu(\|Y\|^2 + \|\dot{x}\|^2)$ .

$$\dot{\varepsilon}_p + \left(a\beta_f - \frac{\rho h + rc}{2}\right) \|X\|^2 + \mu(\|Y\|^2 + \|\dot{x}\|^2) - g \leq 0 \quad (5.13)$$

Prenons  $\beta_f$  tel que  $\forall t > 0, \int_0^t g(s)ds < \int_0^t \left(a\beta_f - \frac{\rho h + rc}{2}\right) \|X(s)\|^2 + \mu(\|Y(s)\|^2 + \|\dot{x}(s)\|^2)ds$ . Il existe un tel  $\beta_f$  car  $\forall t > 0, \lim_{\beta_f \rightarrow 0^+} g(t) = 0$  et  $\forall t > 0, \lim_{\beta_f \rightarrow 0^+} q(\dot{x}(t), Y(t)) > 0$ .

En intégrant (5.13) on obtient que  $0 < \varepsilon_p(t) \leq \varepsilon_p(0)$  soit  $\varepsilon_p$  est bornée et en revenant à la définition de  $\varepsilon_p$  on obtient que :

$$\sup_{t>0} \|x(t) - p\| < \infty \quad (5.14)$$

$$\sup_{t>0} \|x(t) - p + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\| < \infty \quad (5.15)$$

En utilisant l'inégalité triangulaire on obtient que :

$$\sup_{t>0} \|\dot{x}(t) + \frac{\beta_f}{2} \ddot{x}(t)\| < \infty \quad (5.16)$$

En utilisant (5.16) nous en déduisons l'on peut écrire  $\dot{x}(t) = -\frac{\beta_f}{2} \ddot{x}(t) + \xi(t)$  ou  $\sup_{t>0} \|\xi(t)\| < \infty$  et en utilisant (iDINAM2)

$$\ddot{x} \left(1 - \gamma \frac{\beta_f}{2}\right) + \xi(t)\gamma + B \left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) + \nabla f \left(x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\right) = 0 \quad (5.17)$$

Ce qui implique que :

$$\left\| \ddot{x} \left(1 - \gamma \frac{\beta_f}{2}\right) + B \left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) \right\| = \left\| \xi(t)\gamma + \nabla f \left(x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\right) \right\| \quad (5.18)$$

Par (5.15) et comme  $\nabla f$  est continue on sait que  $\exists M > 0, \forall t > 0, M > \|\xi(t)\gamma + \nabla f \left(x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t)\right)\|$ . Or  $1 - \gamma \frac{\beta_f}{2} \neq 0$  donc  $\left\| \ddot{x}(t) \left(1 - \gamma \frac{\beta_f}{2}\right) + B \left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) \right\| < M$ .

Or, en intégrant (5.13) nous avons aussi obtenu que  $\forall t > 0, \int_0^t \left\| B \left(x(s) + \beta_b \dot{x}(s) + \frac{\beta_b^2}{2} \ddot{x}(s)\right) - B(p) \right\|^2 ds < \varepsilon_p(0)$ . On en déduit donc que  $\lim_{t \rightarrow \infty} \left\| B \left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) - B(p) \right\|^2 = 0$  et donc  $\forall \eta > 0, \exists T, \forall t > T, \left\| B \left(x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t)\right) \right\| < \eta$ . Et donc  $\|\ddot{x}(t)\| < (M + \eta) \left(1 - \gamma \frac{\beta_f}{2}\right)^{-1} < \infty$ .

Nous venons de montrer que si il existe  $t^*$  tel que  $\lim_{t \rightarrow t^*} \|\ddot{x}(t)\| = \infty$ , alors  $t^* < \infty$  et que  $\lim_{t \rightarrow t^*} \|\dot{x}(t)\| = \infty$  (selon (5.16)). Nous en déduisons que :

$$\lim_{t \rightarrow t^*} \left\| \int_0^t \dot{x}(s) + \frac{\beta_f}{2} \ddot{x}(s) ds \right\| \leq \lim_{t \rightarrow t^*} \int_0^t \left\| \dot{x}(s) + \frac{\beta_f}{2} \ddot{x}(s) \right\| ds \leq \lim_{t \rightarrow t^*} t \varepsilon_p(0) \leq t^* \varepsilon_p(0) < \infty \quad (5.19)$$

Or nous avons l'égalité suivante :

$$\int_0^t \dot{x}(s) + \frac{\beta_f}{2} \ddot{x}(s) ds = x(t) - x(0) + \frac{\beta_f}{2} (\dot{x}(t) - \dot{x}(0))$$

Donc comme  $\|x(t) - x(0) - \frac{\beta_f}{2} \dot{x}(0)\| < \infty$  selon (5.14). On en déduit que :

$$\lim_{t \rightarrow t^*} \left\| \int_0^t \dot{x}(s) + \frac{\beta_f}{2} \ddot{x}(s) ds \right\| = \lim_{t \rightarrow t^*} \|x(t) - x(0) + \frac{\beta_f}{2} (\dot{x}(t) - \dot{x}(0))\| = \infty$$

Ce qui est impossible selon (5.19), donc on sait que  $\sup_{t>0} \|\ddot{x}(t)\| < \infty$  et par l'inégalité triangulaire on obtient que  $\sup_{t>0} \|\dot{x}(t)\| < \infty$ .

Les résultats précédents sont obtenus pour tout  $\epsilon > 0$ . Nous allons maintenant considérer la suite  $\{x_\epsilon(t)\}_{\epsilon>0}$  obtenue à partir de (iDINAM2e) en faisant tendre  $\epsilon$  vers 0.

L'objectif est de démontrer que  $\{x_\epsilon(t)\}_{\epsilon>0}$  est relativement compacte (c'est à dire qu'elle est incluse dans un compact) en utilisant le théorème d'Ascoli.

Pour cela, il faut que  $\forall t \geq 0, A(t) = \{x_\epsilon(t), \epsilon > 0\}$  soit relativement compacte. Et que  $\forall t, \eta, \epsilon > 0$ , il existe un voisinage  $V$  de  $t$  tel que  $\forall u \in V, \|x_\epsilon(t) - x_\epsilon(u)\| < \eta$ .

On sait que  $\|x_\epsilon(t) - p\| < \varepsilon_p(0)$  qui ne dépend pas de  $\epsilon$ . Donc  $\forall t \geq 0, A(t) = \{x_\epsilon(t), \epsilon > 0\} \subset \{x \in \mathbb{R}^n, \|x - p\| \leq \varepsilon_p(0)\}$ . Ce dernier ensemble est compact (par le théorème de Bolzano-Weierstrass) et donc  $A(t)$  est relativement compacte.

De plus comme  $\exists M, \|\dot{x}_\epsilon(t)\| \leq M$ , on sait que en prenant  $V \subset ]t - \frac{\eta}{M}, t + \frac{\eta}{M}[$ , nous savons que  $\forall u \in V, \|x_\epsilon(t) - x_\epsilon(u)\| < \eta$ .

Donc, nous pouvons appliquer le théorème d'Ascoli et la suite  $\{x_\epsilon(t)\}_{\epsilon>0}$  est relativement compacte. On peut prendre une sous suite  $\phi(n)$ , tel que  $\lim_{n \rightarrow \infty} \phi(n) = 0$  tel que  $\lim_{n \rightarrow \infty} x_{\phi(n)}(t) = x^*(t)$ . De plus  $x^*(t)$  satisfait (iDINAM2). On en déduit donc que (iDINAM2) admet une solution globale (pas forcément unique).

### 5.3 Preuve Théorème 3

Preuve :

Reprenons la preuve du théorème 2, en prenant  $\epsilon = 0$  dans (iDINAM2e). Nous obtenons donc que  $\varepsilon_p(t)$  est bornée et que :

$$\sup_{t>0} \|x(t) - p\| < \infty \quad (5.20)$$

$$\sup_{t>0} \|\dot{x}(t)\| < \infty \quad (5.21)$$

De plus comme  $\forall t > 0, \int_0^t g(s)ds < \int_0^t \left(a\beta_f - \frac{\rho h + rc}{2}\right) \|X(s)\|^2 + \mu(\|Y(s)\|^2 + \|\dot{x}(s)\|^2)ds$  nous savons que :

$$\int_0^\infty \|X(s)\|^2 ds < \infty \quad (5.22)$$

$$\int_0^\infty \|Y(s)\|^2 ds < \infty \quad (5.23)$$

$$\int_0^\infty \|\dot{x}(s)\|^2 ds < \infty \quad (5.24)$$

En réécrivant, (iDINAM2) comme :

$$\ddot{x}(t) = -\gamma \dot{x}(t) - X(t) - Y(t)$$

Nous obtenons que  $\int_0^\infty \|\ddot{x}(s)\|^2 ds < \infty$  car  $\ddot{x}$  est une combinaison linéaire de fonction  $L^2([0; +\infty[, \mathbb{R}^n)$ . Nous pouvons appliquer le lemme de Barlabat à  $\dot{x}(t)$  car  $\dot{x}, \ddot{x} \in L^2([0; +\infty[, \mathbb{R}^n)$ .

$$\lim_{t \rightarrow +\infty} \dot{x}(t) = 0 \quad (5.25)$$

Comme par hypothèse  $\int_0^\infty \|\ddot{x}(t)\|^2 dt < \infty$ . Nous pouvons de nouveau appliquer le lemme de Barlat à  $\ddot{x}(t)$  car  $\ddot{x}, \ddot{x} \in L^2([0; +\infty[, \mathbb{R}^n)$ .

$$\lim_{t \rightarrow +\infty} \ddot{x}(t) = 0 \quad (5.26)$$

Comme  $\nabla f$  est lipschitzienne en dérivant nous trouvons que :  $\left\| \frac{d}{dt} \nabla f \left( x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right) \right\| \leq L \left\| \dot{x}(t) + \beta_f \ddot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right\|$  Et donc :

$$\begin{aligned} \int_0^\infty \left\| \frac{d}{dt} \nabla f \left( x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right) \right\|^2 dt &\leq L \int_0^\infty \left\| \dot{x}(t) + \beta_f \ddot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right\|^2 dt \leq L(Id + 4\sqrt{Id}) \left( \int_0^\infty \|\dot{x}(t)\|^2 dt \right) \\ &+ L(Id + 4\sqrt{Id}) \left( \beta_f^2 \int_0^\infty \|\ddot{x}(t)\|^2 dt \right) + L(Id + 4\sqrt{Id}) \left( \frac{\beta_f^4}{4} \int_0^\infty \|\ddot{x}(t)\|^2 dt \right) < \infty \end{aligned}$$

Soit  $\frac{d}{dt} \nabla f \in L^2([0; +\infty[, \mathbb{R}^n)$  On peut faire de même pour  $\frac{d}{dt} B$ .

Comme  $\frac{d}{dt} \nabla f(p) = 0$  et  $\frac{d}{dt} B(p) = 0$  on obtient que  $\frac{dX}{dt}(t), \frac{dY}{dt}(t) \in L^2([0; +\infty[, \mathbb{R}^n)$  et donc en appliquant le lemme de Barlat que :

$$\lim_{t \rightarrow +\infty} X(t) = 0 \quad (5.27)$$

$$\lim_{t \rightarrow +\infty} Y(t) = 0 \quad (5.28)$$

Et donc que :

$$\lim_{t \rightarrow +\infty} \left\| B \left( x(t) + \beta_b \dot{x}(t) + \frac{\beta_b^2}{2} \ddot{x}(t) \right) - B(p) \right\| = 0 \quad (5.29)$$

$$\lim_{t \rightarrow +\infty} \left\| \nabla f \left( x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) \right) - \nabla f(p) \right\| = 0 \quad (5.30)$$

En utilisant  $\lim_{t \rightarrow +\infty} \dot{x}(t) = 0$ ,  $\lim_{t \rightarrow +\infty} \ddot{x}(t) = 0$  et la continuité de B et  $\nabla f$  :

$$\lim_{t \rightarrow +\infty} \|B(x(t)) - B(p)\| = 0 \quad (5.31)$$

$$\lim_{t \rightarrow +\infty} \|\nabla f(x(t)) - \nabla f(p)\| = 0 \quad (5.32)$$

Prenons maintenant la fonction  $q_p(t) = \frac{1}{2} \|x(t) - p\|^2$ .  
En dérivant  $q_p(t)$  :

$$\dot{q}_p(t) = \langle \dot{x}(t), x(t) - p \rangle, \quad \ddot{q}_p(t) = \|\dot{x}(t)\|^2 + \langle \ddot{x}(t), x(t) - p \rangle \quad (5.33)$$

$$\ddot{q}_p(t) + \gamma \dot{q}_p(t) = \|\dot{x}(t)\|^2 + \langle \ddot{x}(t) + \gamma \dot{x}(t), x(t) - p \rangle = \|\dot{x}(t)\|^2 - \langle X(t) + Y(t), x(t) - p \rangle \quad (5.34)$$

De plus comme  $\nabla f$  et B sont monotones :

$$\left\langle X(t) + Y(t), x(t) + \beta_f \dot{x}(t) + \frac{\beta_f^2}{2} \ddot{x}(t) - p \right\rangle \geq 0 \quad (5.35)$$

$$\langle X(t) + Y(t), x(t) - p \rangle \geq -\beta_f \langle X(t), \dot{x}(t) \rangle - \beta_b \langle Y(t), \dot{x}(t) \rangle - \frac{\beta_f^2}{2} \langle X(t), \ddot{x}(t) \rangle - \frac{\beta_b^2}{2} \langle Y(t), \ddot{x}(t) \rangle \quad (5.36)$$



Et donc :  $\ddot{q}_p(t) + \gamma \dot{q}_p(t) \leq \|\dot{x}(t)\|^2 + \beta_f \langle X(t), \dot{x}(t) \rangle + \beta_b \langle Y(t), \dot{x}(t) \rangle + \frac{\beta_f^2}{2} \langle X(t), \ddot{x}(t) \rangle + \frac{\beta_b^2}{2} \langle Y(t), \ddot{x}(t) \rangle$   
 On en déduit que

$$\ddot{q}_p(t) + \gamma \dot{q}_p(t) \leq g(t) \quad (5.37)$$

De plus selon l'inégalité de Cauchy-Swartz :

$$\langle X(t), \dot{x}(t) \rangle \leq \|X(t)\| \|\dot{x}(t)\|$$

Soit :

$$\int_0^\infty \langle X(t), \dot{x}(t) \rangle dt \leq \int_0^\infty \|X(t)\| \|\dot{x}(t)\| dt \leq \frac{1}{2} \int_0^\infty \|X(t)\|^2 dt + \frac{1}{2} \int_0^\infty \|\dot{x}(t)\|^2 dt \leq \infty$$

Donc  $\langle X(t), \dot{x}(t) \rangle \in L^1([0; +\infty[, \mathbb{R}^n)$ . On peut faire de même avec les autres termes pour en déduire que  $g(t) \in L^1([0; +\infty[, \mathbb{R}^n)$ .

De plus comme  $0 \leq q_p(t)$  et  $\ddot{q}_p(t) + \gamma \dot{q}_p(t) \in L^1([0; +\infty[, \mathbb{R}^n)$  nous savons que  $\lim_{t \rightarrow \infty} q_p(t)$  existe.

Nous pouvons en déduire que  $x(t)$  converge faiblement. Prenons une suite  $\{t_n\}$  tel que  $t_n \rightarrow \infty$ ,  $x(t_n) \rightarrow x^*$ . En utilisant la propriété d'Opial avec (5.31), nous savons que  $\nabla f(x^*) = \nabla f(p)$  et que  $B(x^*) = B(p)$  donc  $B(x^*) + \nabla f(x^*) = 0$ . Ce qui prouve le théorème 3.

## 5.4 Preuve du théorème 4

Preuve :

Soit  $p \in (B + \nabla f)^{-1}(0)$ . Considérons la suite suivante :

$$E_k = \frac{1}{2} \|x_k - p + \frac{\beta_f}{h}(x_k - x_{k-1}) + \frac{\beta_f^2}{2h^2}(x_k - 2x_{k-1} + x_{k-2})\|^2 + \frac{\delta}{2} \|x_k - p\|^2 \quad (5.38)$$

On note  $v_k = x_k - p + \frac{\beta_f}{h}(x_k - x_{k-1}) + \frac{\beta_f^2}{2h^2}(x_k - 2x_{k-1} + x_{k-2})$ . On a :

$$v_{k+1} - v_k = x_{k+1} - x_k + \frac{\beta_f}{h}(x_{k+1} - 2x_k + x_{k-1}) + \frac{\beta_f^2}{2h^2}(x_{k+1} - 3x_k + 3x_{k-1} - x_{k-2}) \quad (5.39)$$

Or :

$$\frac{\beta_f}{h}(x_{k+1} - 2x_k + x_{k-1}) = -\gamma \beta_f (x_{k+1} - x_k) - h \beta_f B \left( x_{k+1} + \frac{\beta_b}{h}(x_{k+1} - x_k) + \frac{\beta_b^2}{2h^2}(x_{k+1} - 2x_k + x_{k-1}) \right) \quad (5.40)$$

$$-h \beta_f \nabla f \left( x_k + \frac{\beta_f}{h}(x_k - x_{k-1}) + \frac{\beta_f^2}{2h^2}(x_k - 2x_{k-1} + x_{k-2}) \right) \quad (5.41)$$

En posant  $y_k = x_k + \frac{\beta_f}{h}(x_k - x_{k-1}) + \frac{\beta_f^2}{2h^2}(x_k - 2x_{k-1} + x_{k-2})$  et  $z_k = x_{k+1} + \frac{\beta_b}{h}(x_{k+1} - x_k) + \frac{\beta_b^2}{2h^2}(x_{k+1} - 2x_k + x_{k-1})$  et  $w_k = x_{k+1} - 3x_k + 3x_{k-1} - x_{k-2}$ , nous obtenons :

$$v_{k+1} - v_k = (1 - \gamma \beta_f)(x_{k+1} - x_k) - \beta_f h B(z_k) - \beta_f h \nabla f(y_k) + \frac{\beta_f^2}{2h^2} w_k \quad (5.42)$$

Nous avons également :

$$\begin{aligned} \frac{1}{2} \|v_{k+1} - v_k\|^2 &= \frac{1}{2} (1 - \gamma\beta_f)^2 \|x_{k+1} - x_k\|^2 + \frac{h^2\beta_f^2}{2} \|B(z_k) + \nabla f(y_k)\|^2 + \frac{\beta_f^4}{8h^4} \|w_k\|^2 \\ &\quad - h\beta_f (1 - \gamma\beta_f) \langle x_{k+1} - x_k, B(z_k) + \nabla f(y_k) \rangle + \frac{\beta_f^2}{2h^2} (1 - \gamma\beta_f) \\ &\quad \langle x_{k+1} - x_k, w_k \rangle - \frac{\beta_f^3}{2h} \langle w_k, B(z_k) + \nabla f(y_k) \rangle \end{aligned}$$

Et donc en prenant  $\delta = -(1 - \gamma\beta_f)$  et  $\sigma = \frac{\beta_f}{h} + \frac{\beta_f^2}{2h^2}$ . Pour que  $\delta > 0$ , il faut que  $\gamma\beta_f > 1$

$$\begin{aligned} \frac{1}{2} \|v_{k+1}\|^2 - \frac{1}{2} \|v_k\|^2 &= -\frac{1}{2} \|v_{k+1} - v_k\|^2 + \langle v_{k+1} - v_k, v_{k+1} \rangle \\ &= -\frac{\delta^2 + \delta\sigma}{2} \|x_{k+1} - x_k\|^2 - \frac{h^2\beta_f^2}{2} \|B(z_k) + \nabla f(y_k)\|^2 - \frac{\beta_f^4}{8h^4} \|w_k\|^2 - h\beta_f (\sigma + \delta) \langle x_{k+1} - x_k, B(z_k) + \nabla f(y_k) \rangle \\ &\quad + \frac{\beta_f^2}{2h^2} (\sigma + \delta) \langle x_{k+1} - x_k, w_k \rangle + \frac{\beta_f^3}{2h} \langle w_k, B(z_k) + \nabla f(y_k) \rangle + \langle x_{k+1} - p, -\beta_f h(B(z_k) + \nabla f(y_k)) + \frac{\beta_f^2}{2h^2} w_k - \\ &\quad \delta(x_{k+1} - x_k) \rangle - \frac{\beta_f^2}{2h^2} \langle x_k - x_{k-1}, -\beta_f h(B(z_k) + \nabla f(y_k)) + \frac{\beta_f^2}{2h^2} w_k - \delta(x_{k+1} - x_k) \rangle \end{aligned}$$

En considérant également l'égalité suivante :

$$\frac{1}{2} \|x_{k+1} - p\|^2 - \frac{1}{2} \|x_k - p\|^2 = -\frac{1}{2} \|x_{k+1} - x_k\|^2 + \langle x_{k+1} - x_k, x_{k+1} - p \rangle$$

Et en utilisant les deux nous en déduisons que :

$$\begin{aligned} E_{k+1} - E_k &= -\frac{\delta^2 + \delta\sigma + \delta}{2} \|x_{k+1} - x_k\|^2 - \frac{h^2\beta_f^2}{2} \|B(z_k) + \nabla f(y_k)\|^2 - \frac{\beta_f^4}{8h^4} \|w_k\|^2 - h\beta_f (\sigma + \delta) \\ &\quad \langle x_{k+1} - x_k, B(z_k) + \nabla f(y_k) \rangle + \frac{\beta_f^2}{2h^2} (\sigma + \delta) \langle x_{k+1} - x_k, w_k \rangle + \frac{\beta_f^3}{2h} \langle w_k, B(z_k) + \nabla f(y_k) \rangle + \langle x_{k+1} - p, \\ &\quad -\beta_f h(B(z_k) + \nabla f(y_k)) + \frac{\beta_f^2}{2h^2} w_k \rangle - \frac{\beta_f^2}{2h^2} \langle x_k - x_{k-1}, -\beta_f h(B(z_k) + \nabla f(y_k)) + \frac{\beta_f^2}{2h^2} w_k - \delta(x_{k+1} - x_k) \rangle \end{aligned}$$

Notons  $Y_k = \nabla f(y_k) - \nabla f(p)$  et  $Z_k = B(y_k) - B(p)$  comme  $B(p) + \nabla f(p) = 0$ ,  $B(y_k) + \nabla f(y_k) = Y_k + Z_k$ . En utilisant la  $\lambda$ -cocoercivité de  $B$  :

$$\begin{aligned} &\langle x_{k+1} - p + \frac{\beta_f}{h}(x_{k+1} - x_k) + \frac{\beta_f^2}{2h^2}(x_{k+1} - 2x_k + x_{k-1}), Z_k \rangle \leq \\ &\lambda \|Z_k\|^2 + \frac{\beta_f - \beta_b}{h} \langle x_{k+1} - x_k, Z_k \rangle + \frac{\beta_f^2 - \beta_b^2}{2h^2} \langle x_{k+1} - 2x_k + x_{k-1}, Z_k \rangle \\ &= \lambda \|Z_k\|^2 + \frac{\beta_f - \beta_b}{h} \langle x_{k+1} - x_k, Z_k \rangle + \frac{\beta_f^2 - \beta_b^2}{2h^2} \langle -(Z_k + Y_k) - \frac{x_{k+1} - x_k}{h}, Z_k \rangle \\ &= \left( \lambda + \frac{\beta_b^2 - \beta_f^2}{2} \right) \|Z_k\|^2 + \frac{1}{h} \left( \beta_f - \beta_b + \frac{\beta_b^2 - \beta_f^2}{2} \right) \langle x_{k+1} - x_k, Z_k \rangle + \frac{\beta_b^2 - \beta_f^2}{2} \langle Y_k, Z_k \rangle \end{aligned}$$

Et en utilisant la  $\frac{1}{L}$ -cocoercivité de  $\nabla f$  :

$$\begin{aligned} & \langle x_{k+1} - p + \frac{\beta_f}{h}(xk + 1 - x_k) + \frac{\beta_f^2}{2h^2}(x_{k+1} - 2x_k + xk - 1), Y_k \rangle \\ &= \langle y_k - p + x_{k+1} - x_k + \frac{\beta_f}{h}(x_{k+1} - 2x_k + xk - 1) + \frac{\beta_f^2}{2h^2}(x_{k+1} - 3x_k + 3x_{k-1} - x_{k-2}), Y_k \rangle \\ &\leq (L - h\beta_f)\|Y_k\|^2 - \delta\langle x_{k+1} - x_k, Y_k \rangle - h\beta_f\langle Z_k, Y_k \rangle + \frac{\beta_f^2}{2}\langle w_k, Y_k \rangle \end{aligned}$$

En injectant ses deux dernières inéquations nous obtenons :

$$\begin{aligned} E_{k+1} - E_k &\leq -\frac{\delta^2 + \delta\sigma + \delta}{2}\|x_{k+1} - x_k\|^2 - \frac{h^2\beta_f^2}{2}\|Z_k + Y_k\|^2 - \frac{\beta_f^4}{8h^4}\|w_k\|^2 \\ &+ \frac{\beta_f^2}{2h^2}(\sigma + \delta)\langle x_{k+1} - x_k, w_k \rangle + \frac{\beta_f^3}{2h}\langle w_k, Z_k + Y_k \rangle + \langle x_{k+1} - p, \frac{\beta_f^2}{2h^2}w_k \rangle \\ &- \frac{\beta_f^2}{2h^2}\langle x_k - x_{k-1}, \frac{\beta_f^2}{2h^2}w_k - \delta(x_{k+1} - x_k) \rangle - \left( \lambda + \frac{\beta_b^2 - \beta_f^2}{2} \right) \|Z_k\|^2 \\ &- (L - h\beta_f)\|Y_k\|^2 - \left( \frac{\beta_f - \beta_b}{2h}(2 - \beta_b - \beta_f) + h\beta_f\delta \right) \langle x_{k+1} - x_k, Z_k \rangle \\ &+ \delta(1 - h\beta_f)\langle x_{k+1} - x_k, Y_k \rangle + \left( h\beta_f - \frac{\beta_b^2 - \beta_f^2}{2} \right) \langle Z_k, Y_k \rangle - \frac{\beta_f^2}{2}\langle w_k, Y_k \rangle \end{aligned}$$

Que l'on peut réécrire :

$$\begin{aligned} E_{k+1} - E_k &\leq -\frac{\delta^2 + \delta\sigma + \delta}{2}\|x_{k+1} - x_k\|^2 - \frac{h^2\beta_f^2}{2}\|Z_k\|^2 - \frac{h^2\beta_f^2}{2}\|Y_k\|^2 - \frac{\beta_f^4}{8h^4}\|w_k\|^2 \\ &+ \frac{\beta_f^2}{2h^2}(\sigma + \delta)\langle x_{k+1} - x_k, w_k \rangle + \frac{\beta_f^3}{2h}\langle w_k, Z_k \rangle + \langle x_{k+1} - p, \frac{\beta_f^2}{2h^2}w_k \rangle \\ &- \frac{\beta_f^2}{2h^2}\langle x_k - x_{k-1}, \frac{\beta_f^2}{2h^2}w_k - \delta(x_{k+1} - x_k) \rangle - \left( \lambda + \frac{\beta_b^2 - \beta_f^2}{2} \right) \|Z_k\|^2 \\ &- (L - h\beta_f)\|Y_k\|^2 - \left( \frac{\beta_f - \beta_b}{2h}(2 - \beta_b - \beta_f) + h\beta_f\delta \right) \langle x_{k+1} - x_k, Z_k \rangle \\ &+ \delta(1 - h\beta_f)\langle x_{k+1} - x_k, Y_k \rangle + \left( h\beta_f - \frac{\beta_b^2 - \beta_f^2}{2} - h^2\beta_f^2 \right) \langle Z_k, Y_k \rangle + \left( \frac{\beta_f^3}{2h} - \frac{\beta_f^2}{2} \right) \langle w_k, Y_k \rangle \end{aligned}$$

De plus nous avons pour tout  $r > 0$

$$-\frac{\beta_f^2}{2h^2}\langle x_k - x_{k-1}, \frac{\beta_f^2}{2h^2}w_k - \delta(x_{k+1} - x_k) \rangle \leq \frac{\beta_f^2 r}{4h^2}\|\frac{\beta_f^2}{2h^2}w_k - \delta(x_{k+1} - x_k)\|^2 + \frac{\beta_f^2}{4h^2 r}\|x_k - x_{k-1}\|^2 \quad (5.43)$$

$$= \frac{\beta_f^2}{4h^2 r}\|x_k - x_{k-1}\|^2 + \frac{\beta_f^2 \delta^2 r}{4h^2}\|x_{k+1} - x_k\|^2 + \frac{\beta_f^6 r}{16h^6}\|w_k\|^2 - \frac{\beta_f^4 r \delta}{4h^4}\langle x_{k+1} - x_k, w_k \rangle \quad (5.44)$$

En notant :

$$a = \frac{\delta^2 + \delta\sigma + \delta}{2} - \frac{\beta_f^2 \delta^2 r}{4h^2} - \frac{\beta_f^2}{4h^2 r} \quad (5.45)$$

$$b = \frac{\beta_f^4}{8h^4} - \frac{\beta_f^6 r}{16h^6} \quad (5.46)$$

$$c = \frac{\beta_f^2}{2h^2}(\sigma + \delta) - \frac{\beta_f^4 r \delta}{4h^4} \quad (5.47)$$

$$d = \frac{\beta_f - \beta_b}{2h}(2 - \beta_b - \beta_f) + h\beta_f \delta \quad (5.48)$$

$$e = h\beta_f - \frac{\beta_b^2 - \beta_f^2}{2} - h^2 \beta_f^2 \quad (5.49)$$

$$g = \frac{\beta_f^3}{2h} - \frac{\beta_f^2}{2} \quad (5.50)$$

$$i = \lambda + \frac{\beta_b^2 + (h^2 - 1)\beta_f^2}{2} \quad (5.51)$$

$$j = L - h\beta_f + \frac{h^2 \beta_f^2}{2} \quad (5.52)$$

Nous obtenons :

$$\begin{aligned} E_{k+1} - E_k &\leq -a\|x_{k+1} - x_k\|^2 - i\|Z_k\|^2 - j\|Y_k\|^2 - b\|w_k\|^2 + c\langle x_{k+1} - x_k, w_k \rangle \\ &+ \frac{\beta_f^3}{2h}\langle w_k, Z_k \rangle - d\langle x_{k+1} - x_k, Z_k \rangle + \delta(1 - h\beta_f)\langle x_{k+1} - x_k, Y_k \rangle + e\langle Z_k, Y_k \rangle + g\langle w_k, Y_k \rangle + \\ &\frac{\beta_f^2}{2h^2}\langle x_{k+1} - p, w_k \rangle + \frac{\beta_f^2}{4h^2 r}(\|x_k - x_{k-1}\|^2 - \|x_{k+1} - x_k\|^2) \end{aligned}$$

Nous pouvons maintenant ré-exprimer cette équation en utilisant des formes quadratiques :

$$\begin{aligned} E_{k+1} - E_k &+ q_1(x_{k+1} - x_k, w_k) + q_2(x_{k+1} - x_k, Z_k) + q_3(Z_k, w_k) + q_4(x_{k+1} - x_k, Y_k) + q_5(Z_k, Y_k) + q_6(Y_k, w_k) \\ &\leq \frac{\beta_f^2}{2h^2}\langle x_{k+1} - p, w_k \rangle + \frac{\beta_f^2}{4h^2 r}(\|x_k - x_{k-1}\|^2 - \|x_{k+1} - x_k\|^2) \end{aligned}$$

Tel que :

$$q_1(x_{k+1} - x_k, w_k) = \alpha_1 a\|x_{k+1} - x_k\|^2 - c\langle x_{k+1} - x_k, w_k \rangle + \chi_1 b\|w_k\|^2 \quad (5.53)$$

$$q_2(x_{k+1} - x_k, Z_k) = \alpha_2 a\|x_{k+1} - x_k\|^2 + d\langle x_{k+1} - x_k, Z_k \rangle + \iota_1 i\|Z_k\|^2 \quad (5.54)$$

$$q_3(w_k, Z_k) = \chi_2 b\|w_k\|^2 - \frac{\beta_f^3}{2h}\langle w_k, Z_k \rangle + \iota_2 i\|Z_k\|^2 \quad (5.55)$$

$$q_4(x_{k+1} - x_k, Y_k) = \alpha_3 a\|x_{k+1} - x_k\|^2 - \delta(1 - h\beta_f)\langle x_{k+1} - x_k, Y_k \rangle + \eta_1 j\|Y_k\|^2 \quad (5.56)$$

$$q_5(Z_k, Y_k) = \iota_3 i\|Z_k\|^2 - e\langle Z_k, Y_k \rangle + \eta_2 j\|Y_k\|^2 \quad (5.57)$$

$$q_6(w_k, Y_k) = \chi_3 b\|w_k\|^2 - g\langle w_k, Y_k \rangle + \eta_3 j\|Y_k\|^2 \quad (5.58)$$

$$\forall i \in \{1, 2, 3\}, \alpha_i \geq 0, \chi_i \geq 0, \iota_i \geq 0, \eta_i \geq 0 \quad (5.59)$$

$$\alpha_1 + \alpha_2 + \alpha_3 = 1 \quad (5.60)$$

$$\chi_1 + \chi_2 + \chi_3 = 1 \quad (5.61)$$

$$\iota_1 + \iota_2 + \iota_3 = 1 \quad (5.62)$$

$$\eta_1 + \eta_2 + \eta_3 = 1 \quad (5.63)$$

Nous allons ajouter des conditions sur les paramètres afin que les formes quadratiques soient toutes définies positives. Pour cela il faut déjà que :

$$\begin{aligned} i > 0 &\Leftrightarrow \lambda > \frac{(1-h^2)\beta_f^2 - \beta_b^2}{2} \\ j > 0 &\Leftrightarrow L > h\beta_f - \frac{h^2\beta_f^2}{2} \\ a > 0 &\Leftrightarrow -\frac{\beta_f^2\delta^2r^2}{4h^2} + \frac{\delta^2 + \delta\sigma + \delta}{2}r - \frac{\beta_f^2}{4h^2} > 0 \end{aligned}$$

Pour cela il faut que ce dernier polynôme admette des racines. Son discriminant est  $\Delta_1 = \frac{\delta^2}{4} \left( \delta^2 + 2\delta(\sigma+1) + (\sigma+1)^2 - \frac{\beta_f^4}{h^4} \right)$ . Pour que celui ci soit positif il faut que  $\delta^2 + 2\delta(\sigma+1) + (\sigma+1)^2 - \frac{\beta_f^4}{h^4} > 0$  ce qui est toujours le cas car en calculant son discriminant on trouve  $\Delta_2 = \frac{4\beta_f^2}{h^4} > 0$ . On trouve donc une condition sur  $r$  :

$$r \in \left[ \frac{2h^2}{\beta_f^2\delta^2} \left( \delta + \sigma + 1 - \frac{\sqrt{\Delta_1}}{2} \right); \frac{2h^2}{\beta_f^2\delta^2} \left( \delta + \sigma + 1 + \frac{\sqrt{\Delta_1}}{2} \right) \right] \quad (5.64)$$

Pour que les formes quadratiques soient définies positives il nous faut encore les six inégalités suivantes :

$$4\alpha_1\chi_1ab - c^2 > 0 \quad (5.65)$$

$$4\alpha_2\iota_1ai - d^2 > 0 \quad (5.66)$$

$$4\chi_2\iota_2bi - \frac{\beta_f^6}{4h^2} > 0 \quad (5.67)$$

$$4\alpha_3\eta_1aj - \delta^2(1 - h\beta_f)^2 > 0 \quad (5.68)$$

$$4\iota_3\eta_2ij - e^2 > 0 \quad (5.69)$$

$$4\chi_3\eta_3bj - g^2 > 0 \quad (5.70)$$

Si celles-ci sont satisfaites alors, il existe  $\mu_1, \mu_2, \mu_3, \mu_4 > 0$  tel que :

$$\begin{aligned} E_{k+1} - E_k + \mu_1\|x_{k+1} - x_k\|^2 + \mu_2\|Z_k\|^2 + \mu_3\|Y_k\|^2 + \mu_4\|w_k\|^2 \\ \leq \frac{\beta_f^2}{2h^2}\langle x_{k+1} - p, w_k \rangle + \frac{\beta_f^2}{4h^2r}(\|x_k - x_{k-1}\|^2 - \|x_{k+1} - x_k\|^2) \end{aligned}$$

Nous savons que  $\forall n > 1, \sum_{k=1}^n \|x_k - x_{k-1}\|^2 - \|x_{k+1} - x_k\|^2 = \|x_1 - x_0\|^2 - \|x_{n+1} - x_n\|^2 < +\infty$  car  $\|x_1 - x_0\|^2 < +\infty$ .

Donc en sommant par hypothèse  $\sum_{k=2}^{\infty} \langle x_{k+1} - p, w_k \rangle < \infty$ , nous en déduisons que  $E_k$  est bornée. Et donc :

$$\sup_{k \geq 0} \left\| x_k - p + \frac{\beta_f}{h}(x_k - x_{k-1}) + \frac{\beta_f^2}{2h^2}(x_k - 2x_{k-1} + x_{k-2}) \right\| < \infty \quad (5.71)$$

$$\sup_{k \geq 0} \|x_k - p\| < \infty \quad (5.72)$$

De plus en sommant nous en déduisons les inégalités suivantes :

$$\sum_{k=0}^{\infty} \|x_{k+1} - x_k\|^2 < \infty \quad \sum_{k=0}^{\infty} \|w_k\|^2 < \infty$$

$$\sum_{k=0}^{\infty} \|\nabla f(y_k) - \nabla f(p)\|^2 < \infty \quad \sum_{k=0}^{\infty} \|B(z_k) - B(p)\|^2 < \infty$$

En utilisant la L-lipschitzianité de  $\nabla f$  et l'inégalité triangulaire :

$$\begin{aligned} \|\nabla f(x_k) - \nabla f(p)\|^2 &\leq (\|\nabla f(y_k) - \nabla f(p)\| + \|\nabla f(x_k) - \nabla f(y_k)\|)^2 \leq 2\|\nabla f(y_k) - \nabla f(p)\|^2 + 2L\|y_k - x_k\|^2 \\ &\leq 2\|\nabla f(y_k) - \nabla f(p)\|^2 + 4L \left( \frac{\beta_f}{h} + \frac{\beta_f^2}{2h^2} \right)^2 \|x_k - x_{k-1}\|^2 + \frac{L\beta_f^4}{h^4} \|x_{k-1} - x_{k-2}\|^2 \end{aligned}$$

En sommant on en déduit que  $\sum_{k=0}^{\infty} \|\nabla f(x_k) - \nabla f(p)\|^2 < \infty$ . Le même raisonnement nous permet de déduire  $\sum_{k=0}^{\infty} \|B(x_k) - B(p)\|^2 < \infty$ .

Donc nous en déduisons que :

$$\lim_{k \rightarrow \infty} \|B(x_k) - B(p)\| = 0 \quad \lim_{k \rightarrow \infty} \|\nabla f(x_k) - \nabla f(p)\| = 0 \quad \lim_{k \rightarrow \infty} \|x_{k+1} - x_k\| = 0$$

En utilisant (5.72) et la limite précédente nous en déduisons que  $x_k$  converge faiblement. De plus en utilisant les limites précédentes  $B(x^*) = B(p)$  et  $\nabla f(x^*) = \nabla f(p)$  soit  $0 = B(x^*) + \nabla f(x^*)$ . en utilisant le lemme d'Opial la démonstration est finie.

## 5.5 Etude numérique

Dans cette partie nous allons analyser les performances en nombre d'itérations de iDINAM2-split en fonction de différents paramètres afin d'analyser leurs influences. Puis nous comparerons iDINAM2-split avec iDINAM-split.

### Analyse de iDINAM2-split

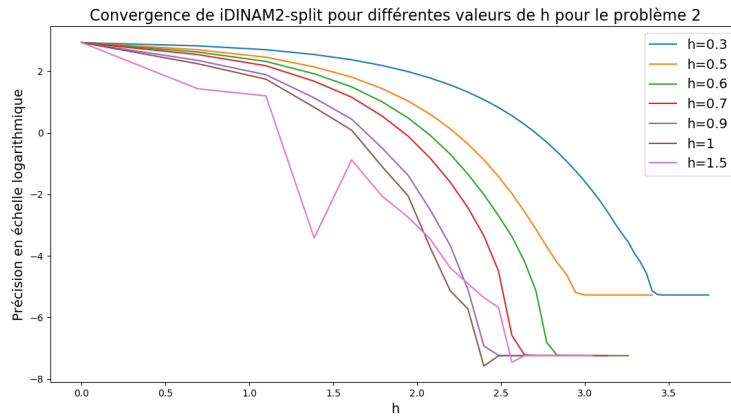
Nous allons considérer les deux problèmes suivants pour les analyses numériques (n=5 dans les deux cas).

$$B(x) = \partial\|x\|_1, \quad f(x) = \sum_{i=0}^n x_i \ln(x_i) + 1 - x_i, \quad x^{(0)} = x^{(1)} = x^{(2)} = (1, 1, \dots, 1) \quad (\text{PB1})$$

$$B(x) = (1 + \ln(x_1), \dots, 1 + \ln(x_n)), \quad f(x) = \|x\|_2^2, \quad x^{(0)} = x^{(1)} = x^{(2)} = (1, 1, \dots, 1) \quad (\text{PB2})$$

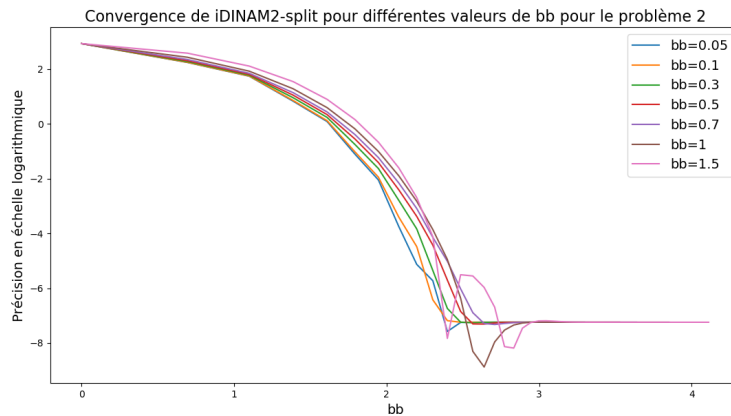
Par la suite, nous allons appeler trajectoires les les normes de  $\nabla f(x_k) + B(x_k)$  en fonction de l'itération k.

Nous pouvons observer sur la figure 5.1 que pour le problème 2, augmenter le pas permet de gagner en vitesse jusqu'à un certain pas ou la vitesse de convergence diminue, puis si le pas augmente encore l'algorithme diverge. De plus augmenter le pas à pour effet de créer de l'instabilité

FIGURE 5.1 – Trajectoires de iDINAM2-split pour différentes valeurs de  $h$  pour le problème 2

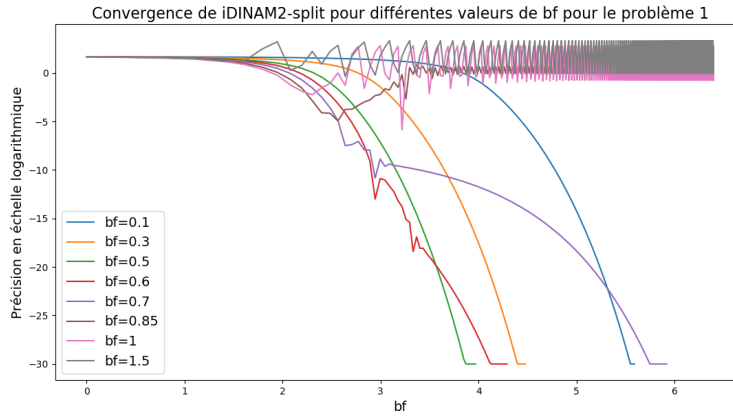
dans la trajectoire (rebond sur le graphique).

Sur le problème 2, augmenter le pas permet d'améliorer la précision cela est dû au problème qui est que si le pas est trop petit, la trajectoire se stabilise trop tôt.

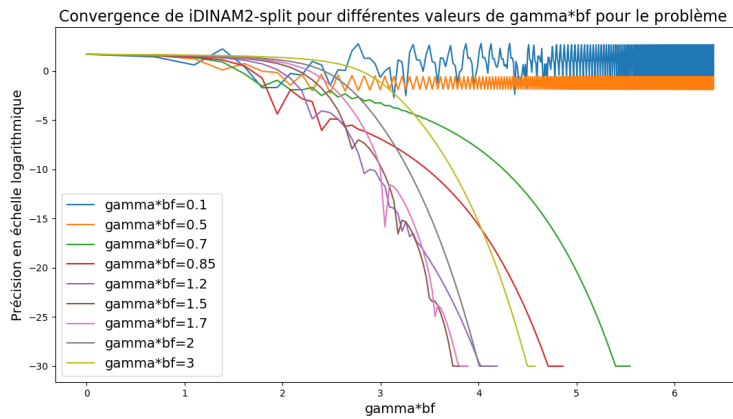
FIGURE 5.2 – Trajectoires de iDINAM2-split pour différentes valeurs de  $\beta_b$  pour le problème 2

Nous pouvons observer sur la figure 5.2 que changer  $\beta_b$  impacte légèrement la vitesse de convergence et la stabilité. Une petite valeur de  $\beta_b$  permet de converger plus rapidement il semblerait. Sur le problème 1, les trajectoires pour tout les  $\beta_b < 1.5$  sont superposées ce qui nous conforte dans l'idée de prendre une petite valeur de  $\beta_b$ .

Nous pouvons observer sur la figure 5.3 que des petites valeurs de  $\beta_f$  entraînent des trajectoires très stables et lentes à converger alors que pour de grandes valeurs l'algorithme est très instable et converge rapidement (si il converge). Il faut donc trouver un compromis entre stabilité


 FIGURE 5.3 – Trajectoires de iDINAM2-split pour différentes valeurs de  $\beta_f$  pour le problème 1

et vitesse pour  $\beta_f$ . Pour les deux problèmes, il semble se trouver entre 0.5 et 0.6.


 FIGURE 5.4 – Trajectoires de iDINAM2-split pour différentes valeurs de  $\beta_f \gamma$  pour le problème 1

Nous pouvons observer sur la figure 5.4 que l'algorithme converge pour des  $\beta_f \gamma \in \{0.5, 0.7, 3\}$  pour lesquelles selon le théorème 3 il ne devrait pas converger vers la solution. De plus, plus  $\beta_f \gamma$  est grand plus la trajectoire est stable. Il peut être une bonne idée de choisir les autres paramètres en premier en cherchant à maximiser la vitesse et de choisir ce paramètre à la fin afin de réduire l'instabilité.

De la même façon que pour  $\beta_f$  il faut trouver un compromis avec  $\beta_f \gamma$  entre la stabilité et la vitesse de convergence.

A l'aide des graphiques nous pourrions en déduire des paramètres "optimaux" pour les deux



problèmes. Nous obtenons :

$$PB1 : h = 1, \beta_b = 0.05, \beta_f = 0.5, \gamma\beta_f = 1.5$$

$$PB2 : h = 1, \beta_b = 0.05, \beta_f = 0.6, \gamma\beta_f = 2$$

Ces paramètres sont très similaires pour les deux problèmes. Ils ne sont pas optimaux (voir figure 5.11 les paramètres précédents sont en rouges).

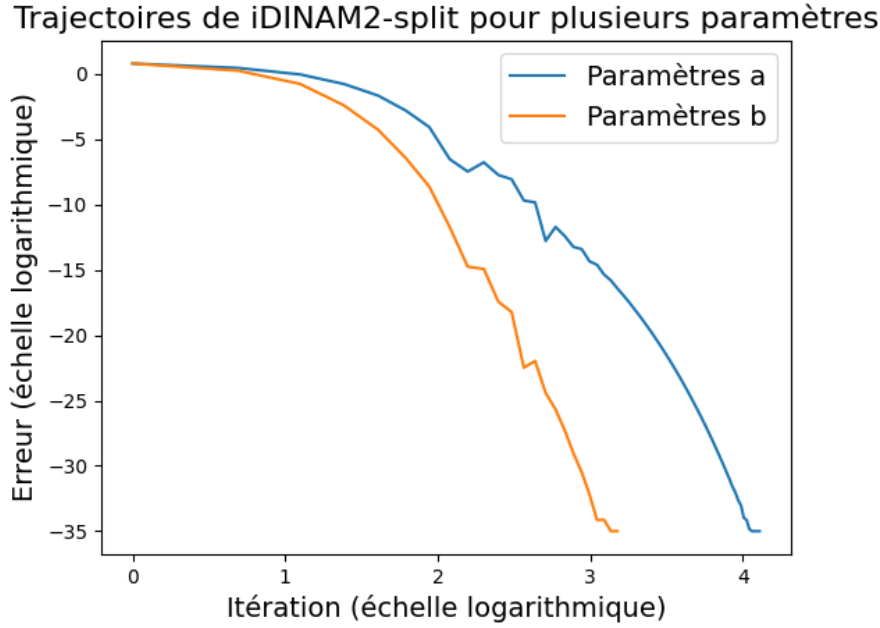


FIGURE 5.5 – Trajectoires de iDINAM2-split pour différents paramètres pour le problème 1

Et de plus si l'on considère le problème suivant :

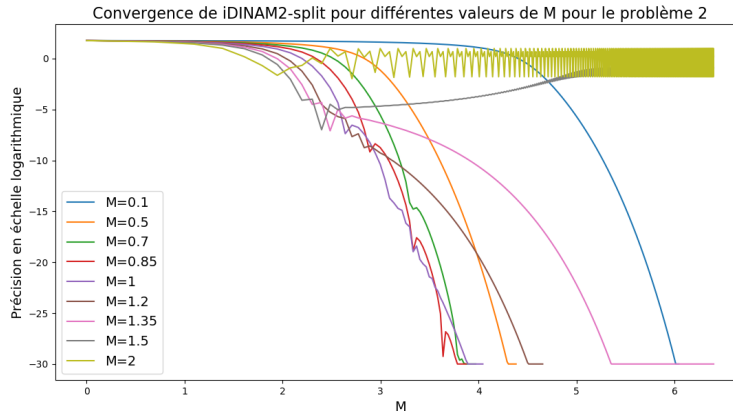
$$B(x) = \|x - 1\|_1, \quad f(x) = \sum_{i=0}^n ix_i + \left( \sum_{i=0}^n x_i \right)^2, \quad x^{(0)} = x^{(1)} = x^{(2)} = (1, 1, \dots, 1) \quad (PB3)$$

Sur le PB3 ces paramètres ne convergent pas. Il faut fixer  $\beta_f \leq 0.2$  pour que l'algorithme converge.

Nous pouvons également ajouter un nouveau paramètre  $M$ , celui sert à remplacer  $\nabla f(x)$  et  $B(x)$  par  $M\nabla f(x)$  et  $MB(x)$ . L'ajout de  $M$  fait que  $MB(x)$  est  $\frac{\lambda}{M}$ -cocoercif et que  $M\nabla f(x)$  est  $LM$ -lipschitzienne, de plus

$$\nabla f(x) + B(x) = 0 \Leftrightarrow M\nabla f(x) + MB(x) = 0$$

Sur 5.6 nous observons que une trop grande valeur de  $M$  empêche la convergence de l'algorithme (présence d'oscillations). Une petite valeur est lente à converger. Les valeurs optimales sont aux alentours de 1. Pour les problèmes 1 et 2 avec les paramètres optimaux, il est possible d'accélérer encore la convergence en prenant  $M = 0.85$ . Le comportement de  $M$  semble relativement similaire à celui de  $\gamma\beta_f$  (en inversant l'effet des petites valeurs et des grandes).


 FIGURE 5.6 – Trajectoires de iDINAM2-split pour différentes valeurs de  $M$  pour le problème 1

Une idée pour choisir les paramètres est de commencer avec de petites valeurs de  $h, \beta_b, \beta_f$  et une grande valeur de  $\beta_f \gamma$  (2 par exemple). Puis d'augmenter progressivement  $h, \beta_b, \beta_f$  afin d'augmenter la vitesse de convergence puis de finalement trouver un bon compromis entre la stabilité et la vitesse en changeant  $\gamma \beta_f$  et  $M$ .

Cet algorithme nous a permis de trouver les paramètres de la courbe bleu de 5.11.

Dans cette analyse nous n'avons pas pu obtenir des paramètres généraux pour tout les problèmes. Nous avons par contre pu établir l'effet des paramètres sur les trajectoires et proposer un algorithme pour trouver des "bons" paramètres.

## Comparaison de iDINAM-split et de iDINAM2-split

Nous allons dans cette partie comparer les vitesses de convergence des algorithmes iDINAM-split et iDINAM2-split.

L'algorithme iDINAM-split est un algorithme obtenu par la discrétisation de (iDINAM). Voici comment sont calculés les itérés.

$$\alpha = 1 + \frac{\beta_f}{h}$$

$$y_k = x_k + \frac{\alpha}{1 + \gamma h} (x_k - x_{k-1}) - \frac{ah^2}{1 + \gamma h} \nabla f \left( x_k + \frac{\beta_f}{h} (x_k - x_{k-1}) \right)$$

$$x_{k+1} = \frac{\alpha - 1}{\alpha} x_k + \frac{1}{\alpha} \left( Id + \frac{ah^2}{1 + \gamma h} B \right)^{-1} (y_k)$$

Pour cela, nous allons d'abord comparer les deux algorithmes pour les mêmes valeurs de paramètres. Puis ensuite pour des paramètres choisis (pour leur vitesse de convergence) indépendamment pour chaque schéma afin de comparer les deux algorithmes si l'on peut choisir les paramètres.

### Mêmes paramètres

Dans cette partie, nous allons afficher des graphes où nous faisons varier un paramètre sur l'axe des abscisses et sur l'axe des ordonnées, nous affichons la première itération  $N$  de l'algorithme tel que  $\forall k > N \|\nabla f(x_k) + B(x_k)\| < 10^{-10}$  ( $5 \times 10^{-6}$  pour le problème 2). Nous observons que pour

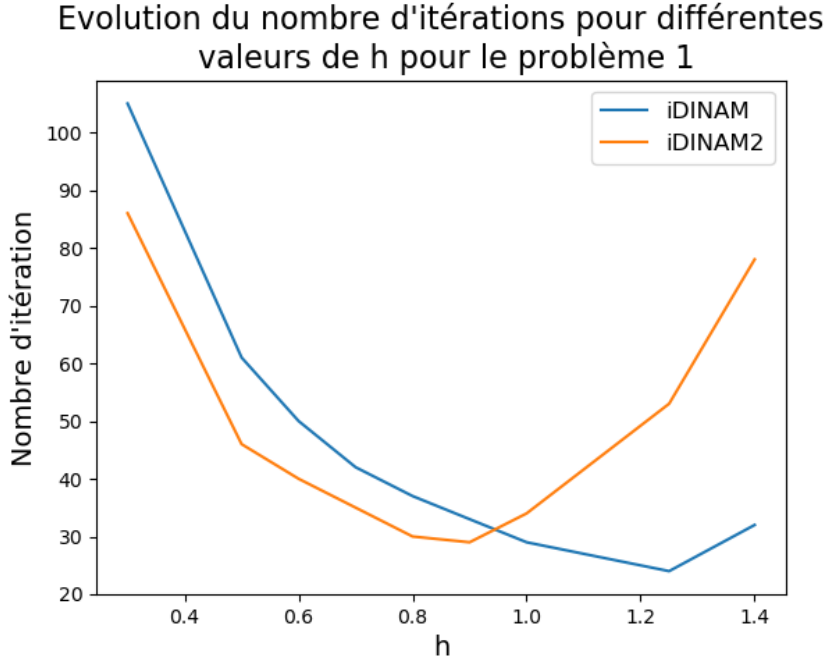


FIGURE 5.7 – Trajectoires de iDINAM2-split et iDINAM2-split pour différents pas pour le problème 1

le problème 1, il faut que le pas soit petit pour qu'iDINAM2-split soit plus rapide qu'iDINAM-split. De plus si le pas devient trop grand alors les deux schémas ne convergeront plus, mais cela arrive plus tôt pour iDINAM2-split ce qui est un point négatif notable.

Pour les deux autres problèmes, les interprétations des courbes sont sensiblement identiques à celles du problème 1.

Pour  $\beta_b$ , pour le problème 1, iDINAM-split est toujours plus rapide qu'iDINAM2-split de quelques itérations alors que pour le problème 2 l'effet inverse est observé. Pour le problème 3, il n'y a pas de phénomène particulier semblant apparaître. L'impact de  $\beta_b$  semble donc dépendre du problème considéré.

Pour  $\beta_f$ , les trois problèmes donnent la même allure de courbe, nous allons donc seulement analyser celle du problème 1. De la même façon que pour  $h$ , iDINAM2-split diverge plus rapidement quand  $\beta_f$  est grand. De plus pour de petites valeurs de  $\beta_f$ , les deux courbes sont superposées et donc le choix de l'algorithme n'importe pas. C'est pour des valeurs médianes de  $\beta_f$  qu'iDINAM2-split est plus rapide, pour les problèmes 1 et 2, il est même plus rapide sur ces valeurs médianes qu'iDINAM-split en son minimum.

Pour  $\gamma\beta_f$ , nous observons que pour iDINAM-split est plus lent quand  $\gamma\beta_f$  augmente alors qu'iDINAM2-split atteint un minimum puis recroît. De plus le minimum d'iDINAM2-split est supérieur au minimum d'iDINAM-split. De même pour de petites valeurs de  $\gamma\beta_f$ , iDINAM2-split

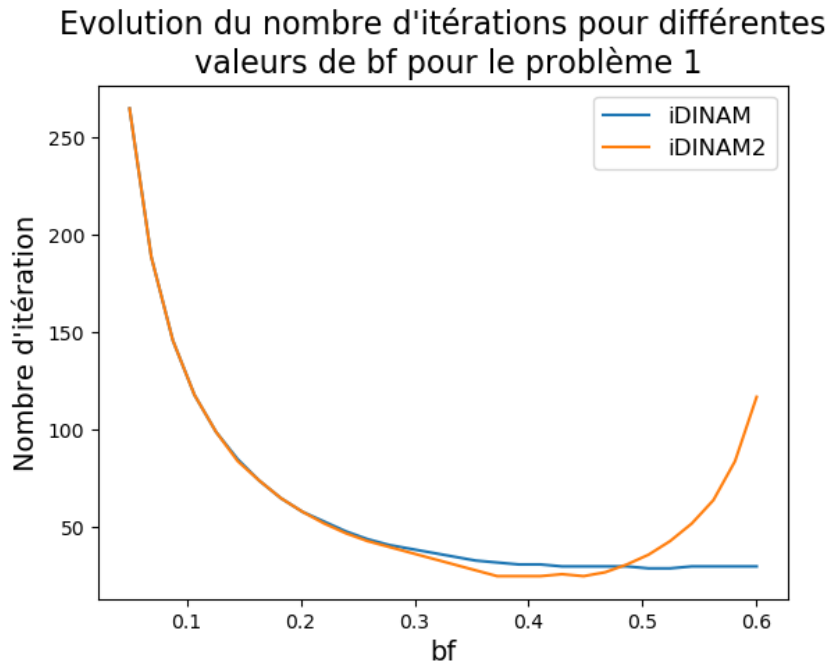


FIGURE 5.8 – Trajectoires de iDINAM2-split et iDINAM2-split pour différents  $\beta_f$  pour le problème 1

ne converge pas contrairement à iDINAM-split.

Pour  $M$ , si la valeur de  $M$  est grande ( $M > 1$ ) alors iDINAM-split est significativement plus rapide qu'iDINAM2-split (qui ne converge plus). Mais pour des valeurs moyennes iDINAM2-split est plus légèrement plus rapide qu'iDINAM-split. Cela est intéressant car pour iDINAM2-split l'on souhaite prendre des valeurs de  $M$  inférieure à 1.

Nous avons pu observer en comparant les deux algorithmes qu'iDINAM-split avait plus de liberté sur le choix des paramètres, et que pour des paramètres quelconques, on ne peut pas savoir à l'avance lequel sera le plus rapide.

### Paramètres différents

Nous avons déterminé des bons paramètres pour iDINAM-split et iDINAM2-split (différents pour les deux algorithmes) pour les trois problèmes. Avec ces paramètres nous pouvons comparer pour chacun de ses problèmes les trajectoires (voir figure 5.11).

Pour les trois problèmes, nous pouvons observer que pour des paramètres bien choisis, iDINAM2-split est toujours plus rapide qu'iDINAM-split. Ce résultat est très encourageant vu que pour des paramètres bien choisis, il vaut mieux prendre iDINAM2-split alors que pour des paramètres quelconques iDINAM-split est plus robuste.

Pour confirmer ce résultat, nous avons comparé des profils de performances. Ceux-ci seront réalisés sur 40 variantes de 5 classes de problèmes soit 200 problèmes. Pour chaque classe de problème, nous avons déterminé pour les deux algorithmes des paramètres permettant une convergence rapide.

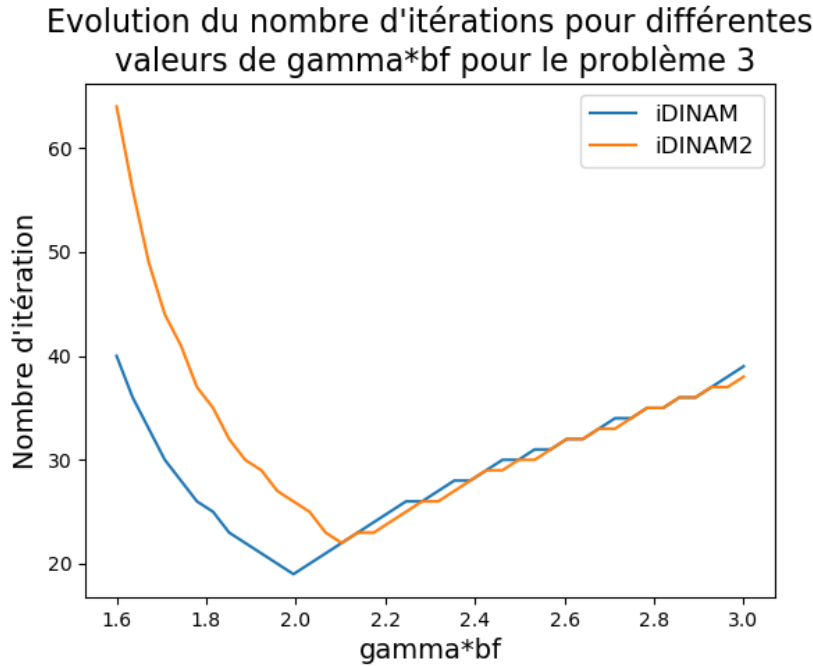


FIGURE 5.9 – Trajectoires de iDINAM2-split et iDINAM2-split pour différents  $\gamma\beta_f$  pour le problème 3

Les profils de performances sont des graphiques permettant de comparer différents algorithmes entre eux pour une heuristique. Nous notons  $E(P, A)$  la valeur de l'heuristique pour le problème  $P$  résolu avec l'algorithme  $A$ . Pour chaque algorithme, sa courbe (selon la variable  $\tau$ ) représente le pourcentage de problème que l'algorithme sur lequel  $E(P, A) < \tau \min_X E(P, X)$ , c'est-à-dire le pourcentage de problème où l'algorithme à une heuristique ayant un rapport maximum de  $\tau$  par rapport au meilleur algorithme (selon cette heuristique). Plus une courbe est à gauche et en haut sur le graphe plus l'algorithme est rapide (par rapport à l'algorithme le plus rapide) en général.

Nous pouvons observer sur la figure 5.12 que pour 75% des problèmes, iDINAM2-split fait moins d'itérations qu'iDINAM-split. Pour les 25% restant, iDINAM2-split fait au maximum 2.1 fois plus d'itérations qu'iDINAM-split. Alors qu'iDINAM-split peut faire jusqu'à 7 fois plus d'itération qu'iDINAM2-split.

Pour générer la figure 5.13, nous avons pris le test utilisant seulement les temps d'exécution supérieurs à 0.15 seconde afin d'éliminer les erreurs dues à la mesure du temps. Nous pouvons observer sur 5.13 qu'iDINAM-split est toujours plus rapide. Mais cette différence est légère (inférieure à 7% dans 99% des cas) et au maximum 70% de temps d'exécution en plus. Cela est en désaccord avec le résultat précédent qui indiquait qu'iDINAM2-split faisait moins d'itérations qu'iDINAM-split et donc devrait être plus rapide. Cela peut être expliqué par la résolution simple de  $(I + \alpha hB)^{-1}$  pour les problèmes étudiés, et comme iDINAM2-split fait plus d'opérations qu'iDINAM-split il serait plus lent malgré qu'il effectue moins d'itérations.

On voit donc avec ce profil de performance qu'iDINAM2-split semble être un meilleur algorithme, car il est en général plus robuste qu'iDINAM-split quand les paramètres sont bien choisis.

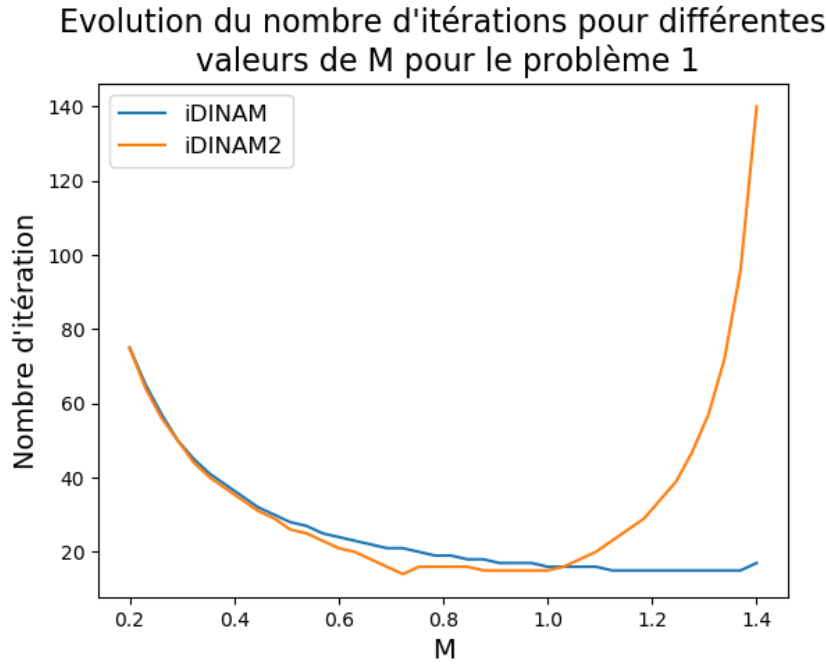


FIGURE 5.10 – Trajectoires de iDINAM2-split et iDINAM2-split pour différents  $M$  pour le problème 1

## iDINAM-split de degré supérieur

Dans cette partie, nous allons chercher à créer un algorithme qui augmente le degré du développement à l'intérieur de  $\nabla f$  et  $B$ . Nous obtenons donc cette équation :

$$x^{(2)}(t) + \gamma x^{(1)}(t) + \nabla f \left( x(t) + \sum_{i=0}^N \frac{\beta_f^i}{i!} x^{(i)}(t) \right) + B \left( x(t) + \sum_{i=0}^N \frac{\beta_b^i}{i!} x^{(i)}(t) \right) = 0 \quad (5.73)$$

Puis en discrétisant 5.73, tel que l'expression soit implicite selon  $B$  et explicite selon  $\nabla f$ . Nous obtenons le schéma suivant :

$$\begin{aligned} \frac{1}{h^2}(x_{k+1} - 2x_k + x_{k-1}) + \frac{\gamma}{h}(x_{k+1} - x_k) + \nabla f \left( x_k + \sum_{n=1}^N \frac{\beta_f^n}{n!h^n} \sum_{i=0}^n (-1)^i x_{k-i} C_n^i \right) \\ + B \left( x_{k+1} + \sum_{n=1}^N \frac{\beta_b^n}{n!h^n} \sum_{i=0}^n (-1)^i x_{k+1-i} C_n^i \right) = 0 \end{aligned}$$

En notant

$$\begin{aligned} q_b^{(i)} &= (-1)^i \frac{\beta_b^i}{i!h^i} \sum_{m=0}^{N-i} \frac{\beta_b^m}{m!h^m} \\ q_f^{(i)} &= (-1)^i \frac{\beta_f^i}{i!h^i} \sum_{m=0}^{N-i} \frac{\beta_f^m}{m!h^m} \\ \alpha &= (1 + \gamma h)^{-1} q_b^{(0)} \end{aligned}$$

Trajectoires de iDINAM2-split pour plusieurs paramètres

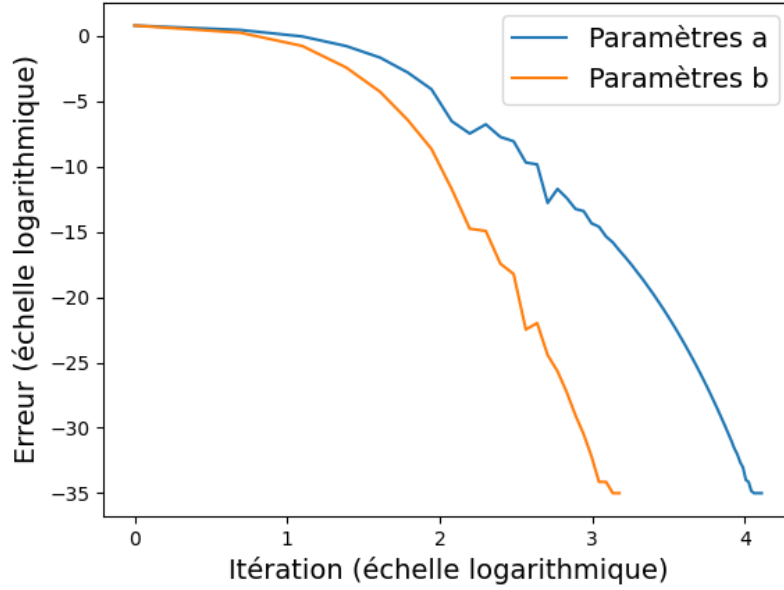


FIGURE 5.11 – Trajectoires de iDINAM2-split et iDINAM-split pour des paramètres "optimaux" pour le problème 3

Nous obtenons :

$$(1 + \gamma h)x_{k+1} + h^2 B \left( \sum_{i=0}^N q_i^{(b)} x_{k+1-i} \right) = (2 + \gamma h)x_k - x_{k-1} - h^2 \nabla f \left( \sum_{i=0}^N q_i^{(f)} x_{k-i} \right)$$

$$(I + \alpha h^2 B) \left( \sum_{i=0}^N q_i^{(b)} x_{k+1-i} \right) = \alpha(2 + \gamma h)x_k - \alpha x_{k-1} + \sum_{i=1}^N q_i^{(b)} x_{k+1-i} - \alpha h^2 \nabla f \left( \sum_{i=0}^N q_i^{(f)} x_{k-i} \right)$$

On en déduit donc l'algorithme suivant :

---

**Algorithm 4** iDINAMN split

---

**Require:**  $x_0, \dots, x_N \in \mathbb{R}^n$ ,  $\epsilon > 0$

$$u_i \leftarrow (-1)^i \frac{\beta_b^i}{i!h^i} \sum_{m=0}^{N-i} \frac{\beta_b^m}{m!h^m}$$

$$v_i \leftarrow (-1)^i \frac{\beta_f^i}{i!h^i} \sum_{m=0}^{N-i} \frac{\beta_f^m}{m!h^m}$$

$$\alpha \leftarrow (1 + \gamma h)^{-1} u_0$$

**while**  $\|x_k - x_{k-1}\| > \epsilon$  **do**

$$z_k \leftarrow \sum_{i=0}^N v_i x_{k-i}$$

$$y_k \leftarrow \alpha(2 + \gamma h)x_k - \alpha x_{k-1} + \sum_{i=1}^N u_i x_{k+1-i} - \alpha h^2 \nabla f(z_k)$$

$$x_{k+1} \leftarrow \frac{1}{u_0} (Id + \alpha h B)^{-1} (y_k) - \sum_{i=1}^N \frac{u_i}{u_0} x_{k+1-i}$$

**end while**

---

Ces algorithmes ont une complexité à chaque opération relativement similaire, car malgré des sommes, en pratique  $N$  va rester petit (voir les analyses suivantes) et la complexité à chaque

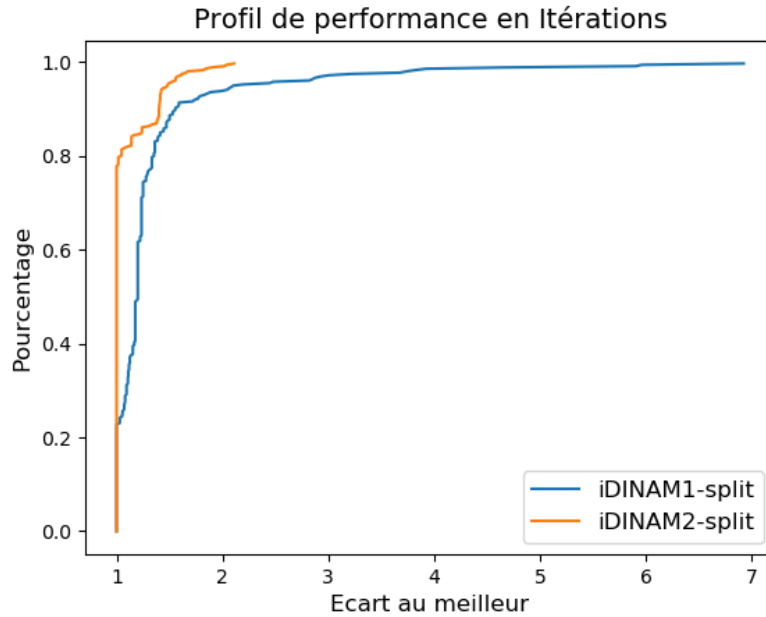


FIGURE 5.12 – Profil de performance du nombre d’itération de iDINAM-split et de iDINAM2-split

opération est majoritairement décidée par celle de  $\nabla f$  et  $(Id + \alpha h B)^{-1}$ .

Afin de comparer ses algorithmes, nous avons réalisé un profil de performance pour plusieurs valeurs de  $N$  :

On observe sur la figure 5.14 directement que pour le nombre d’itération si  $N \neq 2$  alors l’algorithme est moins performant que 2. On préférera donc iDINAM2.

Nous observons dans la figure 5.15 que iDINAM-split est plus performant que les autres algorithmes en temps. Mais cette différence est légère. Elle pourrait être expliquée par le nombre supplémentaire d’étape à effectuer pour les autres algorithmes à chaque itération. Ces résultats dépendent des fonctions choisies en effet pour les tests du profil de performance  $(Id + \alpha h B)^{-1}$  était résolu de façon explicite dans chaque cas et aucun solveur n’a été utilisé pour les résoudre à chaque itération. Si cela avait été le cas, nous n’aurions peut-être pas obtenu les mêmes résultats, car faire un nombre d’itérations plus faible avec plus d’étapes par itérations (dû à un  $N$  plus grand) aurait pu être plus rapide.



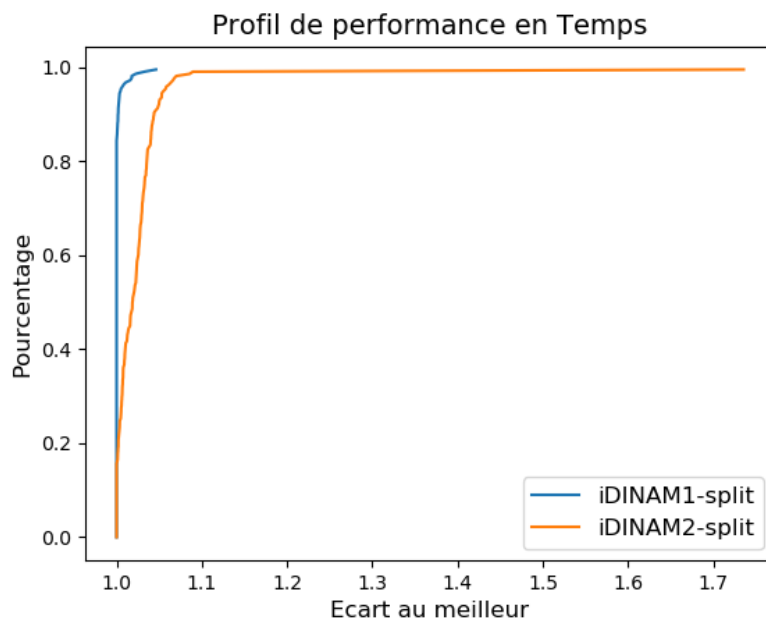


FIGURE 5.13 – Profil de performance du temps d'exécution de iDINAM-split et de iDINAM2-split

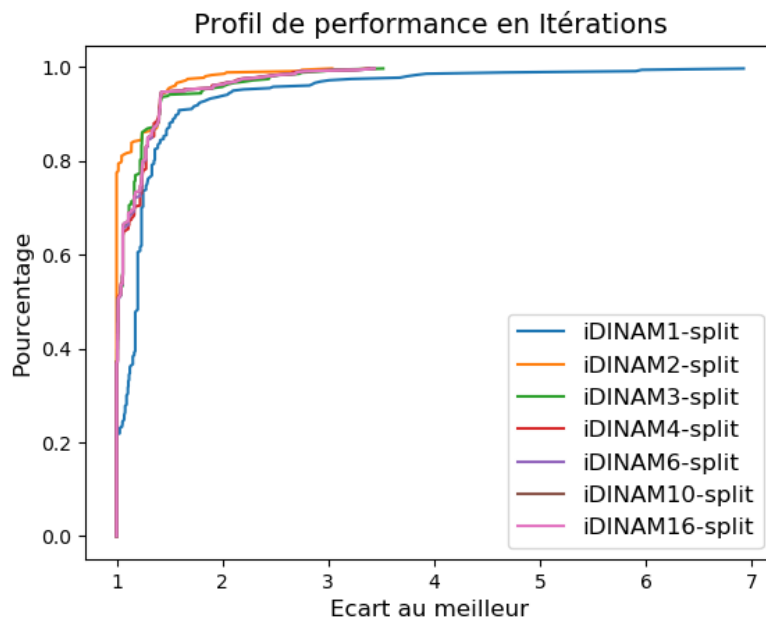


FIGURE 5.14 – Profil de performance du nombre d'itération de iDINAMN-split

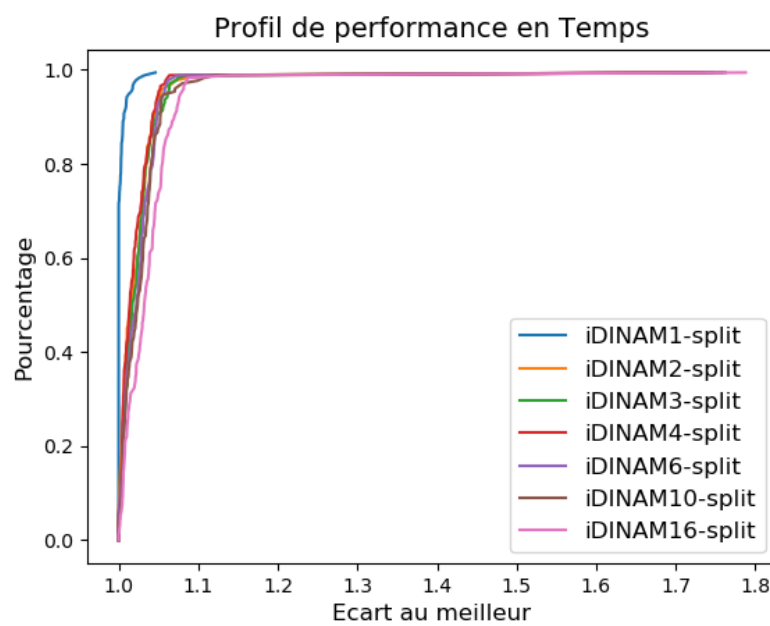


FIGURE 5.15 – Profil de performance du temps d'exécution de iDINAMN-split

# Chapitre 6

## Abstract

Mots clés : Optimisation convexe, algorithme de gradient-proximal, opérateur non-potentiel

Ce rapport de stage présente un nouvel algorithme du premier ordre ayant pour objectif de résoudre le système suivant  $\nabla f(x) + B(x) = 0$  où  $f$  est convexe et  $B$  un opérateur cocoercif. Ce problème peut notamment dériver d'un problème d'optimisation en prenant  $B$  comme le sous-gradient de la fonction à minimiser.

Le nouvel algorithme est basé sur une EDO et le second chapitre propose deux théorèmes démontrant l'existence et la convergence de cette EDO sous de fortes hypothèses. Le troisième chapitre démontre la convergence de l'algorithme sous de fortes hypothèses et effectue ensuite une étude numérique de celui-ci.

Keywords : convex optimization, gradient-proximal algorithm, non-potential operator Convex optimization, gradient-proximal algorithm, non-potential operator

This report presents a new first-order algorithm whose objective is to solve the following system  $\nabla f(x) + B(x) = 0$  where  $f$  is convex and  $B$  is a cocoercive operator. In particular, this problem can be derived from an optimisation problem by taking  $B$  as the subgradient of the function to be minimised. The new algorithm is based on an ODE and the second chapter proposes two theorems proving the existence and convergence of this ODE under strong assumptions. The third chapter demonstrates the convergence of the algorithm under strong assumptions and then carries out a numerical study of it.